Population genetic and phylogenetic insights into the adaptive radiation of Antarctic notothenioid fishes

Inauguraldissertation

zur
Erlangung der Würde eines Doktors der Philosophie
vorgelegt der
Philosophisch-Naturwissenschaftlichen Fakultät
der Universität Basel

von

Michael   **Matschiner**
_____
(Vorname, ausgeschrieben)   (Familienname)

aus **Neumarkt i. d. OPf.**   **Deutschland**
_____
(Heimat)   (Kanton oder Land)

Basel, 2011

Genehmigt von der Philosophisch-Naturwissenschaftlichen Fakultät

auf Antrag von

Ass. Prof. Dr. Walter Salzburger, Dr. Reinhold Hanel, Prof. Thorsten Reusch

(Mitglieder des Dissertationskomitees)

Basel, den _____15. 11. 2011_____

(Datum der Genehmigung durch die Fakultät)

_____Prof. M. Spiess_____

Dekanin/Dekan

(Name der/des amtierenden
Dekanin/Dekans einsetzen)

2

Abstract

Adaptive radiation is the evolution of ecological and phenotypic diversity within a rapidly multiplying lineage, a phenomenon that is considered responsible for a great part of Earth's biodiversity. It occurs as a response to ecological opportunity in the form of competitor-free habitat, extinction of antagonists, or the emergence of a key innovation. One of the most spectacular adaptive radiations in the marine realm is the diversification of notothenioid fishes in the freezing waters of Antarctica. This radiation has led to a unique dominance of the Antarctic marine habitat by notothenioids, and is often assumed to result from the key innovation of freeze resistance. Antifreeze glycoproteins are present in blood and tissue of Antarctic notothenioids and enable them to survive in their sub-zero environment. Notothenioids are further characterized by prolonged pelagic larval stages, that have been suggested to contribute to high levels of inter-population gene flow with oceanic currents, which seems to contradict the high speciation rates observed in the notothenioid adaptive radiation. This doctoral work uses molecular tools to investigate the character of gene flow in notothenioids as well as the origin of their diversification. It is demonstrated that larval dispersal is a common agent of long-distance gene flow in many notothenioid species. The key innovation hypothesis is corroborated by an extensive molecular dating of the divergence events of notothenioids and related acanthomorph fishes. New tools for the analysis of microsatellite markers and for Bayesian divergence date estimation are developed.

Table of Contents

I Introduction

Adaptive radiation is the evolution of ecological and phenotypic diversity within a rapidly multiplying lineage and is commonly claimed responsible for the genesis of a great portion of the diversity of life (Simpson 1953, Schluter 2000). According to Schluter (2000), an adaptive radiation is characterized by rapid speciation, common ancestry, and a phenotype-environment correlation, whereby phenotypes must be beneficial in their respective environments. Adaptive radiation is often considered a consequence of ecological opportunity (Simpson 1953, Schluter 2000) arising through colonization of a new habitat with abundant niche-space, the origin of a key innovation, and/or the extinction of antagonists (Yoder et al. 2010). Prime examples for adaptive radiation include the Darwin's finches of the Galapagos Islands (Grant & Grant 2002, 2011), the Hawaiian *Drosophila* diversification (Kambysellis & Craddock 1997) and the impressive radiations of cichlid fishes in the Great Lakes of East Africa (Salzburger 2009).

Among very few adaptive radiations identified in the marine realm, the most spectacular one is found within the suborder Notothenioidei. Whereas ancestral notothenioid lineages occur in Australia, New Zealand, and South America, the so-called 'Antarctic clade' of notothenioid fishes (including the five highly diverse families Nototheniidae, Harpagiferidae, Artedidraconidae, and Channichthyidae) has radiated in Antarctic waters, and dominates the High Antarctic ichthyofauna in terms of species number (76.6%) and biomass (over 90%) (Eastman 2005).

Antarctic waters are unique marine environments, characterized by sub-zero temperatures and the presence of sea ice. At high latitudes, temperatures constantly remain close to the freezing point of seawater at -1.86 °C (Eastman 1993). Due to the weight of the continental ice cap, the Antarctic shelf is deeper than the world average (Anderson 1999). Many potential shallow water habitats are inaccessible due to ice foots and anchor ice, and gigantic icebergs regularly rework the bottom topography as deep as 550 m below sea level, so that these habitats are constantly in a state of change or recovery (Barnes & Conlan 2007). The Antarctic shelf areas are separated from other continental shelves by the Antarctic Circumpolar Current (ACC), which carries more water than any other ocean current (Tomczak & Godfrey 2003) and reaches the ocean floor (Foster 1982). The Southern Ocean is delimited by the Antarctic Polar Front (APF) (Kock 1992), which, among other oceanic frontal zones, poses a physical barrier to marine organisms and thermally isolates the continent (Shaw et al. 2004). Nevertheless, notothenioid fishes have successfully colonized and radiated in these harsh environments.

During their diversification, notothenioid fishes have acquired a number of exceptional traits, including mitochondrial gene rearrangements (Papetti et al. 2007, Zhuang & Cheng 2010), the loss of hemoglobin in channichthyids, the loss of the otherwise near-universal heat shock response (Hofmann et al. 2000, Place et al. 2004; Hofmann et al, 2005), and the loss of the swim bladder, which may have supported the mostly benthic life style of notothenoids. However several notothenioid lineages have secondarily colonized the water column in a trend termed pelagization (Klingenberg & Ekau 1996). In order to compensate for the lack of a swim bladder, many of these pelagic notothenioids have evolved adaptations to regain neutral buoyancy. These include reduced ossification, weak mineralization of scales, and lipid deposition in large

assemblages of adipose cells (Eastman 1993). In fact, the correlation of habitat (benthic - pelagic) with regained buoyancy strongly supports the 'adaptiveness' of the notothenioid radiation.

The most important innovation for notothenioids may have been blood-borne antifreeze glycoproteins (AFGPs), that are present in members of the Antarctic clade, and enable them to cope with the subzero temperatures of Antarctic waters (Cheng et al. 2003). These proteins evolved from a pancreatic trypsinogen, and and provide the first example of how an existing gene can change to code for a new protein with an entirely different function (Chen et al. 1997). Notothenioid antifreeze glycoproteins evolved only once in notothenioids, prior to the diversification of the Antarctic clade (Chen et al. 1997, Cheng et al. 2003). Consequently, it has often been speculated that antifreeze glyproteins represent a key innovation that has endowed notothenioids with the ability to survive in the cooling waters of Antarctica, and to replace other lineages (Clarke & Johnston 1996). The key innovation hypothesis requires that cooling of the Southern Ocean and extensive sea ice conditions coincided with the emergence of antifreeze glycoproteins. Using a molecular dating of notothenioid fishes and related acanthomorph, this question is elaborated as part of the doctoral work presented here.

The characteristics of the notothenioid diversification have been reviewed in comparison with the adaptive radiations of cichlid fishes of the East African Great Lakes, and with the diversification of reef-dwelling labrid fishes. This review appeared as a book chapter:

1.1   Matschiner M, Hanel R, Salzburger W (2010) Phylogeography and speciation processes in marine fishes and fishes from large freshwater lakes. In: *Phylogeography: concepts, intra-specific patterns and speciation processes* (ed Rutgers DS), pp. 1 - 29. Nova Science Publishers, New York.

References

Anderson JB (1999) Antarctic Marine Geology. Cambridge University Press, Cambridge, UK.

Barnes DKA, Conlan KE (2007) Disturbance, colonization and development of Antarctic benthic communities. *Philos Trans R Soc Lond B Biol Sci* 362: 11–38.

Chen L, DeVries AL, Cheng C-HC (1997) Convergent evolution of antifreeze glycoproteins in Antarctic notothenioid fish and Arcticÿcod. *P Natl Acad Sci USA* 94: 3817–3822.

Cheng C-HC, Chen L, Near TJ, Jin Y (2003) Functional antifreeze glycoprotein genes in temperate-water New Zealand nototheniid fish infer an Antarctic evolutionary origin. *Mol Biol Evol* 20: 1897–1908.

Clarke A, Johnston IA (1996) Evolution and adaptive radiation of antarctic fishes. *Trends Ecol Evol* 11: 212–218.

Eastman JT (1993) Antarctic Fish Biology: Evolution in a Unique Environment. Academic Press, Inc., San Diego, CA.

Eastman JT (2005) The nature of the diversity of Antarctic fishes. *Polar Biol* 28: 93–107.

Foster TD (1982) The marine environment. In: *Antarctic Ecology* (Ed. Laws RM) Academic Press, London, UK.

Grant PR, Grant BR (2011) How and Why Species Multiply: The Radiation of Darwin's Finches. Princeton University Press, Princeton, New Jersey.

Grant PR, Grant BR (2002) Unpredictable evolution in a 30-year study of Darwin's finches. *Science* 296: 707–711.

Hofmann GE, Buckley BA, Airaksinen S, Keen JE, Somero GN (2000) Heat-shock protein expression is absent in the Antarctic fish *Trematomus bernacchii* (family Nototheniidae). *J Exp Biol* 203: 2331–2339.

Hofmann GE, Lund SG, Place SP, Whitmer AC (2005) Some like it hot, some like it cold: the heat shock response is found in New Zealand but not Antarctic notothenioid fishes. *J Exp Mar Biol Ecol* 316: 79–89

Kambysellis MP, Craddock EM (1997) Ecological and reproductive shifts in the diversification of the endemic Hawaiian Drosophila. In: *Molecular Evolution and Adaptive Radiation* (Eds. Givnish TJ, Sytsma KJ), pp. 475–509.

Klingenberg CP, Ekau W (1996) A combined morphometric and phylogenetic analysis of an ecomorphological trend: pelagization in Antarctic fishes (Perciformes: Nototheniidae). *Biol J Linnean Soc* 59: 143–177.

Kock K-H (1992) Antarctic Fish and Fisheries. Cambridge University Press, Cambridge, UK.

Papetti C, Liò P, Rüber L, Patarnello T, Zardoya R (2007) Antarctic fish mitochondrial genomes lack ND6 gene. *J Mol Evol* 65: 519–528

Place SP, Zippay ML, Hofmann GE (2004) Constitutive roles for inducible genes: evidence for the alteration in expression of the inducible hsp70 gene in Antarctic notothenioid fishes. *Am J Physiol Regul Integr Comp Physiol* 287: R429–436–R429–436.

Salzburger W (2009) The interaction of sexually and naturally selected traits in the adaptive radiations of cichlid fishes. *Mol Ecol* 18: 169–185.

Schluter D (2000) The Ecology of Adaptive Radiation. Oxford University Press, New York.

Shaw PW, Arkhipkin AI, Al-Khairulla H (2004) Genetic structuring of Patagonian toothfish populations in the Southwest Atlantic Ocean: the effect of the Antarctic Polar Front and deep-water troughs as barriers to genetic exchange. *Mol Ecol* 13: 3293–3303.

Simpson GG (1953) The Major Features of Evolution. Columbia University Press, New York.

Tomczak M, Godfrey JS (2003) Regional oceanography: an introduction. Daya Publishing House, Delhi, India.

Yoder JB, Clancey E, Roches Des S, et al. (2010) Ecological opportunity and the origin of adaptive radiations. *J Evol Biol* 23: 1581–1598.

Zhuang X, Cheng C-HC (2010) ND6 gene "Lost" and found: evolution of mitochondrial gene rearrangement in Antarctic notothenioids. *Mol Biol Evol* 27: 1–13.

1.1 Phylogeography and speciation processes in marine fishes and fishes from large freshwater lakes

Matschiner M, Hanel R, Salzburger W

1.1.1 Review: p. 12 - 42

Chapter 1

PHYLOGEOGRAPHY AND SPECIATION PROCESSES IN MARINE FISHES AND FISHES FROM LARGE FRESHWATER LAKES

Michael Matschiner[1], Reinhold Hanel[2] and Walter Salzburger[1]*

[1]Zoological Institute, University of Basel, Basel, Switzerland
[2]Institute of Fisheries Ecology, Johann Heinrich von Thünen-Institute, Hamburg, Germany

ABSTRACT

Fishes constitute about half of all known vertebrate species and have colonized nearly all available marine and freshwater habitats. The greatest diversity of fishes is found in the marine realm as well as in large (and often old) freshwater lakes such as the East African Great Lakes. Here, we compare the phylogeographic history of fishes in marine and large freshwater ecosystems, with particular emphasis on groups that underwent adaptive radiation, *i.e.* the emergence of a multitude of species from a single ancestor as a consequence of the adaptation to different ecological niches. Phylogeographic analyses are highly suited to identify and compare causal agents of speciation in rapidly diversifying groups. This is particularly true for fishes, in which distribution ranges and preferred habitat structures can be quantified in a straightforward matter.

Keywords: adaptive radiation, gene flow, cichlids, notothenioids, labrids.

* Author of correspondence.Reinhold Hanel
E-mail: reinhold.hanel@vti.bund.de

PHYLOGEOGRAPHY OF FISHES IN LARGE WATER BODIES

Since Avise *et al.* (1987) first coined the term phylogeography 23 years ago, the field has burgeoned and matured, and became a viable discipline at the intersection of population genetics, phylogenetics and biogeography (Avise 1998; 2009). The field's main concern are the principles and processes that led to contemporary geographic distributions within and between closely related species (Avise 2000). Linking micro- and macroevolutionary approaches, phylogeography has contributed greatly to species conservation, ecology and evolutionary biology. It has been integrated into the concept of 'evolutionary significant unit' (ESU) that classifies distinct populations that merit separate management and are of high priority for conservation (Ryder 1986; Moritz 1994; Crandall *et al.* 2000). Phylogeography has documented the impact of historical events on extant fauna and flora in many instances, and notably so in the case of European Pleistocene glaciations that have shaped the distribution of a wide range of European taxa (see *e.g.* Taberlet *et al.* 1998; Salzburger *et al.* 2003; Debes et al. 2008). It has also provided insights into the process of speciation (Avise 2000) when, for example the spatial simplicity and temporal certainty of volcanic archipelagos like Hawaii and the Canaries allow reconstruction of sequence and timing of speciation events (Shaw *et al.* 1996; Juan *et al.* 1998; Nepokroeff *et al.* 2003; Dimitrov *et al.* 2008; Sequeira *et al.* 2008).

A sizeable body of phylogeographic literature comes from studies conducted on teleost fishes. To some extent, this has been motivated by interest in sustained fisheries management that relies on the conservation of genetic diversity in the targeted species (Bernatchez & Wilson 1998). But fishes have also proven to be particularly informative for phylogeographic investigations. Riverine and especially lacustrine fishes inhabit island-like environments that are analoguous to volcanic archipelagos in respect of datability and spatial arrangement, and thus are similarly suitable for speciation research (Salzburger *et al.* 2005). On the other hand, marine fishes are traditionally characterized by their great diversity, their continuous and temporally stable habitat, large-scale distribution ranges, and high potential for dispersal (Palumbi 1994). Despite these differences, phylogeographic studies of marine fish species yielded important insights into population structures and their causes, the origin of marine diversity and the impact of historic events (Muss *et al.* 2001; Lourie & Vincent 2004; Rocha *et al.* 2007; Rocha *et al.* 2008). It has been shown that Pleistocene glaciations left their mark even in tropical marine settings (due to lowered sea levels; Lourie & Vincent 2004) and the phylogeography of marine species occurring on both sides of the Isthmus of Panama highlights the impact of plate tectonics on speciation over longer time scales (reviewed by Lessios 1998). Similarly, recolonization of the Mediterranean following the reopening of the Straight of Gibraltar 5.2 million years ago (MYA) (Hsü *et al.* 1973; 1977) led to a multitude of cladogenesis events that could be recovered by means of phylogeography (Carreras-Carbonell *et al.* 2005; Paternello *et al.* 2007). Furthermore, comparative phylogeography provides an adequate tool to resolve the relative impact of the many distinct life histories of marine fishes to the distributions of populations and species (Dawson *et al.* 2006). The physical setting of marine habitats also allows conclusions about these traits to be corroborated by incorporation of oceanographic data into phylogeographic analyses, *e.g.* by comparison of gene flow estimates and current speeds (Matschiner *et al.* 2009).

Thus, riverine, lacustrine, as well as marine fishes provide valuable systems for phylogeographic studies. Here, we compare the phylogeographic history of and patterns of speciation in fishes in marine and large freshwater ecosystems, with particular emphasis on groups that underwent adaptive radiation. We also present a literature review, in which we map the geographic patterns of gene flow in fish species from various taxonomic groups living in diverse environments.

THE (PHYLO-)GEOGRAPHY OF SPECIATION

One of the most hotly debated questions in speciation is certainly its geography, and, in particular, whether geographic isolation is required for new biological entities to emerge (Coyne & Orr 2004; Gavrilets 2004). Clearly, speciation can only occur via the evolution of reproductive isolation between diverging lineages. For a long time allopatric speciation[1] has been advanced as major – or even exclusive – mode of speciation (Mayr 1942; Mayr 1963). This is somewhat surprising, given that Darwin himself considered all three modes of speciation plausible (see *e.g.* Coyne & Orr 2004): allopatric[1], sympatric[2], and parapatric[3]. Since sympatric and parapatric speciation has been backed-up with theoretical and empirical evidence over the last two decades (Schliewen *et al.* 1994; Dieckmann & Doebeli 1999; Higashi *et al.* 1999; Kondrashov & Kondrashov 1999; Barluenga *et al.* 2006; Gavrilets *et al.* 2007), the debate has now shifted towards the relative importance of each of these three modes of speciation in nature.

The three possible modes of speciation explicitly impart information about geography, individual migration and gene flow. In allopatric speciation, there is absolutely no migration of individuals between the (isolated) geographic areas occupied by the speciating sub-populations; no gene flow is possible. In sympatric speciation, there is but one place, and all individuals of the speciating entities live there. Thus, there is maximum migration of individuals between the (overlapping) distribution ranges of the diverging sub-populations. This does not mean, however, that individuals belonging to distinct entities interbreed (they may do so occasionally). It simply means that individuals migrate freely in space. In parapatric speciation, a certain degree of migration occurs between the distribution ranges of the speciating sub-populations (Gavrilets 2004), and in this case interbreeding and hybrid zones are an inert feature (see *e.g.* Wu 2001; Gavrilets 2004).

There is thus an obvious and strong link between the study of speciation and phylogeography: Phylogeography provides the concepts and tools to characterize past and ongoing gene flow – and, hence, migration – in the context of geography (see *e.g.* Avise 2009). Intentionally or not intentionally, most speciation research has thus relied on and greatly benefited from phylogeography. And whenever it is necessary to explicitly interlink gene flow and distribution range – for example when testing for sympatric speciation –

---

[1] *Allopatric speciation* describes the situation that there is complete geographic isolation between the speciating entities.

[2] *Sympatric speciation* can best be defined as the emergence of novel species from a population in which mating is random with respect to the birthplace of the mating partners (Gavrilets 2004).

[3] *Parapatric speciation* is everything in between complete geographic isolation and, hence, no migration between the diverging populations (allopatry) and full sympatry; it can also be described as speciation with gene flow (Wu 2001).

phylogeography is the best way to do so (see *e.g.* Barluenga *et al.* 2006; Savolainen *et al.* 2006).

MARINE *VERSUS* LACUSTRINE ADAPTIVE RADIATIONS IN FISHES

Adaptive radiation is a process in which many species evolve in a short period of time by either allopatric, sympatric or parapatric speciation. It is the rapid proliferation of an ecologically and morphologically differentiated species assemblage from one ancestral species as a consequence of the adaptation to various ecological niches (Schluter 2000) – a process that is thought to have shaped much of the diversity of life. According to Schluter (2000), adaptive radiations can be detected by four main criteria: (*i*) common ancestry of the diversifying clade; (*ii*) a correlation between morphological or physiological traits of divergent lineages and their respective environments; (*iii*) evidence for the actual utility of these traits in their environments; and (*iv*) the rapid evolution of reproductive isolation between individuals of the divergent lineages. Often – but not always – adaptive radiations occur after the colonization of a new habitat or the evolution of evolutionary 'key innovations' (Gavrilets & Vose 2005). As a consequence of the rapid cladogenesis at the onset of an adaptive radiation, phylogenies of the radiating groups are typically bottom-heavy (Gavrilets & Vose 2005) and non-bifurcating (Sturmbauer *et al.* 2003). There are not many adaptive radiations, though, for which the fulfillment of all four criteria and bottom-heavy phylogenies has been fully demonstrated.

The most famous textbook examples of adaptive radiations are the Darwin's finches on the Galapagos archipelago (see *e.g.* Grant & Grant 2002; Grant & Grant 2006), the Caribbean *Anoles* lizards (see *e.g.* Losos *et al.* 1998), and the species flocks of cichlid fishes in the Great Lakes of East Africa (Box 1). With an estimated number of at least 1,500 species, the assemblages of cichlid fishes in lakes Victoria, Malawi and Tanganyika constitute the most diverse and species-rich adaptive radiations known (Seehausen 2006; Salzburger 2009). There are, however, at least 20 more lacustrine adaptive radiations in cichlids in Africa (Seehausen 2006); and cichlid adaptive radiations are also known from outside the African continent, *e.g.*, in the Great Lakes of Nicaragua and some smaller crater lakes nearby (Barluenga & Meyer 2004; Barluenga *et al.* 2006). Why cichlid fishes are obviously prone for adaptive radiation and explosive speciation is still under debate. It seems plausible, though, that their evolutionary success rests on a unique interaction of external factors such as habitat structure and ecological opportunity and intrinsic characteristics in form of life-history traits and evolutionary key innovations like a highly adaptable feeding apparatus (Salzburger 2009).

Adaptive radiations in teleost fishes are, in general, quite common in freshwater systems: Three-spined sticklebacks (*Gasterosteus aculeatus*), for example, have repeatedly radiated into benthic and limnetic forms from ancestral marine ecotypes in post-glacial lakes (Schluter & McPhail 1992); lake whitefish (*Coregonus* spp.) have undergone adaptive radiations in post-glacial lakes, too, throughout their distribution range in the Northern hemisphere (Bernatchez *et al.* 1999; Ostbye *et al.* 2005; Vonlanthen *et al.* 2009); in the Malili lake system in Sulawesi, several species of sailfin silversides (*Telmatherina* spp.) have emerged via

adaptive radiation (Herder *et al.* 2006; Roy *et al.* 2007a; Roy *et al.* 2007b); adaptive radiations have also been proposed in African weakly electric fish (*Campylomormyrus* spp.) (Feulner *et al.* 2007), in barbs (*Labeobarbus* spp.) from Lake Tana in Ethiopia (de Graaf *et al.* 2008), in cyprinids from Philippine Lake Lanao (Kornfield & Carpenter 1984), and in cyprinodontids (*Orestias* spp.) from Lake Titicaca in South America (Parenti 1984).

The situation is different in the marine realm, where much fewer cases of adaptive radiations have been described (see *e.g.* Rüber & Zardoya 2005). One of the groups that fulfills all four criteria of an adaptive radiation are the notothenioid fishes that are mainly found in Antarctic waters (Eastman 2005) (Box 2). Several evolutionary key-innovations and adaptiations have been identified (in notothenioids and subgroups thereof) that allow them to cope with the harsh environmental conditions in the Southern Ocean, such as the evolution of antifreeze glycoproteins and the losses of hemoglobin, of parts of the mitochondrial respiratory chain and of the heat-shock response system (Chen *et al.* 1997; di Prisco *et al.* 2002; Papetti *et al.* 2007a; Hofmann *et al.* 2000). However, the radiation of the whole Antarctic clade does not exhibit the bottom-heavy phylogeny (*sensu* Gavrilets & Vose 2005) theoretically expected in adaptive radiations. Instead, the full notothenioid species richness of about 130 species is attained through at least three secondary radiations – those of the artedidraconid genus *Pogonophryne*, the nototheniid subfamily *Trematominae* and the nototheniid genus *Patagonotothen* (Eastman 2005; Sanchez *et al.* 2007; Near & Cheng 2008).

Other radiations in marine fishes are less well documented than the notothenioid one and it remains to be proven whether some of these radiations are 'adaptive' after all. A second teleost radiation may have occurred in the Antarctic region. The deeper parts of the Antarctic shelf are inhabited by 64 species of the scorpaeniform family Liparidae that probably represent a secondary radiation within a larger liparid diversification, centered mainly in the North Pacific region (Eastman & Clarke 1998). The colorful parrotfishes (Scaridae), unambiguously shown to be a specialized lineage deeply nested within the family Labridae (Bellwood 1994, Westneat & Alfaro 2005), inhabit the coral reefs and seagrass beds of tropical waters. Its roughly 90 species have adapted to a variety of habitats as well as social and mating strategies in the course of a radiation that presumably started around 14 MYA in the Tethys Sea (Streelman *et al.* 2002). The overall about 600 labrid species might as well represent an adaptive radiation (Box 3), and it has been argued that – just as in cichlid fishes – a highly adaptable pharyngeal jaw apparatus might have contributed as evolutionary key innovation in that group triggering their radiation (Westneat & Alfaro 2005; Mabuchi *et al.* 2007). Reef-associated gobies, such as the American seven-spined gobies (Gobioseomatini) or the Neotropical reef gobies (*Elacatinus* spp.) apparently underwent adaptive radiations, too (Rüber *et al.* 2003; Taylor & Hellberg 2005). Recently, Puebla and coworkers (Puebla *et al.* 2007; Puebla *et al.* 2008) have highlighted an example of a marine adaptive radiation in its very first stages, once again in colorful coral reef fishes, the hamlets (genus *Hypoplectrus*, family Serranidae). These 13 closely related predatory fish species are widely distributed in the Caribbean Sea.

It is not entirely obvious why adaptive radiation should be less frequent in marine fishes compared to those in (large) freshwater lakes. One reason why there are fewer cases reported in marine fishes might be that adaptive radiations are simply more apparent in geologically young and geographically well-defined areas (Salzburger 2008), and, hence, more easy to investigate. Indeed, the best candidates for adaptive radiations in marine fishes occurred in geographically separated areas such as the Antarctic continent (notothenioids) or the

Caribbean Sea (hamlets). Older radiations, especially in tropical marine perciform families like wrasses, damselfishes, butterflyfishes, angelfishes as well as seabreams and others, date back much longer in time and might also be camouflaged by subsequent geographical separation through climatically and geologically induced range shifts or local extinctions.

THE GEOGRAPHIC SCALE OF GENE FLOW IN FISHES

Because of several reasons, fishes are an ideal group for phylogeographic research: their living space is strictly bordered by migration barriers (*e.g.* land, waterfalls, open water), their habitats are relatively easy to characterize, migration can only follow certain routes (*e.g.* ocean currents, coastlines, rivers), life-history traits (*e.g.* vagility, generation time, number of offspring) are often known, genetic tools are available, *etc*. Thus, it does not come to any surprise that a whole body of literature exists with respect to the phylogeography of various species of fish. For similar reasons, fishes are excellent models for speciation and adaptive radiation research (see *e.g.* Kocher 2004; Rüber & Zardoya 2005; Seehausen 2006; Rocha & Bowen 2008; Salzburger 2009).

Migration, gene flow and genetic differentiation are crucial parameters in both phylogeography and speciation (see above). In order to compare geographic distances over which genetic differentiation takes place in different environments and different groups of fishes, we conducted a literature review and focused on phylogeographic and population genetic studies according to the following criteria: (*i*) either DNA sequences or microsatellite loci were used as molecular markers, (*ii*) sample sizes and sampling locations were specified precisely, (*iii*) pairwise *F*-statistics or similar measures were reported, (*iv*) sequential Bonferroni correction for multiple tests (Rice 1989) or a false discovery rate (Benjamini & Hochberg 1995) was applied to pairwise comparisons, or *p*-values were reported and enabled us to conduct Bonferroni error correction. We ignored studies on populations of unresolved species status, and those that include artificially introduced or cultured populations, as well as studies investigating populations separated by artificial barriers such as river dams. Riverine populations were included only if they were sampled from the same watershed. For every study, we measured both the shortest water connection over which significant genetic differentiation was found (d_{min_s}) and the longest water connection over which no significant gene flow could be detected (d_{max_ns}). All geographic distances were measured using Google Earth®. Exact sampling locations were rarely given for anadromous species from different river systems. In these cases, the distance between river estuaries was taken. We particularly focused on three groups of perciform fishes that underwent adaptive radiations in three distinct environments: cichlids (lacustrine), labrids (tropical to temperate marine), and notothenioids (polar marine).

We based our comparison on 81 articles (marked with * in the References) investigating the population genetic stucture of 114 fish species in environments as diverse as the Arctic and the Great Barrier Reef, the Amazon River and the 34 km long Atsuta River in Japan. A number of species was investigated in more than one study or with both nucleotide and microsatellite markers, so that we ended up with 130 measurements of d_{min_s} and/or d_{max_ns}. In 37 cases, no significant genetic differentiation was found between investigated populations, while all pairwise comparisons were significant in 25 out of the 130 cases. In the most

extreme cases, significant genetic differentiation was found between samples taken at the same location, but in different years (d_{min_s} = 0 km; Zane *et al.* 2006; Lin *et al.* 2008a; Hepburn *et al.* 2009), or no comparison was significant despite a global sampling scheme (d_{max_ns} = 16,309 km; Horne *et al.* 2008).

The shortest geographic distances, over which significant genetic differentiation was found in different taxonomic groups and environments are visualized in Figure 1. Naturally, these measures may depend on parameters such as study design, sample size and number of markers employed. In Figure 2, we plotted d_{min_s} against the sample size of the respective study. Indeed, the result suggests a negative correlation between both values. However, as the average sample sizes were comparable between studies in different fish taxa and environments (with the exception of anadromous fishes: N = 825; others: N = 130-333), the overall picture shown in Figure 1 should not be influenced by the different study practices applied by the different researcher groups.
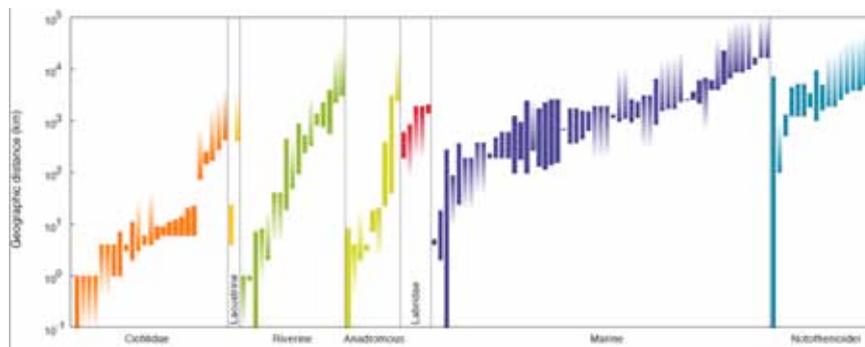
Figure 1. The geographic scale of gene flow in fishes. Shortest geographic distances over which significant genetic population differentiation have been found in different taxonomic groups and environments. Each bar represents one analysis of population differentiation. Bars are drawn between the shortest distance, over which significant differentiation has been found (d_{min_s}), and the longest distance, over which no significant differentiation could be detected (d_{max_ns}). A downward gradient symbolizes that all pairwise comparisons were significant. In these cases, the gradient's top end represents d_{min_s}. This visualizes that significant differentiation could be expected at even shorter, untested distances. Similarly, an upward gradient symbolizes that no pairwise comparison was significant, and that significant differentiation can be expected only at distances greater than those tested (d_{max_ns} is the gradients lower end). All distances were measured as the shortest water connections between fish populations.
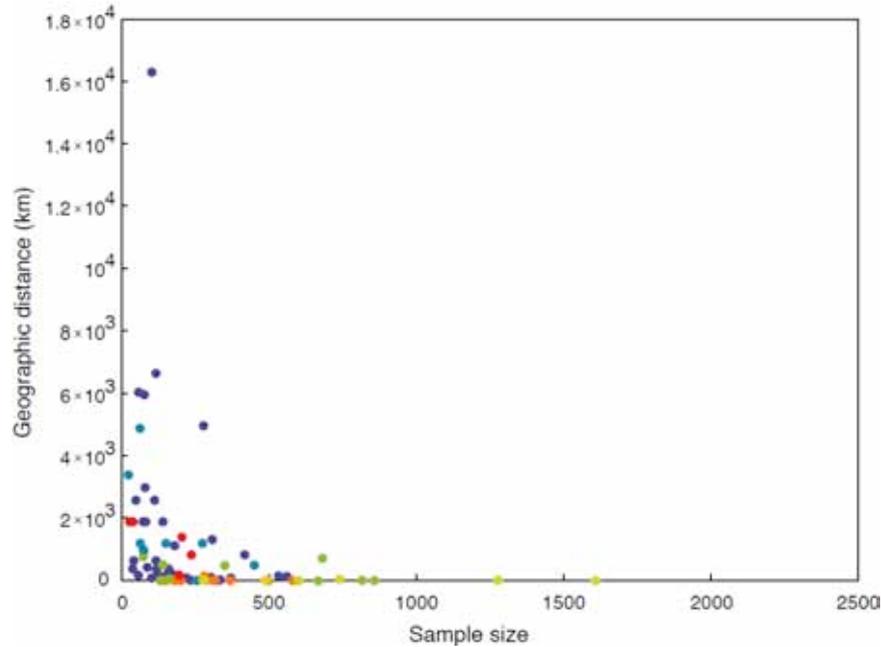
Figure 2. Sample size effects in phylogeographic studies in fishes. The shortest geographic distance over which significant differentiation has been detected plotted against sample size. Color code as in Figure 1.

Lacustrine Fishes

Differentiation over short geographic distances on the order of 10 km and below is commonly found in rock-dwelling cichlids of the East African Great Lakes, and it has been speculated whether their tendency to philopatry and the resulting barriers to gene flow has enabled local adaptation, speciation, and their impressive adaptive radiation (Rico & Turner 2002; Pereyra *et al.* 2004). However, the cichlid radiations also include a number of pelagic species that show genetic homogeneity over hundreds of kilometers, and thus would contradict this hypothesis (see the five bars at the right end of the Cichlidae column in Fig. 1) (Shaw *et al.* 2000; Taylor & Verheyen 2001). We found two studies on non-cichlid lacustrine fishes that matched our criteria: Sailfin silversides of Lake Matano, Indonesia, show significant differentiation at small geographic distances (Walter *et al.* 2009), while large-scale gene flow was observed in the little Baikal oilfish in Lake Baikal, Russia (Teterina *et al.* 2005).

Riverine and Anadromous Fishes

Very variable patterns were found in riverine and anadromous fish species. In the case of the riverine fishes, it appears that river size influences rates of gene flow between populations: Genetic differentiation over short distances was found repeatedly in small river systems such as the Caroni Drainage, Trinidad and Tobago (d_{min_s} = 1 km, all comparisons

being significant; Barson *et al.* 2009), the Amor de Cosmos watershed on Vancouver Island, Canada (d_{min_s}, d_{max_ns} = 1 km; Caldera & Bolnick 2008), and the Novoselka River basin, Sakhalin, Russia (d_{min_s} = 1 km, d_{max_ns} = 7 km; Osinov & Gordeeva 2008). On the other hand, population genetic assessments of fishes of the Amazon River frequently fail to detect significant population structure over the entire sampling area (d_{max_ns} > 2000 km; Batista & Alves-Gomes 2006; Santos *et al.* 2007).

Marine Fishes

In general, marine fishes show great variability in their patterns of differentiation: While reef fishes with low dispersal abilities may exhibit significant population structure at less than 10 km (Miller-Sims *et al.* 2008; Bay *et al.* 2008), most marine fishes display differentiation only at distances of hundreds to thousands of kilometers; no genetic structuring even at a global scale has been observed in lemon sharks (Schultz *et al.* 2008) and two surgeonfishes (Horne *et al.* 2008). Fishes of the family Labridae show comparable patterns of differentiation between the different species. Significant population structure was found between 187 and 1898 km. Fishes of the perciform suborder Notothenioidei show little genetic structuring even compared to other marine fish taxa. One exception aside (significant structure between year-classes sampled at the same location; Zane *et al.* 2006), significant genetic differentiation has been found only over several hundreds or thousands of kilometers, or not at all, as is the case for the majority of studies included in our survey. As the life histories of most notothenioids include long pelagic larval stages of up to one and a half years (Kock & Kellermann 1991; La Mesa & Ashford 2008), it has been speculated that strong oceanic currents, and in particular the Antarctic Circumpolar Current (ACC) may be responsible for gene flow in form of larval dispersal (Zane *et al.* 2006; Jones *et al.* 2008). Using a multidisciplinary approach including oceanographic data and simulations using the isolation-with-migration (IM) model (Hey & Nielsen 2007) to investigate directionality of gene flow in the notothenioid fish *Gobionotothen gibberifrons*, Matschiner *et al.* (2009) indeed found highly asymmetric migration rates between the Antarctic Peninsula and islands of the Scotia Ridge, following the direction of the ACC. As gene flow caused by long-distance migration of adult individuals would be expected to result in roughly symmetric migration rates, this finding corroborates the hypothesis that larval dispersal precludes genetic differentiation in Antarctic waters even across large geographic distances.

THREE ADAPTIVELY RADIATING PERCIFORM GROUPS

At least one in two vertebrate species is a fish and within the fishes at least one third (and more than 10,000 species) belongs to the order Perciformes, making it the largest order of vertebrates. The Perciformes itself is comprised of about 160 families and more than 1500 genera and they dominate vertebrate life in the ocean and in tropical and subtropical freshwaters (Nelson 2006). Much of the diversity of perciforms has arisen through adaptive radiations, of which the ones of the cichlid fishes are the most impressive. Marine (adaptive) radiations within the Perciformes are those of the notothenioids, of the labrids, the gobies, and

the hamlets (Eastman 2005; Westneat & Alfaro 2005; Rüber *et al.* 2003; Puebla *et al.* 2008). Massive bursts of diversification ('explosive speciation') have repeatedly been reported for East African cichlid fishes (*e.g.* McCune 1997; Seehausen 2002; Verheyen *et al.* 2003). In marine fishes, elevated rates of cladogenesis were reported – among others – for *Sebastes* rockfishes, the notothenioid subfamily Trematominae, American seven-spined gobies and sparids (Rüber & Zardoya 2005).

Here, we focus on three groups of Perciformes that apparently underwent adaptive radiations and episodes of explosive speciation in different environments (Eastman & Clarke 1998; Eastman 2005; Seehausen 2006; Mabuchi *et al.* 2007): the cichlids of the tropical Great Lakes in East Africa (Box 1), the notothenioids of the polar marine waters of Antarctica (Box 2), and the labrids of the tropical and subtropical marine waters (Box 3). The adaptive radiations of all three groups have been associated with evolutionary key-innovations (Liem 1973; Chen *et al.* 1997; Hulsey 2006; Mabuchi *et al.* 2007), they all evolved a spectacular diversity of body morphologies and – in the case of cichlids and labrids – color morphs, and members of all three groups dominate their respective fauna.

Phylogeographic and population genetic studies in the three groups cichlids, notothenioids, and labrids reveal substantial differences with respect to the geographic distances over which gene flow could be detected (Figure 1). While in most cichlid species population structure could be detected over small geographic ranges of below or around 10 km, labrids and – with one exception – notothenioids show gene flow over large geographic distances. The latter two groups lie well in the range of other marine fishes, just as a few pelagic cichlid species do (note that the upper geographic limits in these cichlid species is restricted by lake size). This discrepancy between gene flow on a circumantarctic scale in notothenioids and large distances in labrids and the fine-scale genetic structuring in cichlids of the East African Lakes seems puzzling, given that all these clades underwent adaptive radiations in their respective environments, and philopatry has often been proposed as one of the key agents behind local adaptation and, consequently, adaptive radiation (Bouton *et al.* 1999; Rico & Turner 2002; Rico *et al.* 2003; Pereyra *et al.* 2004; Taylor & Hellberg 2005; Gavrilets *et al.* 2007).

Gene flow is generally expected to retard speciation by breaking linkage between genes for local adaptation and those for reproductive isolation (Coyne & Orr 2004). On the other hand, recent theoretical work as well as empirical research (Gavrilets & Vose 2005; Seehausen 2006; Garant *et al.* 2007) has shown that gene flow between populations does not necessarily prevent local adaptation. To the contrary, it can facilitate the spread of beneficial mutations and thus support adaptation under certain circumstances. In the context of adaptive radiation, the individual-based stochastic model of Gavrilets & Vose (2005) predicted that divergence can be maintained for very long periods despite substantial amounts of gene flow, which would lead to a 'porous' genome with low to non-existing differentiation in neutral markers, but divergence at locally selected loci. Evidence for porous genomes has been found in the *Hypoplectrus* complex of coral reef fishes that are supposed to represent an adaptive radiation in its very first stages (Puebla *et al.* 2008).

PHYLOGEOGRAPHY AND SPECIATION IN MARINE *VERSUS* LACUSTRINE FISHES

So what is it that could explain the difference between marine fishes with gene flow over large geographic distances and fishes from large freshwater lakes with often highly structured populations?

Habitat discontinuities, which have been suggested as main reason why rock-dwelling cichlid populations are so structured (Arnegard *et al.* 1999; Rico & Turner 2002; Pereyra *et al.* 2004; Duftner *et al.* 2006; Sefc *et al.* 2007), can only partly explain these differences. Marine reefs are highly fragmented, too. Still, gene flow in reef associated fishes can be observed over large geographic distances, *e.g.* between the West and East Atlantic (Floeter *et al.* 2008; Rocha *et al.* 2008) or between Caribbean islands over hundreds of kilometers (Puebla *et al.* 2008). Habitats of benthic notothenioids are disrupted by iceberg scours (Brenner *et al.* 2001) and open water between island shelves, while the habitat of a limited number of pelagic notothenioids may be assumed continuous over thousands of kilometers (Zane *et al.* 2006). Nevertheless, pelagic and benthic notothenioids alike apparently maintain gene flow over these large distances (Figure 1) (Matschiner *et al.* 2009).

Another extrinsic factor that might explain the observed differences in population structure is *habitat stability*. Large freshwater lakes are very young compared to marine habitats. Lake Tanganyika, for example, the oldest of the East African Great Lakes and second oldest lake in the world, has a maximum age of 12 million years (MY) (Cohen *et al.* 1997); Lakes Malawi and Victoria are considerably younger. More importantly, the lakes have repeatedly undergone dramatic water-level fluctuations of up to several hundred meters. In the case of Lake Victoria, this is equivalent to a complete desiccation, but fish diversity may have survived in tributaries and satellite lakes (Johnson *et al.* 1996; Cohen *et al.* 1997; Mwanja *et al.* 2001; Verheyen *et al.* 2003; Stager & Johnson 2008). It has been argued that these cyclic changes leading to admixis, hybridization, fragmentation of populations, and small founder populations, contributed to the species-richness in the East African lakes (Rossiter 1995; Kornfield & Smith 2000; Sturmbauer *et al.* 2001). It is less apparent, though, how these lake-level fluctuations could account for the structuring in present cichlid populations. Dramatic changes in the environment also characterize the marine habitat of Antarctic notothenioids. During the last two MY, the Antarctic ice sheet has periodically advanced and retreated with each glacial cycle. Presumably it has extended all the way to the shelf edge in glacial maxima (Thatje *et al.* 2005), 'bulldozing the surviving fauna to the deep continental margin' (Barnes & Conlan 2007). Naturally, the associated loss of benthic habitat must place serious constraints on demersal fish communities. There is evidence for at least some refuges in form of ice-free shelf areas (Barnes & Conlan 2007) that could provide analogues to satellite lakes of Lake Victoria during desiccation periods.

The temporal scale of significant and drastic environmental change is clearly different for wrasses and other tropical marine reef fishes and reaches back as far as the Eocene. The split of the most species-rich wrasse lineage, the Julidini, covering about one-third of overall labrid diversity was recently calculated of an age of 36 to 38 MY (Kazancioglu *et al.* 2009) supporting the hypothesis of their Tethyan origin and Indo-Pacific ancestral distribution (Westneat & Alfaro 2005). These estimates imply that by the time the julidine lineage originated, the Antarctic Circumpolar Current was already established, which disrupted the

connection between higher and lower latitudes, and restricted the movement of tropical lineages to the Tethys (Bellwood & Wainwright 2002). A series of diversification events within the Julidini leading to an early burst of diversification and the evolution of the majority of extant julidine lineages nicely coincides with a period of increased diversification and fragmentation of coral reefs, and extensive development of reef communities in the Tethys and the Caribbean (Veron 1995) between 15 to 30 MY (Kazancioglu *et al.* 2009). Habitat fragmentation culminated in the middle Miocene with its rapidly changing paleobiogeographical conditions and strong tectonic activity (Rögl, 1999) that resulted in the final closure of seaway between the Mediterranean and the Indian Ocean some 14 MYA. Hanel *et al.* (2002) correlated the following succession of the Mediterranean with the radiation of the wrasse tribe Labrini, endemic to the northern Atlantic and found striking congruence.

Among the intrinsic (biotic) differences between marine fishes and fishes from large freshwater lakes is the *degree of specialization*. While most lacustrine East African cichlid species are ecologically highly specialized, the majority of marine fishes are not (at least not to the degree observed in cichlids). Rocha & Bowen (2008) attest that most reef fishes are 'neither widely distributed generalists nor ecological specialists'. Clearly, specialization limits gene flow by lowering survival rates and reproductive success of migrants. The question remains whether the much greater degree of specialization is a reason for or the outcome of the limited levels of gene flow between cichlid populations.

Another difference between marine fishes and cichlids is the *breeding behavior*. It is interesting though that in all three groups that underwent adaptive radiations, cichlids, notothenioids and labrids, a certain degree of brood care occurs. The cichlids are famous for their various systems and strategies of brood care behavior ranging from substrate spawning in nests and under custody of the parents to various levels of mouthbrooding (Goodwin *et al.* 1998; Barlow 2000).

Prolonged incubation and pelagic larval duration are common features of most Antarctic notothenioids (Kock & Kellermann 1991, Loeb *et al.* 1993). For example, hatching of larvae of the naked dragonfish *Gymnodraco acuticeps* occurs only about 10 months post-fertilization (Evans *et al.* 2005), while the Scotia Sea icefish *Chaenocephalus aceratus* undergoes an extensive pelagic phase as long as 1.5 years (La Mesa & Ashford 2008). Brood care of demersal eggs has been reported for a number of species and even egg carrying behavior has been observed in one icefish species (*Chionobathyscus dewitti*; Kock *et al.* 2006). However, other nototheniod fishes are open spawners that release their eggs in the open water column, or produce demersal eggs that become pelagic towards the end of their development (Kock 2005; Kellermann 1991). Pelagic eggs and larvae are prone to off-shelf advection and dispersal with strong oceanic currents such as the ACC. While active larval behavior, especially towards the end of the larval phase, may counteract dispersal in many cases (White 1998; Leis 2006), pelagic eggs and larvae have been found hundreds of kilometers away from suitable shelf habitat (Kellermann 1991; Loeb *et al.* 1993). Widespread larval dispersal is further suggested by the fact that only nototheniids and channichthyids with particularly long pelagic larval durations occur at the isolated island of Bouvetøya (Jones *et al.* 2008).

Within the percomorpha, the family Labridae can be considered exceptional in terms of diversity of social and reproductive systems. Most wrasses are sequential hermaphrodites, with a transformation from female to male state being the normal occurrence. The causes and pathways of the evolution of hermphroditism, regularly found in percomorph marine fishes,

as reproductive style have been and are still subject of debate (Atz 1964, Smith 1967, 1975, Ghiselin 1969, Reinboth 1970, Policansky 1982). One clear advantage should be to maximize lifetime reproductive potential (Williams 1966) and hence individual fitness (*sensu* Stearns 1976). However, courtship, spawning, and sex change can be quite varied with mating systems in wrasses including haremic mating groups, promiscuity, lek-like behavior leading to group spawning, and facultative monogamy (pair spawning) (Donaldson, 1995). A change in sex is often associated with a change in color pattern. Broadcast spawning is a general rule in the Labridae, with most species being characterized by planktonic eggs and larvae and therefore a lack of any kind of brood care behavior, a pattern typical for the majority of marine fish species. In contrast, brood care is well developed in the comparatively small wrasse tribe Labrini (Hanel *et al.* 2002). Labrine wrasses show a variety of different brood care strategies, representing evolutionary succession from simple formation of spawning cavities up to the construction of complex nests associated with extensive egg care performed by territorial males and supported by one to several "helpers". Nevertheless, the effect of different brood care strategies on population size and structure as well as on phylogeography has, to our knowledge, not yet been tested.

CONCLUSION

Over the past two decades, fishes have emerged as excellent model groups for the study of phylogeography, speciation and adaptive radiation. This is not least due to their well-defined habitats, the existence of strong migration barriers bordering their living space, their restricted possibilities for migration and dispersal, and the availability of genetic tools. Different groups of fishes vary with respect to phylogeography and population structure: An extensive literature review revealed substantial differences in the geographic distances over which gene flow was detected in various groups of fishes that inhabit diverse environments. Marine fish typically show low to non-existing gene flow over hundreds to thousands of kilometers, while populations of lacustrine fishes, such as the cichlid fishes in the East African Great Lakes, are typically highly structured. Three groups of the highly diverse perciform fishes that underwent adaptive radiations are the cichlids, the notothenioids and the labrids. They radiated in large freshwater lakes, the polar waters of Antarctica, and tropical to temperate marine environments, respectively. Speciation and diversification in all three groups has been connected to external factors such as habitat instability, and paleo-geological and paleo-climatological processes, and all three radiations have been associated with evolutionary key-innovations. Still, they differ in overall within-species phylogeography, in population structure and patterns and levels of gene flow. The marine representatives are also generally less specialized than the cichlids. Whether this is due to differences in life-history traits, such as breeding behavior, would need to be investigated.

Box 1: The adaptive radiations of cichlid fishes

The perciform family of the Cichlidae represents a group of tropical and subtropical freshwater fish that show an Gondwanian distribution with ancestral and relatively species-poor lineages in India, Sri Lanka and Madagscar and two highly diverse clades in South- and Central America and in Africa, respectively (Salzburger & Meyer 2004). The most impressive cichlid adaptive radiations have occurred in the East African Great Lakes where at least 1500 species have evolved in the last few millions to several thousands of years only (Kocher 2004; Seehausen 2006; Salzburger 2009). Various hypotheses exist with respect to the evolutionary success of this group, and it seems likely that a unique combination of intrinsic (biotic) and extrinsic (abiotic) factors have triggered their adaptive radiations (Salzburger 2009). It has long been suggested that the particular architecture of the cichlid's jaw apparatus – with a second set of jaws in the pharynx – has acted as evolutionary key innovation in the adaptive radiations cichlids (Liem 1973). The most species-rich group of cichlids, the haplochromines from East Africa, are characterized by their particular kind of maternal mouthbrooding and egg-dummies on the male anal fins, which mimic real eggs and aid to bring the females mouth close to the male's genital opening. Both maternal mouthbrooding and egg-dummies might have acted as key-innovations, too (Salzburger *et al.* 2005; Salzburger 2009). It appears that both, ecologically relevant and, hence, naturally selected traits (*e.g.* moth morphology, body shape) and sexually selected traits (*e.g.* coloration) are important during cichlid speciation (Salzburger 2009).

Possible extrinsic factors are repeatedly occurring fluctuations of the lake level and the habitat diversity found in the East African lakes (Sturmbauer 1998; Kornfield & Smith 2000; Sturmbauer *et al.* 2001). Habitat discontinuities, together with often philopatric and stenotopic behavior of many of the cichlid species, may be partly responsible for their explosive speciation in lakes Victoria, Malawi and Tanganyika (van Oppen *et al.* 1997; Rico & Turner 2002; Rico *et al.* 2003; Pereyra *et al.* 2004; Duftner *et al.* 2006; Sefc *et al.* 2007).

Number of species (estimated): 3000-5000
Distribution range: Gondwanian (India, Sri Lanka, Madagascar, Africa, South- and Central America)
Habitat: freshwater (lakes and rivers)
Key-innovations (suggested): pharyngeal jaw apparatus, egg-dummies

Box 2: The adaptive radiation of notothenioids

Fishes of the perciform suborder Notothenioidei have successfully colonized the Antarctic waters and radiated under these harsh conditions. Today, the notothenioids dominate the Antarctic continental shelf and upper slope in terms of species number (47%) and biomass (90-95%) (Eastman & Clarke, 1998). Estimates for the onset of the notothenioid radiation range between 24 (Near 2004) and 7-15 MYA (Bargelloni *et al.* 1994; Cheng *et al.* 2003). Today, eight families and at least 130 notothenioid species are known. The three basal families, Bovichtidae, Pseudaphritidae and Eleginopidae comprise 13 species, 12 of which are non-Antarctic and occur in the coastal waters of New Zealand, Australia and around the tip of South America. The five remaining families Nototheniidae, Harpagiferidae, Artedidraconidae, Bathydraconidae and Channichthyidae consist of 116 mainly Antarctic species (Eastman 2005). Typically, only the latter five families (the 'Antarctic clade') are referred to when speaking of the notothenioid radiation.

The remarkable diversification of the Notothenioidei has been accompanied by several innovations in physiology. The most general feature found in all notothenioids, but not in higher-level relatives, is a lack of swim bladders. For this reason, most notothenioids are heavier than seawater and dwell on or near the seafloor. However, several notothenioid lineages have independently colonized the water column in a trend termed pelagization (Klingenberg & Ekau 1996). The expression of heat-shock proteins (HSPs) as a response to elevated temperatures, a feature that is regarded as a universal characteristic of almost all organisms, has been found absent in the highly cold-adapted members of the Antarctic clade (Hofmann *et al.* 2000; Clark *et al.* 2008). Recently, it has been shown that members of the Antarctic clade lack the mitochondrial *ND6* gene (coding for the NADH-Dehydrogenase subunit 6) (Papetti *et al.* 2007a). All members of the most derived notothenioid family, the Channichthyidae, have lost the ability to synthesize hemoglobin (Ruud 1954; Eastman 1993), and thus represent the only vertebrates without oxygen-bearing blood pigments. While the absence of hemoglobin is due to the deletion of the β-globin subunit gene in a single deletion event (di Prisco *et al.* 2002), truncated and inactive remnants of the α-globin gene are retained in channichthyid genomes (Cocca *et al.* 1995; Near *et al.* 2006). Since the oxygen-carrying capacity of the hemoglobinless phenotype is reduced by a factor of ten, the Channichthyidae evolved compensational features such as a blood volume two to four times that of comparable teleosts, a large stroke volume and cardiac output, and relatively large diameters of arteries and capillaries (Eastman 1993).

The most remarkable innovation of notothenioids are special blood-borne antifreeze glycoproteins (AFGPs), that are present in all notothenioids of the Antarctic clade, and enable them to cope with the subzero temperatures of Antarctic waters (Cheng *et al.* 2003). There is evidence that the AFGPs evolved only once in notothenioids from a trypsinogen ancestor gene, and that this happened before the diversification of the Antarctic clade (Chen *et al.* 1997; Cheng *et al.* 2003). It is thus tempting to attribute the notothenioid radiation to the evolution of AFGPs as a key adaptation with respect to the cooling environment. It may have enabled the notothenioids to survive the temperature drop in Antarctic waters from around 20˚C to the current freezing conditions (Clarke & Johnston 1996), and to radiate while most other teleosts could not adapt to the decreasing temperatures.

Number of species: ca. 130
Distribution range: Antractic waters, South Pacific
Habitat: polar marine
Key-innovations (suggested): antifreeze glycoproteins

Box 3: The (adaptive) radiation of labrids

The perciform family Labridae is a diverse group of about 600 mostly reef-dwelling species in 82 genera that exhibit an exceptional diversity in body size, shape, coloration, feeding habits, reproductive behaviors, and life histories (Westneat 1999, Parenti & Randall 2000, Wainwright *et al.* 2004, Westneat & Alfaro 2005). Together with the parrotfishes (Scaridae) as well as the cales and weed-whitings (Odacidae), which were all shown to be deeply nested within the Labridae (Bellwood 1994, Westneat and Alfaro 2005), wrasses comprise the worldwide second largest family of marine fish.

As with many percoid families the fossil record of the Labridae extends back to the Eocene (Lower Tertiary, approx. 54 MYA) (Berg 1958; Patterson, 1993) with *†Phyllopharyngodon longipinnis* Bellwood 1990 being described from a specimen recovered from the Pesciara ("Fish Bowl") in Monte Bolca, Italy (Bellwood 1990). Being dated to topmost Ypresian or lowermost Lutetian (Benton *et al.* 1993), this results in an estimated age of about 48 to 50 MY (Luterbacher *et al.* 2004). Based on the presence of a single predorsal, a well-developed pharyngeal jaw, and the phyllodont form of the teeth found on the pharyngeal jaw, Bellwood (1990) placed the specimen with confidence among the basal wrasse clade Hypsigenyini. However, based on plate tectonics, dating of reef lineages with molecular clocks and patterns of fish otolith preservation, the overall age of the family is estimated to be anywhere between 50 and 90 MY (Bellwood & Wainwright 2002, Westneat & Alfaro 2005).

From an oceanographic point of view, this time period near the end of the Mesozoic and beginning of the Cenozoic was characterized by the continuation of the Gondwana break-up to form present-day shaped continents as well as the central role of the circum-tropical Tethys Sea connecting the Indian with the Atlantic Ocean.

Diversification of the Labridae has often been referred to as a consequence of the evolution of functional novelties in the feeding apparatus that have allowed them to occupy nearly every feeding guild in reef environments (Westneat & Alfaro 2005). Feeding habits in the group are as diverse as in cichlids, including specialized predation on gastropods, bivalves, crustaceans, fishes, coral mucous, zooplankton, ectoparasites, detritus and algae (Randall 1967, Westneat 1997). However, recent investigations point out that territorial behavior and strong sexual dichromatism, as expressed by many wrasse species, may effectively drive sexual selection and are therefore major factors for labrid diversification (Kazancioglu *et al.* 2009).

Number of species (estimated): 600
Distribution range: global
Habitat: tropical to temperate marine
Key-innovations (suggested): pharyngeal jaw apparatus

REFERENCES

*Abila R, Barluenga M, Engelken J, Meyer A, Salzburger W (2004) Population-structure and genetic diversity in a haplochromine fish cichlid of a satellite lake of Lake Victoria. *Molecular Ecology* 13 (9), 2589-2602.

*Appleyard SA, Williams R, Ward RD (2004) Population genetic structure of Patagonian toothfish in the West Indian sector of the Southern Ocean. *CCAMLR Science* 11, 12-32.

*Arnegard ME, Markert JA, Danley PD, Stauffer JR, Ambali AJ, Kocher TD (1999) Population structure and colour variation of the cichlid fish *Labeotropheus fuelleborni* Ahl along a recently formed archipelago of rocky habitat patches in southern Lake Malawi. *Proc R Soc Lond B Biol Sci* 266, 119-130.

Atz JW (1964) Intersexuality in fishes. In: *Intersexuality in vertebrates including man* (eds. Marshall AJ, Armstrong CN) pp. 145-232. Academic Press, NY.

*Aurelle D, Guillemaud T, Afonso P, Morato T, Wirtz P, Santos RS, Cancela ML (2003) Genetic study of *Coris julis* (Osteichtyes, Perciformes, Labridae) evolutionary history and dispersal abilities. *Comptes Rendus Biologies* 326, 771-785.

Avise JC (1998) The history and purview of phylogeography: a personal reflection. *Molecular Ecology* 7, 371-379.

Avise JC (2000) *Phylogeography: the history and formation of species.* Harvard University Press, Cambridge, MA.

Avise JC (2009) Phylogeography: retrospect and prospect. *Journal of Biogeography* 36, 3-15.

Avise JC, Arnold J, Ball RM, Bermingham E, Lamb T, Neigel JE, Reeb CA, Saunders NC (1987) Intraspecific phylogeography: the mitochondrial DNA bridge between population genetics and systematics. *Annual Review of Ecology and Systematics* 18, 489-522.

*Baerwald MR, Feyrer F, May B (2008) Distribution of genetically differentiated splittail populations during the nonspawning season. *Transactions of the American Fisheries Society* 137 (5), 1335-1345.

Bargelloni L, Ritchie PA, Patarnello T, Battaglia B, Lambert DM, Meyer A (1994) Molecular evolution at subzero temperatures: mitochondrial and nuclear phylogenies of fishes from Antarctica (suborder Notothenioidei), and the evolution of antifreeze glycopeptides. *Molecular Biology and Evolution* 11 (6), 854-863.

Barlow GW (2000) *The Cichlid Fishes. Nature's Grand Experiment in Evolution.* Perseus Publishing, Cambridge, MA.

Barluenga M, Meyer A (2004) The Midas cichlid species complex: incipient sympatric speciation in Nicaraguan cichlid fishes? *Mol Ecol* 13, 2061-2076.

Barluenga M, Stolting KN, Salzburger W, Muschick M, Meyer A (2006) Sympatric speciation in Nicaraguan crater lake cichlid fish. *Nature* 439, 719-723.

Barnes DKA, Conlan KE (2007) Disturbance, colonization and development of Antarctic benthic communities. *Philosophical Transactions of the Royal Society of London B: Biological Sciences* 362 (1477), 11-38.

*Barroso RM, Hilsdorf AWS, Modeira HLM, Cabello PH, Traub-Cseko YM (2005) Genetic diversity of wild and cultured populations of *Brycon opalinus* (Cuvier, 1819) (Characiforme, Characidae, Bryconiae) using microsatellites. *Aquaculture* 247, 51-65.

*Barson NJ, Cable J, van Oosterhout C (2009) Population genetic analysis of microsatellite variation of guppies (*Poecilia reticulata*) in Trinidad and Tobago: evidence for a dynamic

source-sink metapopulation structure, founder events and population bottlenecks. *Journal of Evolutionary Biology* 22 (3), 485-497.

*Batista JS, Alves-Gomes JA (2006) Phylogeography of *Brachyplatystoma rousseauxii* (Siluriformes - Pimelodidae) in the Amazon Basin offers preliminary evidence for the first case of "homing" for an Amazonian migratory catfish. *Genetics and Molecular Research* 5 (4), 723-740.

*Bay LK, Caley MJM, Crozier RH (2008) Meta-population structure in a coral reef fish demonstrated by genetic data on patterns of migration, extinction and re-colonisation. *BMC Evolutionary Biology* 8 (1), 248.

Bellwood DR (1990) A new fossil fish *Phyllopharyngodon longipinnis* gen. et sp. nov. (family Labridae) from the Eocene, Monte Bolca, Italy. *Studi e Ricerche sui Giacimenti Terziari di Bolca* 6, 149-160.

Bellwood DR (1994) A phylogenetic study of the parrotfishes family Scaridae (Pisces:Labroidei), with a revision of genera. *Records of the Australian Museum Supplement* 20, 1-86.

Bellwood DR, Wainwright PC (2002) The history and biogeography of fishes on coral reefs. In: *Coral reef fishes: dynamics and diversity in a complex ecosystem* (ed. Sale PF), pp. 5-32. Academic Press San Diego, CA.

Benjamini Y, Hochberg Y (1995) Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Society B: Methodological* 57 (1), 289-300.

Benton MJ (1993) *The Fossil record 2*. Chapman & Hall, London, UK.

Berg LS (1958) *System der rezenten und fossilen Fischartigen und Fische*. VEB Verlag der Wissenschaften, Berlin.

Bernatchez L, Chouinard A, Guoqing L (1999) Integrating molecular genetics and ecology in studies of adaptive radiation: whitefish, *Coregonus* sp., as a case study. *Biological Journal of the Linnean Society* 68, 173-194.

Bernatchez L, Wilson CC (1998) Comparative phylogeography of Nearctic and Palearctic fishes. *Molecular Ecology* 7, 431-452.

Bouton N, Witte F, van Alphen JJ, Seehausen O (1999) Local adaptations in populations of rock-dewlling haplochromines (Pisces: Cichlidae) from southern Lake Victoria. *Proc R Soc Lond B Biol Sci* 266, 366-360.

Brenner M, Buck B, Cordes S, Dietrich L, Jacob U, Mintenbeck K, Schröder A, Brey T, Knust R, Arntz WE (2001) The role of iceberg scours in niche separation within the Antarctic fish genus *Trematomus*. *Polar Biology* 24 (7), 502-507.

*Caldera EJ, Bolnick DI (2008) Effects of colonization history and landscape structure on genetic variation within and among threespine stickleback (*Gasterosteus aculeatus*) populations in a single watershed. *Evolutionary Ecology Research* 10, 575-598.

Carreras-Carbonell J, Macpherson E, Pascual M (2005) Rapid radiation and cryptic speciation in mediterranean triplefin blennies (Pisces: Tripterygiidae) combining multiple genes. *Molecular Phylogenetics and Evolution* 37 (3), 751-761.

*Carvalho-Costa LF, Hatanka T, Galetti Jr. PM (2008) Evidence of lack of population substructuring in the Brazilian freshwater fish *Prochilodus costatus*. *Genetics and Molecular Biology* 31 (1), 377-380.

*Chabot CL, Allen LG (2009) Global population structure of the tope (*Galeorhinus galeus*) inferred by mitochondrial control region sequence data. *Molecular Ecology* 18, 545-552.

*Chauhan T, Lal KK, Mohindra V, Singh R, Punia P, Gopalakrishnan A, Sharma PC, Lakra W (2007) Evaluating genetic differentiation in wild populations of the Indian major carp, *Cirrhinus mrigala* (Hamilton-Buchanan, 1882): Evidence from allozyme and microsatellite markers. *Aquaculture* 269, 135-149.

Chen L, DeVries AL, Cheng, CHC (1997) Evolution of antifreeze glycoprotein gene from a trypsinogen gene in Antarctic notothenioid fish. *PNAS* 94 (8), 3811-3816.

*Chen S, Liu T, Li Z, Gao T (2008) Genetic population structuring and demographic history of red spotted grouper (*Epinephelus akaara*) in South and East China Sea. *African Journal of Biotechnology* 7 (20), 3554-3562.

Cheng CHC, Chen L, Near TJ, Jin Y (2003) Functional antifreeze glycoprotein genes in temperate-water New Zealand nototheniid fish infer an Antarctic evolutionary origin. *Molecular Biology and Evolution* 20 (11), 1897-1908.

*Chow S, Suzuki N, Brodeur RD, Ueno Y (2009) Little population structuring and recent evolution of the Pacific saury (*Cololabis saira*) as indicated by mitochondrial and nuclear DNA sequence data. *Journal of Experimental Marine Biology and Ecology* 369, 17-21.

Clark MS, Fraser KPP, Burns G, Peck LS (2008) The HSP70 heat shock response in the Antarctic fish *Harpagifer antarcticus*. *Polar Biology* 31 (2), 171-180.

Clarke A, Johnston IA (1996) Evolution and adaptive radiation of antarctic fishes. *Trends in Ecology and Evolution* 11 (5), 212-218.

Cocca E, Ratnayake-Lecamwasam M, Parker SK, Camardella L, Ciaramella M, di Prisco G, Detrich III HW (1995) Genomic remnants of alpha-globin genes in the hemoglobinless antarctic icefishes. *PNAS* 92 (6), 1817-1821.

Cohen AS, Lezzar KE, Tiercelin JJ, Soreghan M (1997) New paleogeographic and lake-level reconstructions of Lake Tanganyika: implications for tectonic, climatic and biological evolution in a rift lake. *Basin Research* 7, 107-132.

Coyne JA, Orr HA (2004) *Speciation* Sinauer Associates, Sunderland, Massachusetts.

*Craig MT, Eble JA, Bowen BW, Robertson DR (2007) High genetic connectivity across the Indian and Pacific Oceans in the reef fish *Myripristis berndti* (Holocentridae). *Marine Ecology Progress Series* 334, 245-254.

Crandall KA, Bininda-Emonds ORP, Mace GM, Wayne RK (2000) Considering evolutionary processes in conservation biology. *Trends in Ecology and Evolution* 15 (7), 290-295.

*Curley BG, Gillings MR (2009) Population connectivity in the temperate damselfish *Parma microlepis*: analyses of genetic structure across multiple spatial scales. *Marine Biology* 156 (3), 381-393.

*Danancher D, Izquierdo JI, Garcia-Vazquez E (2008) Microsatellite analysis of relatedness structure in young of the year of the endangered *Zingel asper* (Percidae) and implications for conservation. *Freshwater Biology* 53 (3), 546-557.

Dawson MN, Waples RS, Bernardi (2006) Phylogeography. In: *The Ecology of Marine Fishes* (eds. Allan LG, Pondella DJ, Horn MH), pp. 26-54. University of California Press, CA.

Debes PV, Zachos FE, Hanel R (2008) Mitochondrial phylogeography of the European sprat (*Sprattus sprattus* L., Clupeidae) reveals isolated climatically vulnerable populations in the Mediterranean Sea and range expansion in the northeast Atlantic. *Molecular Ecology* 17, 3873-3888.

de Graaf M, Dejen E, Osse JWM, Sibbing FA (2008) Adaptive radiation of Lake Tana's (Ethiopia) *Labeobarbus* species flock (Pisces, Cyprinidae). *Marine and Freshwater Research* 59, 391-407.

Dieckmann U, Doebeli M (1999) On the origin of species by sympatric speciation. *Nature* 400, 354-357.

*Dillane E, Cross MC, McGinnity P, Coughlan JP, Galvin PT, Wilkins NP, Cross TF (2007) Spatial and temporal patterns in microsatellite DNA variation of wild Atlantic salmon, *Salmo salar*, in Irish rivers. *Fisheries Management and Ecology* 14, 209-219.

*Dillane E, McGinnity P, Coughlan JP, Cross MC, de Eyto E, Kenchington E, Prodöhl P, Cross TF (2008) Demographics and landscape features determine intrariver population structure in Atlantic salmon (Salmo salarL.): the case of the River Moy in Ireland. *Molecular Ecology* 17 (22), 4786-4800.

Dimitrov D, Arnedo MA, Ribera C (2008) Colonization and diversification of the spider genus *Pholcus* Walckenaer, 1805 (Araneae, Pholcidae) in the Macaronesian archipelagos: evidence for long-term occupancy yet rapid recent speciation. *Molecular Phylogenetics and Evolution* 48, 596-614.

di Prisco G, Cocca E, Parker SK, Detrich III HW (2002) Tracking the evolutionary loss of hemoglobin expression by the white-blooded Antarctic icefishes. *Gene* 295 (2), 185-191.

Donaldson TJ (1995). Courtship and spawning of nine species of wrasses (Labridae) from the western Pacific. *Jap J Ichthyol* 42, 311-319.

*Drew J, Allen GR, Kaufmann L, Barber PH (2008) Endemism and regional color and genetic differences in five putatively cosmopolitan reef fishes. *Conservation Biology* 22 (4), 965-975.

*Dudgeon CL, Broderick D, Ovenden JR (2009) IUCN classification zones concord with, but underestimate, the population genetic structure of the zebra shark *Stegostoma fasciatumin* the Indo-West Pacific. *Molecular Ecology* 18 (2), 248-261.

*Duftner N, Sefc KM, Koblmuller S, Nevado B, Verheyen E, Phiri H, Sturmbauer C (2006) Distinct population structure in a phenotypically homogeneous rock-dwelling cichlid fish from Lake Tanganyika. *Mol Ecol* 15, 2381-2395.

Eastman JT (1993) *Antarctic fish biology: evolution in a unique environment*. Academic Press, San Diego, CA.

Eastman JT (2005) The nature of the diversity of Antarctic fishes. *Polar Biology* 28, 93-107.

Eastman JT, Clarke A (1998) A comparison of adaptive radiations of Antarctic fish with those of non-Antarctic fish. In: *Fishes of Antarctica. A biological overview* (eds. di Prisco G, Pisano E, Clarke A), pp. 3-26. Springer, Milan, Italy.

*Eble JA, Toonen RJ, Bowen BW (2009) Endemism and dispersal: comparative phylogeography of three surgeonfishes across the Hawaiian Archipelago. *Marine Biology* 156 (4), 689-698.

*Elmer KR, van Houdt JKJ, Meyer A, Volckaert FAM (2008) Population genetic structure of North American burbot (*Lota lota maculosa*) across the Nearctic and at its contact zone with Eurasian burbot (*Lota lota lota*). *Canadian Journal of Fisheries and Aquatic Sciences* 65 (11), 2412-2426.

Evans CW, Cziko PA, Cheng CHC, DeVries AL (2005) Spawning behaviour and early development in the naked dragonfish *Gymnodraco acuticeps*. *Antarctic Science* 17 (3), 319-327.

Feulner PG, Kirschbaum F, Mamonekene V, Ketmaier V, Tiedemann R (2007) Adaptive radiation in African weakly electric fish (Teleostei: Mormyridae: Campylomormyrus): a combined molecular and morphological approach. *J Evol Biol* 20, 403-414.

Floeter SR, Rocha LA, Robertson DR*, et al.* (2008) Atlantic reef fish biogeography and evolution. *Journal of Biogeography* 35, 22-47.

*Fontaine PM, Dodson JJ, Bernatchez L, Slettan A (1997) A genetic test of metapopulation structure in Atlantic salmon (*Salmo salar*) using microsatellites. *Canadian Journal of Fisheries and Aquatic Sciences* 54 2434-2442.

*Francisco SM, Cabral H, Vieira MN, Almada VC (2006) Contrasts in genetic structure and historical demography of marine and riverine populations of *Atherina* at similar geographical scales. *Estuarine, Coastal and Shelf Science* 69, 655-661.

*Froukh T, Kochzius M (2007) Genetic population structure of the endemic fourline wrasse (*Larabicus quadrilineatus*) suggests limited larval dispersal distances in the Red Sea. *Molecular Ecology* 16, 1359-1367.

*Galarza JA, Carbonell-Carreras J, Macpherson E, Pascual M, Roques S, Turner G, Rico C (2009) The influence of oceanographic fronts and early-life-history traits on connectivity among littoral fish species. *PNAS* 106 (5), 1473-1478.

Garant D, Forde SE, Hendry AP (2007) The multifarious effects of dispersal and gene flow on contemporary adaptation. *Functional Ecology* 21 (3), 434-443.

Gavrilets S (2004) *Fitness landscapes and the origin of species.* Princeton University Press, Princeton, New Jersey.

Gavrilets S, Vose A (2005) Dynamic patterns of adaptive radiation. *Proc Natl Acad Sci U S A* 102, 18040-18045.

Gavrilets S, Vose A, Barluenga M, Salzburger W, Meyer A (2007) Case studies and mathematical models of ecological speciation. 1. Cichlids in a crater lake. *Mol Ecol* 16, 2893-2909.

Ghiselin MT (1969) The evolution of hermaphroditism among animals. *Q Rev Biol* 44, 189-208.

Goodwin NB, Balshine-Earn S, Reynolds JD (1998) Evolutionary transitions in parental care in cichlid fish. *Proc R Soc Lond B Biol Sci* 265, 2265-2272.

*Goswami M, Thangaraj K, Chaudhary BK, Bjaskar, LVSK, Gopalakrishnan A, Joshi MB, Singh L, Lakra WS (2009) Genetic heterogeneity in the Indian stocks of seahorse (*Hippocampus kuda* and *Hippocampus trimaculatus*) inferred from mtDNA cytochrome b gene. *Hydrobiologia* 621, 213-221.

Grant PR, Grant BR (2002) Unpredictable evolution in a 30-year study of Darwin's finches. *Science* 296, 707-711.

Grant PR, Grant BR (2006) Evolution of character displacement in Darwin's finches. *Science* 313, 224-226.

*Grunwald C, Stabile J, Waldman JR, Gross R, Wirgin I (2002) Population genetics of shortnose sturgeon *Acipenser brevirostrum* based on mitochondrial DNA control region sequences. *Molecular Ecology* 11, 1885-1898.

*Grunwald C, Maceda L, Waldman JR, Stabile J, Wirgin I (2008) Conservation of Atlantic sturgeon *Acipenser oxyrinchus oxyrinchus*: delineation of stock structure and distinct population segments. *Conservation Genetics* 9, 1111-1124.

*Guy TJ, Gresswell RE, Banks MA (2008) Landscape-scale evaluation of genetic structure among barrier-isolated populations of coastal cutthroat trout, *Oncorhynchus clarkii clarkii*. *Canadian Journal of Fisheries and Aquatic Sciences* 65, 1749-1762.

Hanel R, Westneat MW, Sturmbauer C (2002)Phylogenetic relationships, evolution of broodcare behavior, and geographic speciation in the Wrasse tribe Labrini. *Journal of Molecular Evolution* 55, 776-789.

*Hatanaka T, Henrique-Silva F, Galetti Jr. PM (2006) Population Substructuring in a Migratory Freshwater Fish Prochilodus argenteus (Characiformes, Prochilodontidae) from the São Francisco River. *Genetica* 126, 153-159.

*Hepburn RI, Sale PF, Dixon B, Heath DD (2009) Genetic structure of juvenile cohorts of bicolor damselfish (*Stegastes partitus*) along the Mesoamerican barrier reef: chaos through time. *Coral Reefs* 28 (1), 277-288.

Herder F, Nolte AW, Pfaender J, Schwarzer J, Hadiaty RK, Schliewen UK (2006) Adaptive radiation and hybridization in Wallace's Dreamponds: evidence from sailfin silversides in the Malili Lakes of Sulawesi. *Proc Biol Sci* 273, 2209-2217.

Hey J, Nielsen R (2007) Integration within the Felsenstein equation for improved Markov chain Monte Carlo methods in population genetics. *PNAS* 104 (8), 2785-2790.

*Hickey AJR, Lavery SD, Hannan DA, Baker CS, Clements KD (2009) New Zealand triplefin fishes (family Tripterygiidae): contrasting population structure and mtDNA diversity within a marine species flock. *Molecular Ecology* 18, 680-696.

Higashi M, Takimoto G, Yamamura N (1999) Sympatric speciation by sexual selection. *Nature* 402, 523-526.

Hofmann GE, Buckley BA, Airaksinen S, Keen JE, Somero GN (2000) Heat-shock protein expression is absent in the antarctic fish Trematomus bernacchii (family Nototheniidae). *Journal of Experimental Biology* 203 (15), 2331-2339.

*Horne JB, van Herwerden L, Choat JH, Robertson DR (2008) High population connectivity across the Indo-Pacific: Congruent lack of phylogeographic structure in three reef fish congeners. *Molecular Phylogenetics and Evolution* 49 (2), 629-638.

*Hrbek T, Farias IP, Crossa M, Sampaio I, Porto JIR, Meyer A (2005) Population genetic analysis of *Arapaima gigas*, one of the largest freshwater fishes of the Amazon basin: implications for its conservation. *Animal Conservation* 8, 297-308.

Hsü KJ, Montadert L, Bernoulli D, Cita MB, Erickson A, Garrison RE, Kidd RB, Mèlierés F, Müller C, Wright R (1977) History of the Mediterranean salinity crisis. *Nature* 267, 399-403.

Hsü KJ, Ryan WBF, Cita MB (1973) Late Miocene desiccation of the Mediterranean. *Nature* 242, 240-244.

*Hubert N, Duponchelle F, Nuñez J, Rivera R, Bonhomme F, Renno JF (2007) Isolation by distance and Pleistocene expansion of the lowland populations of the white piranha *Serrasalmus rhombeus*. *Molecular Ecology* 16, 2488-2503.

Hulsey (2006) Function of a key morphological innovation: fusion of the cichlid pharyngeal jaw. *Proc Biol Sci* 263, 669-675.

Janko K, Lecointre G, DeVries AL, Couloux A, Cruaud C, Marshall C (2007) Did glacial advances during the Pleistocene influence differently the demographic histories of benthic and pelagic Antarctic shelf fishes? - Inferences from intraspecific mitochondrial and nuclear DNA sequence diversity. *BMC Evolutionary Biology* 7, 220.

Johnson TC, Scholz CA, Talbot MR, Kelts K, Ricketts RD, Ngobi G, Beuning K, Ssemmanda II, McGill JW (1996) Late Pleistocene Desiccation of Lake Victoria and Rapid Evolution of Cichlid Fishes. *Science* 273, 1091-1093.

*Jones CD, Anderson ME, Balushkin AV, Duhamel G, Eakin RR, Eastman JT, Kuhn KL, Lecointre G, Near TJ, North AW, Stein DL, Vacchi M, Detrich III HW (2008) Diversity, relative abundance, new locality records and population structure of Antarctic demersal fishes from the northern Scotia Arc islands and Bouvetøya. *Polar Biology* 31, 1481-1497.

Juan C, Ibrahim KM, Oromí P, Hewitt GM (1998) The phylogeography of the darkling beetle, *Hegeter politus*, in the eastern Canary Islands. *Proceedings of the Royal Society of London B: Biological Sciences* 265 (1391), 135-140.

Kazancioglu E, Near TJ, Hanel R, Wainwright PC (2009) Influence of feeding functual morphology and sexual selection on diversification rate of parrotfishes (Scaridae). Submitted to *Proceedings of the Royal Society B*.

Kellermann AK (1991) Egg and larval drift of the Antarctic fish *Notothenia coriiceps*. *Cybium* 15 (3), 199-210.

*Kitanishi S, Yamamoto T, Higashi S (2009) Microsatellite variation reveals fine-scale genetic structure of masu salmon, *Oncorhynchus masou*, within the Atsuta River. *Ecology of Freshwater Fish* 18, 65-71.

Klingenberg CP, Ekau W (1996) A combined morphometric and phylogenetic analysis of an ecomorphological trend: pelagization in Antarctic fishes (Perciformes: Nototheniidae). *Biological Journal of the Linnean Society* 59, 143-177.

Kocher TD (2004) Adaptive evolution and explosive speciation: the cichlid fish model. *Nature Reviews Genetics* 5, 288-298.

Kock KH (2005) Antarctic icefishes (Channichthyidae): a unique family of fishes. A review, part I. *Polar Biology* 28 (11), 862-895.

Kock KH, Kellermann AK (1991) Reproduction in Antarctic notothenioid fish. *Antarctic Science* 3 (2), 125-150.

Kock KH, Pshenichnov LK, DeVries AL (2006) Evidence for egg brooding and parental care in icefish and other notothenioids in the Southern Ocean. *Antarctic Science* 18 (2), 223-227.

Kondrashov AS, Kondrashov FA (1999) Interactions among quantitative traits in the course of sympatric speciation. *Nature* 400, 351-354.

Kornfield I, Carpenter KE (1984) Cyprinids of Lake Lanao, Philippines: Taxonomic validity, evolutionary rates and speciation scenarios. In: *Evolution of Fish Species Flocks* (eds. Echelle AA, Kornfield I), pp. 69-84. University of Maine at Orono Press, Orono, Maine.

Kornfield I, Smith PF (2000) African Cichlid Fishes: Model systems for evolutionary biology. *Annu. Rev. Ecol. Syst.* 31, 163-196.

*Kuhn KL, Gaffney PM (2006) Preliminary assessment of population structure in the mackerel icefish (*Champsocephalus gunnari*). *Polar Biology* 29, 927-935.

La Mesa M, Ashford J (2008) Age and early life history of juvenile scotia sea icefish, *Chaenocephalus aceratus*, from Elephant and the South Shetland Islands. *Polar Biology* 31 (2), 221-228.

Leis JM (2006) Are larvae of demersal fishes plankton or nekton? *Advances in Marine Biology* 51, 57-141.

Lessios HA (1998) The first stage of speciation as seen in organisms separated by the Isthmus of Panama. In: *Endless Forms: Species and Speciation* (eds. Howard DH, Berlocher SH), pp. 186-201. Oxord University Press, Oxford, UK.

Liem KF (1973) Evolutionary strategies and morphological innovations: cichlid pharyngeal jaws. *Systematic Zoology* 22, 425-441.

*Lin J, Quinn TP, Hilborn R, Hauser L (2008a) Fine-scale differentiation between sockeye salmon ecotypes and the effect of phenotype on straying. *Heredity* 101 (4), 341-350.

*Lin J, Zigeler E, Quinn TP, Hauser L (2008b) Contrasting patterns of morphological and neutral genetic divergence among geographically proximate populations of sockeye salmon *Oncorhynchus nerka* in Lake Aleknagik, Alaska. *Journal of Fish Biology* 73 (8), 1993-2004.

Loeb VJ, Kellermann AK, Koubbi P, North AW, White MG (1993) Antarctic larval fish assemblages: a review. *Bulletin of Marine Science* 53 (2), 416-449.

Lourie SA, Vincent ACJ (2004) A marine fish follows Wallace's Line: the phylogeography of the three-spot seahorse (*Hippocampus trimaculatus*, Syngnathidae, Teleostei) in Southeast Asia. *Journal of Biogeography* 31, 1975-1985.

Losos JB, Jackmann TR, Larson A, De Queiroz K, Rodrigues-Schettino L (1998) Contingency and determinism in replicated adaptive radiations of island lizards. *Science* 279, 2115-2118.

Luterbacher HP, Ali JR, Brinkhuis H, Gradstein FM, Hooker JJ, Monechi S, Ogg JG, Powell J, Röhl U, Sanfilippo A, Schmitz B (2004) The Paleogene period. In: *A geologic time scale* (ed. Gradstein F, Ogg J, Smith A), pp. 384-408. Cambridge University Press, Cambridge.

Mabuchi K, Miya M, Azuma Y, Nishida M (2007) Independent evolution of the specialized pharyngeal jaw apparatus in cichlid and labrid fishes. *BMC Evol Biol* 7, 10.

*Markert JA, Arnegard ME, Danley PD, Kocher TD (1999) Biogeography and population genetics of the Lake Malawi cichlid *Melanochromis auratus*: habitat transience, philopatry and speciation. *Molecular Ecology* 8, 1013-1026.

*Matschiner M, Hanel R, Salzburger W (2009) Gene flow by larval dispersal in the Antarctic notothenioid fish *Gobionotothen gibberifrons*. *Molecular Ecology*, in press.

Mayr E (1942) *Systematics and the Origin of Species* Columbia University Press, New York.

Mayr E (1963) *Animal Species and Evolution* Harvard University Press, Cambridge.

McCune AR (1997) How fast is speciation? Molecular, geological and phylogenetic evidence from adaptive radiations of fishes. In: *Molecular evolution and adaptive radiation* (eds. Givinish TJ, Sytsma KJ), pp. 585-610. Cambridge University Press, Cambridge, UK.

*Miller-Sims VC, Gerlach G, Kingsford MJ, Atema J (2008) Dispersal in the spiny damselfish, *Acanthochromis polyacanthus*, a coral reef fish species without a larval pelagic stage. *Molecular Ecology* 17 (23), 5036-5048.

Moritz C (1994) Defining 'Evolutionary Significant Units' for conservation. *Trends in Ecology and Evolution* 9 (10), 373-375.

Muss A, Robertson DR, Stepien CA, Wirtz P, Bowen BW (2001) Phylogeography of *Ophioblennius*: the role of ocean currents and geography in reef fish evolution. *Evolution* 55 (3), 561-572.

Mwanja WW, Armoudlian AS, Wandera SB, Kaufmann L, Wu L, Booton GC, Fuerst PA (2001) The bounty of minor lakes: the role of small satellite water bodies in evolution

and conservation of fishes in the Lake Victoria Region, East Africa. *Hydrobiologia* 458, 55-62.

Near TJ (2004) Estimating divergence times of notothenioid fishes using a fossil-calibrated molecular clock. *Antarctic Science* 16, 37-44.

Near TJ, Cheng CHC (2008) Phylogenetics of notothenioid fishes (Teleostei: Acanthomorpha): inferences from mitochondrial and nuclear gene sequences. *Mol Phylogenet Evol* 47, 832-840.

Near TJ, Parker SK, Detrich III HW (2006) A genomic fossil reveals key steps in hemoglobin loss by the antarctic icefishes. *Molecular Biology and Evolution* 23 (11), 2008-2016.

*Neethling M, Matthee CA, Bowie RCK, von der Heyden S (2008) Evidence for panmixia despite barriers to gene flow in the southern African endemic, *Caffrogobius caffer* (Teleostei: Gobiidae). *BMC Evolutionary Biology* 8, 325.

Nelson JS (2006) *Fishes of the World*.

Nepokroeff M, Sytsma KJ, Wagner WL, Zimmer EA (2003) Reconstructing ancestral patterns of colonization and dispersal in the Hawaiian understory tree genus *Psychotria* (Rubiaceae): A comparison of parsimony and likelihood approaches. *Systematic Biology* 52 (6), 820-838.

*Osinov AG, Gordeeva NV (2008) Variability of the microsatellite DNA and genetic differentiation of the populations of residual form of Dolly varden trout *Salvelinus malma krascheninnikovi* of Sakhalin. *Journal of Ichthyology* 48 (9), 691-706.

Ostbye K, Bernatchez L, Naesje TF, Himberg KJ, Hindar K (2005) Evolutionary history of the European whitefish Coregonus lavaretus (L.) species complex as inferred from mtDNA phylogeography and gill-raker numbers. *Mol Ecol* 14, 4371-4387.

*Pálsson S, Källman T, Paulsen J, Árnason E (2009) An assessment of mitochondrial variation in Arctic gadoids. *Polar Biology* 32 (3), 471-479.

Palumbi SR (1994) Genetic divergence, reproductive isolation, and marine speciation. *Annual Review of Ecology and Systematics* 25, 547-572.

Papetti C, Liò P, Rüber L, Patarnello T, Zardoya R (2007a) Antarctic fish mitochondrial genomes lack ND6 gene. *Journal of Molecular Evolution* 65 (5), 519-528.

*Papetti C, Susana E, La Mesa M, Kock KH, Patarnello T, Zane L (2007b) Microsatellite analysis reveals genetic differentiation between year-classes in the icefish *Chaenocephalus aceratus* at South Shetlands and Elephant Island. *Polar Biology* 30 (12), 1605-1613.

Parenti LR (1984) Biogeography of the Andean killifish genus *Orestias* with comments on the species flock concept. In: *Evolution of Fish Species Flocks* (eds. Echelle AA, Kornfield I), pp. 85-92. University of Maine at Orono Press, Orono, Maine.

Parenti P, Randall JE (2000) An annotated checklist of the species of the labroid fish families Labridae and Scaridae. *Ichthyol. Bull. (J.L.B. Smith Inst.)* 68, 1–97.

*Patarnello T, Marcato S, Zane L, Varotto V, Bargelloni L (2003) Phylogeography of the Chionodraco genus (Perciformes, Channichthydae) in the Southern Ocean. *Molecular Phylogenetics and Evolution* 28, 420-429.

Patarnello T, Volckaert FJ, Castilho R (2007) Pillars of Hercules: is the Atlantic-Mediterranean transition a phylogeographical break? *Molecular Ecology* 16, 4426-4444.

Patterson C (1993) Vertebrates, Osteichthyes: Teleostei. In: *The fossil record 2* (ed. Benton MJ), pp. 621-657. Chapman & Hall, London, UK.

*Pereyra R, Taylor MI, Turner GF, Rico C (2004) Variation in habitat preference and population structure among three species of the Lake Malawi cichlid genus Protomelas. *Mol Ecol* 13, 2691-2697.

Policansky D (1982) Sex change in animals and plants. *Ann Rev Ecol Syst* 13, 471-495.

*Prodocimo V, Tschá MK, Pie MR, Oliveira-Neto JF, Ostrensky A, Boeger WA (2008) Lack of genetic differentiation in the fat snook *Centropomus parallelus* (Teleostei: Centropomidae) along the Brazilian coast. *Journal of Fish Biology* 73 (8), 2075-2082.

Puebla O, Bermingham E, Guichard F (2008) Population genetic analyses of Hypoplectrus coral reef fishes provide evidence that local processes are operating during the early stages of marine adaptive radiations. *Mol Ecol* 17, 1405-1415.

Puebla O, Bermingham E, Guichard F, Whiteman E (2007) Colour pattern as a single trait driving speciation in Hypoplectrus coral reef fishes? *Proc Biol Sci* 274, 1265-1271.

*Ramon ML, Nelson PA, Martini E, Walsh WJ, Bernardi G (2008) Phylogeography, historical demography, and the role of post-settlement ecology in two Hawaiian damselfish species. *Marine Biology* 153, 1207-1217.

Randall JE (1967) Food habits of reef fishes of the West Indies. *Stud Trop Oceanog (Miami)* 5, 655–847.

Reinboth R (1970) Intersexuality in fishes. *Mem Soc Endocr* 18, 515-543.

Rice WR (1989) Analysing tables of statistical tests. *Evolution* 43, 223-225.

Rico C, Bouteillon P, van Oppen MJ, Knight ME, Hewitt GM, Turner GF (2003) No evidence for parallel sympatric speciation in cichlid species of the genus *Pseudotropheus* from north-western Lake Malawi. *J Evol Biol* 16, 37-46.

*Rico C, Turner GF (2002) Extreme microallopatric divergence in a cichlid species from Lake Malawi. *Mol Ecol* 11, 1585-1590.

*Rivera MAJ, Kelley CD, Roderick GK (2004) Subtle population genetic structure in the Hawaiian grouper, *Epinephelus quernus* (Serranidae) as revealed by mitochondrial DNA analyses. *Biological Journal of the Linnean Society* 81, 449-468.

*Rocha LA, Bass AL, Robertson DR, Bowen BW (2002) Adult habitat preferences, larval dispersal, and the comparative phylogeography of three Atlantic surgeonfishes (Teleostei: Acanthuridae). *Molecular Ecology* 11, 243-252.

Rocha LA, Bowen BW (2008) Speciation in coral-reef fishes. *Journal of Fish Biology* 72, 1101-1121.

Rocha LA, Craig MT, Bowen BW (2007) Phylogeography and the conservation of coral reef fishes. *Coral Reefs* 26 (3), 501-512.

Rocha LA, Rocha CR, Robertson DR, Bowen BW (2008) Comparative phylogeography of Atlantic reef fishes indicates both origin and accumulation of diversity in the Caribbean. *BMC Evol Biol* 8, 157.

*Rogers AD, Morley S, Fitzcharles E, Jarvis K, Belchier M (2006) Genetic structure of Patagonian toothfish (*Dissostichus eleginoides*) populations on the Patagonian Shelf and Atlantic and western Indian Ocean Sectors of the Southern Ocean. *Marine Biology* 149, 915-924.

Rögl F (1999) Mediterranean and Paratethys. Facts and hypotheses of an Oligocene to Miocene palaeogeography (short overview). *Geologica Carpathica* 50 (4), 339-349.

Rossiter A (1995) The cichlid fish assemblages of Lake Tanganyika: ecology, behavior and evolution of its species flocks. *Advances in Ecological Research* 26, 157-252.

Roy D, Docker MF, Haffner GD, Heath DD (2007a) Body shape vs. colour associated initial divergence in the Telmatherina radiation in Lake Matano, Sulawesi, Indonesia. *J Evol Biol* 20, 1126-1137.

Roy D, Paterson G, Hamilton PB, Heath DD, Haffner GD (2007b) Resource-based adaptive divergence in the freshwater fish Telmatherina from Lake Matano, Indonesia. *Mol Ecol* 16, 35-48.

Rüber L, Van Tassell JL, Zardoya R (2003) Rapid speciation and ecological divergence in the American seven-spined gobies (Gobiidae, Gobiosomatini) inferred from a molecular phylogeny. *Evolution* 57, 1584-1598.

Rüber L, Zardoya R (2005) Rapid cladogenesis in marine fishes revisited. *Evolution Int J Org Evolution* 59, 1119-1127.

Ruud JT (1954) Vertebrates without erythrocytes and blood pigment. *Nature* 173 (4410), 848-850.

Ryder OA (1986) Species conservation and systematics: the dilemma of subspecies. *Trends in Ecology and Evolution.* 1, 9-10.

*Saito T, Washio S, Dairiki K, Shimojo M, Itoi S, Sugita H (2008) High gene flow in *Girella punctata* (Perciformes, Kyphosidae) among the Japanese Islands inferred from partial sequence of the control region in mitochondrial DNA. *Journal of Fish Biology* 73 (8), 1937-1945.

Salzburger W (2008) To be or not to be a hamlet pair in sympatry. *Mol Ecol* 17, 1397-1399.

Salzburger W (2009) The interaction of sexually and naturally selected traits in the adaptive radiations of cichlid fishes. *Mol Ecol* 18, 169-185.

Salzburger W, Brandstätter A, Gilles A, Parson W, Hempel M, Sturmbauer C, Meyer A (2003) Phylogeography of the vairone (*Leuciscus souffia*, Risso 1826) in Central Europe. *Molecular Ecology* 12 (9), 2371-2386.

Salzburger W, Mack T, Verheyen E, Meyer A (2005) Out of Tanganyika: Genesis, explosive speciation, key-innovations and phylogeography of the haplochromine cichlid fishes. *BMC Evolutionary Biology* 5, 17.

Salzburger W, Meyer A (2004) The species flocks of East African cichlid fishes: recent advances in molecular phylogenetics and population genetics. *Naturwissenschaften* 91, 277-290.

Sanchez S, Dettai A, Bonillo C, Ozouf-Costaz C, Detrich III HW, Lecointre G (2007) Molecular and morphological phylogenies of the Antarctic teleostean family Nototheniidae, with emphasis on the Trematominae. *Polar Biology* 30, 155-166.

*Santos MCF, Ruffino ML, Farias IP (2007) High levels of genetic variability and panmixia of the tambaqui *Colossoma macropomum* (Cuvier, 1816) in the main channel of the Amazon River. *Journal of Fish Biology* 71, 33-44.

Savolainen V, Anstett MC, Lexer C*, et al.* (2006) Sympatric speciation in palms on an oceanic island. *Nature* 441, 210-213.

Schliewen UK, Tautz D, Paabo S (1994) Sympatric speciation suggested by monophyly of crater lake cichlids. *Nature* 368, 629-632.

Schluter D (2000) *The Ecology of Adaptive Radiation* Oxford University Press, New York.

Schluter D, McPhail JD (1992) Ecological character displacement and speciation in sticklebacks. *American Naturalist* 140, 85-108.

*Schultz JK, Feldheim KA, Gruber SH, Ashley MV, McGovern TM, Bowen BW (2008) Global phylogeography and seascape genetics of the lemon sharks (genus *Negaprion*). *Molecular Ecology* 17 (24), 5336-5348.

Seehausen O (2002) Patterns in fish radiation are compatible with Pleistocene desiccation of Lake Victoria and 14,600 year history for its cichlid species flock. *Proc R Soc Lond B Biol Sci* 269, 491-470.

Seehausen O (2006) African cichlid fish: a model system in adaptive radiation research. *Proc Biol Sci* 273, 1987-1998.

*Sefc KM, Baric S, Salzburger W, Sturmbauer C (2007) Species-specific population structure in rock-specialized sympatric cichlid species in Lake Tanganyika, East Africa. *J Mol Evol* 64, 33-49.

Sequeira AS, Lanteri AA, Albelo LR, Bhattacharya S, Sijapati M (2008) Colonization history, ecological shifts and diversification in the evolution of endemic Galápagos weevils. *Molecular Ecology* 17 (4), 1089-1107.

Shaw KL (1996) Sequential Radiations and Patterns of Speciation in the Hawaiian Cricket Genus Laupala Inferred from DNA Sequences. *Evolution* 50 (1), 237-255.

*Shaw PW, Arkhipkin A, Al-Khairulla A (2004) Genetic structuring of Patagonian toothfish populations in the Southwest Atlantic Ocean: the effect of the Antarctic Polar Front and deep-water troughs as barriers to genetic exchange. *Molecular Ecology* 13, 3293-3303.

*Shaw PW, Turner GF, Idid MR, Robinson RL, Carvalho GR (2000) Genetic population structure indicates sympatric speciation of Lake Malawi pelagic cichlids. *Proc R Soc Lond B Biol Sci* 267, 2273-2280.

*Silva-Oliveira GC, do Rêgo PS, Schneider H, Sampaio I, Vallinoto M (2008) Genetic characterisation of populations of the critically endangered Goliath grouper (*Epinephelus itajara*, Serranidae) from the Northern Brazilian coast through analyses of mtDNA. *Genetics and Molecular Biology* 31 (4), 988-994.

Smith CL (1967) Contribution to a theory of hermaphroditism. *Journal of Theoretical Biology* 17, 76-90.

Smith CL (1975) The evolution of hermaphroditism in fishes. In: *Intersexuality in the Animal Kingdom* (ed. Reinboth R), pp. 295-310. Springer Verlag, New York.

*Smith PJ, Gaffney PM (2005) Low genetic diversity in the Antarctic toothfish (*Dissostichus mawsoni*) observed with mitochondrial and intron DNA markers. *CCAMLR Science* 12, 43-51.

Stager JC, Johnson TC (2008) The late Pleistocene desiccation of Lake Victoria and the origin of its endemic biota. *Hydrobiologia* 596, 5-16.

Stearns SC (1976) Life-history tactics: a review of the ideas. *Q Rev Biol* 51, 3-45.

*Stepien CA, Murphy DJ, Strange RM (2007) Broad- to fine-scale population genetic patterning in the smallmouth bass *Micropterus dolomieu* across the Laurentian Great Lakes and beyond: an interplay of behaviour and geography. *Molecular Ecology* 16, 1605-1624.

*Streelman JT, Albertson RC, Kocher TD (2007) Variation in body size and trophic morphology within and among genetically differentiated populations of the cichlid fish, *Metriaclima zebra*, from Lake Malawi. *Freshwater Biology* 52 (3), 525-538.

Streelman JT, Alfaro M, Westneat MW, Bellwood DR, Karl SA (2002) Evolutionary history of the parrotfishes: biogeography, ecomorphology, and comparative diversity. *Evolution* 56, 961-971.

Sturmbauer C (1998) Explosive speciation in cichlid fishes of the African Great Lakes: a dynamic model of adaptive radiation. *Journal of Fish Biology* 53 (Supplement A), 18-36.

Sturmbauer C, Baric S, Salzburger W, Rüber L, Verheyen E (2001) Lake level fluctuations synchronize genetic divergences of cichlid fishes in African lakes. *Mol Biol Evol* 18, 144-154.

Sturmbauer C, Hainz U, Baric S, Verheyen E, Salzburger W (2003) Evolution of the tribe Tropheini from Lake Tanganyika: synchronized explosive speciation producing multiple evolutionary parallelism. *Hydrobiologia* 500, 51-64.

Taberlet P, Fumagalli L, Wust-Saucy A, Cosson J (1998) Comparative phylogeo-graphy and postglacial colonization routes in Europe. *Molecular Ecology* 7, 453-464.

*Taylor MI, Verheyen E (2001) Microsatellite data reveals weak population substructuring in *Copadichromis sp.* 'virginalis kajose', a demersal cichlid from Lake Malawi, Africa. *Journal of Fish Biology* 59, 593-604.

Taylor MS, Hellberg ME (2005) Marine radiations at small geographic scales: speciation in neotropical reef gobies (*Elacatinus*). *Evolution* 59, 374-385.

*Teterina VI, Sukhanova LV, Bogdanov BE, Anoshko PN, Kirilchik SV (2005) Genetic polymorphism of a pelagic fish species, little Baikal oilfish *Comephorus dybowski*, deduced from analysis of microsatellite loci. *Animal Genetics* 41 (7), 750-754.

Thatje S, Hillenbrand CD, Larter R (2005) On the origin of Antarctic marine benthic community structure. *Trends in Ecology and Evolution* 20 (10), 534-540.

*Tinti F, di Nunno C, Guarniero I, Talenti M, Tommasini S, Fabbri E, Picinetti C (2002) Mitochondrial DNA sequence variation suggests the lack of genetic heterogeneity in the Adriatic and Ionian stocks of *Sardina pilchardus*. *Marine Biotechnology* 4, 163-172.

*van Oppen MJ, Turner GF, Rico C, Deutsch JC, Ibrahim KM, Robinson RL, Hewitt GM (1997) Unusually fine-scale genetic structuring found in rapidly speciating Malawi cichlid fishes. *Proc R Soc Lond B Biol Sci* 264, 1803-1812.

*Vasconcellos AV, Vianna P, Paiva PC, Schama R, Solé-Cava A (2008) Genetic and morphometric differences between yellowtail snapper (*Ocyurus chrysurus*, Lutjanidae) populations of the tropical West Atlantic. *Genetics and Molecular Biology* 31, 308-316.

Verheyen E, Salzburger W, Snoeks J, Meyer A (2003) Origin of the superflock of cichlid fishes from Lake Victoria, East Africa. *Science* 300, 325-329.

Veron JEN (1995) *Corals in space and time: biogeography and evolution of the Scleractinia*. Comstock/Cornell, Ithaca, NY.

*Vonlanthen P, Excoffier L, Bittner D, Persat H, Neuenschwander S, Largiadèr CR (2007) Genetic analysis of potential postglacial watershed crossings in Central Europe by the bullhead (*Cottus gobio* L.). *Molecular Ecology* 16, 4572-4584.

Vonlanthen P, Roy D, Hudson AG, Largiader CR, Bittner D, Seehausen O (2009) Divergence along a steep ecological gradient in lake whitefish (*Coregonus* sp.). *J Evol Biol* 22, 498-514.

Wainwright PC, Bellwood DR, Westneat MW, Grubich JR, Hoey AS (2004) A functional morphospace for the skull of labrid fishes: patterns of diversity in a complex biomechanical system. *Biological Journal of the Linnean Society* 82, 1-25.

*Walter RP, Haffner GD, Heath DD (2009) Dispersal and population genetic structure of Telmatherina antoniae, an endemic freshwater Sailfin silverside from Sulawesi, Indonesia. *Journal of Evolutionary Biology* 22 (2), 314-323.

Westneat MW (1997) Phylogenetic relationships of labrid fishes: an analysis of morphological characters. *Am Zool* 37, 198A.

Westneat MW (1999) The living marine resources of the Western Central Pacific: FAO species identification sheets for fishery purposes. Family Labridae. *Food and Agriculture Organization of the United Nations* 6, 3381–3467.

Westneat MW, Alfaro ME (2005) Phylogenetic relationships and evolutionary history of the reef fish family Labridae. *Molecular Phylogenetics and Evolution* 36, 370-390.

White MG (1998) Development, dispersal and recruitment - a paradox for survival among Antarctic fish. In: *Fishes of Antarctica. A biological overview* (eds. di Prisco G, Pisano E, Clarke A). Springer-Verlag, Milano, Italy.

Williams GC (1966) *Adaptation and natural selection. A critique of some current evolutionary thought*. Princeston University Press, Princeton.

Wu C-I (2001) The genic view of the process of speciation. *J Evol Biol* 14, 851-865.

*Wu GCC, Chiang HC, Chen KS (2009) Population structure of albacore (*Thunnus alalunga*) in the Northwestern Pacific Ocean inferred from mitochondrial DNA. *Fisheries Research* 95, 125-131.

*Xiao Y, Takahashi M, Yanagimotot T, Zhang Y, Gao T, Yabe M, Sakurai Y (2008) Genetic variation and population structure of willowy flounder *Tanakius kitaharai* collected from Aomori, Ibaraki and Niigata in Northern Japan. *African Journal of Biotechnology* 7 (21), 3836-3844.

*Zane L, Bargelloni L, Bortolotto E, Papetti C, Simonato M, Varotto V, Patarnello T (2006) Demographic history and population structure of the Antarctic silverfish *Pleuragramma antarcticum*. *Molecular Evolution,* 15 (14), 4499-4511.

II   Population Structure

Most notothenioids are characterized by a benthic sedentary life style, and long range adult migration is rarely observed. Using mark-recapture experiments, individuals of *Dissostichus eleginoides* were found to travel almost 2 000 km in the southern Indian Ocean (Williams et al. 2002). A single stray specimen of the same species was even caught in the northwest Atlantic off Greenland, and provides the first example of transequatorial migration in notothenioid fishes (Møller et al. 2003). Given that even low levels of gene flow may offset the effects of genetic drift or natural selection (Slatkin 1987), the occasional long distance migrant might be sufficient to homogenize populations. However, *D. eleginops* is one of few notothenioids with a predominantly pelagic life style (Eastman 1993) and an enormous depth range of 3850 m (Laptikhovsky et al, 2006), and it is therefore likely that *D. eleginops* migration patterns differ substantially from those of the otherwise mostly benthic notothenioids.

The generally benthic life style of many notothenioid fishes, however, includes a prolonged pelagic stage of eggs, larvae, and juveniles (Kellermann 1986, Kellermann & Kock 1988, Kock & Kellermann 1991, Loeb et al. 1993). For example, larvae of the bathydraconid *Gymnodraco acuticeps* hatch only about 10 months post-fertilization (Evans et al. 2005), whereas the pelagic phase of *Chaenocephalus aceratus* may last up to 1.5 years (La Mesa & Ashford 2008). The contribution of larval dispersal to gene flow in marine systems has long been debated. Traditionally, marine populations have been considered demographically open, and interconnected by larval dispersal with oceanic currents (Caley et al. 1996). This view has changed in recent years as evidence for larval retention has accumulated (Swearer et al. 2002). Many larvae are capable of active vertical migration, which could, in combination with vertically stratified flows, suffice to avoid advection. Furthermore, many larvae, especially during later developmental stages, are able to swim against the current (Leis 2006).

Notothenioid fishes are ideal model systems to investigate the effect of larval dispersal on genetic population structure (Loeb et al. 1993). The populations of shelf-dwelling benthic notothenioids are often disjunct by deep water trenches, and oceanic currents that could potentially transport larvae are well-mapped and can be analysed in more detail with GPS-tagged drifting buoys (Lumpkin & Pazos 2007). Evidence for larval retention comes from systematic sampling schemes, that showed larvae of oceanic fish species to be present both over shelf areas and in the open ocean, whereas those of demersal fishes were only observed in the vicinity of shelf habitats (White 1998). On the other hand, larvae of *Notothenia coriiceps*, a benthic nototheniid, were found at a number of Scotia Sea sampling locations between the AP and South Georgia shelves, whereby positive catches occurred earlier in the southern Scotia Sea than near South Georgia, which suggests that larvae hatched near the AP and drifted towards South Georgia with the ACC (Kellermann 1991). An important contribution of larval dispersal to notothenioid biogeography is further suggested by the fact that only nototheniids and channichthyids with particularly long pelagic larval durations occur at the isolated island of Bouvetøya (Jones et al. 2008).

Numerous publications have investigated notothenioid population structure with molecular markers (reviewed in Matschiner et al. 2009; see below). In summary, these studies suggest

that geographic distance alone cannot account for the observed population structures and that instead the presence or absence of oceanographic discontinuities and currents (Clement et al. 1998; Shaw et al. 2004) has a measurable effect, thus suggesting a contribution of larval dispersal to inter-population gene flow.

If indeed larval advection with the ACC is the main reason for gene flow in the Scotia Sea, then gene flow should be unidirectional and with ocean currents. This hypothesis has been tested as part of this doctoral work. The migration patterns of seven notothenioid species were investigated with mitochondrial sequences and microsatellite markers, and the observed directionality was compared with the trajectories of passively floating GPS-tracked buoys (Lumpkin & Pazos 2007). The results strongly support larvae as the agent of gene flow in benthic notothenioid species. The analysis of microsatellite markers was improved and facilitated with the development of TANDEM, a user-friendly software for automated binning of microsatellite allele lengths.

The population genetic analyses of this doctoral work resulted in two articles, of which the first appeared in *Molecular Ecology*, and the second has been submitted to *Polar Biology*. An "Application Note" in *Bioinformatics* describes the software TANDEM:

2.1 Matschiner M, Hanel R, Salzburger W (2009) Gene flow by larval dispersal in the Antarctic notothenioid fish *Gobionotothen gibberifrons*. *Molecular Ecology* 18: 2574-2587.

2.2 Damerau M, Matschiner M, Salzburger W, Hanel R: Comparative population genetics of seven notothenioid fish species reveals high levels of gene flow along ocean currents in the southern Scotia Arc, Antarctica. Submitted to *Polar Biology*.

2.3 Matschiner M, Salzburger W (2009) TANDEM: integrating automated allele binning into genetics and genomics workflows. *Bioinformatics* 25: 1982-1983.

References

Caley MJ, Carr MH, Hixon MA, et al. (1996) Recruitment and the local dynamics of open marine populations. *Annu Rev Ecol Syst* 27: 477–500.

Clement O, Ozouf-Costaz C, Lecointre G, Berrebi P (1998) Allozymic polymorphism and phylogeny of the family Channichthyidae. In: *Fishes of Antarctica. A Biological Overview* (Eds. di Prisco G, Pisano E, Clarke A) pp. 299–309. Springer-Verlag, Milan, Italy.

Eastman JT (1993) Antarctic Fish Biology: Evolution in a Unique Environment. Academic Press, Inc., San Diego, CA.

Evans CW, Cziko PA, Cheng C-HC, DeVries AL (2005) Spawning behaviour and early development in the naked dragonfish *Gymnodraco acuticeps*. *Antarct Sci* 17: 319-327.

Jones CD, Anderson ME, Balushkin AV, et al. (2008) Diversity, relative abundance, new locality records and population structure of Antarctic demersal fishes from the northern Scotia Arc islands and Bouvetøya. *Polar Biol* 31: 1481-1497.

Kellermann AK (1986) Zur Biologie der Jugendstadien der Notothenioidei (Pisces) an der Antarktischen Halbinsel. *Ber Polarforsch* 31: 155 pp.

Kellermann AK (1991) Egg and larval drift of the Antarctic fish *Notothenia coriiceps*. *Cybium* 15: 199–210.

Kellermann AK, Kock K-H (1988) Pattern of spatial and temporal distribution and their variation in early life stages of Antarctic fish in the Antarctic Peninsula region. In*: Antarctic Ocean and Resources Variability* (Ed. Sahrhage D). Springer-Verlag, Berlin, Germany.

Kock K-H, Kellermann AK (1991) Reproduction in Antarctic notothenioid fish. *Antarct Sci* 3: 125–150.

La Mesa M, Ashford J (2008) Age and early life history of juvenile scotia sea icefish, *Chaenocephalus aceratus*, from Elephant and the South Shetland Islands. *Polar Biol* 31: 221–228.

Laptikhovsky V, Arkhipkin AI, Brickle P (2006) Distribution and reproduction of the Patagonian toothfish *Dissostichus eleginoides* Smitt around the Falkland Islands. *J Fish Biol* 68: 849–861.

Leis JM (2006) Are larvae of demersal fishes plankton or nekton? *Advances in Marine Biology* 51: 57–141.

Loeb VJ, Kellermann AK, Koubbi P, North AW, White MG (1993) Antarctic larval fish assemblages: a review. *B Mar Sci* 53: 416–449.

Lumpkin R, Pazos M (2007) Measuring surface currents with Surface Velocity Program drifters: the instrument, the data, and some recent results. In: *Lagrangian Analysis and Prediction of Coastal and Ocean Dynamics* (Eds. Mariano A, Rossby T, Kirwan D). Cambridge University Press, Cambridge*.* pp 39–67.

Møller PR, Nielsen JG, Fossen I (2003) Fish migration: Patagonian toothfish found off Greenland. *Nature* 421: 599–599.

Shaw PW, Arkhipkin AI, Al-Khairulla H (2004) Genetic structuring of Patagonian toothfish populations in the Southwest Atlantic Ocean: the effect of the Antarctic Polar Front and deep-water troughs as barriers to genetic exchange. *Mol Ecol* 13: 3293–3303.

Slatkin M (1987) Gene flow and the geographic structure of natural populations. *Science* 236: 787–792.

Swearer SE, Shima JS, Hellberg ME, et al. (2002) Evidence of self-recruitment in demersal marine populations. *B Mar Sci* 70: 251–271.

White MG (1998) Development, dispersal and recruitment - a paradox for survival among Antarctic fish. In: *Fishes of Antarctica. A biological overview* (Eds. di Prisco G, Pisano E, Clarke A). Springer-Verlag, Milan, Italy.

Williams RC, Tuck GN, Constable AJ, Lamb T (2002) Movement, growth and available abundance to the fishery of *Dissostichus eleginoides* Smitt, 1898 at Heard Island, derived from tagging experiments. *CCAMLR Science* 9: 33–48.

2.1 Gene flow by larval dispersal in the Antarctic notothenioid fish *Gobionotothen gibberifrons*

Matschiner M, Hanel R, Salzburger W

Gene flow by larval dispersal in the Antarctic notothenioid fish *Gobionotothen gibberifrons*

MICHAEL MATSCHINER,* REINHOLD HANEL† and WALTER SALZBURGER*

*Zoological Institute, University of Basel, Vesalgasse 1, CH-4051 Basel, Switzerland, †Institute of Fisheries Ecology, Johann Heinrich von Thünen-Institut, Federal Research Institute for Rural Areas, Forestry and Fisheries, Palmaille 9, D-22767 Hamburg, Germany

Abstract

The diversification of the teleost suborder Notothenioidei (Perciformes) in Antarctic waters provides one of the most striking examples of a marine adaptive radiation. Along with a number of adaptations to the cold environment, such as the evolution of antifreeze glycoproteins, notothenioids diversified into eight families and at least 130 species. Here, we investigate the genetic population structure of the humped rockcod (*Gobionotothen gibberifrons*), a benthic notothenioid fish. Six populations were sampled at different locations around the Scotia Sea, comprising a large part of the species' distribution range (*N* = 165). Our analyses based on mitochondrial DNA sequence data (352 bp) and eight microsatellite markers reveal a lack of genetic structuring over large geographic distances ($\Phi_{ST} \leq 0.058$, $F_{ST} \leq 0.005$, *P* values nonsignificant). In order to test whether this was due to passive larval dispersal, we used GPS-tracked drifter trajectories, which approximate movement of passive surface particles with ocean currents. The drifter data indicate that the Antarctic Circumpolar Current (ACC) connects the sampling locations in one direction only (west–east), and that passive transport is possible within the 4-month larval period of *G. gibberifrons*. Indeed, when applying the isolation-with-migration model in IMA, strong unidirectional west-east migration rates are detected in the humped rockcod. This leads us to conclude that, in *G. gibberifrons*, genetic differentiation is prevented by gene flow via larval dispersal with the ACC.

Keywords: adaptive radiation, population genetics, isolation-with-migration model, drifters

Received 15 December 2008; revised 25 March 2009; accepted 27 March 2009

Introduction

Adaptive radiation is the evolution of ecological and morphological diversity within a rapidly multiplying lineage (Schluter 2000). Only very few adaptive radiations are known from the marine realm, which is surprising given the numerous examples of adaptive radiations in freshwater systems (Salzburger 2009). One explanation for this observation could be that adaptive radiations are simply more apparent in geographically well-defined areas such as islands or lakes and less detectable in open systems such as oceans (Salzburger 2008). Some adaptive radiations in marine fishes are indeed characterized by their geographic circumscription. The colourful hamlet species complex (*Hypoplectrus*; family Serranidae), for example, is confined

Correspondence: W. Salzburger, Fax: +41 61 267 03 01, E-mail: walter.salzburger@unibas.ch

to the Caribbean Sea, where about a dozen of species have rather recently emerged (Puebla *et al.* 2008). Among the most species-rich marine adaptive radiations in teleosts is the one of notothenioid fishes in Antarctic waters that diversified into at least 130 species (Eastman 2005; Cziko & Cheng 2006). Today, the notothenioids dominate the Antarctic continental shelf and upper slope in terms of species number (47%) and fish biomass (90–95%) (Eastman & Clarke 1998). Antarctic shelf areas are separated from other continental shelves by the Antarctic Circumpolar Current (ACC) that reaches the ocean floor (Foster 1984) and transports more water than any other ocean current (Tomczak & Godfrey 2003). The Antarctic Polar Front (APF), among other oceanic frontal zones, delimits the Southern Ocean (Kock 1992), posing an oceanographic barrier to marine organisms and thermally isolating the continent (Shaw *et al.* 2004).

The remarkable diversification of the Notothenioidei has been accompanied by several morphological adaptations

and evolutionary innovations. Presumably most important for the adaptation to the Antarctic environment was the evolution of antifreeze glycoproteins from a trypsinogen progenitor (Chen *et al.* 1997). In the freezing waters of Antarctica, these proteins bind to the surface of forming ice crystals in blood and tissue and thus inhibit their further growth (DeVries 1988). On the other hand, some notothenioids lack otherwise common features. For example, the channichthyid family is characterized by the inability to synthesize haemoglobin, which is unique among vertebrates (Kock 2005). Channichthyidae have adapted to the lack of respiratory pigments with increased blood volume and cardiac output while at the same time maintaining a low metabolic rate. In addition, the mitochondrial ND6 gene (coding for NADH subunit 6) went amiss in notothenioids of the 'Antarctic clade' (Papetti *et al.* 2007a), and swim bladders are absent in all notothenioids (Eastman 1993). Reasons for the persistence of these presumably deleterious traits are difficult to interpret, and their influence on the notothenioid radiation is not yet known (Sidell & O'Brien 2006; Papetti *et al.* 2007a).

Here, we investigate the population structure of the humped rockcod (*Gobionotothen gibberifrons*), a benthic notoheniid with a depth range down to 750 m (Eastman 2005; Kompowski 1985). It is distributed along the northwestern Antarctic Peninsula (AP), around the South Shetland Islands including Elephant Island, the South Orkney Islands, and the islands and sea mounts of the Scotia Ridge (SR), including South Georgia (Fig. 1; DeWitt *et al.* 1990). The species spawns small eggs of around 2 mm in diameter in July and August during the austral winter. After 2 to 3

months of incubation, hatching occurs in October when larvae are *c.* 8 mm in length. Larvae become pelagic and feed mainly on copepods in the upper 100 m of the water column. The end of the larval phase is reached by mid-January to early February, at a standard length of 25 mm when most fin-rays are developed. The early juvenile stages return to a demersal lifestyle before the first winter (North 2001).

Characteristics of *G. gibberifrons* habitats differ between the AP and the SR. Due to its lower latitude, South Georgia water temperatures are higher than those at the AP over at least a part of the *G. gibberifrons* depth range. Furthermore, temperature variability decreases with latitude (Barnes *et al.* 2006). As a result of the later onset of the production cycle, *G. gibberifrons* spawning and hatching times at AP locations are delayed by about 1 month compared to South Georgia (Kock & Kellermann 1991). In addition, nutrient content of seawater differs between the two locations, with higher levels of nitrate and silicic acid being available at the AP (Silva S. *et al.* 1995; Whitehouse *et al.* 1996). It could thus be expected that local adaptation led to genetic differentiation between AP and SR populations. To test this hypothesis, we analysed the population genetic structure of six *G. gibberifrons* populations around the Scotia Sea using mitochondrial and nuclear DNA markers. Combining molecular and oceanographic data, we then evaluate whether adult migration or larval dispersal are agents of gene flow in Antarctic waters. We were particularly interested in the question whether neutral drift and/or gene flow affect local adaptation in notothenioids. This was based on the observation that in many adaptive radiations, the interplay of barriers
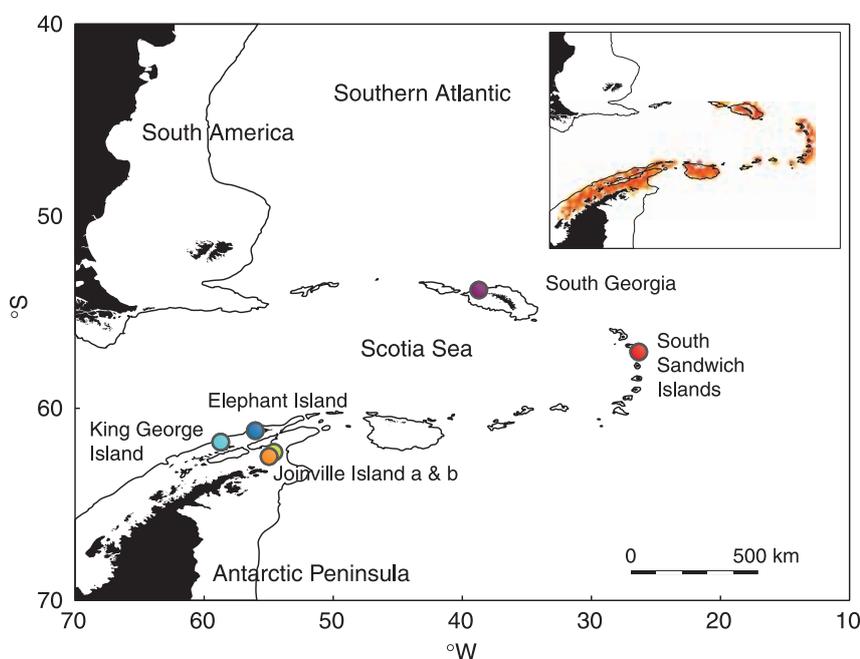

Fig. 1 *Gobionotothen gibberifrons* sampling sites at the tip of the AP, South Georgia, and the South Sandwich Islands. The solid line indicates the 1000-m depth contour. Inset: *G. gibberifrons* distribution range.

Table 1 Sampling sites for *Gobionotothen gibberifrons* around the Scotia Sea. Mean values are given for latitude, longitude and depth. *n*, sample size

Location	Latitude	Longitude	Depth	*n*
Elephant Island	61 13′S	55 53′W	144 m	49
Joinville Island A	62 15′S	55 18′W	356 m	30
Joinville Island B	62 26′S	55 37′W	240 m	33
King George Island	61 51′S	59 14′W	267 m	35
South Georgia	53 48′S	38 43′W	255 m	8
South Sandwich Islands	57 04′S	26 47′W	118 m	10

to gene flow and local adaptation are driving forces for allopatric (or parapatric) speciation (see, e.g. Mayr 1984; Schluter 2000; Rico & Turner 2002).

Materials and methods

Sample collection and DNA extraction

Sampling of *Gobionotothen gibberifrons* specimen was undertaken as part of the ICEFISH 2004 cruise with RV Nathaniel B. Palmer (Jones *et al*. 2008), and during expedition ANT-XXIII/8 with RV Polarstern in the austral summer 2006/2007. In total, 165 specimens were available from six locations around the Scotia Sea (Fig. 1 and Table 1). Muscle or fin tissue samples were taken from all specimens and preserved in 95% ethanol. Genomic DNA was extracted from *c*. 25 mm³ of muscle or fin tissue using the BioSprint 96 workstation (QIAGEN) according to the manufacturer's guidelines.

Mitochondrial DNA: D-loop

The hypervariable 3′ end of the mitochondrial control region was amplified in polymerase chain reactions (PCRs) using primers LPR-02 and HDL2 (Derome *et al*. 2002). The PCR mixture contained 2 L template DNA, 3.5 mM MgCl₂, 1.0 mM of each nucleotide, 0.2 M of each primer, 2.5 U *Taq* polymerase (QIAGEN) in 25 L 1× PCR buffer (QIAGEN) containing Tris-Cl, KCl and (NH₄)₂SO₄ and adjusted to pH 8.7 (20 C). Amplifications were performed in a Veriti thermal cycler (Applied Biosystems) with a cold-start PCR profile consisting of an initial predenaturation phase (3 min, 94 C), followed by 35 cycles of denaturing (30 s, 94 C), annealing (30 s, 52 C) and elongation phase (90 s, 72 C), and a final extended elongation phase (7 min, 72 C). PCR purification was done using the GenElute PCR Clean-Up kit (Sigma-Aldrich), following the manufacturer's protocol. Cycle sequencing was performed in forward direction using the BigDye Terminator version 3.1 Cycle Sequencing kit (Applied Biosystems), according to the manufacturer's instructions. Sequencing products were purified by sodium

acetate precipitation and run on a 3130 Genetic Analyzer (Applied Biosystems).

Sequence analysis resulted in an alignment of 162 sequences. The alignment was collapsed, but information about the frequency of haplotypes was kept. MODELTEST 3.7 (Posada & Crandall 1998) was run on the collapsed alignment to determine the best-fitting model of sequence evolution. Phylogenetic tree reconstruction was done using the maximum-likelihood method implemented in PAUP* 4.0b10 (Swofford 2003), and the model of sequence evolution selected by likelihood ratio test, HKY + Γ (Hasegawa *et al*. 1985). On the basis of the inferred phylogenetic tree, a haplotype genealogy was constructed.

A distance matrix of all haplotypes was calculated from this genealogy, and used in a hierarchical analysis of molecular variance (AMOVA) in order to compare genetic variation within populations, within predefined groups and among groups. Two different weighting schemes of transitions and transversions were applied: (i) using even weights, and (ii) taking into account the observed ratio of 2:1. In both cases, all possible groupings were assessed and ranked by among group variation. In addition, analogues to Wright's (1978) *F*-statistics were calculated in pairwise comparisons of all populations. Both analyses were done with ARLEQUIN 3.11 (Excoffier *et al*. 2005) and running 10 000 permutations.

The statistical power of the mitochondrial DNA (mtDNA) data set was assessed using POWSIM 4.0 (Ryman & Palm 2006). This software estimates the probability of false negatives for population differentiation, given an expected degree of divergence. Simulations were run with various combinations of N_e (effective population size) and *t* (time since divergence) to yield F_{ST} values of 0.001, 0.0025, 0.005, 0.01, and 0.02, both on a global level including all populations, and between pooled AP and SR populations. For every simulation, 1000 replicates were run, and default parameters were used for the number of dememorizations, batches, and iterations per batch.

The demographic history of *G. gibberifrons* was tested at the species level, with a mismatch analysis (Li 1977) of all 162 D-loop sequences. The distribution of pairwise mutational distances was fitted to a model of instantaneous population expansion by a generalized nonlinear least-square procedure as implemented in ARLEQUIN, taking into account the observed transition to transversion ratio. The validity of this model was tested by a parametric bootstrap approach running 10 000 bootstrap replicates. Time of population expansion (scaled by mutation rate) was estimated directly from the mismatch distribution and translated into absolute time in years (t_e), using the equation $t_e = \tau/2$, where is the mutation rate per locus per year. We used a mutation rate of 6.5–8.8% per million years (Myr) that was found in perciform fishes for the 3′ end of the mitochondrial control region (Sturmbauer *et al*. 2001).

Similar rates have been found in damselfishes (6.9–7.8%; Domingues *et al.* 2005), sculpins (9%; Volckaert *et al.* 2002), and salmonids (5–10%; Brunner *et al.* 2001).

Nuclear DNA: microsatellites

Nine microsatellite loci were cross-amplified using primers isolated from other notothenioid species (Table S1, Supporting information). Loci Trne35, Trne37, Trne53 and Trne66 were isolated from a nototheniid relative, *Trematomus newnesi* (van Houdt *et al.* 2006), while Cr38, Cr127, Cr170, Cr259 and Ca26 have been isolated from channichthyid notothenioids, *Chionodraco rastrospinosus* and *Chaenocephalus aceratus* (Papetti *et al.* 2006; Susana *et al.* 2007). With the exception of Cr127 and Cr259 (Papetti *et al.* 2007b), none of the loci have been cross-amplified before. All forward primers were fluorescently labelled. Amplifications were done in total volumes of 10 L using the QIAGEN Multiplex PCR kit.

Individual amplification volumes contained 0.8 L template DNA, 0.2 M forward and reverse primers in 1× QIAGEN Multiplex PCR Master Mix comprising HotStar *Taq* DNA Polymerase, nucleotides and 3 mM $MgCl_2$. DNA polymerase was activated in an initial activation step (15 min, 95 C), followed by 31–37 thermocycles (see Table S1) of denaturation (30 s, 94 C), annealing (90 s, 59 or 60 C) and extension phase (90 s, 72 C), and a final extension (10 min, 72 C). Amplified products were processed on a 3130 Genetic Analyzer (Applied Biosystems) with LIZ500 size standard (Applied Biosystems).

Microsatellite data were further analysed using GENE-MAPPER, version 4.0 (Applied Biosystems). All fragment sizes were automatically pre-analysed by the software and checked by eye. Data of 164 individuals met the quality criteria. We used TANDEM, version 0.9 (Matschiner & Salzburger 2009) for automated binning of allele sizes.

Binned alleles were statistically analysed using ARLEQUIN. Locus Cr170 was found to be monomorphic and was excluded from all further analyses. Pairwise tests of linkage disequilibrium (Slatkin 1994; Slatkin & Excoffier 1996) were performed on the eight remaining loci, running 1000 permutations. The software MICROCHECKER (Van Oosterhout *et al.* 2004) was used to test for null alleles, stuttering and large allele dropout. In addition, an analysis of molecular variance (AMOVA) was conducted, and *F*-statistics were calculated, again as implemented in ARLEQUIN. Ten thousand permutations were performed in both cases. Again, all possible groupings of the hierarchical AMOVA were assessed.

A population assignment test was carried out using the Bayesian model-based clustering method implemented in the software STRUCTURE (Falush *et al.* 2007). The admixture model with standard settings was applied and 100 000 Markov chain Monte Carlo steps, with a burn-in period of 10 000, were used. Six runs were done to test for the number of genetic clusters, *K*, in the data set ($1 \leq K \leq 6$). Every run

was repeated three times to assess convergence. Resulting log-likelihoods were compared between values of *K* to determine the actual number of population partitions.

Statistical power analyses were conducted with POWSIM 4.0, using the same settings as for the mtDNA data set.

Drifter analysis

In order to investigate possible means of gene flow in the Scotia Sea, the trajectories of satellite-tracked drifting buoys (hereafter called drifters) of the Global Drifter Program (Lumpkin & Pazos 2007) were analysed. This program is conducted by the US National Oceanic and Atmospheric Administration (NOAA). Drifters consist of a surface float equipped with a Global Positioning System (GPS) device, and a drogue centred at 15 m depth to ensure drifter movement along with ocean surface currents (Lumpkin & Pazos 2007). Interpolated data of all drifters passing 40–70 S 10–70 W between 15 February 1979 and 31 July 2007 were downloaded from http://www.aoml.noaa.gov/phod/dac/gdp.html. Three elliptical regions were defined to encompass the main shelf habitats of AP and South Shetland Islands, South Georgia, and South Sandwich Islands. Chosen diameters were 4 in latitudinal direction and 6 in longitudinal direction, which resulted in radii between 155 and 222 km, depending on latitude. Ellipses were centred at 61.84 S 56.33 W, 54.39 S 36.95 W, and 57.76 S 26.42 W, respectively. Drifter data was filtered to exclude all drifters that did not pass any of the three defined regions. To simulate dispersal of pelagic larvae from and to the shelf habitats, trajectories of the remaining drifters were plotted (i) over a period of 4 months, starting the day of departure from one of the regions, and (ii) for 4 months before arrival at one of the areas. The 4-month period was chosen to reflect the duration of the *G. gibberifrons* pelagic larval stage (North 2001). To account for potential surface current differences during the *G. gibberifrons* hatching period (North 2001), a third plot was produced using only drifters that left one of the regions between August and November.

Isolation-with-migration model

The isolation-with-migration (IM) model, as implemented in IMA (Hey & Nielsen 2007), was applied to determine directionality of gene flow between the AP shelf and the SR island shelves. To this end, all AP samples were grouped into one population, while South Georgia and South Sandwich Islands samples constituted the second population. Two possible scenarios of gene flow were considered: (i) unidirectional larval dispersal with ocean currents as approximated by drifter trajectories and (ii) bidirectional stepwise migration of adults along the SR. In both scenarios, we expected gene flow between South Georgia

and South Sandwich Islands to be direct, i.e. not via AP populations. For feasible run durations, the AP population was reduced to two independent subsets of 60 individuals each, drawn evenly from the four AP locations. Mitochondrial D-loop sequences and seven out of the eight microsatellite loci were included in the model. Trne35 was excluded, as null alleles were indicated by a departure from Hardy–Weinberg expectations, and TANDEM analysis revealed poor binning quality of this locus. To adjust for expected effective population sizes, inheritance scalars of 0.25 and 1 were assigned to mtDNA and microsatellite loci, respectively. The Hasegawa–Kishino–Yano (HKY) model of sequence evolution was applied to mitochondrial sequences, and a stepwise mutation model (SMM) was assumed for all microsatellite loci. Parameter ranges for uniform priors were empirically determined in a series of initial runs, and set to $m_1, m_2 \in (0,50]$, $t \in (0,10]$, and scalars $\Theta_1, \Theta_2, \Theta_A \in (0,3]$. Five final runs were conducted using both AP sample subsets. Each run included 40 Metropolis-coupled Markov chains and a geometric heating scheme. The first 1 million updates were discarded as burn-in. In each run, 100 000 genealogies were sampled from 10 million updates. Saved genealogies from all runs were combined according to AP subset for two subsequent analyses in IMA's 'Load-Trees' mode. Likelihood-ratio tests of nested models were conducted to assess whether unidirectional AP to SR gene flow can be rejected. Statistical significance was approximated using a chi-square (χ^2) distribution, following Hey & Nielsen (2007).

In order to exclude bias in gene flow directionality caused by unequal sample sizes, we ran an additional analysis with only 18 AP individuals to match the size of the combined SR population. AP individuals were drawn evenly from the four populations. In yet another approach to test for directionality bias, we used IMA to assess migration rates between sets of 18 and 60 individuals that were both drawn evenly from all AP populations, without inclusion of SR individuals. This was carried out three times, whereby the smaller subsets were non-overlapping between runs. For all runs, above settings for parameter ranges, number of Markov chains, heating scheme, and run duration were applied.

Results

Mitochondrial DNA: D-loop

Alignment of D-loop sequences from 162 individuals yielded a consensus sequence of 352 bp (GenBank Accession nos FJ528746–FJ528907). No gaps were found in the alignment. The ratio of transitions to transversions was 2.0. Sequences collapsed into 32 unique haplotypes (Table S2, Supporting information). DNA extraction, amplification and sequencing were repeated for 46 randomly chosen samples, confirming previous results. The resulting haplotype genealogy (Fig. 2) shows no obvious structure between populations.

Power analysis with POWSIM showed that the mtDNA data set had enough statistical power to detect global population structure with high probability when true F_{ST} values were as low as 0.02 (> 97%). With pooled AP and SR populations, this degree of divergence would be detected with a probability of 88% when using the χ^2 test (Fig. S1).
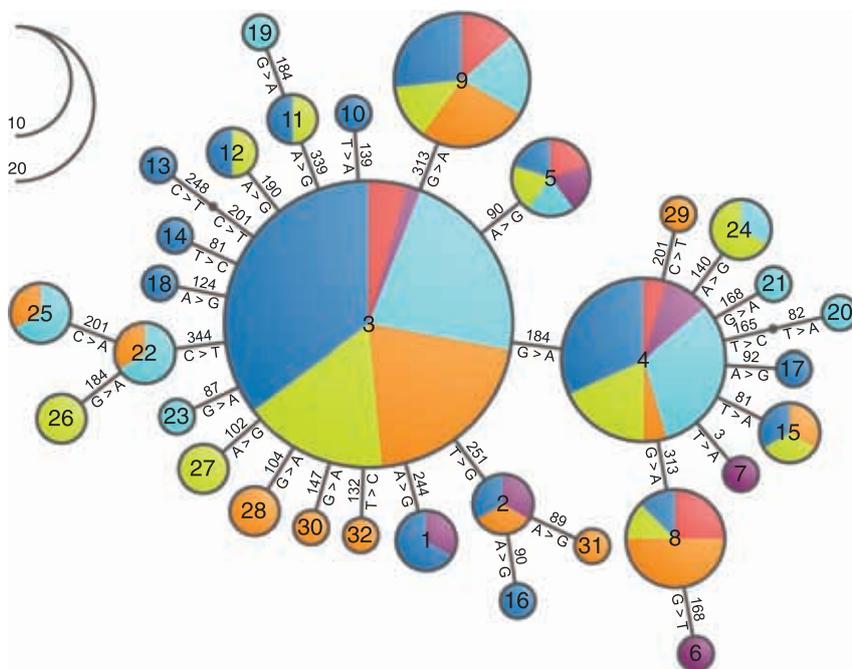
Fig. 2 Unrooted haplotype genealogy based on 162 D-loop sequences (352 bp). Radii reflect number of individuals.

Table 2 Population pairwise Φ_{ST} and F_{ST} values, based on mtDNA (below diagonal) and microsatellites (above diagonal) respectively. EI, Elephant Island; JIa, Joinville Island A; JIb, Joinville Island B; KGI, King George Island; SG, South Georgia; SSI, South Sandwich Islands

	EI	JIa	JIb	KGI	SG	SSI
Elephant Island		0.001	−0.001	0.001	0.001	−0.003
Joinville Island A	0.009		0.000	−0.001	0.000	−0.003
Joinville Island B	0.005	0.014		0.004	0.003	−0.004
King George Island	0.008	−0.017	0.011		0.005	0.001
South Georgia	0.042	−0.008	0.035	0.010		−0.009
South Sandwich Islands	0.058	0.029	−0.010	0.047	−0.001	

Of all possible groupings for the hierarchical AMOVA (Tables S3 and S4, Supporting information), the scheme with pooled AP populations and separate groups for SR populations produced high Φ_{ST} values relative to other groupings, and is therefore reported in more detail (Table S5, Supporting information). Here, 97% of variation occurred within populations, while only 2.5% and 0.4% were attributed to variation among groups and to variation within groups, but among populations. All Φ-statistics were low, and none were significant at the 95% confidence interval.

Population differentiation was further examined using pairwise comparisons (below diagonal in Table 2). Negative Φ_{ST} values are probably caused by rounding errors and are not significantly different from zero (Long 1986). The lowest Φ_{ST} values were found among the AP populations, while relatively higher fixation indices were detected between AP and SR populations. However, none of the pairwise Φ_{ST} values were significant at the 95% confidence level (below diagonal in Table S6, Supporting information).

Pairwise mutational distances were calculated over all 162 sequences and summarized in a coalescence-based mismatch analysis (Rogers & Harpending 1992). The model of sudden population expansion could not be rejected (sum of square deviation $P = 0.44$) (Fig. 3). As implemented in ARLEQUIN, the best-fitting model of population growth was calculated. The resulting model was characterized by expansion time parameter $\tau = 1.496$ (95% confidence interval: 1.199–1.961) and population size parameters $\Theta_0 = 0$ (0–0.098) and $\Theta_1 = 99\,999.0$ (6.871–99 999). Using a mutation rate of 6.5–8.8% per Myr and taking into account the sequence length of 352 bp, τ was translated to absolute time in years. Assuming this mutation rate applies for *Gobionotothen gibberifrons*, a sudden population expansion should have occurred 24 148–32 692 years before present (BP) (95% confidence interval: 19 353–42 854 BP).

Nuclear DNA: microsatellites

A total of 164 individuals from six different sampling locations were scored for nine microsatellite loci. Analyses with TANDEM revealed effective repeat sizes between 1.83 and 2.15 bp, and average rounding errors between 0.04 and
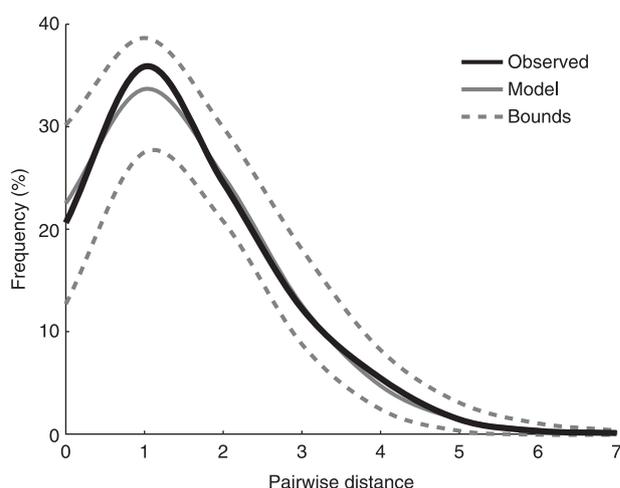
Fig. 3 Mismatch distribution over 162 *Gobionotothen gibberifrons* individuals, based on D-loop sequences (352 bp).

0.38 bp. The largest average rounding error was associated with locus Trne35. The number of alleles per locus, allelic size range, as well as observed and expected heterozygosities are reported in Table S7, Supporting information. Up to 61 alleles were found for single loci. Locus Cr170 was monomorphic in all populations, as was Cr127 in all specimens from South Georgia. Trne35 featured the widest range of fragment sizes, and the largest number of alleles. However, for Trne35, the presence of null alleles was indicated by a significant ($P < 0.001$) departure from Hardy–Weinberg expectations (Table S7) in all AP populations (O'Connell & Wright 1997). Analysis with the software MICROCHECKER confirmed that null alleles are the causes of all departures from Hardy–Weinberg equilibrium. No tests for linkage disequilibrium were significant after Bonferroni correction (Rice 1989; Slatkin 1994; Slatkin & Excoffier 1996).

Statistical power analysis indicated high probabilities (> 99%) to detect global population structure when true F_{ST} values are as low as 0.005. Between AP and SR populations, the same degree of divergence would be detected with probabilities exceeding 92% and 98%, using Fisher's exact test and χ^2 test, respectively (Fig. S2).

Hierarchical AMOVA tests indicated that almost all variation (~100%) occurred within populations. Hardly any genetic differentiation was attributed to population or group identities, irrespective or grouping scheme (Table S8, Supporting information). Detailed results are reported for the grouping chosen for mtDNA sequences in Table S9, Supporting information.

Applying *F*-statistics to the same microsatellite data set confirmed genetic homogeneity between populations (above diagonal in Table 2). F_{ST} values range between –0.009 and 0.005. Given that the greatest absolute value and the average F_{ST} were negative, it is likely that rounding errors are responsible for all departures from zero. None of the associated *p*-values (above diagonal in Table S6) were significant after Bonferroni correction.

Above findings were corroborated by the clustering method implemented in STRUCTURE. Log-likelihood values were calculated for the existence of $1 \leq K \leq 6$ clusters within the microsatellite data set. The highest log-likelihood value was scored for the assignment of all individuals to a single cluster ($K = 1$) in three independent run replicates. Log-likelihood values decreased with increasing number of assumed clusters (Table S10, Supporting information).

Drifter analysis

A total of 661 drifters crossed 40–70 S latitude and 10–70 W longitude between 15 February 1979 and 31 July 2007, and 140 of them entered one of the predefined areas around the AP and SR populations. Out of 52 drifters that left the AP shelf area, 13 drifters reached South Georgia within four months. Of those leaving the South Georgia or South Sandwich Island areas (21 and 8 drifters, respectively), none arrived at a different area within the 4-month period (Fig. 4A).

When including only drifters that left during the *G. gibberifrons* hatching period between August and November (North 2001), a similar picture arises. A single drifter left the AP area within this period and arrived at the South Georgia area after three and a half months. To the contrary, five drifters were advected off the South Georgia shelf, and dispersed into the Southwest Atlantic without crossing other shelf areas (Fig. 4B). A multivariate analysis of variation (MANOVA) was run to compare trajectory endpoints of drifters that left the South Georgia shelf between August and November with those of drifters leaving between December and July (R Development Core Team 2008). No significant difference was found ($F_{2,18} = 1.89$, $P = 0.18$). Thus, using year-round drifters as proxies for *G. gibberifrons* larval dispersal seems reasonable.

Finally, we analysed drifter histories over 4 months prior to entrance of predefined shelf areas. Three drifters entered the AP area from the west; none of them had come close to other areas. Out of 25 drifters that entered the South
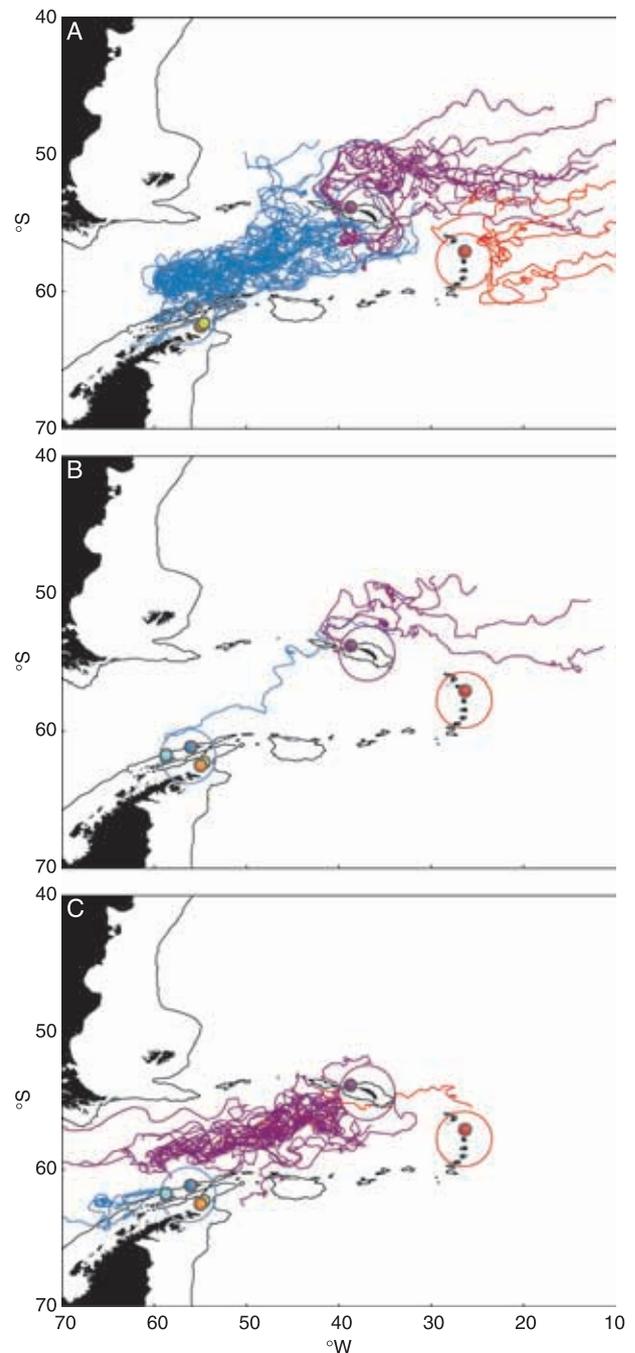
Fig. 4 Trajectories of Surface Velocity Program Drifters in the Scotia Sea and the Southwest Atlantic between 1979 and 2007. (A) Four-month drifter trajectories after departure from shelf areas approximated by three elliptical regions. (B) As in (A), but using only drifters that left one of the regions between August and November. (C) Four-month drifter trajectories before entering one of the three regions.

Georgia area, one had left the AP shelf area 4 months earlier, and several more passed this area within short distance. A single drifter entered the South Sandwich Islands area coming from South Georgia (Fig. 4C).

© 2009 Blackwell Publishing Ltd
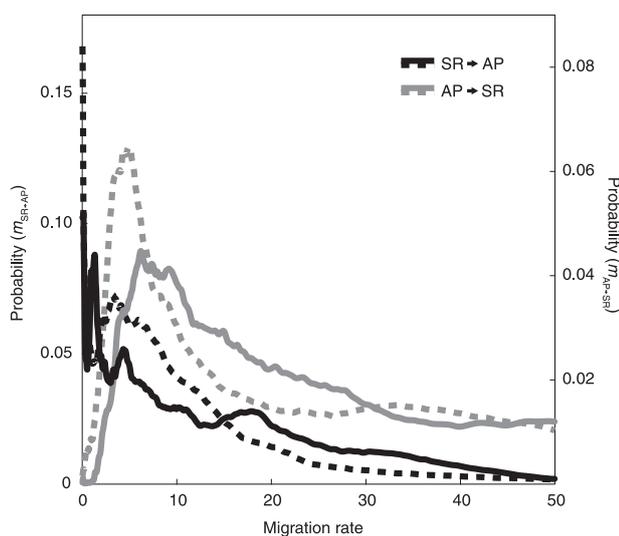
Fig. 5 Posterior probabilities of migration rates determined with IMA. Black: SR to AP migration rates, grey: AP to SR migration rates. Solid line: using the first AP subset of 60 *Gobionotothen gibberifrons* individuals, dashed line: using the second AP subset.

Isolation-with-migration model

Our aim was to discriminate between two contrasting scenarios of gene flow across the Scotia Sea: bidirectional adult migration or unidirectional larval dispersal from the AP to the SR. Replicate runs of the IMA program revealed asymmetric migration rates (Fig. 5). In two independent sample subsets, the highest posterior probabilities for SR to AP migration rates (scaled for mutation rate) were consistently found close to zero (0.025 and 0.075). Out of 1000 bins distributed evenly over the whole parameter range, these values corresponded to the two bins closest to zero. On the other hand, nonzero rates were inferred for AP to SR migration (4.875 and 6.225). One of two sample subsets produced sharp peaks for all three population size parameters (Θ_1: 27.166, Θ_2: 24.476, Θ_A: 79.007). The second subset failed to produce a clear peak for the SR population size parameter, but AP and ancient population size parameters were congruent with the first subset (Θ_1: 30.470, Θ_A: 87.763). Population size parameters of the first subset were used to convert migration parameters into per-generation population migration rates ($M = \Theta \times m / 2$). Peak locations corresponded to 59.659 and 76.180 migration events per generation. Taking into account a *G. gibberifrons* generation time of 6–8 years (Kock & Kellermann 1991), this translates to 7.45–12.70 migration events per year. Likelihood ratio tests did not reject unidirectional AP to SR gene flow in either subset ($P_1 = 0.19$, $P_2 = 0.23$), whereas unidirectional gene flow in the opposite direction, from the SR to the AP, was clearly rejected in one out of two subsets ($P_1 = 0.18$, $P_2 < 0.0001$).

In order to test for directionality bias, we used IMA to estimate migration rates (i) between equally sized sets of 18 AP and 18 SR individuals, and (ii) between sets of 60 and 18 individuals that were both drawn evenly from all AP populations. In the first case, reduction of the AP sample size did not influence directionality of migration rates. The highest posterior probability was assigned to a near-zero SR to AP migration rate (0.025), while a substantially higher rate was found in the opposite direction (3.225). In the second case of unequal sample sizes, we found mostly balanced gene flow between the larger and the smaller subset of AP individuals (m_1: 4.475, 7.075, 0.925; m_2: 5.625, 4.225, 6.325; all values scaled for mutation rate). These results suggest that estimation of migration rates, as implemented in IMA, is robust to unequal sample sizes. Taken together, our analyses using the IM model indicate unidirectional gene flow across the Scotia Sea, from the AP shelf to South Georgia and the South Sandwich Islands.

Discussion

Demographic history of Gobionotothen gibberifrons

We investigated the demographic history of the total *Gobionotothen gibberifrons* population using a coalescent-based mismatch analysis, as implemented in ARLEQUIN. In order to date the observed sudden population expansion, we assumed a control region mutation rate of 6.5–8.8% per Myr that has been found in Lake Malawi cichlid fishes, another perciform group that underwent adaptive radiation (Sturmbauer *et al*. 2001). Applying this rate, our results suggest an expansion 24 148–32 692 years ago, at the height of the last ice age (EPICA community members 2004). Presumably, the Antarctic ice sheet extended all the way to the shelf in glacial cycles (Thatje *et al*. 2008), 'bulldozing the surviving fauna to the deep continental margin' (Barnes & Conlan 2007). It seems difficult to imagine how extensive glaciation of the Antarctic shelf, the *G. gibberifrons* habitat, may have contributed to increasing population size. However, time estimates based on molecular clocks should in general be treated with caution, and serve as rough approximations only. Published estimates of control region mutation rates in bony fishes vary on two orders of magnitude. For example, mutation rates as low as 2.2% per Myr were inferred for haplochromine cichlids (Sato *et al*. 2003), while rates up to 108% per Myr were estimated for Indo-Pacific butterflyfishes (McMillan & Palumbi 1997; see Bowen *et al*. 2006 for a list of published estimates). It has recently been shown that variation in mitochondrial mutation rates can partly be explained by metabolic rate and generation time (Nabholz *et al*. 2008). The metabolic rate hypothesis states that the mitochondrial mutation rate is linked to metabolic rate and production of reactive oxygen species (ROS) (Martin & Palumbi 1993, but see Lanfear

et al. 2007). Given the low metabolic rate of notothenioids (Clarke & Johnston 1999), a low mutation rate could be expected (Bargelloni *et al.* 1994). In addition, the generation time of *G. gibberifrons* (6–8 years, Kock & Kellermann 1991) is higher than in cichlid fishes (1–3 years, Won *et al.* 2005). Therefore, the assumed mutation rate of 6.5–8.8% could be an overestimate, and the population expansion might be older than inferred. Should the *G. gibberifrons* substitution rate be substantially lower, the expansion could date back to the beginning of the last interglacial 180 000 years ago (EPICA community members 2004) when the Antarctic ice sheet disconnected from the shelf floor, and suitable shelf habitat became available. On the other hand, most Antarctic notothenioids lack NADH 6 dehydrogenase, which is part of the mitochondrial electron transport chain. It has been hypothesized that this loss allows heat production through proton leakage across the inner mitochondrial membrane (Papetti *et al.* 2007a). If so, possible consequences on ROS production, and thus on mutation rate, cannot be excluded. Unusually high mutation rates of the mitochondrial control region have previously been observed in butterflyfishes (McMillan & Palumbi 1997; 33–108%). Of interesting note, very similar shapes of mismatch distributions have been found in another Antarctic notothenioid, *Pleuragramma antarcticum*, as well as in Antarctic krill, *Euphausia superba* Dana (Zane *et al.* 2006, 1998). Taken together, these findings suggest that periodical glaciations may have affected a large part of the Antarctic marine fauna in one way or another. However, more precise estimates of notothenioid mutation rates will be needed in order to correlate demographic histories of different species, and to shed light on potential geological and/or climatological causes of population expansions.

Larval dispersal across the Scotia Sea

Our genetic analyses based on neutral mitochondrial and nuclear markers show no significant population structure between *G. gibberifrons* populations around the Scotia Sea suggesting ongoing gene flow between sampling sites. Between-population fixation indices were close to zero and AMOVA tests attributed ~100% of genetic differentiation to variation within populations. Although genetic homogeneity could theoretically be explained solely by ancestral polymorphism, our analyses using the isolation-with-migration model confirm that large amounts of gene flow do occur between *G. gibberifrons* populations. Moreover, our results suggest that gene flow is highly unidirectional, following the direction of the ACC. However, our results may be affected by departures from the strict IM model. Since this model only considers pairs of populations, estimates of migration rates can be distorted by unsampled populations that exchange migrants with the two sampled populations (Won & Hey 2005). In the Scotia Sea, the shelf

area surrounding the South Orkney Islands harbours a large *G. gibberifrons* population (Kock & Jones 2005). Given its geographic location between the AP and SR sampling locations, it seems possible that gene flow between the AP and the SR sampling locations occurs via the South Orkney shelf. The effect of so-called 'ghost populations' is difficult, if not impossible to quantify, in particular if migration is asymmetric (Slatkin 2005). Furthermore, population subdivisions would violate the IM model and affect migration rate estimates (Wakeley 2000). However, given our above results of mtDNA and microsatellite data, we consider both AP and SR populations as panmictic, and thus in agreement with the model. Based on additional analyses using equally sized AP and SR subsets as well as AP subsets of different sizes, we conclude that unequal sample sizes apparently do not affect migration rate estimates of the IMA program. Similar results were found for the software IM, that implements the same model and is structurally related to IMA (Hey & Nielsen 2004; Rosenblum 2006). Taken together, we found a clear signal of unidirectional gene flow that is not affected by unequal sample sizes, and thus must be inherent to the data.

Analyses of surface drifter trajectories show that passive particles cross the Scotia Sea between AP and South Georgia in less than 4 months, the pelagic larval duration of *G. gibberifrons*. We therefore conclude that larval dispersal along the ACC is the main agent of gene flow in *G. gibberifrons*. Given the extended pelagic phases of many notothenioid fishes, larval dispersal with the ACC has been suggested for a number of notothenioid species (Loeb *et al.* 1993). In a recent survey of the ichthyofauna of Bouvetøya, a small volcanic island within the ACC, about 2500 km east of South Georgia, all detected nototheniid and channichthyid species had long larval durations of 1–2 years (Jones *et al.* 2008). In contrast to similarly isolated islands (e.g. Easter Island), not a single endemic fish species was found at Bouvetøya. Jones *et al.* (2008) conclude that the Bouvetøya ichthyofauna is primarily derived from South Georgia, through dispersal of pelagic larvae with the ACC.

Marine populations have long been considered demographically open, and generally interconnected by larval dispersal. It was believed that virtually all fish larvae would be advected from the local sources to settle in downstream habitats (Caley *et al.* 1996). However, this view has shifted in recent years as evidence for larval retention has accumulated (reviewed in Swearer *et al.* 2002). Many larvae are capable of active vertical migration, which in combination with vertically stratified flows may suffice to avoid advection, and especially towards the end of their pelagic phase, larvae are able to swim even faster than ambient currents in many cases (Leis 2006). In notothenioids, larvae of many species, including *G. gibberifrons*, are known to undergo vertical migration (North & Murray 1992). Thus, active retention mechanisms could potentially counteract advection with the

Table 3 Summary of published studies on the population structure of notothenioids. These include studies employing allozyme electrophoresis (the number of analysed protein-coding loci is given), restriction fragment length polymorphisms (RFLP, number of informative restriction enzymes), randomly amplified polymorphic DNA (RAPD, number of polymorphic primers), mitochondrial (mtDNA) and nuclear DNA (nDNA) sequence analysis (fragment length is given) and microsatellite analysis (STR, number of polymorphic loci). The shortest distance over which significant differentiation was found (d_s) is given, as well as the longest distance over which no significant divergence could be detected (d_{ns}). *Nonsignificant differentiation within the Weddel Sea. †Details not reported in publication. ‡Differentiation found between collections of different years

Organism	N	Allozymes	RFLP	RAPD	mtDNA	nDNA	STR	d_s (km)	d_{ns} (km)	Reference
Champsocephalus gunnari	53		7				—		400	Williams *et al.* (1994)
Lepidonotothen squamifrons	215	5						6400	1300	Schneppenheim *et al.* (1994)
Champsocephalus gunnari	86	13					—		6400	Duhamel *et al.* (1995)
Notothenia rossii	76	13					—		400	Duhamel *et al.* (1995)
Chionodraco myersi	65	10					16		< 1000*	Clement *et al.* (1998)
Neopagetopsis ionah	35	10					4600		< 1000*	Clement *et al.* (1998)
Dissostichus eleginoides	32						5	60	—	Reilly & Ward (1999)
Dissostichus eleginoides	196–230	7					8	2000	6000	Smith & McVeagh (2000)
Dissostichus eleginoides	? †	?		?			?	~500	8300	Smith & Gaffney (2000)
Dissostichus eleginoides	439–623		2				7	5200	—	Appleyard *et al.* (2002)
Dissostichus mawsoni	42			12				4700	—	Parker *et al.* (2002)
Chionodraco hamatus	74				302 bp			1000	9300	Patarnello *et al.* (2003)
Dissostichus eleginoides	113–136		2				7	—	2600	Appleyard *et al.* (2004)
Dissostichus eleginoides	396–450		2				5	500	1300	Shaw *et al.* (2004)
Dissostichus mawsoni	24–57		4		1304 bp		5	—	5000	Smith & Gaffney (2005)
Champsocephalus gunnari	63				1817 bp	3037 bp		1200	4400	Kuhn & Gaffney (2006)
Dissostichus eleginoides	151–274				249 bp		7	1200	5100	Rogers *et al.* (2006)
Pleuragramma antarcticum	256				277 bp			0‡	7000	Zane *et al.* (2006)
Trematomus bernacchii	61				468 bp	307 bp		4900	1600	Janko *et al.* (2007)
Trematomus newnesi	36				483 bp	299 bp		—	4900	Janko *et al.* (2007)
Chaenocephalus aceratus	247						11	—	100	Papetti *et al.* (2007b)
Dissostichus mawsoni	4–68				~4000 bp	~11 500 bp		1300	7800	Kuhn & Gaffney (2008)
Chaenocephalus aceratus	23				1047 bp			—	3900	Jones *et al.* (2008)
Lepidonotothen squamifrons	23				1047 bp			3400	1800	Jones *et al.* (2008)
Notothenia coriiceps	21				1047 bp			—	3900	Jones *et al.* (2008)
Lepidonotothen larseni	23				1047 bp			—	3400	Jones *et al.* (2008)
G. gibberifrons	162–164				352 bp		8	—	1900	This study

ACC. At South Georgia, off-shelf dispersal was observed for pelagic fish eggs, but not for notothenioid larvae, suggesting active larval behaviour (White 1998). On the other hand, Kellermann (1991) found larvae of *Notothenia coriiceps*, another benthic nototheniid, at a number of Scotia Sea sampling locations between the AP and South Georgia shelves. Moreover, positive catches occurred earlier in the southern Scotia Sea than near South Georgia, leading Kellermann (1991) to conclude that larvae hatched near the AP and reach South Georgia with the ACC. Less is known about the distribution of *G. gibberifrons* larvae. However, large numbers of larvae were occasionally found in offshore waters around South Georgia (Loeb *et al.* 1993). These observations corroborate our results, showing that advection of notothenioid larvae away from their local shelf habitats does indeed occur, and that larvae can travel hundreds of kilometres, surfing the ACC. In *G. gibberifrons*, we detected no significant population structure across

distances as large as 1900 km, comprising a large part of the species' distribution range (Fig. 1). These results are comparable to findings of previous studies on genetic differentiation in notothenioids (Table 3).

Our analyses based on the IM model suggest around 10 migration events per year between the AP and the SR populations of South Georgia and the South Sandwich Islands. In order to calculate per-generation migration rates, we used the equation $M = \Theta \times m/2$, which assumes genetic equilibrium of populations. We note that this cannot be the case for *G. gibberifrons*, as we detected a recent population expansion. Therefore, the detected number of migration events should be treated as a rough estimate. However, to maintain genetic homogeneity between populations, even lower rates would be sufficient (Slatkin 1987). As the population size of *G. gibberifrons* is large (Kock & Jones 2005), detected migration rates are negligible for demographic processes. This would mean that most of

the South Georgia population recruits locally. It remains unclear, whether this is due to active retention mechanisms, mortality of advected larvae, or low mating success of migrants. But over time periods important for evolutionary processes, AP populations represent source populations, while SR populations constitute sinks. Travelling with the ACC, larvae advected from SR shelves would fail to find suitable habitat. Given this risk of losses due to advection, the presence of extended pelagic larval phases in many Antarctic fishes (1–2 years, Loeb *et al.* 1993) seems puzzling, but may be balanced by retention mechanisms (White 1998) that save most larvae from advection. Although highly speculative, protracted larval phases might be explained by selection for dispersal subsequent to range expansions (Thomas *et al.* 2001). As mentioned above, many shelf habitats close to the Antarctic continent were covered by the Antarctic ice cap during ice ages. Following glacial retreat, habitats became available and presumably caused range expansions, whereby longer larval phases could have been favoured.

Overall, our results compare well with those of similar studies. Table 3 summarizes, to the best of our knowledge, all published population genetic studies of notothenioid fishes. Despite large variation in sample size and markers used, nonsignificant differentiation is commonly found across thousands of kilometres (d_{ns}) and even between populations at opposite sides of Antarctica (Patarnello *et al.* 2003). Significant differentiation over less than 100 km was found only in very few cases. Of particular note are findings on the population structure of *Chionodraco myersi*, a benthic channichthyid (Iwami & Kock 1990) with long pelagic larval phase (Kock & Kellermann 1991), and *P. antarcticum*, one of few notothenioids that adapted a truly pelagic lifestyle (Eastman 1993). Using allozyme markers, Clement *et al.* (1998) detected significant differentiation between *C. myersi* Weddel Sea populations no more than 16 km apart. The authors attribute this differentiation to a rapid geotrophic stream running between both populations. Perhaps more surprising were findings in *P. antarcticum*. Using comparatively large sample sizes and mitochondrial control region sequences, Zane *et al.* (2006) found significant differentiation between samples taken at Halley Bay, Weddell Sea at intervals of 2 years, but no significant differentiation between samples taken on opposite sides of the continent. These exceptions aside, it seems that gene flow across large distances is a common feature in notothenioids.

Acknowledgements

References

Appleyard SA, Ward RD, Williams R (2002) Population structure of the Patagonian toothfish around Heard, McDonald and Macquarie Islands. *Antarctic Science*, **14**, 364–373.

Appleyard SA, Williams R, Ward RD (2004) Population genetic structure of Patagonian toothfish in the West Indian sector of the Southern Ocean. *CCAMLR Science*, **11**, 12–32.

Bargelloni L, Ritchie PA, Patarnello T *et al.* (1994) Molecular evolution at subzero temperatures: mitochondrial and nuclear phylogenies of fishes from Antarctica (suborder Notothenioidei), and the evolution of antifreeze glycopeptides. *Molecular Biology and Evolution*, **11**, 854–863.

Barnes DKA, Conlan KE (2007) Disturbance, colonization and development of Antarctic benthic communities. *Philosophical Transactions of the Royal Society B: Biological Sciences*, **362**, 11–38.

Barnes DKA, Fuentes V, Clarke A, Schloss IR, Wallace MI (2006) Spatial and temporal variation in shallow seawater temperatures around Antarctica. *Deep Sea Research Part II: Topical Studies in Oceanography*, **53**, 853–865.

Bowen BW, Muss A, Rocha LA, Grant WS (2006) Shallow mtDNA coalescence in Atlantic pygmy angelfishes (genus *Centropyge*) indicates a recent invasion from the Indian Ocean. *Journal of Heredity*, **97**, 1–12.

Brunner PC, Douglas MR, Osinov A, Wilson CC, Bernatchez L (2001) Holarctic phylogeography of Arctic charr (*Salvelinus alpinus* L.) inferred from mitochondrial DNA sequences. *Evolution*, **53**, 573–586.

Caley MJ, Carr MH, Hixon MA *et al.* (1996) Recruitment and the local dynamics of open marine populations. *Annual Review of Ecology and Systematics*, **27**, 477–500.

Chen L, DeVries AL, Cheng CHC (1997) Evolution of antifreeze glycoprotein gene from a trypsinogen gene in Antarctic notothenioid fish. *Proceedings of the National Academy of Sciences, USA*, **94**, 3811–3816.

Clarke A, Johnston NM (1999) Scaling of metabolic rate with body mass and temperature in teleost fish. *Journal of Animal Ecology*, **68**, 893–905.

Clement O, Ozouf-Costaz C, Lecointre G, Berrebi P (1998) Allozymic polymorphism and phylogeny of the family Channichthyidae. In: *Fishes of Antarctica. A Biological Overview* (eds di Prisco G, Pisano E, Clarke A), pp. 299–309. Springer-Verlag Publishers, Milan, Italy.

Cziko PA, Cheng CHC (2006) A new species of nototheniid (Perciformes: Notothenioidei) fish from McMurdo Sound, Antarctica. *Copeia*, **2006**, 752–759.

Derome N, Chen WJ, Dettaï A, Bonillo C, Lecointre G (2002) Phylogeny of Antarctic dragonfishes (Bathydraconidae, Notothenioidei, Teleostei) and related families based on their anatomy and two mitochondrial genes. *Molecular Phylogenetics and Evolution*, **24**, 139–152.

DeVries AL (1988) The role of antifreeze glycopeptides and peptides in the freezing avoidance of Antarctic fishes. *Comparative Biochemistry and Physiology*, **90B**, 611–621.

DeWitt HH, Heemstra PC, Gon O (1990) Nototheniidae. In: *Fishes of the Southern Ocean* (eds Gon O, Heemstra PC), pp. 279–332 J.L.B. Smith Institute of Ichthyology, Grahamstown, South Africa.

Domingues VS, Bucciarelli G, Almada VC, Bernardi G (2005) Historical colonization and demography of the Mediterranean damselfish, *Chromis chromis*. *Molecular Ecology*, **14**, 4051–4063.

Duhamel G, Ozouf-Costaz C, Cattaneo-Berrebi G, Berrebi P (1995) Interpopulation relationships in two species of Antarctic fish *Notothenia rossii* and *Champsocephalus gunnari* from the Kerguelen Islands: an allozyme study. *Antarctic Science*, **7**, 351–356.

Eastman JT (1993) *Antarctic Fish Biology: Evolution in a Unique Environment*. Academic Press, San Diego, California.

Eastman JT (2005) The nature of the diversity of Antarctic fishes. *Polar Biology*, **28**, 93–107.

Eastman JT, Clarke A (1998) A comparison of adaptive radiations of Antarctic fish with those of non-Antarctic fish. In: *Fishes of Antarctica. A Biological Overview* (eds di Prisco G, Pisano E, Clarke A), pp. 3–26. Springer-Verlag Publishers, Milan, Italy.

EPICA community members (2004) Eight glacial cycles from an Antarctic ice core. *Nature*, **429**, 623–628.

Excoffier L, Laval G, Schneider S (2005) ARLEQUIN version 3.0: an integrated software package for population genetics data analysis. *Evolutionary Bioinformatics Online*, **1**, 47–50.

Falush D, Stephens M, Pritchard JK (2007) Inference of population structure using multilocus genotype data: dominant markers and null alleles. *Molecular Ecology Notes*, **7**, 574–578.

Foster TD (1984) The marine environment. In: *Antarctic Ecology* (ed. Laws RM) **2**, pp. 345–371. Academic Press, London.

Hasegawa M, Kishino H, Yano T (1985) Dating of the human–ape splitting by a molecular clock of mitochondrial DNA. *Journal of Molecular Evolution*, **22**, 160–174.

Hey J, Nielsen R (2004) Multilocus methods for estimating population sizes, migration rates and divergence time, with applications to the divergence of *Drosophila pseudoobscura* and *D. persimilis*. *Genetics*, **167**, 747–760.

Hey J, Nielsen R (2007) Integration within the Felsenstein equation for improved Markov chain Monte Carlo methods in population genetics. *Proceedings of the National Academy of Sciences, USA*, **104**, 2785–2790.

Iwami T, Kock KH (1990) Channichthyidae. In: *Fishes of the Southern Ocean* (eds Gon O, Heemstra PC), pp. 279–332 J.L.B. Smith Institute of Ichthyology, Grahamstown, South Africa.

Janko K, Lecointre G, DeVries AL *et al.* (2007) Did glacial advances during the Pleistocene influence differently the demographic histories of benthic and pelagic Antarctic shelf fishes? — Inferences from intraspecific mitochondrial and nuclear DNA sequence diversity. *BMC Evolutionary Biology*, **7**, 220–220.

Jones CD, Anderson ME, Balushkin AV *et al.* (2008) Diversity, relative abundance, new locality records and population structure of Antarctic demersal fishes from the northern Scotia Arc islands and Bouvetøya. *Polar Biology*, **31**, 1481–1497.

Kellermann AK (1991) Egg and larval drift of the Antarctic fish *Notothenia coriiceps*. *Cybium*, **15**, 199–210.

Kock KH (1992) *Antarctic Fish and Fisheries*. Cambridge University Press, Cambridge, UK.

Kock KH (2005) Antarctic icefishes (Channichthyidae): a unique family of fishes. A review, part I. *Polar Biology*, **28**, 862–895.

Kock KH, Jones CD (2005) Fish stocks in the Southern Scotia Arc region — a review and prospects for future research. *Reviews in Fisheries Science*, **13**, 75–108.

Kock KH, Kellermann AK (1991) Reproduction in Antarctic notothenioid fish. *Antarctic Science*, **3**, 125–150.

Kompowski A (1985) Studies on food and fecundity of *Notothenia gibberifrons* Lönnberg, 1905 (Pisces, Notothenioidei) off South Georgia. *Acta Ichthyologica et Piscatoria*, **15**, 55–73.

Kuhn KL, Gaffney PM (2006) Preliminary assessment of population structure in the mackerel icefish (*Champsocephalus gunnari*). *Polar Biology*, **29**, 927–935.

Kuhn KL, Gaffney PM (2008) Population subdivision in the Antarctic toothfish (*Dissostichus mawsoni*) revealed by mitochondrial and nuclear single nucleotide polymorphisms (SNPs). *Antarctic Science*, **20**, 12.

Lanfear R, Thomas JA, Welch JJ, Brey T, Bromham L (2007) Metabolic rate does not calibrate the molecular clock. *Proceedings of the National Academy of Sciences, USA*, **104**, 15388–15393.

Leis JM (2006) Are larvae of demersal fishes plankton or nekton? *Advances in Marine Biology*, **51**, 57–141.

Li WH (1977) Distribution of nucleotide differences between two randomly chosen cistrons in a finite population. *Genetics*, **85**, 331–337.

Loeb VJ, Kellermann AK, Koubbi P, North AW, White MG (1993) Antarctic larval fish assemblages: a review. *Bulletin of Marine Science*, **53**, 416–449.

Long JC (1986) The allelic correlation structure of Gainj- and Kalam-speaking people. I. the estimation and interpretation of Wright's *F*-statistics. *Genetics*, **112**, 629–647.

Lumpkin R, Pazos M (2007) Measuring surface currents with surface velocity program drifters: the instrument, the data, and some recent results. In: *Lagrangian Analysis and Prediction of Coastal and Ocean Dynamics* (eds Mariano A, Rossby T, Kirwan D), pp. 39–67. Cambridge University Press, Cambridge, UK.

Martin AP, Palumbi SR (1993) Body size, metabolic rate, generation time, and the molecular clock. *Proceedings of the National Academy of Sciences, USA*, **90**, 4087–4091.

Matschiner M, Salzburger W (2009) TANDEM: integrating automated allele binning into genetics and genomics workflows. *Bioinformatics*, [Epub ahead of print] doi: 10.1093/bioinformatics/btp303.

Mayr E (1984) Evolution of fish species flocks: a commentary. In: *Evolution of Fish Species Flocks* (eds Echelle AT, Kornfield A), pp. 3–12. University of Maine at Orono Press, Orono, Maine.

McMillan WO, Palumbi SR (1997) Rapid rate of control-region evolution in Pacific butterflyfishes (Chaetodontidae). *Journal of Molecular Evolution*, **45**, 473–484.

Nabholz B, Glemin S, Galtier N (2008) Strong variations of mitochondrial mutation rate across mammals — the longevity hypothesis. *Molecular Biology and Evolution*, **25**, 120–130.

North AW (2001) Early life history strategies of notothenioids at South Georgia. *Journal of Fish Biology*, **58**, 496–505.

North AW, Murray AWA (1992) Abundance and diurnal vertical distribution of fish larvae in early spring and summer in a fjord at South Georgia. *Antarctic Science*, **4**, 405–412.

O'Connell M, Wright JM (1997) Microsatellite DNA in fishes. *Reviews in Fish Biology and Fisheries*, **7**, 331–363.

Papetti C, Zane L, Patarnello T (2006) Isolation and characterization of microsatellite loci in the icefish *Chionodraco rastrospinosus* (Perciformes, Notothenioidea, Channichthyidae). *Molecular Ecology Notes*, **6**, 207–209.

Papetti C, Liò P, Rüber L, Patarnello T, Zardoya R (2007a) Antarctic fish mitochondrial genomes lack ND6 gene. *Journal of Molecular Evolution*, **65**, 519–528.

Papetti C, Susana E, Mesa ML *et al.* (2007b) Microsatellite analysis reveals genetic differentiation between year-classes in the icefish *Chaenocephalus aceratus* at South Shetlands and Elephant Island. *Polar Biology*, **30**, 1605–1613.

Parker RW, Paige KN, DeVries AL (2002) Genetic variation among populations of the Antarctic toothfish: evolutionary insights and implications for conservation. *Polar Biology*, **25**, 256–261.

Patarnello T, Marcato S, Zane L, Varotto V, Bargelloni L (2003) Phylogeography of the *Chionodraco* genus (Perciformes, Channichthydae) in the Southern Ocean. *Molecular Phylogenetics and Evolution*, **28**, 420–429.

Posada D, Crandall KA (1998) MODELTEST: testing the model of DNA substitution. *Bioinformatics*, **14**, 817–818.

Puebla O, Bermingham E, Guichard F (2008) Population genetic analyses of *Hypoplectus* coral reef fishes provide evidence that local processes are operating during the early stages of marine adaptive radiations. *Molecular Ecology*, **17**, 1405–1415.

R Development Core Team (2008) R: *A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0. www. R-project.org.

Reilly A, Ward RD (1999) Microsatellite loci to determine population structure of the Patagonian toothfish *Dissostichus eleginoides*. *Molecular Ecology*, **8**, 1753–1754.

Rice WR (1989) Analysing tables of statistical tests. *Evolution*, **43**, 223–225.

Rico C, Turner GF (2002) Extreme microallopatric divergence in a cichlid species from Lake Malawi. *Molecular Ecology*, **11**, 1585–1590.

Rogers AR, Harpending H (1992) Population growth makes waves in the distribution of pairwise genetic differences. *Molecular Biology and Evolution*, **9**, 552–569.

Rogers AD, Morley S, Fitzcharles E, Jarvis K, Belchier M (2006) Genetic structure of Patagonian toothfish (*Dissostichus eleginoides*) populations on the Patagonian Shelf and Atlantic and western Indian Ocean Sectors of the Southern Ocean. *Marine Biology*, **149**, 915–924.

Rosenblum EB (2006) Convergent evolution and divergent selection: lizards at the white sands ecotone. *The American Naturalist*, **167**, 1–15.

Ryman N, Palm S (2006) POWSIM: a computer program for assessing statistical power when testing for genetic differentiation. *Molecular Ecology Notes*, **6**, 600–602.

Salzburger W (2008) To be or not to be a hamlet pair in sympatry. *Molecular Ecology*, **17**, 1397–1399.

Salzburger W (2009) The interaction of sexually and naturally selected traits in the adaptive radiations of cichlid fishes. *Molecular Ecology*, **18**, 169–185.

Sato A, Takezaki N, Tichy H et al. (2003) Origin and speciation of haplochromine fishes in East African crater lakes investigated by the analysis of their mtDNA, Mhc genes, and SINEs. *Molecular Biology and Evolution*, **20**, 1448–1462.

Schluter D (2000) *The Ecology of Adaptive Radiation*. Oxford University Press, New York.

Schneppenheim R, Kock KH, Duhamel G, Janssen G (1994) On the taxonomy of the *Lepidonotothen squamifrons* group (Pisces, Perciformes, Notothenioidei). *Archive of Fishery and Marine Research*, **42**, 137–148.

Shaw PW, Arkhipkin A, Al-Khairulla H (2004) Genetic structuring of Patagonian toothfish populations in the Southwest Atlantic Ocean: the effect of the Antarctic Polar Front and deep-water troughs as barriers to genetic exchange. *Molecular Ecology*, **13**, 3293–3303.

Sidell BD, O'Brien K (2006) When bad things happen to good fish: the loss of hemoglobin and myoglobin expression in Antarctic icefishes. *The Journal of Experimental Biology*, **209**, 1791–1802.

Silva S. N, Helbling EW, Villafañe VE, Amos AF, Holm-Hansen O (1995) Variability in nutrient concentrations around Elephant Island, Antarctica, during 1991–93. *Polar Research*, **14**, 69–82.

Slatkin M (1987) Gene flow and the geographic structure of natural populations. *Science*, **236**, 787–792.

Slatkin M (1994) Linkage disequilibrium in growing and stable populations. *Genetics*, **137**, 331–336.

Slatkin M (2005) Seeing ghosts: the effect of unsampled populations on migration rates estimated for sampled populations. *Molecular Ecology*, **14**, 67–73.

Slatkin M, Excoffier L (1996) Testing for linkage disequilibrium in genotypic data using the Expectation-Maximization algorithm. *Heredity*, **76**, 377–383.

Smith P, Gaffney PM (2000) Toothfish stock structure revealed with DNA methods. *Water and Atmosphere*, **8**, 17–18.

Smith PJ, Gaffney PM (2005) Low genetic diversity in the Antarctic toothfish (*Dissostichus mawsoni*) observed with mitochondrial and intron DNA markers. *CCAMLR Science*, **12**, 43–51.

Smith P, McVeagh M (2000) Allozyme and microsatellite DNA markers of toothfish population structure in the Southern Ocean. *Journal of Fish Biology*, **57**, 72–83.

Sturmbauer C, Baric S, Salzburger W, Rüber L, Verheyen E (2001) Lake level fluctuations synchronize genetic divergences of cichlid fishes in African lakes. *Molecular Biology and Evolution*, **18**, 144–154.

Susana E, Papetti C, Barbisan F et al. (2007) Isolation and characterization of eight microsatellite loci in the icefish *Chaenocephalus aceratus* (Perciformes, Notothenioidei, Channichthyidae). *Molecular Ecology Notes*, **7**, 791–793.

Swearer SE, Shima JS, Hellberg ME et al. (2002) Evidence of self-recruitment in demersal marine populations. *Bulletin of Marine Science*, **70**, 251–271.

Swofford DL (2003) *PAUP*. Phylogenetic Analysis Using Parsimony (*and Other Methods)*, ed. 4.04a. Sinauer Associates, Sunderland, Massachusetts.

Thatje S, Hillenbrand CD, Mackensen A, Larter R (2008) Life hung by a thread: endurance of Antarctic fauna in glacial periods. *Ecology*, **89**, 682–692.

Thomas CD, Bodsworth EJ, Wilson RJ et al. (2001) Ecological and evolutionary processes at expanding range margins. *Nature*, **411**, 577–581.

Tomczak M, Godfrey JS (2003) *Regional Oceanography: An Introduction*. Daya Publications House, Delhi, India.

van Houdt JKJ, Hellemans B, van de Putte A, Koubbi P, Volckaert FAM (2006) Isolation and multiplex analysis of six polymorphic microsatellites in the Antarctic notothenioid fish, *Trematomus newnesi*. *Molecular Ecology Notes*, **6**, 157–159.

Van Oosterhout C, Hutchinson WF, Wills DPM, Shipley P (2004) Micro-Checker: software for identifying and correcting genotyping errors in microsatellite data. *Molecular Ecology Notes*, **4**, 535–538.

Volckaert FAM, Hänfling B, Hellemans B, Carvalho GR (2002) Timing of the population dynamics of bullhead *Cottus gobio* (Teleostei: Cottidae) during the Pleistocene. *Journal of Evolutionary Biology*, **15**, 930–944.

Wakeley J (2000) The effects of subdivision on the genetic divergence of populations and species. *Evolution*, **54**, 1092–1101.

White MG (1998) Development, dispersal and recruitment — a paradox for survival among Antarctic fish. In: *Fishes of Antarctica. A Biological Overview* (eds di Prisco G, Pisano E, Clarke A), pp. 53–62. Springer-Verlag Publishers, Milan, Italy.

Whitehouse MJ, Priddle J, Symon C (1996) Seasonal and annual change in seawater temperature, salinity, nutrient and chlorophyll a distributions around South Georgia, South Atlantic. *Deep Sea Research Part I: Oceanographic Research Papers*, **43**, 425–443.

Supporting Information

Gene flow by larval dispersal in the Antarctic notothenioid fish *Gobionotothen gibberifrons*

Michael Matschiner[*], **Reinhold Hanel**[†], **and Walter Salzburger**[*]

Addresses:
[*]Zoological Institute, University of Basel, Vesalgasse 1, CH-4051 Basel, Switzerland
[†]Institute of Fisheries Ecology, Johann Heinrich von Thünen-Institute, Federal Research Institute for Rural Areas, Forestry and Fisheries, Palmaille 9, D-22767 Hamburg, Germany

Correspondence:
Walter Salzburger, Fax: +41 61 267 03 01, E-mail: walter.salzburger@unibas.ch

1 Figures

Fig. S1: Statistical power tests for the mitochondrial control region data. Expected F_{ST} values of 0.001, 0.0025, 0.005, 0.01, and 0.02 were tested using the software POWSIM.
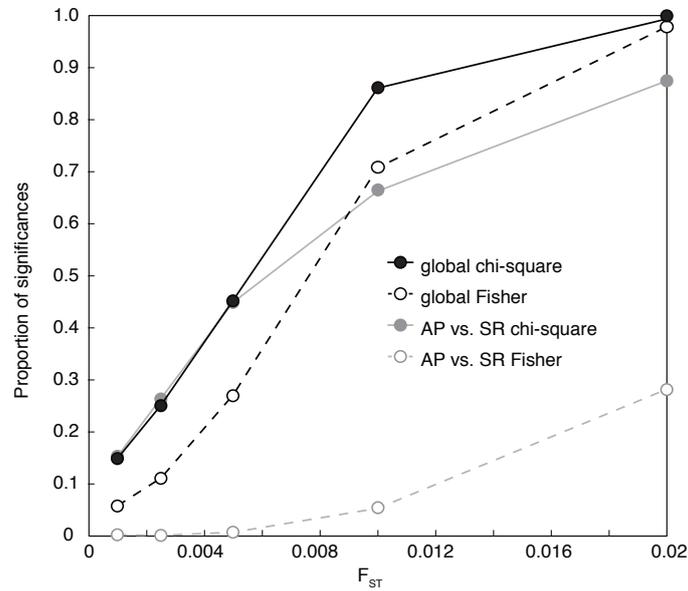
Fig. S2: Statistical power tests for the microsatellite data set. Expected F_{ST} values of 0.001, 0.0025, 0.005, and 0.01 were tested using the software POWSIM.
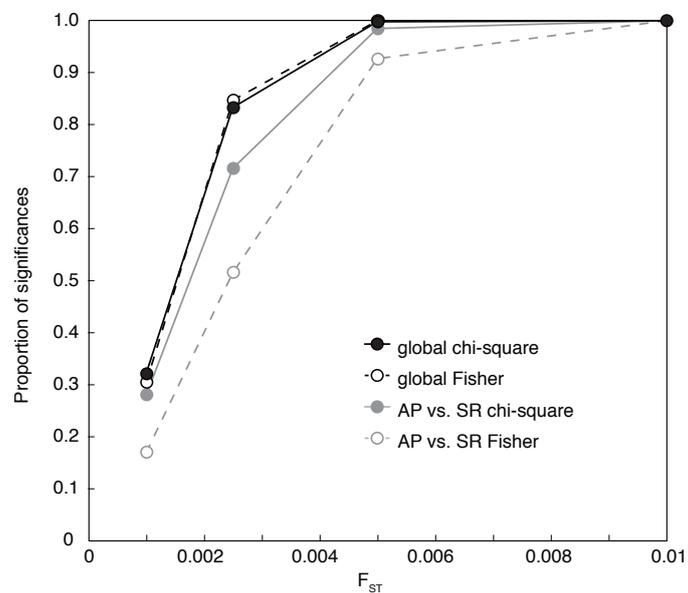
2   Tables

Tab. S1: Locus name, repeat motif, forward (F) and reverse (R) primer sequences, fluorescent dye as well as PCR protocol details (Number of thermocycles and annealing temperatures, T_A) are reported for nine microsatellite loci. Loci Trne35, Trne37, Trne53 and Trne66 were isolated from another nototheniid species, *T. newnesi*, while Cr38, Cr127, Cr170 Cr259 and Ca26 were isolated from channichthyid notothenioids, *C. rastrospinosus* and *C. aceratus*.

Locus	Repeat motif	Primer sequence (5' → 3')	Dye	Cycles	T_A
Trne35	$(TG)_n$	F: ACTGAAGCATGCTGGGAACT	HEX	37	59°C
		R: CGTGTTGAGGCCCGTCAG			
Trne37	$(TG)_n$	F: AGGTGAGTGCTTGCGTGTCAG	NED	32	60°C
		R: GCACTCCATACAGACAAGCACGCT			
Trne53	$(AC)_n$	F: ACACTCCCACCAGCAACC	6-FAM	37	59°C
		R: GCCTTGTGACAGCCTGGAC			
Trne66	$(CA)_n$	F: TGCTTGGACAGACTCCAGC	6-FAM	32	60°C
		R: TGGTAGTGGAGACATGCACAC			
Cr38	$(AC)_n$	F: ACGCCATGCTAATCAGAATC	NED	31	60°C
		R: GAGTCCCCACACATGACTGT			
Cr127	$(GT)_n ATAATGA(GT)_n$	F: CGTATAGGGCCGTACCTCA	HEX	31	60°C
		R: GCTCCATCATAGATCCAGTCA			
Cr170	$(AC)_n(GCAC)_n$	F: AGTACTATTACGCCTGGGTCT	6-FAM	31	60°C
		R: ACTCTCCTCCACTTTATTGTTG			
Cr259	$(AG)_n GG(AG)_n$	F: TGATTACTTCCATCTTCACACATA	VIC	37	59°C
		R: CACAAAGAATTCTGGGAACAG			
Ca26	$(TGCGTG)_n$	F: AAGGTGGGCAACAGGTTAGAGT	NED	37	59°C
		R: ATGAACACATACAAGTGGTCACAT			

Tab. S2: Frequencies of haplotypes among populations. All D-loop sequences were submitted to GenBank (Accession numbers FJ528746-FJ528907).

Haplotype	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
Elephant Island	2	1	24	7	1	0	0	1	4	1	1	1	1	1	1	1
Joinville Island A	0	0	11	4	1	0	0	1	2	0	1	1	0	0	1	0
Joinville Island B	0	1	14	1	0	0	0	4	4	0	0	0	0	0	1	0
King George Island	0	0	15	7	1	0	0	0	3	0	0	0	0	0	0	0
South Georgia	1	1	1	2	1	1	1	0	0	0	0	0	0	0	0	0
S. Sandwich Islands	0	0	3	1	1	0	0	2	2	0	0	0	0	0	0	0
Total	3	3	68	22	5	1	1	8	15	1	2	2	1	1	3	1

Haplotype	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32
Elephant Island	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Joinville Island A	0	0	0	0	0	0	0	2	0	2	2	0	0	0	0	0
Joinville Island B	0	0	0	0	0	1	0	0	1	0	0	2	1	1	1	1
King George Island	0	0	1	1	1	2	1	1	2	0	0	0	0	0	0	0
South Georgia	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
S. Sandwich Islands	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Total	1	1	1	1	1	3	1	3	3	2	2	2	1	1	1	1

Tab. S3: Among group genetic differentiation, based on mtDNA and calculated by AMOVA. A transition/transversion ratio of 1:1 was used for calculations. All possible groupings were tested, but only the ten best groupings are shown, sorted by Φ_{CT}. EI: Elephant Island, JIa: Joinville Island a, JIb: Joinville Island b, KGI: King George Island, SG: South Georgia, SSI: South Shetland Islands.

Grouping	Sum of squares (d.f.)	Variance component	Percentage of variation	Φ-statistics	P
(JIa,KGI)(EI)(JIb)(SG)(SSI)	3.922 (4)	Va = 0.0204	3.03	$\Phi_{CT} = 0.030$	0.066
(EI,JIa,JIb,KGI)(SG)(SSI)	1.928 (2)	Va = 0.0173	2.54	$\Phi_{CT} = 0.025$	0.069
(EI,JIa,JIb,KGI,SG)(SSI)	0.981 (1)	Va = 0.0155	2.27	$\Phi_{CT} = 0.023$	0.168
(JIa,KGI,SG)(JIb,SSI)(EI)	2.511 (2)	Va = 0.0145	2.15	$\Phi_{CT} = 0.021$	0.017
(EI,JIa,JIb,KGI,SSI)(SG)	0.931 (2)	Va = 0.0142	2.08	$\Phi_{CT} = 0.021$	0.330
(EI,JIa,KGI)(JIb,SSI)(SG)	2.309 (2)	Va = 0.0138	2.04	$\Phi_{CT} = 0.020$	0.032
(EI,JIa,JIb,SSI)(SG,SSI)	1.113 (1)	Va = 0.0132	1.95	$\Phi_{CT} = 0.019$	0.065
(JIa,KGI,SG)(EI,JIb)(SSI)	2.371 (2)	Va = 0.0125	1.86	$\Phi_{CT} = 0.019$	0.050
(EI,JIa,KGI)(JIb)(SG)(SSI)	2.912 (3)	Va = 0.0120	1.78	$\Phi_{CT} = 0.018$	0.100
(JIb,SSI)(EI)(JIa)(KGI)(SSI)	3.634 (4)	Va = 0.0118	1.76	$\Phi_{CT} = 0.018$	0.139

Tab. S4: Among group genetic differentiation, based on mtDNA and calculated by AMOVA. The observed transition/transversion ratio of 2:1 was used for calculations. All possible groupings were tested, but only the ten best groupings are shown, sorted by Φ_{CT}. Abbreviations as in Table S8.

Grouping	Sum of squares (d.f.)	Variance component	Percentage of variation	Φ-statistics	P
(EI,JIa,JIb,KGI,SSI)(SG)	1.249 (1)	Va = 0.0298	3.78	$\Phi_{CT} = 0.038$	0.162
(EI,JIa,JIb,KGI)(SG)(SSI)	2.262 (2)	Va = 0.0224	2.88	$\Phi_{CT} = 0.029$	0.065
(JIa,KGI)(EI)(JIb)(SG)(SSI)	4.406 (4)	Va = 0.0216	2.83	$\Phi_{CT} = 0.028$	0.067
(JIb,SSI)(EI)(JIa)(KGI)(SG)	4.163 (4)	Va = 0.0165	2.17	$\Phi_{CT} = 0.022$	0.130
(EI,JIa,KGI)(JIb,SSI)(SG)	2.662 (2)	Va = 0.0165	2.15	$\Phi_{CT} = 0.021$	0.016
(EI,JIb,SSI)(JIa,KGI)(SG)	2.686 (2)	Va = 0.0146	1.90	$\Phi_{CT} = 0.019$	0.117
(EI,JIa,JIb,SSI)(SG,SSI)	1.248 (1)	Va = 0.0144	1.87	$\Phi_{CT} = 0.019$	0.068
(EI,JIa,KGI)(JIb)(SG)(SSI)	3.304 (3)	Va = 0.0136	1.77	$\Phi_{CT} = 0.018$	0.050
(EI,JIb)(JIa,KGI)(SG,SSI)	2.613 (2)	Va = 0.0121	1.59	$\Phi_{CT} = 0.016$	0.067
(JIa,KGI,SG)(JIb,SSI)(EI)	2.613 (2)	Va = 0.0114	1.49	$\Phi_{CT} = 0.015$	0.048

Tab. S5: Levels of genetic differentiation, as calculated by hierarchical AMOVA of mitochondrial D-loop sequences (10 000 permutations). AP populations were pooled, and SR populations constituted separate groups.

Source of variation	Sum of squares (d.f.)	Variance component	Percentage of variation	Φ-statistics		P
among groups	1.928 (2)	Va = 0.0173	2.54	Φ_{CT} =	0.025	0.069
among populations	2.309 (3)	Vb = 0.0030	0.44	Φ_{SC} =	0.004	0.231
within populations	103.411 (156)	Vc = 0.6629	97.03	Φ_{ST} =	0.023	0.131

Tab. S6: p values of pairwise Φ_{ST}s and F_{ST}s, based on mtDNA (below diagonal) and microsatellites (above diagonal). After Bonferroni correction, none of the p values are significant at the table-wide 5% significance level.

	EI	JIa	JIb	KGI	SG	SSI
Elephant Island		0.475	0.752	0.327	0.462	0.848
Joinville Island A	0.205		0.536	0.777	0.597	0.862
Joinville Island B	0.263	0.168		0.048	0.273	0.850
King George Island	0.196	0.898	0.199		0.237	0.511
South Georgia	0.129	0.455	0.159	0.304		0.906
S. Sandwich Islands	0.067	0.210	0.493	0.123	0.410	

Tab. S7: Diversity indices for nine microsatellite loci. Reported are number of alleles n_A, fragment size range, observed heterozygosity H_O, expected heterozygosity H_E and probability p (based on 10 000 permutations) of observing even larger differences between H_O and H_E under the assumption of Hardy-Weinberg equilibrium. The latter three are given individually for the six investigated populations. No tests were done for monomorphic alleles. After Bonferroni correction, Trne35 significantly departs from Hardy-Weinberg equilibrium at the table-wide 0.001% significance level in the total sample set.

	n_A	Size range	Elephant Island H_O	H_E	P	Joinville Island A H_O	H_E	P	Joinville Island B H_O	H_E	P
Trne35	61	292-538	0.837	0.974	0.000	0.900	0.972	0.047	0.906	0.968	0.218
Trne37	35	130-252	0.898	0.910	0.077	0.900	0.889	0.583	0.906	0.896	0.215
Trne53	43	312-422	0.959	0.973	0.567	0.900	0.964	0.229	0.879	0.968	0.020
Trne66	34	272-364	0.898	0.953	0.037	0.800	0.946	0.002	0.875	0.935	0.167
Cr38	35	186-276	0.755	0.840	0.052	0.900	0.875	0.949	0.871	0.877	0.466
Cr127	5	116-128	0.061	0.060	1.000	0.100	0.155	0.163	0.091	0.144	0.155
Cr170	1	198	-	-	-	-	-	-	-	-	-
Cr259	38	222-362	0.959	0.952	0.834	1.000	0.964	1.000	1.000	0.961	1.000
Ca26	31	152-230	0.938	0.936	0.290	0.900	0.929	0.740	0.970	0.951	0.770

	King George Island H_O	H_E	P	South Georgia H_O	H_E	P	S. Sandwich Islands H_O	H_E	P	Total H_O	H_E	P
Trne35	0.882	0.963	0.000	0.875	0.958	0.355	0.800	0.968	0.057	0.871	0.972	0.000
Trne37	0.943	0.900	0.930	1.000	0.917	1.000	0.800	0.895	0.630	0.909	0.898	0.083
Trne53	0.914	0.969	0.078	0.875	0.933	0.583	1.000	0.963	1.000	0.921	0.967	0.060
Trne66	0.886	0.945	0.335	1.000	0.958	1.000	0.900	0.947	0.556	0.878	0.949	0.279
Cr38	0.857	0.795	0.800	1.000	0.967	1.000	0.900	0.905	0.379	0.846	0.854	0.158
Cr127	0.057	0.056	1.000	-	-	-	0.200	0.195	1.000	0.079	0.099	0.078
Cr170	-	-	-	-	-	-	-	-	-	-	-	-
Cr259	0.800	0.942	0.000	1.000	0.967	1.000	1.000	0.979	1.000	0.945	0.955	0.842
Ca26	0.914	0.929	0.192	1.000	0.933	1.000	1.000	0.953	0.500	0.939	0.938	0.532

Tab. S8: Among group genetic differentiation, based on eight microsatellites and calculated by AMOVA. All possible groupings were tested, but only the ten best groupings are shown, sorted by Φ_{CT}. Abbreviations as in Table S8.

Source of variation	Sum of squares (d.f.)	Variance component	Percentage of variation	Φ-statistics	P
(SG,SSI)(EI)(JIa)(JIb)(KGI)	13.850 (4)	Va = 0.0256	0.78Φ	$_{CT} = 0.008$	0.068
(JIb,SSI)(EI)(JIa)(KGI)(SG)	13.853 (4)	Va = 0.0157	0.48Φ	$_{CT} = 0.005$	0.134
(JIa,SSI)(EI)(JIb)(KGI)(SG)	13.663 (4)	Va = 0.0095	0.29Φ	$_{CT} = 0.003$	0.199
(EI,SSI)(JIa)(JIb)(KGI)(SG)	13.681 (4)	Va = 0.0094	0.29Φ	$_{CT} = 0.003$	0.267
(EI,JIb)(JIa,KGI)(SG,SSI)	7.710 (2)	Va = 0.0086	0.26Φ	$_{CT} = 0.003$	0.067
(EI,JIb,SSI)(JIa,KGI)(SG)	7.626 (2)	Va = 0.0084	0.25Φ	$_{CT} = 0.003$	0.016
(EI,JIb,SSI)(JIa)(KGI)(SG)	10.700 (3)	Va = 0.0083	0.25Φ	$_{CT} = 0.003$	0.051
(JIa,SG,SSI)(EI)(JIb)(KGI)	10.465 (3)	Va = 0.0079	0.24Φ	$_{CT} = 0.002$	0.100
(JIb,SG,SSI)(EI)(JIa)(KGI)	10.470 (3)	Va = 0.0079	0.24Φ	$_{CT} = 0.002$	0.158
(EI,JIb,SG,SSI)(JIa,KGI)	4.201 (1)	Va = 0.0073	0.22Φ	$_{CT} = 0.002$	0.063

Tab. S9: Levels of genetic differentiation in eight microsatellite loci, calculated by means of hierarchical AMOVA (10 000 permutations). AP populations were pooled, and SR populations constituted separate groups.

Source of variation	Sum of squares (d.f.)	Variance component	Percentage of variation	Φ-statistics	P
among groups	6.426 (2)	Va = -0.0033	-0.10Φ	$_{CT} = -0.001$	0.802
among populations	10.275 (3)	Vb = 0.0020	0.06Φ	$_{SC} = 0.001$	0.427
within populations	1063.166 (324)	Vc = 3.2813	100.04Φ	$_{ST} = 0.000$	0.583

Tab. S10: Estimated log probabilities of the microsatellite data, given the number of assumed genetic clusters, K, in the data set.

	$\ln \Pr(X|K)$		
K	Replicate 1	Replicate 2	Replicate 3
1	-7 495.0	-7 494.7	-7 495.5
2	-7 571.5	-7 532.2	-7 577.7
3	-7 650.2	-7 625.4	-7 856.8
4	-8 096.5	-7 752.3	-8 030.3
5	-8 470.2	-8 892.7	-7 961.9
6	-8 896.2	-8 896.2	-9 014.4

2.2 Comparative population genetics of seven notothenioid fish species reveals high levels of gene flow along ocean currents in the southern Scotia Arc, Antarctica

Damerau M, Matschiner M, Salzburger W, Hanel R
Submitted to *Polar Biology*

Personal contribution:

In the study of Damerau et al., I contributed to study design, population genetic analyses, construction of haplotype genealogies, analysis of drifter trajectories, application of the isolation-with-migration model, and writing of the manuscript.

Comparative population genetics of seven notothenioid fish species reveals high levels of gene flow along ocean currents in the southern Scotia Arc, Antarctica

Malte Damerau • Michael Matschiner • Walter Salzburger • Reinhold Hanel

R. Hanel (✉)

Institute of Fisheries Ecology, Johann Heinrich von Thünen-Institute, Federal Research Institute for Rural Areas, Forestry and Fisheries, Palmaille 9, D-22767 Hamburg, Germany

email: reinhold.hanel@vti.bund.de

tel: (+49) 40 38905290

fax: (+49) 40 38905261

M. Damerau

Institute of Fisheries Ecology, Johann Heinrich von Thünen-Institute, Federal Research Institute for Rural Areas, Forestry and Fisheries, Palmaille 9, D-22767 Hamburg, Germany

M. Matschiner • W. Salzburger

Zoological Institute, University of Basel, Vesalgasse 1, CH-4051 Basel, Switzerland

1

Abstract

The Antarctic fish fauna is characterized by high endemism and low species diversity with one perciform suborder, the Notothenioidei, dominating the whole species assemblage on the shelves and slopes. Notothenioids diversified *in situ* through adaptive radiation and show a variety of life history strategies as adults ranging from benthic to pelagic modes. Their larval development is unusually long, lasting from a few months to more than a year, and generally includes a pelagic larval stage. Therefore, the advection of eggs and larvae with ocean currents is a key factor modulating population connectivity. Here, we compare the genetic population structures and gene flow of seven ecologically distinct notothenioid species of the southern Scotia Arc based on nuclear microsatellites and mitochondrial DNA sequences (D-loop/cytochrome *b*). The seven species belong to the families Nototheniidae (*Gobionotothen gibberifrons, Lepidonotothen squamifrons, Trematomus eulepidotus, T. newnesi*) and Channichthyidae (*Chaenocephalus aceratus, Champsocephalus gunnari, Chionodraco rastrospinosus*). Our results show low population differentiation and high gene flow for all investigated species independent of their adult life history strategies. In addition, gene flow is primarily in congruence with the prevailing ocean current system, highlighting the role of larval dispersal in population structuring of notothenioids.

Keywords

Introduction

The Southern Ocean surrounding the Antarctic continent is a unique marine environment and its fish fauna is characterized by a high degree of endemism at low species diversity (Andriashev 1987; Eastman 1993, 2005). The northern boundary of the Southern Ocean is delimited by the Antarctic Convergence at about 50-60°S, which is marked by a sharp decrease of surface temperature from north to south and constitutes a thermal barrier for many marine organisms existing since approximately 22-25 My (Dayton et al. 1994; Eastman & McCune 2000). Beside its thermal isolation, the formation of deep circum-polar currents like the Antarctic Circumpolar Current (ACC) as well as large distances and deep ocean basins

2

between the Antarctic continental shelf and those of adjacent continents form additional oceanographic, geographic and bathymetric barriers to migration and dispersal. The Antarctic ichthyofauna as known today consists of 322 species from 50 families, with about 88% being endemic to the waters south of the Antarctic Convergence. A single group of fish, the perciform suborder Notothenioidei, dominates the species assemblage on the shelves and slopes (Andriashev 1987; Eastman 2005).

Notothenioids consist of 131 species in 8 families, with 104 species out of 5 of these families being endemic to the Antarctic region, where they constitute up to 77% of species diversity and 91% of biomass on the shelves and slopes of the continent and nearby islands (Eastman 2005). Together with members of the families Zoarcidae (24 species) and Liparidae (70 species), they comprise 88% of the Antarctic fish fauna (Eastman & McCune 2000). But unlike the latter two families that probably invaded the area from North Pacific waters, notothenioids diversified *in situ* in the course of an adaptive radiation (Eastman 1993; Clarke & Johnston 1996; Eastman & McCune 2000; Matschiner et al. 2011). This radiation is thought to have been triggered by the acquisition of antifreeze glycoproteins (AFGPs) that keep body fluids from freezing in the ice-laden waters of Antarctica (Cheng 1998; Matschiner et al. 2011). While the cooling of Antarctic waters as well as repeated expansions and retreats of the Antarctic ice sheet forced most Antarctic species of the Oligocene to either shift their distribution northwards or into deeper waters, or otherwise led to their extinction (Briggs 2003; Barnes & Conlan 2007), notothenioids radiated in the absence of competitors and filled vacant ecological niches (Eastman 1991). Although about one-half of today's species show a demersal life-style (as is also presumed for their ancestors), the notothenioid radiation is largely based on diversification related to niches in the water column (Eastman 1993). Since notothenioids lack a swim bladder, buoyancy for pelagization is gained through extended lipid depositions and reduced ossification (Eastman 1993), resulting in a variety of epibenthic, semipelagic, cryopelagic and pelagic life strategies. Adaptive radiations in the marine realm are rare compared to those known from freshwater systems like e.g. cichlid fishes in the Great Lakes of East Africa (Seehausen 2006; Salzburger 2009; Matschiner et al. 2010) or are camouflaged by subsequent dispersal in the course of evolution. The notothenioids therefore constitute a prime example for a marine adaptive radiation, making their ecological and morphological diversification a highly interesting target for evolutionary studies (Eastman 2000).

In contrast to the variety of adult life history strategies, the early larval development in notothenioids is always pelagic. Depending on the species and locality, the larval stage may be completed

3

within two months after hatching or last more than one year (Kellermann 1986, 1989; North 2001). During this stage, strong currents like the clockwise ACC (Westwind Drift) or the counterclockwise Eastwind Drift along the Antarctic continent are likely to modulate larval dispersal away from the shelves into the open ocean. This may cause substantial losses to spawning populations and can lead to source-sink relationships by transporting larvae downstream towards distant shelf habitats. As White (1998) pointed out, the prolonged pelagic early life history strategy in notothenioids is at odds with a successful larval survival strategy. However, ichthyoplankton studies have shown that larval abundances for demersal species are surprisingly high on the shelf areas and decrease with increasing distance to the coast, despite the fact that their distributions are generally influenced by bathymetry, hydrography and seasonal events (Loeb et al. 1993; White 1998). Local retention mechanisms, such as gyres formed behind islands or shelf-break frontal systems limit offshore transport of larvae (White 1998) and should increase genetic heterogeneity between populations of different shelves, thereby fostering speciation.

Population genetic studies in notothenioids provide evidence that the oceanography of the Southern Ocean indeed has large influence on the genetic structure of populations. Populations of species with circumpolar distributions as for example the pelagic Antarctic toothfish *Dissostichus mawsoni* and the more sedentary benthopelagic Patagonian toothfish *D. eleginoides* are not significantly differentiated over large parts of their distribution range (Smith & Gaffney 2005; Rogers et al. 2006). However, these results do not imply complete absence of genetic heterogeneity, as other genetic markers were able to resolve differentiations on varying geographic scales (Parker et al. 2002; Shaw et al. 2004; Kuhn & Gaffney 2008). Populations connected along currents like the ACC are often found to be more closely related than those located in proximity but separated across frontal systems like e.g. the Polar Front (Shaw et al. 2004; Rogers et al. 2006). Even strictly benthic species like the humped rockcod *Gobionotothen gibberifrons* show no signs of differentiation among populations separated geographically by nearly 2000 km and bathymetrically by deep basins (Matschiner et al. 2009). By combining oceanographic data with population genetic signatures, Matschiner et al. (2009) showed that dispersal of pelagic larvae in *G. gibberifrons* is most probably the major means of gene flow in this otherwise benthic species. The contribution of larval dispersal to population structure and in the long term on species' biogeography is still lively debated not only in notothenioids, but also in fish from warmer waters with distinctly shorter pelagic early life stages (e.g. Taylor & Hellberg 2003; Bay et al. 2006; Cowen & Sponaugle 2009). In particular, it is unclear what influence a prolonged pelagic early life stage and the

4

existence of strong currents (which together should result in high levels of gene flow among populations) had on the adaptive radiation in notothenioids.

In this study, we compare the genetic signatures derived from microsatellites and mitochondrial (mt) DNA sequences of seven notothenioid species with different life history strategies and larval durations inhabiting the southern Scotia Arc (Table 1). It is the first time that the genetic population structures based on two types of genetic markers are compared between multiple notothenioid species. We also included data obtained from drifting buoys to infer the influence of larval dispersal with oceanic currents on gene flow. Our study area is the southern Scotia Arc, consisting of the tip of the Antarctic Peninsula (AP), South Shetland Islands (SSh) including Elephant Island (EI) located about 200 km north of the Peninsula and the South Orkney Islands (SO) approximately 420 km further east (Fig. 1). The shelves of the AP and SSh/EI are separated by trenches of more than 500 m depth, whereas the SO shelf is separated by depths of 2000-3000 m. This region of the Seasonal Pack-Ice Zone, which is ice-free during the austral summer, is largely influenced by two water regimes: the ACC flowing eastward through the Scotia Sea in the north, and water originating from the Weddell Sea in the south (Whitworth et al. 1994).

The species investigated in this study comprise the three channichthyids *Chaenocephalus aceratus*, *Champsocephalus gunnari* and *Chionodraco rastrospinosus* as well as the four nototheniids *Gobionotothen gibberifrons*, *Lepidonotothen squamifrons*, *Trematomus eulepidotus* and *T. newnesi*, which are all among the most abundant species in the southern Scotia Arc. Their life histories differ in a variety of traits (Table 1): *C. aceratus*, *G. gibberifrons* and *L. squamifrons* are benthic species of which the two former ones spend most of their time resting on the bottom (Fanta et al. 1994; Kock & Jones 2005). *C. gunnari, C. rastrospinosus* and *T. eulepidotus* show a benthopelagic life style preying for food in the water column (Rutschmann et al. 2011). Vertical migrations between near bottom layers during the day and sub-surface waters during the night are known from several notothenioids including *C. gunnari* (Kock & Everson 1997). *T. newnesi* shows a remarkable feeding plasticity and is considered a bentho-cryo-pelagic species (La Mesa et al. 2000). It is generally benthivorous, but carries out vertical migrations during summer feeding on pelagic organisms. In winter, when there is sea ice cover, *T. newnesi* switches to a cryo-pelagic mode and feeds on organisms under the ice (Daniels 1982; Casaux et al. 1990; La Mesa et al. 2000). We find that all seven species show none or only weak population differentiation in the southern Scotia Arc.

5

Methods

Sampling

The specimens analysed in this study were collected during expedition ANT-XXIII/8 aboard RV Polarstern in December 2005-January 2006 and U.S. AMLR (United States Antarctic Marine Living Resources) survey in February-March 2009 aboard RV Yuzhmorgeologiya. Sampling sites were located on the shelves at the tip of the AP, EI (the most easterly island of the SSh), and SO to their east (Fig. 1). Muscle tissue of the specimens was stored in 95% ethanol. DNA was extracted using two different protocols depending on the cruise. All 2005-2006 samples were extracted with the BioSprint 96 workstation (QIAGEN) following the manufacturer's protocol, whereas DNA from the AMLR 2009 samples was extracted by incubating muscle tissue in 300 µl 5%-Chelex solution containing 12 µl Proteinase K (20 mg/ml) for 3 hours at 55 °C, followed by a denaturation step of 25 minutes at 98 °C in a thermomixer.

MtDNA sequencing and data analysis

Depending on the amplification success, mitochondrial gene sequences were either generated from the control region/D-loop (*C. aceratus, C. gunnari, C. rastrospinosus* and *G. gibberifrons*) or cytochrome *b* (cyt *b*; L. *squamifrons, T. eulepidotus* and *T. newnesi*). Partial D-loop or cyt *b* were amplified with the primers LPR-02 and HDL2 (Derome et al. 2002) or NotCytbF and H15915n (Matschiner et al. 2011) respectively. For amplification of the D-loop region 2 µl template DNA were mixed with 7.5 µl Taq PCR Master Mix (QIAGEN), 0.5 µl of each 10 µM primer, 1 µl bovine serum albumin and 14.5 µl sterile water. A simplified hot start at 94 °C for 2 min initiated the PCR profile followed by 35 cycles of 94 °C for 30 s, 52 °C for 30 s and 72 °C for 90 s. Thermocycling finished with a final elongation step at 72 °C for 7 min. Cytochrome *b* sequences were amplified using Phusion polymerase (Finnzymes) following the manufacturers manual at 57 °C annealing temperature. PCR products were purified by adding 2 µl ExoSAP-IT (USB Corporation) to 5 µl PCR product following the manufacturer's instructions.

Sequencing PCR with forward primers was performed using the BigDye Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems). After purification with BigDye XTerminator (Applied

Biosystems), sequencing products were run on an AB3130*xl* Genetic Analyzer (Applied Biosystems). Sequences were automatically aligned with CodonCode Aligner (CodonCode Corp.), inspected by eye and corrected manually if necessary.

Basic sequence properties as well as intraspecific sequence polymorphisms measured as nucleotide diversity (π) and haplotype diversity (h) were examined with DNASP 5.10 (Librado & Rozas 2009). Population structure among sampling localities was assessed by analysis of molecular variance (AMOVA) calculated with 16000 permutations as implemented in ARELQUIN 3.5 (Excoffier & Lischer 2010).

Phylogenetic trees were inferred with the Maximum Likelihood (ML) method implemented in PAUP* 4.0a112 (Swofford 2003), whereby models of sequence evolution were selected according to BIC (Posada 2008). On the basis of these phylogenies, haplotype genealogies were constructed following the method described in Salzburger et al. (2011).

The statistical power of both mtDNA and microsatellites to detect significant genetic differentiation between populations was tested with POWSIM 4.0 (Ryman & Palm 2006) using both the Chi^2 test and Fishers exact test. Various levels of differentiation (measured as F_{ST} in the range from 0.001 to 0.08) were tested by combining different effective population sizes (N_e) and times since divergence (t). In addition, POWSIM allows calculating the α error (type I error), which is the probability of rejecting the null hypothesis of genetic homogeneity although it was true by drawing the alleles directly from the base population (t=0).

Microsatellite genotyping and data analysis

In addition to mtDNA sequences we included data of twelve previously published microsatellites in our analyses. Microsatellites Cr15, Cr38, Cr127, Cr236, Cr259 were originally isolated from *Chionodraco rastrospinosus* (Papetti et al. 2006), Trne20, Trne35, Trne37, Trne53, Trne55, Trne66 from *Trematomus newnesi* (Van Houdt et al. 2006) and Ca26 from *Chaenocephalus aceratus* (Susana et al. 2007). Marker sets for each species were composed of 8 to 10 microsatellites, depending on the amplification success (Online Resource 1).

All amplification reactions contained 5 µl Multiplex Master Mix (QIAGEN), 0.2 µl of each 10 µM primer, 0.8 µl template DNA and water added to a final volume of 10 µl. All reactions contained

primers for up to 3 microsatellites of which the forward primers were fluorescently labelled. The PCR profile was 95 °C for 15 min followed by 35 cycles of 94 °C for 30 s, 59 °C for 90 s, 72 °C for 90 s and final elongation at 72 °C for 10 min. Fragment lengths were determined with GeneScan LIZ500 size standard (Applied Biosystems) on an AB3130xl Genetic Analyzer (Applied Biosystems) and scored with GENEMAPPER 4.0 (Applied Biosystems).

Alleles were automatically binned with TANDEM (Matschiner & Salzburger 2009) and subsequently converted with CONVERT (Glaubitz 2004). We used a 3D factorial correspondence analysis as implemented in GENETIX (Belkhir et al. 2001) to visualize outliers in the data. Suspicious individuals with potential errors in the data were either corrected, re-genotyped or otherwise completely removed from the dataset.

Microsatellites were tested for the presence of null alleles, stuttering and large allele dropout with MICRO-CHECKER 2.2.3 (Van Oosterhout et al. 2004). Allele size ranges and genotypic linkage-disequilibrium between loci were examined with 1000 iterations in GENEPOP 4.0.10 (Raymond & Rousset 1995).

The number of alleles per sample and locus were calculated with FSTAT (Goudet 1995, 2001) using the implemented rarefaction method to account for differences in sample sizes.

Population structure was assessed performing an AMOVA as implemented in ARLEQUIN 3.5 (Excoffier & Lischer 2010). To account for biases attributed to null alleles present in the data, F_{ST} values were also calculated excluding null alleles with FREENA (Chapuis & Estoup 2007). ARLEQUIN was further used to test loci for Hardy-Weinberg-equilibrium (HWE) in each population. Molecular diversities were measured as mean number of pairwise differences and average gene diversity. The Garza-Williamson index was calculated as indicator of recent demographic history. This statistic is sensitive to population size reductions since a recent bottleneck usually reduces the number of alleles more than the allele size range hence leaving "vacant" positions in between. The index is supposed to be very small in populations having experienced a recent bottleneck and close to one in stationary populations (Garza & Williamson 2001). In a similar approach populations were tested for a recent reduction in effective population size with BOTTLENECK 1.2.02 (Piry et al. 1999) using the two-phase model (TPM) of mutation with 10% infinite allele model (IAM) and 90% single step mutation model (SMM) with a variance of 15% and 1000 iterations. Significance was tested with the Wilcoxon signed-rank test. In addition, we used the mode-shift indicator as qualitative descriptor of allele frequency distribution. A normal L-shaped

8

distribution indicates populations in mutation-drift equilibrium whereas a shifted mode is a sign for recent bottlenecks. We considered that a population truly underwent a bottleneck if this was indicated by all three measurements (low Garza-Williamson index, significant Wilcoxon signed-rank test and mode shift).

To further analyse the structure of populations and to identify clusters of individuals, we used a Bayesian approach based on the genotypes of microsatellites as implemented in STRUCTURE 2.3.1 (Pritchard et al. 2000). For every data set we ran simulations for up to 6 clusters (k) with 20 iterations each. The parameters were set to 10000 steps of burn-in period and 100000 MCMC replications thereafter. The admixture model was used as we expected a weak population structure as often encountered in marine fishes with long larval phases (and indicated by our calculated F_{ST} values). Alpha was inferred from an initial value of 1.0 and correlated allele frequencies with lambda set to 1.0. In a second approach we ran the program with the same settings but incorporated *a priori* information about the sampling sites to help with clustering. We followed the method of Evanno et al. (2005) to calculate Δk as indicator of the most likely number of clusters. The power of microsatellites to detect significant population structure was tested in the same way as the mtDNA sequences using POWSIM 4.0 (Ryman & Palm 2006) but in the F_{ST} range of 0.001-0.01.

Isolation-with-migration analyses

The directionality and extent of gene flow between AP/EI and SO populations was examined for the combined data set of mtDNA sequences and nuclear microsatellites with the isolation-with-migration (IM) model as implemented in IMA2 (Hey & Nielsen 2007). The Hasegawa-Kishino-Yano (HKY) model of sequence evolution was applied to mitochondrial sequences, and a stepwise mutation model (SMM) was assumed for all microsatellite loci. Inheritance scalars of 0.25 and 1 were assigned to mtDNA and microsatellite loci, respectively. Appropriate prior parameter ranges were determined in a series of initial runs. We chose wide population size parameter ranges Θ_1, Θ_2, Θ_A ∈ (0,500], a divergence time prior t ∈ (0,10], and exponential migration rate priors m_1, m_2 with distribution means of 5.0. Each run included 80 Metropolis-coupled Markov chains. Geometric heating scheme parameters were chosen to optimize chain swap rates, and set to ha = 0.97 and hb = 0.86. Per population comparison, ten replicate runs were conducted for 4.5 million generations, discarding the first 500 000 generations as burn-in. Genealogies of run replicates were jointly analysed in IMA2's 'Load Trees'-mode. The migration rates per year (M) were

calculated from the resulting parameters $M = \Theta \times m/2$ under consideration of the species' generation times.

Drifter analysis

In order to compare directionality of gene flow and ocean currents, we also analysed trajectories of satellite-tracked drifting buoys (hereafter called drifters) of the Global Drifter Program (Lumpkin & Pazos 2007), following established protocols (Matschiner et al. 2009). Interpolated drifter data was downloaded from http://www.aoml.noaa.gov/phod/dac/gdp.html for all drifters passing the AP/SO region (55-65˚S, 40-60˚W) between 15 February 1979 and 31 December 2009. Three polygons were mathematically defined to encompass the AP, EI, and SO shelf areas at 500m depth. Polygon vertices were 65.0˚S, 60.0˚W; 65.0˚S, 54.5˚W; 63.9˚S, 54.1˚W; 63.3˚S, 52.3˚W; 62.2˚S, 54.3˚W; 62.1˚S, 55.4˚W; and 63.3˚S, 60˚W for the AP, 62.7˚S, 60˚W; 61.4˚S, 54.1˚W; 61.1˚S, 53.9˚W; 60.8˚S, 55.7˚W; and 61.9˚S, 60.0˚W for EI, and 62.3˚S, 44.8˚W; 61.5˚S, 44.0˚W; 61.2˚S, 42.4˚W; 60.8˚S, 42.8˚W; 60.3˚S, 46.6˚W; 60.5˚S, 47.3˚W; and 61.6˚S, 46.9˚W for SO. Trajectories of drifters passing these polygons were plotted for 90 days, starting with the day of departure from one of the polygons (Fig. 2).

Results

mtDNA – genetic diversity and demographic history

The number of individuals successfully sequenced varied between species and ranged from 49 (*L. squamifrons*) to 194 (*G. gibberifrons*; Online Resource 2). The D-loop region could be amplified in 4 out of 7 species (*C. aceratus, C. gunnari, C. rastrospinosus and G. gibberifrons*). For the remaining three species *L. squamifrons*, *T. eulepidotus* and *T. newnesi* amplification of the D-loop region consistently failed. Therefore, we amplified a part of the mt cyt *b* gene as an alternative population genetic marker. Although we are aware that the use of two mtDNA markers is not optimal and hinders a direct comparison between datasets without restrictions, we found both markers to resolve the population genetic structures in a similar fashion. All sequences generated were submitted to GenBank (Accession

10

nos JN241690-JN241831) and a list of haplotypes per shelf area can be found in Online Resource 3. Although a clear relationship between sample size or sequence length and the number of haplotypes could generally be expected, no such trend was found (R^2=0.06 and R^2=0.18, respectively). Similarly, the genetic diversities varied between species, but irrespective of the locus or phylogenetic relationship. Remarkably, the three high-Antarctic species *C. rastrospinosus*, *T. eulepidotus* and *T. newnesi* had the highest nucleotide diversities (each >0.004) and a similar pattern arose from the haplotype diversities. On the population level, most samples from the AP region had higher diversities than SO samples. Unique to *C. rastrospinosus* the diversity for SO was higher than for AP.

mtDNA – genetic population structure

Power analyses revealed poor capabilities of the mtDNA sequences to detect subtle differentiations among populations for both D-loop and cyt *b*. High probability to detect true differentiation as low as $F_{ST} = 0.01$ was only evident for *G. gibberifrons* (Online Resource 4). Finescale genetic differentiations measured as pairwise F_{ST} between sampling localities (AP, EI and SO) were neither high nor significant for any species (data not shown). We therefore combined the AP and EI samples to test differentiation along the prevailing current from the AP/EI region towards SO. After pooling, differentiation between sample localities remained non significant and ranged between -0.003 (p = 0.48) for *C. gunnari* and 0.04 (p = 0.06) for *C. rastrospinosus* (Table 2). The constructed haplotype genealogies support these findings and reflect genetic diversity rather than differentiations between localities (Online Resource 5).

Microsatellites – genetic diversity and demographic history

Microsatellites were successfully genotyped for 56 (*C. rastrospinosus*) to 125 individuals (*C. aceratus*) per species (Online Resource 1). Individuals with missing data at one or more loci were excluded from the analyses. Significant genotypic linkage disequilibrium between loci was limited to the pair of Cr236 and Trne20 in *C. gunnari* (p = 0.03) (data not shown). Null alleles might be present at least at one locus in every species with the exception of *T. newnesi* (Online Resource 6). Also, in all species except *T. newnesi*, the hypothesis of HWE could be rejected. On the population level, HWE could also be detected

in the EI samples of *C. aceratus* (Online Resource 7). Among all species, *G. gibberifrons* had the highest (24.0±14.0) and *C. gunnari* the lowest (10.7±7.7) allelic richness (data not shown).

Reductions in population size were examined for species and populations. On the species level, none showed signs of a bottleneck concurrently in all three indicators examined (Online Resource 8). The probability of heterozygosity excess was significant only in *C. gunnari* and *C. rastrospinosus* (both p<0.01), while the Garza-Williamson index was low (0.48±0.21) only in *T. newnesi*. For every species, the frequency distributions of alleles were normal L shaped. On the population level, the only sample, which showed evidence for a recent bottleneck in all tests, was *T. newnesi* from the SO shelf. In *C. gunnari* from SO heterozygosity excess and mode shift were evident, but the Garza-Williamson index was relatively high (0.63±0.20).

Microsatellites – genetic population structure

The power of the microsatellites to detect significant population differentiation was generally much higher than for mtDNA sequences. Simulations suggest an average probability of 97% to detect a true differentiation of $F_{ST} = 0.01$ resulting from both the Chi^2 (SD = 0.03) and Fisher exact (SD = 0.02) tests (Online Resource 4).

Population structure based on microsatellites was assessed with AMOVA (Table 2) and in a Bayesian approach (Table 3). Similar to mtDNA sequences, we found no significant differentiation between sampling localities of the AP and EI region (data not shown) and therefore combined these samples to test the genetic structure along the current system. In congruence with mtDNA data, differentiations between AP/EI and SO populations were minor in every species. Indeed, F_{ST} values were mostly one order of magnitude lower than for mtDNA except for *C. aceratus* and *C. gunnari*. These two species were also the only ones showing significant differentiation. Excluding null alleles from analyses did not alter the previous findings of low differentiation in any species, although it also changed F_{ST} values in some cases by one order of magnitude. The overall genetic variation can rather be explained by larger differences between individuals than between populations.

12

Drifter analyses

Between 15 February 1979 and 31 December 2009, a total of 73 drifters crossed the AP, EI, and SO shelf areas (Fig. 2). Out of 64 drifters leaving the AP and EI shelf areas, one reached the SO shelf after 46 days, which is shorter than most notothenioid larval stages. This drifter had left the AP shelf at its easternmost end on 30 January 2008 and crossed the Philip Passage perpendicular to Weddell Sea Deep Water outflow (Heywood et al. 2002). This may have been facilitated by wind-driven surface currents, and indicates that dispersal of passive particles from AP/EI to SO shelf areas is possible albeit comparatively rare. Most other drifters leaving the AP/EI shelf areas took a more northerly route in the ACC's main current and missed the SO shelf. Drifters leaving the SO shelf area dispersed in a north-eastern direction, and none of them reached the AP/EI region.

Microsatellites and mtDNA – isolation-with-migration model

The IM model was used to test whether gene flow is unidirectional with the prevailing current as expected by gene flow through passive larval dispersal or bidirectional through gene flow by adult migration. The migration parameters m_1 (from SO to AP/EI) and m_2 (from AP/EI to SO) derived from the models indicate asymmetric gene flow with the current only in *C. aceratus* and *C. gunnari*, while the majority of gene flow in *T. newnesi* is against the current (Table 4; Online Resource 9). For the remaining species the migration parameters are close to zero in both directions. However, different migration patterns arise when population migration rates (effective rates at which genes come into populations per generation; M_1, M_2) are considered. In this case, low gene flow in both directions remains for *C. rastrospinosus*, *G. gibberifrons* and *L. squamifrons*, while *T. eulepidotus* shows a higher migration rate with the prevailing current. A distinct pattern of gene flow with the current from AP/EI to SO remains apparent in *C. aceratus* and *C. gunnari*, but it is negligible from SO to AP/EI. For *T. newnesi* gene flow remains high against the current from SO to AP/EI, but is lower from AP/EI to SO than expected from migration parameters m_1 and m_2 alone. With regard to private alleles in microsatellites, which may indicate a source-sink relationship by holding more unshared alleles in populations that act as sinks (and are usually found downstream), there is no uniform coherence between the direction of gene flow according to IMA2 and the number of private alleles per population.

13

Discussion

Genetic structure and diversity

The results obtained in the present study show low or non-existing genetic population differentiation for seven of the most abundant notothenioid species in the southern Scotia Arc. These results seem to be independent of the adult life history strategies but are accordant to expectations for marine species with long pelagic larval stages. Both types of genetic markers used in this study show either no or only weak genetic structure among populations for all studied species. None of the obtained haplotype genealogies based on mtDNA shows a clear separation between localities (Fig. 2) and *F*-statistics for both markers revealed, if at all, only minor differentiations. The only significant differentiations were found within the channichthyids *C. aceratus*, *C. gunnari* and *C. rastrospinosus*, while the nototheniid populations of *G. gibberifrons*, *L. squamifrons*, *T. eulepidotus* and *T. newnesi* generally lack a clear genetic structure. However, the channichthyid differentiations are not congruent among marker types and can therefore be used to discriminate between short and long term population dynamics. Microsatellites evolve faster than mtDNA and their higher diversities allow to infer present connectivity patterns better than mtDNA, which carries a longer-persisting signature of past events (Selkoe & Toonen 2006). Historical or long existing barriers to gene flow are hence more likely to be detected with mtDNA sequences. In our data, significant differentiations with microsatellite markers between samples from the AP/EI region and SO were only detected in the benthic *C. aceratus* and the benthopelagic *C. gunnari*. However, the differentiation observed in both species are only minor and cluster analyses suggest in both cases the existence of only one population in the study area. The small differentiations detected with mtDNA indicate homogenizing gene flow between AP/EI and SO. The highest differentiation of mtDNA sequences was found in the benthopelagic *C. rastrospinosus* ($F_{ST} = 0.04$), but even for this species, the differentiation was not significant. Although the differentiation in microsatellite allele sizes was close to zero in *C. rastrospinosus*, which indicates that the true differentiation might not be as high as measured with mtDNA, their power to detect subtle population differences was rather low in this case. Differences in the cluster analyses between runs including and excluding information about sampling localities, which result in one and two clusters respectively, suggest that a possible differentiation in *C. rastrospinosus* is more likely based on unrecognized factors (as e.g. sampling of cohorts) than geographically separated

14

populations. In the four nototheniid species *G. gibberifrons, L. squamifrons, T. eulepidotus* and *T. newnesi*, the genetic differentiations between localities are generally low and not significant giving little evidence for the existence of barriers to gene flow between shelf areas in the southern Scotia Arc.

Our results add new information on the connectivity of notothenioid populations in the area. Previous studies on notothenioid population structures along the Scotia Arc were based on parasite infestation rates (Kock & Möller 1977; Siegel 1980a), morphometric characters (Kock 1981) and more recently on a variety of genetic markers (e.g. Papetti et al. 2009). Studies based on parasite infestation rates revealed differences between "populations" north and south of the ACC in *C. gunnari* and *C. aceratus*, but not among populations of *C. aceratus* and *C. rastrospinosus* along the southern Scotia Arc (Kock & Möller 1977; Siegel 1980a, 1980b). For *C. gunnari*, four different populations had been identified based on morphometric characters from South Georgia, SO, SSh and EI (Kock 1981). Significant differentiation was also confirmed with genetic marker sets for *C. gunnari* populations north and south of the ACC (Kuhn & Gaffney 2006), but the island shelves along the southern Scotia Arc had not yet been compared. A recent publication on the genetic population structure of *C. aceratus* from the southern Scotia Arc based on microsatellites is in agreement with our results and shows that the populations on both sides of the Philip Passage are weakly, but significantly differentiated, while migration is still evident (Papetti et al. 2009). In our study, significant genetic differentiation in *C. aceratus* was detected with microsatellites, but not with mtDNA sequences. Hence, in the long run the migration rates in this sedentary species seem to be high enough (e.g. >1 individual per generation (Mills & Allendorf 1996)) to counteract genetic drift and population differentiation. Microsatellite fragment lengths are susceptible to changes in frequency with every generation and *F*-statistics rather show a captured moment of population structure than history. Overall, the generally low genetic differentiation suggests high connectivity between populations in the study area.

With regard to the mutation rates of microsatellites and mtDNA, it seems to be counterintuitive that genetic differentiations between populations measured with fast evolving microsatellites are smaller than for slower evolving mtDNA as observed in *C. rastrospinosus* and to a lesser extent in *G. gibberifrons* and *T. newnesi*. A similar pattern was already observed in populations of *D eleginoides* from Heard and McDonald Island, Macquarie Island and South Georgia (Appleyard et al. 2002). Different genetic patterns between maternally inherited markers such as mtDNA and bi-parentally inherited markers such as nuclear microsatellites can arise from sexual differences in spawning behaviour or

simply by genetic drift and population bottlenecks. Maternally inherited mtDNA is more affected by the latter two than nuclear DNA, since its effective population size is only one quarter that of nuclear DNA, which may result in higher divergences for mtDNA. Of the three species showing this discrepancy between markers in our study, spawning migrations are only known for *C. rastrospinosus*, which migrates to shelf waters of 200-300 m depth to spawn (Kock 1989). However, it is currently unknown whether *C. rastrospinosus* prefers specific spawning grounds and whether sexual differences in migration behaviour exist (Kock 2005a), leaving this explanation to speculation. By contrast, for all three species, at least one bottleneck indicator suggests that populations might have undergone a reduction in population size. We therefore cannot exclude bottlenecks and genetic drift as possible reasons for this pattern. However, these unusual differences between markers do not affect our general finding of low or absent population structure in the southern Scotia Arc.

According to genetic studies from the last decade, the levels of population differentiation in notothenioids vary widely and do not show a universal pattern. These studies primarily focused on single species targeted by the fisheries industry in the Southern Ocean like *C. gunnari*, *D. eleginoides* or *D. mawsoni* (Appleyard et al. 2002; Kuhn & Gaffney 2006; 2008; Parker et al. 2002; Rogers et al. 2006; Shaw et al. 2004; Smith & McVeagh 2000; Smith & Gaffney 2005) and their results depend on the types of genetic markers used (allozymes, RAPDs, microsatellites, mt and nuclear DNA sequences). Although most species show genetic differences between single populations, a lack of differentiation over large parts of even circum-Antarctic distribution ranges is evident. In this regard, the adult life history strategy and habit seems to play only a minor role for the genetic structuring of populations, since this pattern can be found not only in active pelagic swimmers as the Antarctic silverfish *Pleuragramma antarcticum* (Zane et al. 2006), but also in strictly benthic species as *G. gibberifrons* (Matschiner et al. 2009). It seems plausible that gene flow between populations of notothenioids is primarily based on larval dispersal. Hence, oceanography is a key factor influencing the structure of notothenioid populations.

Migration and gene flow

The genetic population structures found in this study revealed ongoing gene flow in the southern Scotia Arc for all study species, regardless of their adult habit or larval stage duration. However, the isolation-with-migration models from which the amount and directionality of gene flow were inferred did not result in a uniform pattern among species. A unidirectional gene flow in congruence with the prevailing current between AP/EI and SO, indicative of gene flow by larval dispersal, is only evident in *C. aceratus* and *C. gunnari*. In contrast, in *C. rastrospinosus*, *G. gibberifrons*, *L. squamifrons* and *T. eulepidotus* the amount of gene flow is close to zero in both directions. This is in clear contrast to our population genetic data. A possible explanation is that one of the assumptions underlying the isolation-with-migration model is violated. The software IMA2 calculates migration rates between two populations that derived from one ancestral population. Panmictic populations as found in this study may violate the model to the extent that a proper calculation of migration rates is not feasible. Indeed, in these four species, the estimated sizes of the ancestral populations are smaller than at least one of the two derived populations, and the time since population divergence was estimated near zero in *C. rastrospinosus*, *G. gibberifrons* and *L. squamifrons*. The only species in our data set showing bidirectional gene flow is *T. newnesi* with a far higher migration rate against the current. Although we cannot exclude the possibility, that this cryopelagic species utilizes the underside of ice during the austral winter to traverse areas that are normally inaccessible to them, it is more likely that the high migration rate in *T. newnesi* is an artefact caused by unsampled 'ghost' populations from the Weddell Sea that contribute larvae to the populations in our study area. As shown earlier, large immigration rates from ghost populations can result in overestimated migration rates between sampled populations (Beerli 2004). To clarify this issue, it would be necessary to conduct a study including samples from Weddell Sea populations. To this point, reliable estimates of the amount and direction of gene flow in the southern Scotia Arc were hence only derived for *C. aceratus* and *C. gunnari*. Both show unidirectional gene flow with the current indicating that larval dispersal in notothenioids is a likely key feature connecting populations from separated habitats.

In *C. aceratus* and *C. gunnari*, which both showed signs of subtle differentiation, the unidirectional gene flow from AP/EI to SO coincides with the current pattern of the ACC in the study area. This indicates that gene flow at least in these two species is likely to be caused by advection of larvae with the current. However, it remains unclear why the migration rate in the egg-guarding, benthic

17

88

C. aceratus is about twice as high as in the benthopelagic *C. gunnari*, which scatters its eggs freely into the water, where they may be dispersed already during the egg stage. Both species have an absolute fecundity in the same order of magnitude, so that the amount of larvae reaching the other shelf should also be comparable.

We conclude that the general pattern of weak or absent genetic population structure found for notothenioids with differing adult life history strategies is primarily based on the characteristic they all have in common, the pelagic larval phase. With regard to the adaptive radiation of notothenioids it remains unclear how species evolve rapidly while differentiation is counteracted by high gene flow through larval dispersal. It seems likely that notothenioid speciation events are restricted to periods when larval dispersal is hindered as e.g. during extended ice-coverages during glacial maxima. To further examine the role of larval dispersal on population structure and adaptive radiation in notothenioids it is necessary to conduct further comparative population genetic studies over wider geographic scales including hydrographic features like the ACC.

18

Table 1 Distribution area and selected life history characteristics for all seven study species. Spawning and hatching times are for the region of South Shetland Islands and the Antarctic Peninsula

Family - Species	Distribution range	Depths range [m]	Adult habit	Eggs	Spawning time	Hatching time	Abs. fecund.	Gen. time [years]
Channichthyidae								
C. aceratus	Scotia Arc region, Bouvet Is[a]	5-770[b]	benthic[c]	bottom, guarded[d]	May-Jun[e]	Aug-Nov[e,f]	3082-22626[c,g]	6-8[h]
C. gunnari	Scotia Arc region, Bouvet Is, Kerguelen Is, Heard Is[a]	0-700[a]	bentho-pelagic[c]	bentho-pelagic or pelagic[c], scattered	Jun-Jul[e]	Jan-Mar[e,i]	1294-31045[j,k]	3[h]
C. rastrospinosus	South Orkney Is, South Shetland Is, Antarctic Peninsula[a]	0-1000[b]	bentho-pelagic[b]	demersal[c], scattered	Mar-May[e,l]	Sep-Oct[j,m]	1464-5136[e]	4-8[n,o]
Nototheniidae								
G. gibberifrons	Scotia Arc region[p]	5-750[p]	benthic[a,f]	demersal[q] scattered	Aug-Sep[e]	Nov-?[e]	21699-143620[e]	6-8[h]
L. squamifrons	Sub-Antarctic Islands, intervening seamounts of the Indian Ocean sector, Scotia Arc region, Bouvet Is[p,r]	5-670[s,t]	benthic[a,f]	demersal[q] scattered	Feb-Mar[l,u]	Apr-Jun[v]	38000-280000[w,x]	7-9[v]
T. eulepidotus	Circum-Antarctic: nearshore and continental shelf and nearby islands[p]	70-550[p]	bentho-pelagic[a,b]	substrate[y], non-guarded	Apr[l]	Sep[y]	1400-12854[h,z]	7[y]
T. newnesi	Circum-Antarctic: shallow shelf waters of the continent and adjacent islands[p]	0-400[aa]	bentho-cryo-pelagic[bb,cc]	?	Mar-May[dd,ee]	Sep-Nov[h,cc]	2300-12200[dd]	?

[a]Iwami & Kock 1990, [b]Hureau 1985, [c]Permitin 1973, [d]Detrich et al. 2005, [e]Kock 1989, [f]Ślósarczyk 1987, [g]Lisovenko & Sil'yanova 1991, [h]Kock & Kellermann 1991, [i]Kellermann 1989, [j]Kock 1981, [k]Lisovenko & Zakharov 1988, [l]Kock et al. 2001, [m]Kock & Jones 2005, [n]Kock 2005b, [o]La Mesa & Ashford 2008, [p]DeWitt et al. 1990, [q]Permitin & Sil'yanova 1971, [r]Schneppenheim et al. 1994, [s]Duhamel 1981, [t]Ekau 1990, [u]Kellermann 1986, [v]Duhamel & Ozouf-Costaz 1985, [w]Lisovenko & Sil'yanova 1979, [x]Kock 1992, [y]Ekau 1989, [z]Ekau 1991, [aa]Tiedtke & Kock 1989, [bb]Andriashev 1987, [cc]Radtke et al. 1989, [dd]Shust 1987, [ee]Jones & Kock 2006

Table 2 Population differentiation (*F*-statistics) between AP and SO samples based on mtDNA and microsatellites. ENA = excluding null alleles. ENA results without significances, AMOVA significances: *p≤0.05, **p≤0.01

	Species						
	C. aceratus	*C. gunnari*	*C. rastrospinosus*	*G. gibberifrons*	*L. squamifrons*	*T. eulepidotus*	*T. newnesi*
mtDNA							
F_{ST}	-0.0056	-0.0027	0.0440	0.0027	-0.0087	-0.0053	-0.0146
Microsatellites							
F_{ST}	0.0088*	0.0226**	-0.0004	0.0002	-0.0015	-0.0035	0.0067
F_{ST} ENA	0.0088	0.0219	0.0021	0.0002	-0.0006	-0.0018	0.0027
F_{IS}	0.022*	0.034*	0.182**	0.062**	0.074**	0.161**	0.006
F_{IT}	0.03*	0.06*	0.18*	0.06*	0.07*	0.16*	0.01

Table 3 Number of clusters with highest mean posterior probability inferred from Bayesian analyses as indicated by ΔK (maximum ΔK in parenthesis). INA = including null alleles, ENA = excluding null alleles (according to MicroChecker 2.2.3), locprior = a priori information of sampling sites incorporated in analysis

	Species						
	C. aceratus	*C. gunnari*	*C. rastrospinosus*	*G. gibberifrons*	*L. squamifrons*	*T. eulepidotus*	*T. newnesi*
INA	1 (30.6)	1 (48.8)	2 (147.3)	1 (130.9)	1 (4.5)	2 (31.1)	1 (20.0)
INA locprior	1 (8.1)	1 (63.2)	1 (93.6)	1 (5.7)	1 (7.2)	1 (23.1)	1 (44.1)
ENA	1 (12.9)	1 (7.3)	2 (52.4)	1 (157.8)	1 (9.0)	1 (11.1)	no NA
ENA locprior	1 (88.7)	1 (69.3)	1 (63.0)	1 (9.8)	2 (1.0)	1 (12.9)	no NA

Table 4 Isolation-with-migration results reflecting parameter bins with highest posterior probabilities (High Points). t0 = time since divergence, Θ1 = effective AP/EI populations size, Θ2 = effective SO populations size, ΘA = effective ancestral populations size, m1 = migration rate from SO to AP/EI, m2 = migration rate from AP/EI to SO, M1 & M2 = accordant population migration rates

Species	Parameter							
	t0	Θ1	Θ2	ΘA	m1	m2	M1	M2
C. aceratus	0.129	7.25	14.74	141.20	0.05	4.55	0.18	33.56
C. gunnari	0.443	4.75	2.75	52.25	0.05	2.65	0.12	3.64
C. rastrospinosus	0.001	74.75	23.75	47.25	0.05	0.05	1.87	0.59
G. gibberifrons I	0.001	71.25	17.25	21.25	0.05	0.05	1.78	0.43
G. gibberifrons II	0.001	125.80	21.25	46.75	0.05	0.05	3.15	0.53
L. squamifrons	0.001	35.00	15.00	55.00	0.05	0.05	0.875	0.375
T. eulepidotus	0.193	25.00	375.00	35.00	0.05	0.05	0.625	9.375
T. newnesi	0.002	37.50	16.50	360.50	3.85	0.05	72.188	0.413

20

Acknowledgments

References

Andriashev AP (1987) A general review of the Antarctic bottom fish fauna. In: Kullander SO, Fernholm B (eds) Proceedings, fifth congress of European ichthyologists, Stockholm 1985. Swedish Museum of Natural History, Stockholm, pp 357-372

Appleyard S, Ward R, Williams R (2002) Population structure of the Patagonian toothfish around Heard, McDonald and Macquarie Islands. Antarct Sci 14:364-373

Barnes DKA, Conlan KE (2007) Disturbance, colonization and development of Antarctic benthic communities. Phil Trans R Soc B 362:11-38

Bay LK, Crozier RH, Caley MJ (2006) The relationship between population genetic structure and pelagic larval duration in coral reef fishes on the Great Barrier Reef. Mar Biol 149:1247-1256

Beerli P (2004) Effect of unsampled populations on the estimation of population sizes and migration rates between sampled populations. Mol Ecol 13:827-836

Beerli P, Felsenstein J (1999) Maximum-likelihood estimation of migration rates and effective population numbers in two populations using a coalescent approach Genetics 152:763-773

Beerli P, Felsenstein J (2001) Maximum likelihood estimation of a migration matrix and effective population sizes in n subpopulations by using a coalescent approach. Proc Natl Acad Sci USA 98:4563-4568

Belkhir K, Borsa P, Chikhi et al. (2001) Genetix 402, logiciel sous Windows TM pour la génétique des populations.

Briggs J (2003) Marine centres of origin as evolutionary engines. J Biogeogr 30:1-18

Casaux R, Mazzotta A, Barrera-Oro ER (1990) Seasonal aspects of the biology and diet of nearshore nototheniid fish at Potter Cove, South Shetland Islands, Antarctica. Polar Biol 11:63-72

Chapuis M, Estoup A (2007) Microsatellite Null Alleles and Estimation of Population Differentiation. Mol Biol Evol 24:621-631

Cheng CC (1998) Origin and mechanism of evolution of anti-freeze glycoproteins in polar fishes. In: Di Prisco G, Pisano E, Clarke A (eds) Fishes of Antarctica. A biological overview. Springer, Milan, pp 311-328

Clarke A, Johnston I (1996) Evolution and adaptive radiation of Antarctic fish. Trends Ecol Evol 11:212-218

Cowen R, Sponaugle S (2009) Larval dispersal and marine population connectivity. Annu Rev Mar Sci

22

1:443-466

Daniels R (1982) Feeding ecology of some fishes of the Antarctic Peninsula. Fish Bull US 80:575-588

Dayton P, Mordida B, Bacon F (1994) Polar marine communities. Am Zool 34:90-99

Derome N, Chen W, Dettai A et al. (2002) Phylogeny of Antarctic dragonfishes (Bathydraconidae, Notothenioidei, Teleostei) and related families based on their anatomy and two mitochondrial genes. Mol Phylogenet Evol 24:139-152

Detrich H, Jones C, Kim S et al. (2005) Nesting behavior of the icefish *Chaenocephalus aceratus* at Bouvetøya Island, Southern Ocean. Polar Biol 28:828-832

DeWitt HH, Heemstra PC, Gon O (1990) Nototheniidae. In: Gon O, Heemstra PC (eds) Fishes of the Southern Ocean. JLB Smith Institute of Ichthyology, Grahamstown, pp 279-380

Duhamel G (1981) Caracteristiques biologiques des principales especes de poissons du plateau continental des Iles Kerguelen. Cybium 5:19-32

Duhamel G, Ozouf-Costaz C (1985) Age, growth and reproductive biology of Notothenia squamifrons Gunther, 1880 from the Indian sector of the Southern Ocean. Polar Biol 4:143-153

Eastman JT (1991) Evolution and diversification of antarctic notothenioid fishes. Am Zool 31:93-109

Eastman JT (1993) Antarctic fish biology: evolution in a unique environment. Academic Press, San Diego

Eastman JT (2000) Antarctic notothenioid fishes as subjects for research in evolutionary biology. Antarct Sci 12:276-287

Eastman JT (2005) The nature of the diversity of Antarctic fishes. Polar Biol 28:93-107

Eastman JT, McCune A (2000) Fishes on the Antarctic continental shelf: evolution of a marine species flock. J Fish Biol 57 (Suppl A):84-102

Ekau W (1989) Egg development of *Trematomus eulepidotus* Regan, 1914 (Nototheniidae, Pisces) from the Weddell Sea, Antarctica. Cybium 13:213-219

Ekau W (1990) Demersal fish fauna of the Weddell Sea, Antarctica. Antarct Sci 2:129-137

Ekau W (1991) Reproduction in high Antarctic fishes (Notothenioidei). Meeresforsch 33:159-167

Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software structure: a simulation study. Mol Ecol 14:2611-2620

Excoffier L, Lischer HEL (2010) Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. Mol Ecol Resour 10:564-567

23

Fanta E, Meyer AA, Grötzner SR, Luvizotto MF (1994) Comparative study on feeding strategy and activity patterns of two Antarctic fish: *Trematomus newnesi* Boulenger, 1902 and *Gobionotothen gibberifrons* (Lonnberg, 1905) (Pisces, Nototheniidae) under different light conditions. Antarct Rec 38:13-29

Garza J, Williamson E (2001) Detection of reduction in population size using data from microsatellite loci. Mol Ecol 10:305-318

Glaubitz JC (2004) Convert: A user-friendly program to reformat diploid genotypic data for commonly used population genetic software packages. Mol Ecol Notes 4:309-310

Goudet J (1995) FSTAT (Version 1.2): A computer program to calculate F-statistics. J Hered 86:485-486

Goudet J (2001) FSTAT, a program to estimate and test gene diversities and fixation indices, version 2.9.3. [http:// www2.unil.ch/popgen/softwares/fstat.htm].

Hey J, Nielsen R (2007) Integration within the Felsenstein equation for improved Markov chain Monte Carlo methods in population genetics. Proc Natl Acad Sci USA 104:2785-2790

Heywood KJ, Garabato A, Stevens D (2002) High mixing rates in the abyssal Southern Ocean. Nature 415:1011-1014

Hureau JC (1985) Channichthyidae. In: Fischer W, Hureau JC (eds) FAO species identification sheets for fishery purposes. Southern Ocean (Fishing areas 48, 58 and 88), Vol. 2. FAO, Rome, pp 261-277

Iwami T, Kock KH (1990) Channichthyidae. In: Gon O, Heemstra PC (eds) Fishes of the Southern Ocean. JLB Smith Institute of Ichthyology, Grahamstown, pp 381-399

Jones CD, Kock KH (2006) Standing stock, spatial distribution, and biological features of demersal finfish from the 2006 US AMLR bottom trawl survey of the northern Antarctic Peninsula and Joinville-D'Urville Islands (Subarea 48.1). WG-FSA-06/14. CCAMLR, Hobart

Kellermann AK (1986) Zur Biologie der Jugendstadien der Notothenioidei (Pisces) an der Antarktischen Halbinsel. Ber Polarforsch 31:1-155

Kellermann AK (1989) The larval fish community in the zone of seasonal pack-ice cover and its seasonal and interannual variability. Arch Fisch Wiss 39 (Beih 1):81-109

Kock KH (1981) Fischereibiologische Untersuchungen an drei antarktischen Fischarten: *Champsocephalus gunnari* Lönnberg, 1905, *Chaenocephalus aceratus* (Lönnberg, 1906) und *Pseudochaenichthys georgianus* Norman, 1937 (Notothenioidei, Channichthyidae). Mitt Inst Seefisch Hambg 32:1-226

24

Kock KH (1989) Reproduction in fish around Elephant Island. Arch Fisch Wiss 39 (Beih 1):171-210

Kock KH (1992) Antarctic fish and fisheries. Cambridge University Press, Cambridge

Kock KH (2005a) Antarctic icefishes (Channichthyidae): a unique family of fishes. A review, Part I. Polar Biol 28:862-895

Kock KH (2005b) Antarctic icefishes (Channichthyidae): a unique family of fishes. A review, Part II. Polar Biol 28:897-909

Kock KH, Everson I (1997) Biology and ecology of mackerel icefish, *Champsocephalus gunnari*: An Antarctic fish lacking hemoglobin. Comp Biochem Physiol A 118:1067-1077

Kock KH, Jones CD, Wilhelms S (2001) Biological characteristics of Antarctic fish stocks in the southern Scotia Arc region. Ccamlr Science 7:1-41

Kock KH, Jones CD (2005) Fish stocks in the southern Scotia Arc region - A review and prospects for future research. Rev Fish Sci 13:75-108

Kock KH, Kellermann AK (1991) Reproduction in Antarctic notothenioid fish. Antarct Sci 3:125-150

Kock KH, Möller H (1977) On the occurrence of the parasite copepod *Eubrachiella antarctica* on some Antarctic fish. Arch Fisch Wiss 28:149-156

Kuhn KL, Gaffney PM (2006) Preliminary assessment of population structure in the mackerel icefish (*Champsocephalus gunnari*). Polar Biol 29:927-935

Kuhn KL, Gaffney PM (2008) Population subdivision in the Antarctic toothfish (*Dissostichus mawsoni*) revealed by mitochondrial and nuclear single nucleotide polymorphisms (SNPs). Antarct Sci 20:327-338

La Mesa M, Ashford J (2008) Age and growth of ocellated icefish, *Chionodraco rastrospinosus* DeWitt and Hureau, 1979, from the South Shetland Islands. Polar Biol 31:1333-1342

La Mesa M, Vacchi M, Sertorio T (2000) Feeding plasticity of *Trematomus newnesi* (Pisces, Nototheniidae) in Terra Nova Bay, Ross Sea, in relation to environmental conditions. Polar Biol 23:38-45

Librado P, Rozas J (2009) DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. Bioinforma 25:1451-1452

Lisovenko LA, Sil'yanova ZS (1979) The fecundity of some species of the family Nototheniidae in the atlantic sector of the Southern Ocean. J Ichthyol 19:79-85

Lisovenko LA, Sil'yanova ZS (1980) The reproduction and fecundity of fish of the family

25

Chaenichthyidae. In: An Ecological and Biological Description of Some Species of Antarctic Fishes. Trudy All-Union Institute for Fisheries Research and Oceanography, Moscow pp 38-52

Lisovenko LA, Zakharov GP (1988) On fecundity of the striped pike glassfish, *Champsocephalus gunnari*, in the region of South Georgia Island. J Ichthyol 27:131-134

Loeb VJ, Kellermann AK, Koubbi P et al. (1993) Antarctic larval fish assemblages: a review. Bull Mar Sci 53:416-449

Lumpkin R, Pazos M (2007) Measuring surface currents with surface velocity program drifters: the instrument, the data, and some recent results. In: Mariano, A et al. (eds) Lagrangian Analysis and Prediction of Coastal and Ocean Dynamics. Cambridge University Press, Cambridge, pp 39-67

Matschiner M, Salzburger W (2009) TANDEM: integrating automated allele binning into genetics and genomics workflows. Bioinformatics 25:1982-1983

Matschiner M, Hanel R, Salzburger W (2009) Gene flow by larval dispersal in the Antarctic notothenioid fish *Gobionotothen gibberifrons*. Mol Ecol 18:2574-2587

Matschiner M, Hanel R, Salzburger W (2010) Phylogeography and speciation processes in marine fishes and fishes from large freshwater lakes. In: Rutgers, DS (ed) Phylogeography. Concepts, Intraspecific Patterns and Speciation Processes. Nova Science Publishers, pp 1-29

Matschiner M, Hanel R, Salzburger W (2011) On the origin and trigger of the notothenioid adaptive radiation. PLoS one 6:e18911

Mills L, Allendorf F (1996) The one-migrant-per-generation rule in conservation and management. Conserv Biol 10:1509-1518

North AW (2001) Early life history strategies of notothenioids at South Georgia. J Fish Biol 58:496-505

Papetti C, Susana E, Patarnello T, Zane LL (2009) Spatial and temporal boundaries to gene flow between *Chaenocephalus aceratus* populations at South Orkney and South Shetlands. Mar Ecol Prog Ser 376:269-281

Papetti C, Zane LL, Patarnello T (2006) Isolation and characterization of microsatellite loci in the icefish *Chionodraco rastrospinosus* (Perciformes, Notothenioidea, Channichthyidae). Mol Ecol Notes 6:207-209

Parker R, Paige K, De Vries A (2002) Genetic variation among populations of the Antarctic toothfish: evolutionary insights and implications for conservation. Polar Biol 25:256-261

Permitin Y (1973) Fecundity and reproductive biology of icefish (Chaenichthyidae), fish of the family

Muarenolepidae and dragonfish (Bathydraconidae) of the Scotia Sea (Antarctica). J Ichthyol 13:204-215

Permitin Y (1977) Species composition and zoogeographical analysis of the bottom fish fauna of the Scotia Sea. J Ichthyol 17:710-726

Permitin Y, Sil'yanova Z (1971) New data on the reproductive biology and fecundity of fishes of the genus *Notothenia* Rich. in the Scotia Sea (Antarctica). J Ichthyol 11:693-705

Piry S, Luikart G, Cornuet J (1999) BOTTLENECK: A computer program for detecting recent reductions in the effective population size using allele frequency data. J Hered 90:502-503

Posada D (2008) jModelTest: phylogenetic model averaging. Mol Biol Evol 25:1253-1256

Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. Genetics 155:945-959

Radtke R, Targett T, Kellermann AK et al. (1989) Antarctic fish growth: profile of *Trematomus newnesi*. Mar Ecol Prog Ser 57:103-117

Raymond M, Rousset F (1995) Genepop (Version 1.2): population genetics software for exact tests and ecumenicism. J Hered 86:248-249

Rogers AD, Morley S, Fitzcharles et al. (2006) Genetic structure of Patagonian toothfish (*Dissostichus eleginoides*) populations on the Patagonian Shelf and Atlantic and western Indian Ocean Sectors of the Southern Ocean. Mar Biol 149:915-924

Rutschmann S, Matschiner M, Damerau M et al. (2011) Parallel ecological diversification in Antarctic notothenioid fishes as evidence for adaptive radiation. Mol Ecol in print.

Ryman N, Palm S (2006) POWSIM: a computer program for assessing statistical power when testing for genetic differentiation. Mol Ecol Notes 6:600-602

Salzburger W (2009) The interaction of sexually and naturally selected traits in the adaptive radiations of cichlid fishes. Mol Ecol 18:169-185

Salzburger W, Ewing GB, von Haeseler A (2011) The performance of phylogenetic algorithms in estimating haplotype genealogies. Mol Ecol 20:1952-1963

Schneppenheim R, Kock KH, Duhamel G, Janssen G (1994) On the taxonomy of the Lepidonotothen squamifrons group (Pisces, Perciformes, Notothenioidei). Arch Fish Mar Res 42:137-148

Seehausen O (2006) African cichlid fish: a model system in adaptive radiation research. Proc R Soc B 273:1987-1998

Selkoe K, Toonen R (2006) Microsatellites for ecologists: a practical guide to using and evaluating microsatellite markers. Ecol Lett 9:615-629

Shaw P, Arkhipkin A, Al-Khairulla H (2004) Genetic structuring of Patagonian toothfish populations in the Southwest Atlantic Ocean: the effect of the Antarctic Polar Front and deep-water troughs as barriers to genetic exchange. Mol Ecol 13:3293-3303

Shust KV (1987) Distribution and important biological aspects of abundant Antarctic notothenioid species. In: Biological Resources of the Arctic and Antarctic. Moscow, pp 296-320

Siegel V (1980a) Quantitative investigations on parasites of Antarctic channichtyid and nototheniid fishes. Meeresforsch 28:146-156

Siegel V (1980b) Parasite tags for some antarctic channichthyid fish. Arch Fisch Wiss 31:97-103

Ślósarczyk W (1987) Contribution to the early life history of Channichthyidae from the Bransfield Strait and South Georgia (Antarctica). In: Kullander SO, Fernholm B (eds) Proceedings, fifth congress of European ichthyologists, Stockholm 1985. Swedish Museum of Natural History, Stockholm, pp 427-433

Smith PJ, Gaffney PM (2005). Low genetic diversity in the Antarctic toothfish (*Dissostichus mawsoni*) observed with mitochondrial and intron DNA markers. Ccamlr Science 12:43-51

Smith P, McVeagh M (2000) Allozyme and microsatellite DNA markers of toothfish population structure in the Southern Ocean. J Fish Biol 57:72-83

Susana E, Papetti C, Barbisan F et al. (2007) Isolation and characterization of eight microsatellite loci in the icefish *Chaenocephalus aceratus* (Perciformes, Notothenioidei, Channichthyidae). Mol Ecol Notes 7:791-793

Swofford DL (2003) PAUP*. Phylogenetic Analysis Using Parsimony (*and Other Methods). ed. 4.04a. Sinauer Associates, Sunderland, Massachusetts.

Taylor M, Hellberg ME (2003) Genetic evidence for local retention of pelagic larvae in a Caribbean reef fish. Science 299:107-109

Tiedtke JE, Kock K (1989) Structure and composition of the demersal fish fauna around Elephant Island. Arch Fisch Wiss 39 (Beih 1):143-169

Van Houdt JKJ, Hellemans B, Van De Putte AP et al. (2006) Isolation and multiplex analysis of six polymorphic microsatellites in the Antarctic notothenioid fish, *Trematomus newnesi*. Mol Ecol Notes 6:157-159

28

Van Oosterhout C, Hutchinson W, Wills D, Shipley P (2004) Micro-Checker: software for identifying and correcting genotyping errors in microsatellite data. Mol Ecol Notes 4:535-538

White M (1998) Development, dispersal and recruitment: A paradox for survival among Antarctic fish. In: . Di Prisco G, Pisano E, Clarke A (eds) Fishes of Antarctica. A biological overview. Springer, Milan, pp 53-62

Whitworth T, Nowlin WD, Orsi A et al. (1994) Weddell Sea shelf water in the Bransfield Strait and Weddell-Scotia Confluence. Deep-Sea Res I 41:629-641

Fig. 1 Study area and sampling localities in the southern Scotia Arc, Antarctica. Open circles = stations sampled during ANT-XXIII/8 2006, crosses = stations sampled during US AMLR 2009
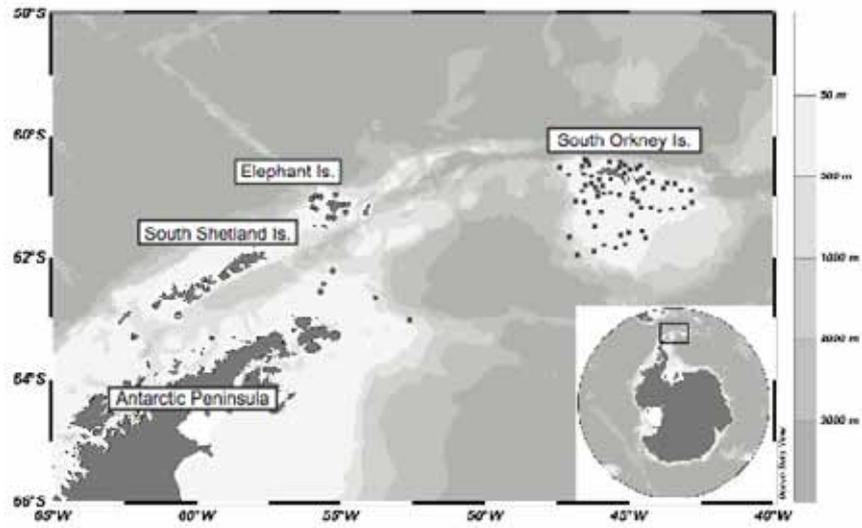
Fig. 2 Trajectories of surface drifters 45 and 90 days after leaving the shelves (500m isobath contour line) of South Shetland/Antarctic Peninsula (blue) and South Orkney Islands (yellow)
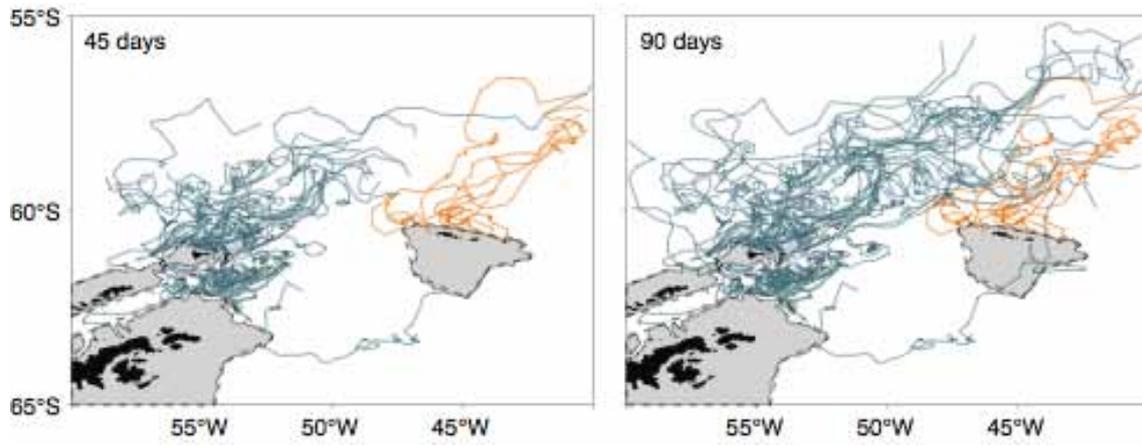
Supporting Information

Comparative population genetics of seven notothenioid fish species reveals high levels of gene flow along ocean currents in the southern Scotia Arc, Antarctica

Malte Damerau[*]**, Michael Matschiner**[†]**, Walter Salzburger**[†]**, and Reinhold Hanel**[*]

Addresses:
[*]Institute of Fisheries Ecology, Johann Heinrich von Thünen-Institute, Federal Research Institute for Rural Areas, Forestry and Fisheries, Palmaille 9, D-22767 Hamburg, Germany
[†]Zoological Institute, University of Basel, Vesalgasse 1, CH-4051 Basel, Switzerland

Correspondence:
Reinhold Hanel, Fax: +49 40 38 90 52 61, E-mail: reinhold.hanel@vti.bund.de

1 Online resources

Online resource 1: Microsatellite allele size ranges per species. n = number of genotyped individuals, # = number of polymorphic marker. Number of alleles in parenthesis.

	Species			
	C. aceratus	*C. gunnari*	*C. rastrospinosus*	*G. gibberifrons*
n	125	92	56	114
#	7	8	8	10
Cr15	147–153 (4)	161–247 (29)	127–159 (9)	-
Cr38	175 (1)	175-185 (4)	168-194 (7)	185-277 (33)
Cr127	103-127 (9)	101-129 (11)	103-169 (24)	123-125 (2)
Cr236	-	117-137 (9)	112-154 (14)	-
Cr259	213-261 (22)	201-251 (19)	186-230 (21)	220-298 (34)
Trne20	158-230 (25)	136-142 (3)	136-144 (4)	181-331 (49)
Trne35	-	-	-	289-505 (55)
Trne37	-	-	-	131-199 (23)
Trne53	305-473 (57)	351-417 (34)	346-480 (46)	312-422 (43)
Trne55	-	-	-	111-121 (2)
Trne66	269-373 (43)	275-311 (10)	283-401 (43)	270-368 (30)
Ca26	147-233 (31)	144-234 (26)	171-267 (36)	165-275 (48)

	Species		
	L. squamifrons	*T. eulepidotus*	*T. newnesi*
n	59	69	62
#	8	7	7
Cr15	-	-	-
Cr38	-	-	-
Cr127	-	-	-
Cr236	120 (1)	120 (1)	121 (1)
Cr259	201-213 (4)	191-193 (2)	201-251 (20)
Trne20	197-265 (25)	136-156 (5)	132-208 (28)
Trne35	218-252 (13)	210-366 (58)	224-314 (26)
Trne37	135-175 (18)	134-188 (18)	154-216 (11)
Trne53	310-358 (22)	362-422 (27)	286-412 (24)
Trne55	176-218 (19)	178 (1)	180-294 (13)
Trne66	292-400 (34)	300-376 (23)	299-431 (24)
Ca26	-	-	-

Online resource 2: Mitochondrial DNA sequence properties and genetic diversities for each species and population. Cace = *C. aceratus*, Cgun = *C. gunnari*, Cras = *C. rastrospinosus*, Ggib = *G. gibberifrons*, Lsqu = *L. squamifrons*, Teul = *T. eulepidotus*, Tnew = *T. newnesi*, n = number of sequences, bp = basepairs, π = nucleotide diversity, h = haplotype diversity, AP/EI = Antarctic Peninsula/Elephant Island, SO = South Orkney Islands.

				Species			
	Cace	Cgun	Cras	Ggib	Lsqu	Teul	Tnew
All							
Locus	D-loop	D-loop	D-loop	D-loop	cyt *b*	cyt *b*	cyt *b*
n	105	88	60	194	49	71	61
bp	356	357	323	352	708	667	723
No. haplotypes	6	10	13	32	11	41	29
π ($\pm$SD)	0.0012	0.0014	0.0048	0.0031	0.0009	0.0043	0.0044
	(0.0002)	(0.0003)	(0.0004)	(0.0002)	(0.0002)	(0.0004)	(0.0005)
h ($\pm$SD)	0.35	0.38	0.83	0.72	0.49	0.95	0.89
	(0.05)	(0.06)	(0.03)	(0.03)	(0.09)	(0.02)	(0.03)
AP/EI							
n	36	64	30	121	26	38	34
No. haplotypes	3	9	8	27	9	26	21
π ($\pm$SD)	0.0013	0.0017	0.0036	0.0035	0.0013	0.004	30.0049
	(0.0004)	(0.0004)	(0.0005)	(0.0003)	(0.0003)	(0.0005)	(0.0007)
h ($\pm$SD)	(0.08)	(0.08)	(0.05)	(0.04)	(0.11)	(0.03)	(0.04)
SO							
n	69	24	30	73	23	33	27
No. haplotypes	5	3	10	15	5	22	13
π ($\pm$SD)	0.0011	0.0007	0.0058	0.0025	0.0005	0.0044	0.0039
	(0.0003)	(0.0003)	(0.0310)	(0.0003)	(0.0002)	(0.0005)	(0.0007)
h ($\pm$SD)	0.32	0.24	0.88	0.64	0.32	0.95	0.85
	(0.07)	(0.11)	(0.03)	(0.06)	(0.12)	(0.02)	(0.06)

Online resource 3 (next page): Number of mtDNA haplotypes per species and area. Cace = *C. aceratus*, Cgun = *C. gunnari*, Cras = *C. rastrospinosus*, Ggib = *G. gibberifrons*, Lsqu = *L. squamifrons*, Teul = *T. eulepidotus*, Tnew = *T. newnesi*, AP = Antarctic Peninsula, SO = South Orkney Islands. Note that haplotype numbers are labeled for each species individually according to GenBank Accession nos JN241690–JN241831.

Gene	D-loop								cyt *b*					
Species	Cace		Cgun		Cras		Ggib		Lsqu		Teul		Tnew	
Area	AP	SO	AP	SO	AP	SO	AP	SO	AP	SO	AP	SO	AP	SO
Haplotype 1	27	56	48	21	8	5	57	43	16	19	8	6	1	2
Haplotype 2	8	10	4		1		11	2	1	1	1		1	
Haplotype 3	1		1		12	6	1		2	1	1	1	1	
Haplotype 4		1	6	2	3	7	3		1		1		2	3
Haplotype 5		1	1		1		12	9	1		1		3	
Haplotype 6		1	1		1		2	2	1		1		1	
Haplotype 7			1		3	2	2	5	2		1		9	10
Haplotype 8			1		1	3	1	1	1		1		3	2
Haplotype 9			1			1	2	2	1		1		1	
Haplotype 10				1		2		1		1	4	4	1	
Haplotype 11						2	6	2		1	2	1	1	
Haplotype 12						1		1			1	1	1	
Haplotype 13						1		1			1		1	2
Haplotype 14							1	1			1		1	
Haplotype 15							3	1			1		1	
Haplotype 16								1			1		1	
Haplotype 17								1			1		1	
Haplotype 18							1				1		1	
Haplotype 19							2				1		1	
Haplotype 20							1				1		1	
Haplotype 21							1				2		1	
Haplotype 22							1				1			1
Haplotype 23							1				1			1
Haplotype 24							2				1			1
Haplotype 25							2				1	1		1
Haplotype 26							2				1	3		1
Haplotype 27							2					1		1
Haplotype 28							1					1		1
Haplotype 29							1					2		1
Haplotype 30							1					1		
Haplotype 31							1					1		
Haplotype 32							1					1		
Haplotype 33												1		
Haplotype 34												1		
Haplotype 35												1		
Haplotype 36												1		
Haplotype 37												1		
Haplotype 38												1		
Haplotype 39												1		
Haplotype 40												1		
Haplotype 41												1		

Online resource 4: Statistical power tests for microsatellite (a, b) and mtDNA sequence (c, d) data sets inferred with POWSIM 4.0.
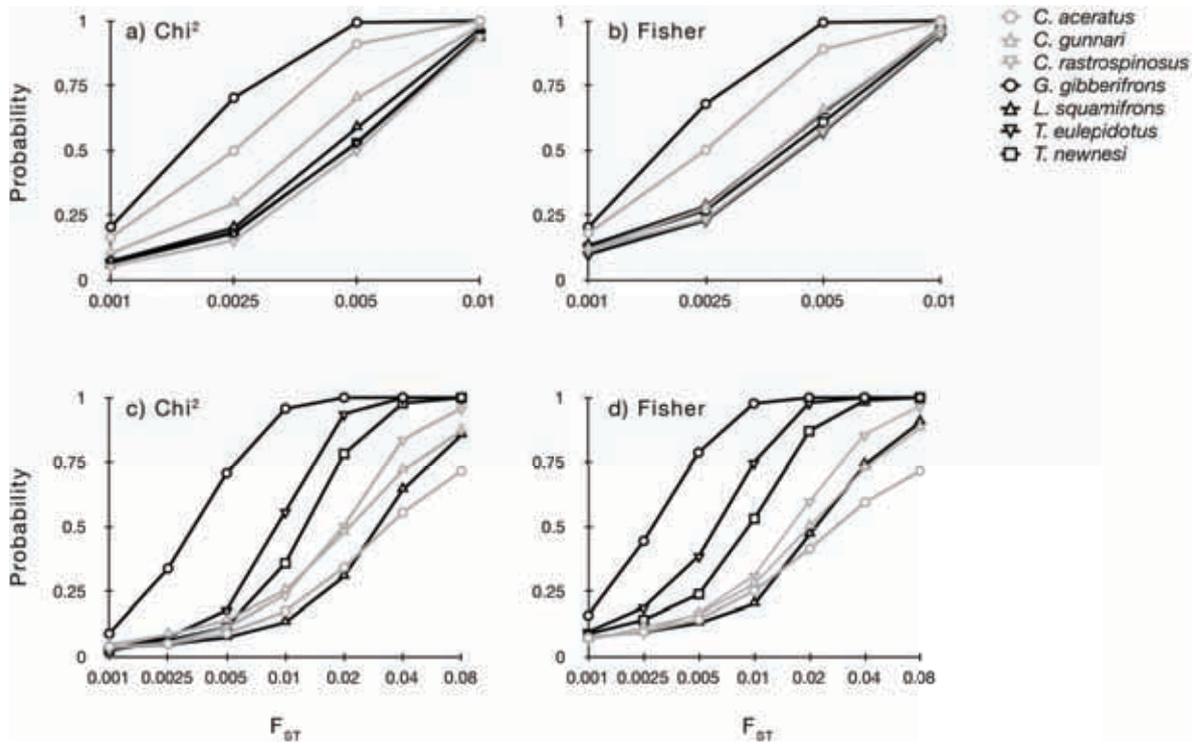
Online resource 5: Unrooted haplotype genealogies based on D-loop (a–d) and cyt *b* sequences (e–g). Radii reflect number of individuals.
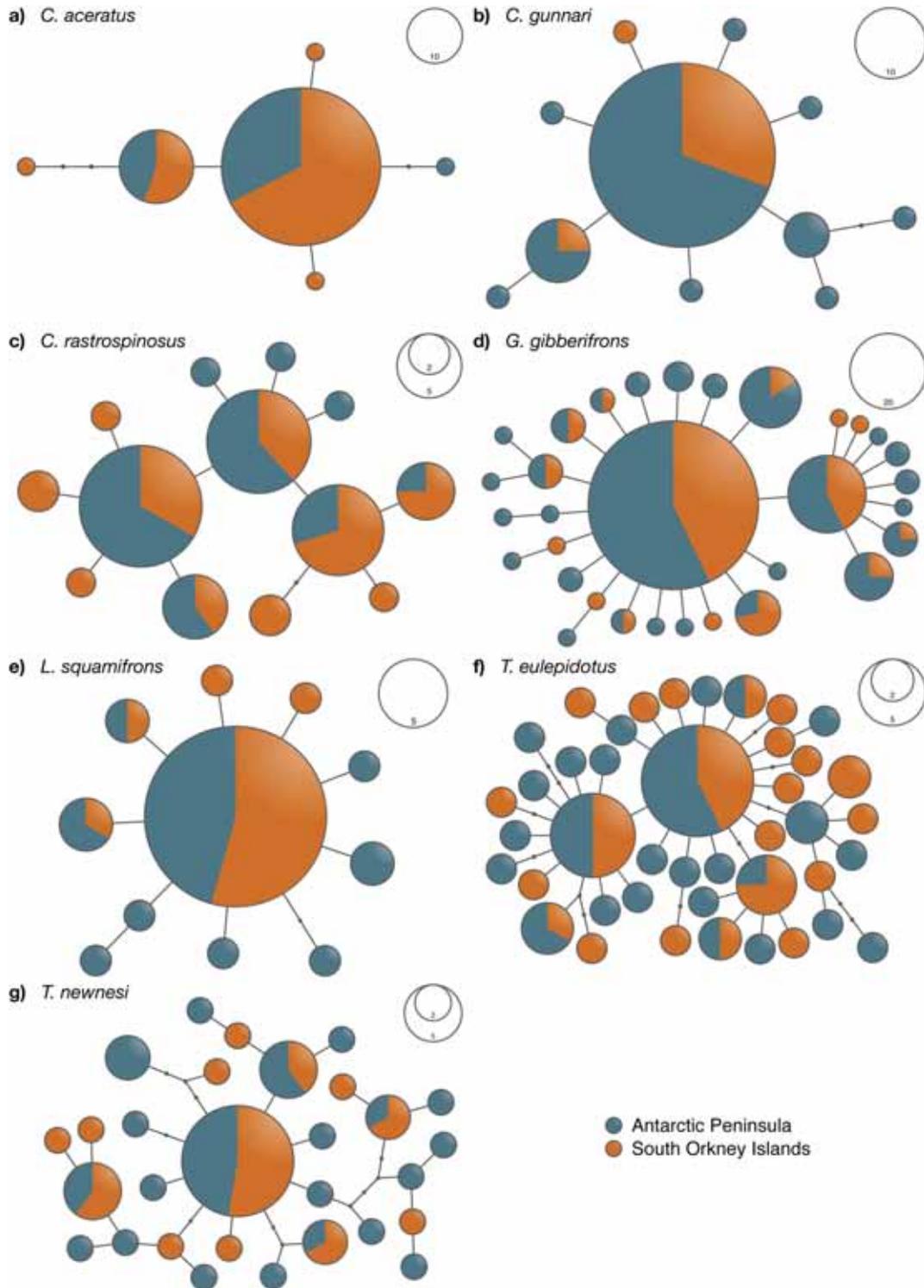
Online resource 6: Microsatellites allelic richness per locus and species. * = null alleles might be present according to the software BOTTLENECK. Cace = *Chaenocephalus aceratus*, Cgun = *Champsocephalus gunnari*, Cras = *Chionodraco rastrospinosus*, Ggib = *Gobionotothen gibberifrons*, Lsqu = *Lepidonotothen squamifrons*, Teul = *Trematomus eulepidotus*, Tnew = *Trematomus new-nesi*.

Locus	Cace	Cgun	Cras	Species Ggib	Lsqu	Teul	Tnew	Mean (±SD)
Cr15	3.2*	16.9	6.6*					8.9 (7.2)
Cr38		3.3	6.0	26.4				11.9 (12.6)
Cr127	7.8	7.6	19.7*	1.5				9.1 (7.6)
Cr236		7.4	11.7*					9.6 (3.1)
Cr259	18.2	15.4	17.6	29.1	3.5	2.0*	16.9	14.7 (9.3)
Trne20	18.2	2.3*	3.5*	39.8*	20.6	4.1*	21.6	15.7 (13.6)
Trne35				41.4*	10.4	37.6*	19.4	27.2 (14.8)
Trne37				18.4	15.6*	13.8	9.5	14.3 (3.7)
Trne53	36.6	24.6	32.5*	35.5	17.7	23.6	17.4	26.9 (8.1)
Trne55				1.5	15.0		11.3	9.2 (7.0)
Trne66	31.4	7.8	29.7	24.9	25.5	18.8	16.5	22.1 (8.3)
Ca26	18.0*			21.2	26.7	35.0*		25.2 (7.4)
Mean	19.1		15.9	24.0	16.9	19.3	16.1	
(±SD)	(11.8)		(11.0)	(14.0)	(7.7)	(13.9)	(4.3)	

Online resource 7: Microsatellite properties per species populations. AP = Antarctic Peninsula and Elephant Island region, SO = South Orkney Islands, n = number of samples, NA = number of alleles [number of private alleles], N_S = number of alleles standardized to the smallest sample size for each species [number of private alleles], H_E = expected heterozygosity, H_O = observed heterozygosity, F_{IS} = inbreeding coefficient. Significant departures from Hardy-Weinberg Equilibrium are indicated as bold F_IS values.

Species Pop (n)		*C. aceratus* Antarctic Peninsula (36)				South Orkneys (89)				
	N_A	N_S	H_E	H_O	F_{IS}	N_A	N_S	H_E	H_O	F_{IS}
Cr15	2 [0]	2 [0]	0.22	0.14	0.38	4 [2]	3.4 [0.8]	0.3	0.21	0.31
Cr127	7 [0]	7 [0]	0.78	0.81	-0.05	9 [2]	7.9 [0.8]	0.78	0.82	-0.05
Cr259	15 [0]	15 [0]	0.87	0.89	-0.03	22 [7]	19.2[2.8]	0.91	0.99	-0.09
Trne20	18 [2]	18 [2]	0.09	0.94	-0.06	23 [7]	17.9 [2.8]	0.92	0.94	-0.02
Trne53	37 [6]	37 [6]	0.98	0.92	0.06	51 [20]	36.3 [8.1]	0.98	0.94	0.04
Trne66	32 [3]	32 [3]	0.97	0.97	0	40 [11]	29.0 [4.5]	0.95	0.93	0.03
Ca26	19 [5]	19 [5]	0.89	0.86	0.03	26 [12]	17.7 [4.9]	0.87	0.75	0.14
Mean	18.6	18.6	0.80	0.79		25.0	18.8	0.81	0.79	
(SD)	(12.5)	(12.5)	(0.26)	(0.29)		(16.4)	(11.3)	(0.24)	(0.27)	

Online resource 7 (continued)

Species					*C. gunnari*					
Pop (n)	Antarctic Peninsula (66)					South Orkneys (26)				
	N_A	N_S	H_E	H_O	F_{IS}	N_A	N_S	H_E	H_O	F_{IS}
Cr15	24 [15]	17.4 [5.9]	0.91	0.91	0	14 [5]	14 [5]	0.82	0.73	0.11
Cr38	4 [2]	3.4 [0.8]	0.6	0.59	0.01	2 [0]	2 [0]	0.5	0.35	0.32
Cr127	9 [3]	7.4 [1.2]	0.8	0.83	-0.04	8 [2]	8 [2]	0.68	0.54	0.21
Cr236	7 [1]	6.9 [0.4]	0.82	0.88	-0.07	8 [2]	8 [2]	0.73	0.73	0
Cr259	19 [3]	15.5 [1.2]	0.92	0.96	-0.04	16 [0]	16 [0]	0.93	0.92	0.01
Trne20	2 [0]	1.9 [0]	0.06	0	1	3 [1]	3 [1]	0.39	0.04	0.9
Trne53	32 [7]	24.2 [2.8]	0.96	0.86	0.1	27 [2]	27 [2]	0.97	1	-0.04
Trne66	9 [1]	6.9 [0.4]	0.78	0.8	-0.03	9 [1]	9 [1]	0.78	0.85	-0.1
Mean	13.3	10.5	0.73	0.73		10.9	10.9	0.72	0.66	
(SD)	(10.6)	(7.7)	(0.29)	(0.32)		(8.1)	(8.1)	(0.20)	(0.30)	

Species					*C. rastrospinosus*					
Pop (n)	Antarctic Peninsula (26)					South Orkneys (30)				
	N_A	N_S	H_E	H_O	F_{IS}	N_A	N_S	H_E	H_O	F_{IS}
Cr15	7 [3]	7 [3]	0.7	0.69	0	6 [2]	5.7 [1.7]	0.65	0.37	0.44
Cr38	6 [1]	6 [1]	0.51	0.42	0.17	6 [1]	5.8 [0.9]	0.57	0.57	0
Cr127	16 [3]	16 [3]	0.94	0.81	0.15	21 [8]	20.0 [6.9]	0.95	0.77	0.19
Cr236	13 [3]	13 [3]	0.9	0.73	0.19	11 [1]	10.7 [0.9]	0.82	0.7	0.15
Cr259	17 [2]	17 [2]	0.94	0.92	0.02	19 [4]	18.0 [3.5]	0.92	0.97	-0.06
Trne20	3 [0]	3 [0]	0.62	0.04	0.94	4 [1]	3.9 [0.9]	0.6	0.07	0.89
Trne53	31 [12]	31 [12]	0.98	0.96	0.02	34 [15]	31.6 [13]	0.98	0.83	0.15
Trne66	28 [9]	28 [9]	0.97	0.96	0.01	34 [15]	31.3 [13]	0.98	0.9	0.08
Mean	15.1	15.1	0.82	0.70		16.9	15.9	0.81	0.64	
(SD)	(10.2)	(10.2)	(0.19)	(0.31)		(12.2)	(11.2)	(0.18)	(0.30)	

Species					*G. gibberifrons*					
Pop (n)	Antarctic Peninsula (52)					South Orkneys (62)				
	N_A	N_S	H_E	H_O	F_{IS}	N_A	N_S	H_E	H_O	F_{IS}
Cr38	27 [7]	27 [7]	0.92	0.92	-0.01	26 [6]	24.6 [5.0]	0.92	0.82	0.11
Cr127	2 [1]	2 [1]	0.02	0.02	0	1 [0]	1.0 [0]	-	-	-
Cr259	30 [5]	30 [5]	0.96	0.96	0	29 [4]	27.8 [3.4]	0.95	0.92	0.03
Trne20	40 [5]	40 [5]	0.97	0.9	0.07	44 [9]	40.9 [7.5]	0.97	0.71	0.27
Trne35	38 [10]	38 [10]	0.96	0.81	0.16	45 [17]	41.8 [14.3]	0.97	0.82	0.15
Trne37	20 [3]	20 [3]	0.85	0.83	0.03	19 [4]	18.0 [3.4]	0.88	0.84	0.05
Trne53	35 [5]	35 [5]	0.97	0.94	0.03	38 [8]	36.2 [6.7]	0.97	0.98	-0.01
Trne55	2 [1]	2 [1]	0.02	0.02	0.16	1 [0]	1.0 [0]	-	-	-
Trne66	26 [4]	26 [4]	0.95	0.9	0.04	26 [4]	24.7 [3.4]	0.95	0.97	-0.02
Ca26	22 [7]	22 [7]	0.93	0.87	0.07	19 [4]	18.4 [3.4]	0.93	0.94	-0.01
Mean	24.2	24.2	0.92	0.86		24.8	23.4	0.94	0.87	
(SD)	(13.4)	(13.4)	(0.99)	(0.10)		(15.52)	(14.5)	(0.03)	(0.09)	

Online resource 7 (continued)

Species										
					L. squamifrons					
Pop (n)	Antarctic Peninsula (27)					South Orkneys (32)				
	N_A	N_S	H_E	H_O	F_{IS}	N_A	N_S	H_E	H_O	F_{IS}
Cr259	4 [1]	4 [1]	0.67	0.59	0.11	3 [0]	3.0 [0]	0.66	0.69	-0.04
Trne20	21 [4]	21 [4]	0.95	0.93	0.03	21 [4]	20.2 [3.4]	0.95	0.81	0.15
Trne35	10 [1]	10 [1]	0.85	0.93	-0.09	12 [3]	11.3 [2.5]	0.79	0.69	0.13
Trne37	14 [2]	14 [2]	0.92	0.48	0.48	16 [4]	15.2 [3.4]	0.91	0.66	0.28
Trne53	18 [4]	18 [4]	0.94	0.85	0.1	18 [4]	17.0 [3.4]	0.91	0.91	0
Trne55	17 [4]	17 [4]	0.91	0.85	0.07	15 [2]	14.0 [1.7]	0.9	0.94	-0.05
Trne66	24 [5]	24 [5]	0.95	1	-0.05	29 [10]	26.7 [8.4]	0.96	0.91	0.05
Ca26	28 [9]	28 [9]	0.97	0.89	0.09	27 [8]	25.2 [6.8]	0.95	1	-0.05
Mean	17.0	17.0	0.87	0.80		17.6	16.6	0.88	0.82	
(SD)	(7.7)	(7.7)	(0.11)	(0.20)		(8.3)	(7.7)	(0.10)	(0.13)	

Species										
					T. eulepidotus					
Pop (n)	Antarctic Peninsula (38)					South Orkneys (31)				
	N_A	N_S	H_E	H_O	F_{IS}	N_A	N_S	H_E	H_O	F_{IS}
Cr259	2 [0]	2.0 [0]	0.14	0	1	2 [0]	2 [0]	0.228	0	1
Trne20	5 [2]	4.8 [1.6]	0.44	0.08	0.83	3 [0]	3 [0]	0.41	0.161	0.61
Trne35	42 [23]	37.3 [18.8]	0.98	0.76	0.22	35 [16]	35 [16]	0.976	0.871	0.109
Trne37	13 [3]	12.2 [2.4]	0.83	0.77	0.09	15 [5]	15 [5]	0.857	0.935	-0.093
Trne53	25 [4]	23.7 [3.3]	0.96	0.92	0.04	23 [2]	23 [2]	0.957	0.935	0.023
Trne66	20 [5]	18.9 [4.1]	0.92	0.92	-0.01	18 [3]	18 [3]	0.923	0.903	0.022
Ca26	41 [14]	37.0 [11.4]	0.98	0.87	0.12	34 [7]	34 [7]	0.976	0.774	0.21
Mean	21.1	19.4	0.75	0.62		18.6	18.6	0.76	0.65	
(SD)	(16.0)	(14.3)	(0.33)	(0.40)		(13.3)	(13.3)	(0.31)	(0.40)	

Species										
					T. newnesi					
Pop (n)	Antarctic Peninsula (34)					South Orkneys (28)				
	N_A	N_S	H_E	H_O	F_{IS}	N_A	N_S	H_E	H_O	F_{IS}
Cr259	19 [5]	18.0 [4.1]	0.923	0.882	0.045	15 [1]	15 [1]	0.912	0.893	0.021
Trne20	25 [11]	23.4 [9.1]	0.956	0.941	0.016	17 [3]	17 [3]	0.877	0.786	0.106
Trne35	19 [7]	17.7 [5.8]	0.911	0.941	-0.033	19 [7]	19 [7]	0.918	1	-0.092
Trne37	11 [3]	10.6 [2.5]	0.869	0.941	-0.085	8 [0]	8 [0]	0.827	0.893	-0.081
Trne53	20 [7]	18.3 [5.8]	0.913	0.912	0.001	17 [4]	17 [4]	0.885	0.929	-0.05
Trne55	12 [2]	11.6 [1.6]	0.857	0.735	0.144	11 [1]	11 [1]	0.794	0.75	0.056
Trne66	18 [7]	15.8 [5.8]	0.791	0.735	0.072	17 [6]	17 [6]	0.88	0.929	-0.056
Mean	17.7	16.5	0.91	0.81		14.9	14.9	0.87	0.88	
(SD)	(4.8)	(4.4)	(0.05)	(0.12)		(3.9)	(3.9)	(0.04)	(0.09)	

Online resource 8: Bottleneck results based on microsatellite data per species populations. Probability for mutation-drift equilibrium derived from Wilcoxon signed-ranked test (one-tail for heterozygosity excess), mode-shift indicator and non-modified Garza-Williamson index.

| | Antarctic Peninsula | | | South Orkneys | | |
Species	p (H_e excess)	Mode-shift	Garza-Williamson	p (H_e excess)	Mode-shift	Garza-Williamson
C. aceratus	0.234	no	0.71 (0.17)	0.148	no	0.74 (0.15)
C. gunnari	0.014	no	0.75 (0.19)	0.004	yes	0.63 (0.20)
C. rastrospinosus	0.002	no	0.65 (0.20)	0.004	no	0.67 (0.17)
G. gibberifrons	0.348	no	0.63 (0.21)	0.012	no	0.63 (0.14)
L. squamifrons	0.273	no	0.71 (0.13)	0.148	no	0.68 (0.12)
T. eulepidotus	0.148	no	0.71 (0.21)	0.188	no	0.73 (0.22)
T. newnesi	0.531	no	0.46 (0.19)	0.008	yes	0.44 (0.21)

Online resource 9: Posterior probabilities of migration rates simulated with IMA2. a) migration from AP/EI to SO, b) migration from SO to AP/EI.
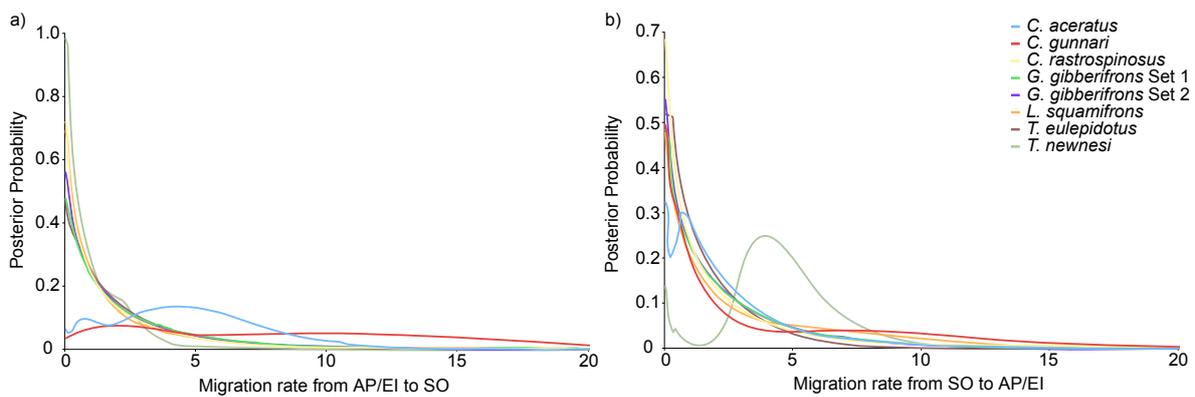
2.3 TANDEM: integrating automated allele binning into genetics and genomics workflows

Matschiner M, Salzburger W

Genetics and population analysis

TANDEM: integrating automated allele binning into genetics and genomics workflows

Michael Matschiner* and Walter Salzburger

Zoological Institute, University of Basel, Vesalgasse 1, 4051 Basel, Switzerland

ABSTRACT

Summary: Computer programs for the statistical analysis of microsatellite data use allele length variation to infer, e.g. population genetic parameters, to detect quantitative trait loci or selective sweeps. However, observed allele lengths are usually inaccurate and may deviate from the expected periodicity of repeats. The common practice of rounding to the nearest whole number frequently results in miscalls and underestimations of allelic richness. Manual sorting of allele lengths into discrete classes, a process called binning, is tedious and error-prone. Here, we present a new program for the automated binning of microsatellite allele lengths to overcome these problems and to facilitate high-throughput allele binning.

Availability: www.evolution.unibas.ch/salzburger/software.htm

Contact: michael.matschiner@unibas.ch

Supplementary information: Supplementary data are available at *Bioinformatics* online.

1 INTRODUCTION

Microsatellites are among the most important molecular markers for a wide range of biological questions (Schlötterer, 2004). Despite the recent trend towards single nucleotide polymorphisms (SNPs), microsatellites still enjoy many advantages, in particular for studies of nonmodel organisms, where substantial effort is required for the development of a sufficient number of biallelic markers (Ryynanen *et al.*, 2007). Among the main problems associated with microsatellites are effort and error rates of genotyping. Scoring of microsatellite alleles is typically performed using commercial software such as GENEMAPPER (Applied Biosystems) or GENEMARKER (Softgenetics). Both programs calculate allele lengths through comparison with internal size standards run alongside PCR amplified fragments in capillary electrophoresis. This calculation is based on the assumption of equal migration rates of DNA fragments of the same length. However, migration rates depend not only on fragment length, but also on DNA sequence motifs (Rosenblum *et al.*, 1997) and fluorescent labels (Wenz *et al.*, 1998). Therefore, measured fragment lengths are often inaccurate. As microsatellite variability mainly results from slippage synthesis (Schlötterer and Tautz, 1992) allele sizes are expected to conform to the periodicity of the repeated motif. However, it has been observed that the effective spacing between peaks of observed allele sizes varies between 1.77 and 2.23 base pairs (bp) (Amos *et al.*, 2007). This so-called 'allelic drift' renders automated binning of alleles

---

*To whom correspondence should be addressed.

a nontrivial task (Idury and Cardon, 1997). Simple rounding of alleles to the nearest whole number will lead to inconsistencies, such as the presence of even and odd alleles for the same marker. Rounding to the nearest number conforming the expected periodicity will merge alleles and may cause underestimates of allelic richness and heterozygosity when the effective spacing between peaks of dinucleotide repeat loci is >2.0 bp. Manual binning of alleles is time-consuming, error-prone and often arbitrary. GENEMAPPER's built-in binning method requires reference data that is usually not available for nonmodel organisms. When GENEMAPPER is run without exhaustive reference data, new alleles that fall outside established bins are placed with poor accuracy (Amos *et al.*, 2007). Automated binning without reference data has been addressed by the software packages ALLELOBIN (Idury and Cardon, 1997) and FLEXIBIN (Amos *et al.*, 2007), using least-squares minimization procedures and allowing for allelic drift. Here, we present a new program called TANDEM that is specifically designed for seamless integration into population genetic and genomic workflows as it requires no additional reformatting of data files. It is freely available in two versions: (i) a Macintosh version, which is equipped with a basic graphical user interface (GUI) and (ii) Ruby source code, which is compatible with Macintosh, Windows and Linux systems.

2 METHODS

For minimum configuration effort, TANDEM has been designed to accept files in the format of the programs MSA (Dieringer and Schlötterer, 2003) and CONVERT (Glaubitz, 2004), that are both commonly used starting points for population genetic and genomic workflows (Excoffier and Heckel, 2006; Teschke *et al.*, 2008; Zhang *et al.*, 2007). These two programs are able to convert spreadsheet data into input files for a large number of downstream applications, such as ARLEQUIN (Excoffier *et al.*, 2005), STRUCTURE (Pritchard *et al.*, 2000), GENEPOP (Raymond and Rousset, 1995), MIGRATE (Beerli, 2005) and IM (Hey and Nielsen, 2004). MSA as well as CONVERT expect alleles that have already been binned, while TANDEM uses the same formats with unbinned alleles.

In order to compensate for allelic drift and compression at large fragment sizes, TANDEM transforms all allele sizes before rounding. To this end, TANDEM optimizes all parameters of the power function

$$\text{transformed allele size} = a + b \times \text{observed allele size}^c \qquad (1)$$

so that rounding errors of transformed allele sizes become minimal. TANDEM applies a least-squares minimization of rounding errors. Parameter optimization is performed using an exhaustive search or, optionally, using a heuristic search with the Nelder–Mead Downhill Simplex algorithm (Nelder and Mead, 1965). Prior parameter bounds and step sizes for the exhaustive search are listed in Supplementary Table 1. By default, transformed allele
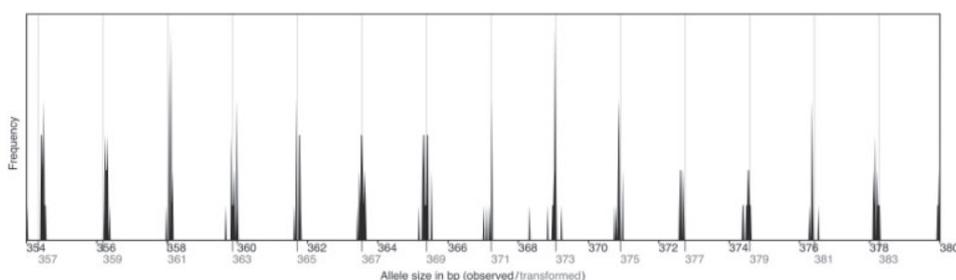
Fig. 1. Frequency histogram for one marker from data set 1. Black labels refer to observed allele sizes, while transformed allele sizes are shown with grey labels (in bp). Observed alleles do not conform to dinucleotide periodicity, but transformed alleles match the expected pattern.

sizes are adjusted so that the shortest observed allele size per locus presents a fixpoint. However, TANDEM allows the user to fix other points based on prior information about actual allele sizes. Subsequent to transformation, allele sizes are rounded to the nearest whole number, whereby the repeat size of the microsatellite is taken into account (e.g. repeat size = 2 for dinucleotide repeats). Where repeat sizes of loci are not specified by the user, they are estimated by TANDEM, based on observed allele sizes. Rounded allele sizes are written to a separate file in the same format as the input file. The average rounding error is calculated over all alleles and is included in the output to serve as a marker-specific quality indicator. The transformation of allele sizes is visualized by frequency histograms as shown in Figure 1. From these histograms, the overall quality of each marker, and the validity of the performed allele transformation become apparent. Outlier alleles with large rounding errors are highlighted to alert the user of problematic samples that should possibly be removed from the data set. Bin sets are automatically exported.

3 RESULTS AND DISCUSSION

In order to verify TANDEM results, we sequenced the locus depicted in Figure 1 from five individuals following Matschiner *et al.* (2009). Results show close agreement of fragment lengths observed by genotyping and by sequencing (Supplementary Table 2). The best fit was found for shorter allele sizes, thus justifying our approach to use the shortest fragment as a fixpoint. However, if the user is interested in absolute allele sizes, we generally recommend to sequence the respective locus in at least one individual, and to specify fixpoints accordingly. We also recommend specification of repeat sizes whenever known.

We benchmarked TANDEM's exhaustive search algorithm against ALLELOBIN and FLEXIBIN using four different data sets, containing 8–23 microsatellite loci and varying numbers of diallelic individuals (Supplementary Table 3). We found TANDEM to perform favorably compared to both other programs. Especially when data sets included tri- and tetranucleotide repeats, TANDEM performed substantially better than FLEXIBIN. In conclusion, we present a user-friendly and versatile program for the automatic binning of microsatellite alleles that performs better than alternative software. Moreover, TANDEM is the first such program that does not require tedious and error-prone reformatting of allelic data, and thus integrates well into existing population genetic and genomic workflows.

ACKNOWLEDGEMENTS

We thank O. Balmer for discussion and permission to use his unpublished data, F. Münzel and four reviewers for valuable comments.

REFERENCES

Amos,W. *et al.* (2007) Automated binning of microsatellite alleles: problems and solutions. *Mol. Ecol. Notes*, 7, 10–14.

Beerli,P. (2005) Comparison of Bayesian and maximum-likelihood inference of population genetic parameters. *Bioinformatics*, **22**, 341–345.

Dieringer,D. and Schlötterer,C. (2003) Microsatellite Analyser (MSA): a platform independent analysis tool for large microsatellite data sets. *Mol. Ecol. Notes*, **3**, 167–169.

Excoffier,L. and Heckel,G. (2006) Computer programs for population genetics data analysis: a survival guide. *Nat. Rev. Gen.*, **7**, 745–758.

Excoffier,L. *et al.* (2005) Arlequin ver. 3.0: an integrated software package for population genetics data analysis. *Evol. Bioinf. Online*, **1**, 47–50.

Glaubitz,J.C. (2004) Convert: a user-friendly program to reformat diploid genotypic data for commonly used population genetic software packages. *Mol. Ecol. Notes*, **4**, 309–310.

Hey,J. and Nielsen,R. (2004) Multilocus methods for estimating population sizes, migration rates and divergence time, with applications to the divergence of *Drosophila pseudoobscura* and *D. persimilis*. *Genetics*, **167**, 747–760.

Idury,R.M. and Cardon,L.R. (1997) A simple method for automated allele binning in microsatellite markers. *Genome Res.*, **7**, 1104–1109.

Matschiner,M. *et al.* (2009) Gene flow by larval dispersal in the Antarctic notothenioid fish *Gobionotothen gibberifrons*. *Mol. Ecol.*, doi: 10.1111/j.1365-294X.2009.04220.x.

Nelder,J.A. and Mead,R. (1965) A simplex method for function minimization. *Comput. J.,* **7**, 308–313.

Pritchard,J.K. *et al.* (2000) Inference of population structure using multilocus genotype data. *Genetics*, **155**, 945–959.

Raymond,M. and Rousset,F. (1995) Genepop (version 1.2): population genetics software for exact tests and ecumenicism. *J. Hered.*, **86**, 248–249.

Rosenblum,B.B. *et al.* (1997) Improved single-strand DNA sizing accuracy in capillary electrophoresis. *Nucl. Acids Res.*, **25**, 3925–3929.

Ryynanen,H.J. *et al.* (2007) A comparison of biallelic markers and microsatellites for the estimation of population and conservation genetic parameters in Atlantic salmon (*Salmo salar*). *J. Hered.*, **98**, 692–704.

Schlötterer,C. (2004) The evolution of molecular markers-just a matter of fashion? *Nat. Rev. Gen.*, **5**, 63–69.

Schlötterer,C. and Tautz,D. (1992) Slippage synthesis of simple sequence DNA. *Nucleic Acids Res.*, **20**, 211–215.

Teschke,M. *et al.* (2008) Identification of selective sweeps in closely related population of the house mouse based on microsatellite scans. *Genetics*, **180**, 1537–1545.

Wenz,H.M. *et al.* (1998) High-precision genotyping by denaturing capillary electrophoresis. *Genome Res.*, **8**, 69–80.

Zhang,J. *et al.* (2007) Genetic analysis and linkage mapping in a resource pig population using microsatellite markers. *J. Genet. Genomics,* **34**, 10–16.

Supporting Information

TANDEM: integrating automated allele binning into population genetic workflows

Michael Matschiner[*] **and Walter Salzburger**[*]

Address:

[*]Zoological Institute, University of Basel, Vesalgasse 1, CH-4051 Basel, Switzerland

Correspondence:

Michael Matschiner, Fax: +41 61 267 03 01, E-mail: michael.matschiner@unibas.ch

1   Tables

Tab. S1: Prior parameter bounds and step sizes used for the exhaustive optimization of parameters a, b, c of Equation 1. AS_{MIN}: size of the shortest allele found per locus; RS: repeat size (e.g. RS = 2 for dinucleotide repeats).

Parameter	Lower bound	Upper bound	Step size	Number of steps
a	$AS_{MIN} - 0.3 \times RS^*$	$AS_{MIN} + 0.3 \times RS^*$	$0.05 \times RS$	13
b	$0.85^\dagger$	$1.15^\dagger$	0.01	31
c	0.985	1.015	0.005	7

* By default, a is centered around AS_{MIN}, and AS_{MIN} is subtracted from all observed allele sizes before transformation in order to treat the shortest allele size as a fixpoint that is transformed onto itself before rounding. This is done to account for our results that reveal a close agreement between short fragment lengths observed by genotyping and those observed by sequencing (see Results section and Supplementary Table 3). If another fixpoint is specified by the user (see Methods section and TANDEM manual) the above bounds for a are still used, but an additional parameter is added to all allele sizes after transformation for results to conform with the specified fixpoint.
We chose a parameter range of $2 \times 0.3 \times RS = 0.6 \times RS$ to cover more than $0.5 \times RS$, instead of covering a full RS. This is sufficient because rounding is always tested both to the nearest even and to the nearest odd integer (in the case of dinucleotide repeats, and accordingly for other repeat sizes).
$\dagger$ Prior parameter bounds for b are based on the results of Amos et $al.$ (2007), who found effective spacings between allele peaks that ranged from 1.77 bp to 2.23 bp in the case of dinucleotide repeats.

Tab. S2: Comparison of fragment lengths observed by both genotyping and sequencing, and of length estimates calculated by TANDEM. Specimen 1 sequencing results were used as a reference point for TANDEM calculations. Sequencing led to ambiguous results for specimens 4 and 5 due to stutter bands.

Specimen	Observed fragment length		TANDEM
	Genotyping	Sequencing	estimate
1	330.15	330	330
2	356.26	358	358
3	356.28	358	358
4	359.98	360-362	362
5	369.03	370-372	372

Tab. S3: Performance comparison of ALLELOBIN, FLEXIBIN, and TANDEM. Four microsatellite data sets were tested, including the one supplied with the FLEXIBIN software (data set 4). The average rounding error is given as a criterion for binning quality.

Data set	Locus	Sample size	Repeat size (bp)	Number of alleles			Average rounding error		
				ALLELOBIN	FLEXIBIN	TANDEM	ALLELOBIN	FLEXIBIN	TANDEM
1	1	163	2	61	62	57	0.45	0.40	0.37
	2	164	2	36	33	33	0.39	0.16	0.17
	3	165	2	44	47	47	0.46	0.22	0.14
	4	164	2	34	34	34	0.26	0.24	0.24
	5	163	2	34	34	34	0.24	0.18	0.19
	6	165	2	5	5	5	0.05	0.05	0.05
	7	165	2	1	1	1	0.04	0.04	0.04
	8	165	2	38	38	38	0.16	0.08	0.08
	9	165	2	30	30	30	0.33	0.16	0.15
	All						**0.22**	**0.17**	**0.16**
2	1	524	2	16	16	16	0.21	0.08	0.05
	2	524	2	17	17	17	0.22	0.22	0.22
	3	524	2	9	8	8	0.11	0.08	0.08
	4	514	2	10	10	10	0.25	0.21	0.15
	5	521	2	10	11	10	0.08	0.09	0.07
	6	383	2	15	15	15	0.13	0.08	0.09
	7	322	2	17	17	16	0.26	0.13	0.14
	8	522	2	11	11	11	0.17	0.05	0.05
	All						**0.18**	**0.12**	**0.11**
3	1	219	2	16	16	16	0.19	0.15	0.14
	2	222	2	18	19	20	0.49	0.42	0.39
	3	110	2	38	38	39	0.24	0.16	0.16
	4	43	2	18	18	18	0.13	0.18	0.10
	5	117	2	32	32	32	0.21	0.19	0.13
	6	224	2	19	19	19	0.22	0.16	0.15
	7	230	2	22	23	21	0.25	0.29	0.24
	8	254	2	24	24	24	0.14	0.14	0.13
	9	232	2	17	17	16	0.36	0.15	0.13
	10	47	2	19	19	20	0.24	0.13	0.11
	11	211	2	20	20	21	0.13	0.22	0.11
	12	126	4	11	11	11	0.43	0.70	0.40
	13	275	2	25	25	25	0.29	0.16	0.15
	14	90	2	29	30	30	0.29	0.18	0.14
	15	228	3	17	17	17	0.27	0.18	0.18
	16	119	3	16	15	15	0.17	0.14	0.13
	17	137	3	17	17	16	0.28	0.30	0.10
	18	35	3	8	8	8	0.63	0.22	0.17
	19	15	2	15	15	15	0.28	0.28	0.26
	20	139	3	12	12	14	0.45	0.45	0.29
	21	113	2	20	20	20	0.17	0.15	0.15
	22	41	2	10	13	13	0.37	0.31	0.30
	23	110	4	9	10	10	0.32	0.55	0.16
	All						**0.30**	**0.24**	**0.20**

Tab. S3 (continued)

Data set	Locus	Sample size	Repeat size (bp)	Number of alleles			Average rounding error		
				ALLELOBIN	FLEXIBIN	TANDEM	ALLELOBIN	FLEXIBIN	TANDEM
4	d13s153	219	2	18	19	18	0.30	0.27	0.27
	d13s158	222	2	15	15	15	0.16	0.10	0.10
	d13s159	110	2	16	17	17	0.38	0.30	0.33
	d13s170	43	2	14	14	14	0.17	0.16	0.15
	d13s171	117	2	13	13	13	0.23	0.12	0.13
	d13s173	224	2	12	12	12	0.25	0.25	0.25
	d13s175	230	2	10	10	10	0.22	0.16	0.17
	d13s217	254	2	11	11	11	0.35	0.31	0.33
	d13s218	232	2	8	8	8	0.10	0.10	0.10
	d13s263	47	2	12	12	12	0.26	0.20	0.21
	d13s265	211	2	11	10	10	0.27	0.17	0.17
	d13s285	126	2	13	14	14	0.34	0.20	0.21
	All						**0.25**	**0.20**	**0.20**

tandem
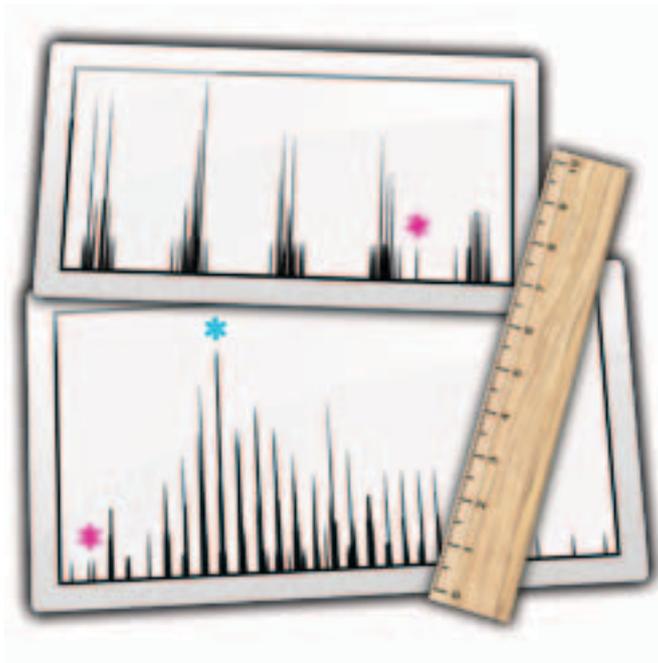
Summary

All microsatellite analysis software expects allele sizes given in integer numbers, while allele scoring produces allele sizes with two decimals that are dependent not only on fragment length, but also on fluorescent dye, and GC content. Therefore, allele binning is not a trivial task. tandem fills a gap of the microsatellite workflow by rounding allele sizes to valid integers, depending on the microsatellite repeat units. Publish-ready vector graphics output shows allele size distribution and visualizes the rounding method. The average rounding error is given and indicates the overall quality of microsatellite data.
tandem runs natively on Macintosh computers. Source code is written in Ruby and works on Mac, Windows, and Linux computers. tandem is easy to use. All you need is an input file.

Problem

When analyzing microsatellites, one typically uses software like Genemapper (Applied Biosystems) to score the sizes of fluorescently labelled PCR products. These products are a combination of forward primer, flanking region, microsatellite, flanking region, and reverse primer. Of these, both primer sizes are known, and flanking region sizes are assumed to be constant among all individuals. Thus, variation in PCR product sizes among individuals should directly reflect different numbers of microsatellite repeats. In the case of dinucleotide repeats, alleles are expected to be either only even or only odd integers. However, Genemapper calculates allele sizes from comparison with labeled size standards of known size that are added to all samples before running them on a capillary sequencer. Errors are introduced by minor differences between runs and capillaries, the precision limit of sequencers, and imperfect linear regressions between size standard and PCR product run lengths. As a result, calculated allele sizes hardly ever are integers. Instead Genemapper measures allele sizes to two

decimals. Its built-in automated binning method requires reference data, which is often not available in population genetic studies and has a number of problems associated with it (Amos *et al.* 2007). Thus, unbinned allele size data is commonly exported in tables like the one shown in fig. 1.

◇	A	B	C	D	E
1		Locus 1		Locus 2	
2	sample 1	182.02	187.73	189.6	189.6
3	sample 2	183.93	183.93	181.33	189.59
4	sample 3	180.07	187.67	189.66	191.64
5	sample 4	183.88	197.04	191.68	210.18
6	sample 5	183.94	185.82	189.66	189.66
7	sample 6	183.93	185.83	189.69	191.77
8	sample 7	182.04	183.93	193.77	199.9
9	sample 8	183.84	197.07	191.67	210.08
10	sample 9	183.88	185.78	189.56	191.64
11	sample 10	183.89	183.89	189.63	189.63
12	sample 11	183.88	191.44	189.66	199.9
13	sample 12	185.74	193.3	191.71	199.9

Fig. 1: *Allele size table, as given by Genemapper.*

However, all microsatellite analysis software, including Arlequin (Excoffier *et al.* 2005), Structure (Falush *et al.* 2003), Migrate (Beerli 2006), IM (Hey & Nielsen 2004), and others, expect integer allele sizes. Of course, you could use MS Excel to either cut off all decimals, or to round all values to the nearest integer. You could even program Excel to round only to even or odd numbers in every case. And you could easily find out whether rounding errors are larger to the even, or to the odd numbers. But this still has some problems associated with it. One of them is demonstrated in fig. 2. This figure is part of the tandem output and shows the relative densities of allele sizes for a single microsatellite locus (don't confuse this with stutter bands!). As you can see, allele sizes do not peak at integer numbers. Instead, a consistent negative shift of, say 0.3 bp can be found throughout all values. In this case, rounding to the nearest even number would be fine for most allele sizes, but not for some outliers. The allele marked with an asterisk would be rounded down to 324, and not to the de facto closer peak at ~325.7. It's easy to imagine more serious situations, where a larger fraction of allele sizes is rounded the wrong way.

Worse than this, peaks may not only be shifted, but the average distance between peaks can be slightly less or more than the microsatellite repeat size. Thus you may find dinucleotide loci with average peak distances of, say 1.9, or 2.1, a phenomenon called 'allelic drift' (Idury & Cardon 1997). This is because Genemapper expects collinearity between size standard and PCR product mobility, which is not necessarily the case. Migration rates of alleles depend not only on its length, but also on GC content (Amos *et al.* 2007) and fluorescent labels (Wenz *et al.* 1998). Fig. 3 shows an example where peak distance is less than 2.0. The two peaks marked with asterisks would be rounded to 368 and 370 bp, while the peak in between would be split in two, and merged with the 368 and 370 bp peaks. This will change peak patterns and underestimate allelic variation. Similar problems arise if peak distances are larger than the original repeat size.
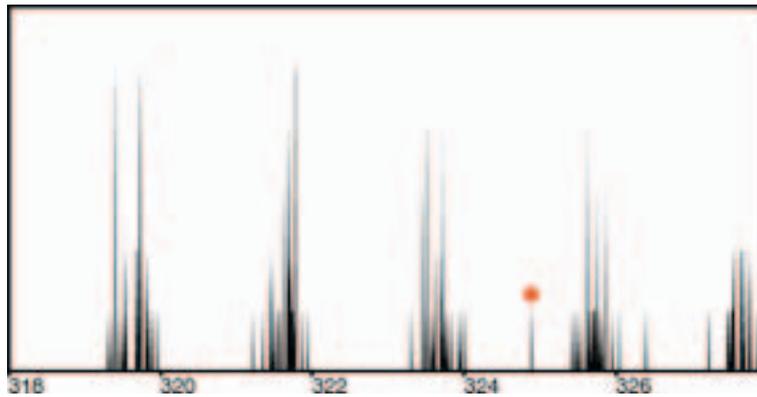
Fig. 2: *Allele size peaks with a small shift to the left.*
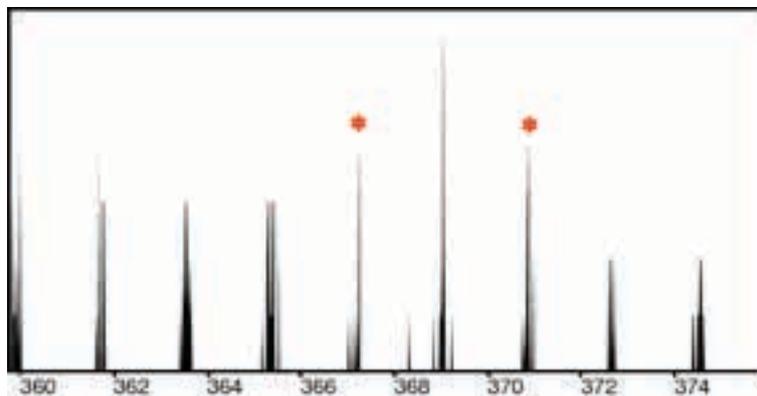
Fig. 3: *Allele size peaks with distances of less than 2.0 bp.*

Solution

This is where tandem comes in. tandem goes through tab delimited versions of Excel sheets like the one shown in fig. 1, and rounds all allele sizes to integer numbers. But instead of simply rounding to the nearest even or odd number (or other numbers following tri-, tetra-, etc-, nucleotide repeat patterns) tandem finds the most consistent way of rounding. It transforms all observed allele sizes using the power function

$$\text{(transformed allele size)} = a + b \times \text{(observed allele size)}^c$$

and exhaustively optimizes parameters a, b, and c, so that rounding errors are minimal when rounding transformed allele sizes to integers that fit the expected nucleotide repeat patterns (*e.g.* when rounding transformed allele sizes only to even, or only to odd integers).

More precisely, the following values of a, b, and c are tested:

a: -0.3 × repeat size ⸺ +0.3 × repeat size; steps of 0.05 × repeat size (*i.e.* -0.6, -0.5, -0.4,...+0.4, +0.5, +0.6 for dinucleotide repeats)

b: 0.85 ⸺ 1.15; steps of 0.01 (*i.e.* 0.85, 0.86, 0.87, ... 1.13, 1.14, 1.15)

c: 0.985 ⸺ 1.015; steps of 0.005 (*i.e.* 0.985, 0.990, 0.995, 1.000, 1.005, 1.010, 1.015)

Boundaries were chosen based on empirical tests. Also see Amos *et al.* (2007) for boundaries of b. Despite not covering a whole repeat size, boundaries of a are sufficient as rounding of transformed allele sizes is always tested both to odd, and to even integers (for dinucleotide repeats, and in a similar manner for tri-, tetra-, etc. repeats.)

tandem's way of rounding is made transparent by HTML output including publish-ready SVG vector graphic plots that show, per locus, the original allele size distributions, as well as the fitting of the data to the according repeat size pattern (see fig. 4). Two plots are given per locus, one for the full range of allele sizes, and one focussing on the part with the highest density. Per locus, the optimized parameters a, b, c, and the average rounding error are reported alongside other relevant information (see below). Rounding error outliers are highlighted and indicate individuals that should be removed from the data set, or problems with the specified repeat size. The consistency of tandem's rounding scheme immediately becomes obvious from these plots. Grey vertical lines indicate bin centers after allele transformation. If these match the peaks of the allele size distribution (shown in black), as they do in fig. 4, tandem successfully optimized all parameters of the power function, so that rounding errors are minimized. This means that tandem was able to bin alleles in the most consistent way *relative to each other*. If you've got your own microsatellite data set that you're not going to combine with data sets from other laboratories, and if your next step would be a population genetic analysis with software like Arlequin, Structure, Migrate, or IM, you're fine with that, as these programs only use relative distances, and never the absolute values. Some applications, like IM, even expect only the number of repeats (not the allele size in bp), whereby it is commonly assumed that the shortest allele in the data set has a repeat number of 0. In these cases, relative consistency is all you need to worry about, and you don't need to read the rest of this chapter.

However, there are cases when you need to know the true *absolute* fragment length, for example when combining data sets scored in different laboratories, or with different fluorescent dyes that differently affect capillary migration rates. In these cases, you'd like to know, for example, whether the peak on the far left in fig. 4 really is produced by alleles of 362 bp, or rather of 360, or 364 bp. As allelic drift commonly produces between-peak distances that deviate from the actual repeat size (as shown in fig. 3), one could assume that the observed fragment lengths are far from the actual fragment lengths. For example, if you find between-peak distances of roughly 1.8 bp instead of 2.0 bp, as in fig. 3, you could assume that allelic migration is slower than size standard migration by a factor of 1.8/2.0 = 0.9. Thus, you would expect that a fragment of 400 bp migrates as fast as the size standard of 360 bp (400 bp × 0.9), and therefore, that the peak observed at 360 bp actually corresponds to fragments of 400 bp. In order to test this assumption, we've sequenced a microsatellite locus that showed between-peak distances of roughly 1.8 bp. We did this for six specimens. Surprisingly, we did not find the expected large differences between observed and actual fragment lengths. Instead, observed fragment lengths reflected the actual fragment lengths rather precisely, with minor differences of 2-4 bp. This leads us to conclude that the relation between fragment lengths and migration rates is nonlinear, and may for example be affected by changing GC content with length when the microsatellite motif is GC or AT. Taken together, the absolute fragment length is difficult to predict, and tandem cannot guarantee to find it. Given our sequencing results, the above power function was extended to be conservative about absolute fragment lengths. tandem by default fixes the shortest observed allele size per locus to equal the transformed allele size:

(transformed allele size) = a + b × (observed allele size - shortest observed allele size)^c,

where a is at least the shortest observed allele size (± 0.3 × repeat size; see above). This should result in estimates of fragment lengths that are at least very close to the actual fragment lengths. However, if you want to be sure about absolute fragment lengths, you will need to sequence the locus in one of the specimens (choose a homozygote specimen). tandem provides a way to include your sequencing result in its calculations, and adjusts its estimates accordingly (see chapter 'Advanced settings').
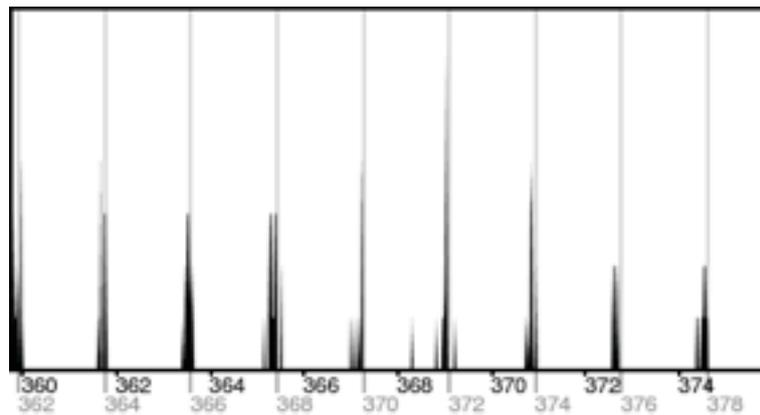
Fig. 4: *Values in grey show allele sizes after transformation.*

Input file

Since no standard file format exists for the microsatellite workflow, and most downstream applications use their own specific format, there's a number of programs that mainly do conversion from spreadsheet format to specific formats and thus provide excellent starting points for microsatellite analysis. These conversion programs include MSA (Dieringer & Schlötterer 2003), Convert (Glaubitz 2004), and Create (Coombs et al. 2007). While the format of Create is very flexible, those of MSA and Convert are more strict, but still very easy to build. tandem is designed for minimal configuration effort, and thus supports both MSA and Convert format. This means you can create your MSA or Convert input file as usual, but without the need of rounding allele sizes. tandem does that for you.

If using Convert format, there's one minor detail to add: Since Convert format does not include information about the repeat size, but this information is important for tandem, please add the repeat size (i.e. whether its a di-, tri-, etc- nucleotide repeat) in the cells directly above the cells containing the names of the loci. While tandem nevertheless tries to estimate the repeat size from the data, and alerts you if there's a conflict between its estimate and the repeat size you specified, this estimation doesn't always work. You're on the safe side if you specify a repeat size.

MSA and Convert input files are shown in fig. 5 and fig. 6. Examples for both MSA and Convert format are included in the tandem distribution package as Excel (xls) files. These files include comments on how to enter information according to the respective formats. **In order to run these example files with tandem, you need to open them in Excel and save them again as tab-delimited files!** (Click 'Save as' and choose 'Tab Delimited Text (.txt)' format). **This applies not only to the example files, but to all input files.**

124

Fig. 5: *MSA input format, unbinned.*

Fig. 6: *Convert input format, unbinned.*

Running tandem

Mac version (GUI)

Running the Mac version of tandem could hardly be easier. Simply double-click tandem, and you'll be prompted to choose an input file. Do so, and tandem immediately starts running. Runs take around twenty seconds per locus, for data sets containing a few hundred individuals, on a 2 GHz iMac. At the end of the run, tandem gives you a short message, telling you which file format it recognized (MSA or Convert), and where it wrote the output files (which is in the folder of the input file).

Source code

If using the source code version, I recommend to copy 'tandem.rb' as well as the folder 'tandem_resources' into the directory of the input file. And if you're using Windows, I also recommend to choose, as a directory for all these files something as simple as the Desktop,

because longer paths are difficult to navigate to with the Windows command line.

While the programming language Ruby is installed by default on Macintosh computers, you'll first need to get it if you're on Windows. There's a number of packages available, but I recommend the one-click, self-contained installer found at http://rubyinstaller.rubyforge.org/wiki/wiki.pl. Once you've got Ruby, open the Terminal (on a Mac) or the command prompt (on Windows; click Start > Run, then type 'command'). Test whether Ruby is installed by typing ruby -v, and you should get a version number as reply. Navigate into the directory of the input file by typing 'cd' (on Mac) or 'chdir' (on Windows), followed by a space and the path to this directory (*e.g.* 'chdir DESKTOP'). Once you're there, type

ruby tandem.rb -i <input_file_name>

to start the run, where <input_file_name> needs to be replaced with the actual name of your input file. The output will be written to the same directory.

Output files

tandem outputs a file which looks exactly like your input file, but with rounded allele sizes. It is called '**<input_file_name>_tandem.txt**'. tandem also writes an HTML file with embedded SVG vector graphics, called '**<input_file_name>_tandem.htm**'. The SVG images plot the frequency of allele sizes, first over the full allele size range, then over the shortest range containing 50% of all allele sizes. You can edit these SVG images for publication using software such as Adobe Illustrator or a free alternative, such as Inkscape. You'll find the SVG plots in folder '**<input_file_name>_tandem_resources**'.

In addition, tandem writes a bin file and a panel file that can be imported into Genemapper: '**<input_file_name>_tandem_bins.txt**' and ' **<input_file_name>_tandem_panel.txt**'. These files enable you to use tandem results as reference data for future analyses with Genemapper. This makes sense if you have a large data set that probably includes all the existing alleles for a particular locus, and if the average rounding error is small. If you intend to use tandem results as Genemapper reference data, you should also consider sequencing each locus in one individual, and using a fixpoint file (see 'Advanced settings') to improve reliability in absolute allele size estimates. The bin file contains, per locus, all allele sizes after transformation and rounding (*i.e.* integer numbers) as Genemapper bin names, and the corresponding untransformed alleles as Genemapper bin centers. This reference data allows you to use Genemappers automated binning function, and to export binned alleles directly from Genemapper. This may safe some time in your microsatellite workflow, however, if your reference data wasn't exhaustive, you risk missing new alleles that fall outside of established bins. Also note that per locus, you should use the same fluorescent dye in future analyses as you did when you created the reference data with tandem. This is important because fluorescent dyes may affect migration rates of alleles.

Furthermore, the HTML file contains the following information for every locus:

- **Specified repeat size:** this is what you specified in the cells above loci names. If your specification is readable for tandem (i.e. if you entered an integer number), it will be used for rounding. The maximum repeat size is currently set to 8. If you seriously think your repeat size is larger, contact me, and I can change this setting. It is recommended that you specify the repeat size, if you know it, as tandem cannot guarantee to estimate the repeat size correctly.

- **Estimated repeat size:** tandem's best guess for the actual repeat size, given the data. If this estimate is not what you specified, it will be written in red to alert you. You may want to

check the respective SVG plots in this case. If you did not specify a repeat size, tandem uses its own estimate for rounding. If tandem is not able to estimate the repeat size, a question mark will be given. If no repeat size is specified, and tandem cannot estimate it, the default repeat size of 2 will be used for rounding.

- **Individuals:** The number of individuals for which allele size information is found. This is calculated as the total number of allele sizes divided by two, as two alleles are expected per individual.

- **Allele range:** The smallest and the largest allele size of this locus are given.

- **a, b, c:** Parameters of the power function
  (transformed allele size) = a + b × (observed allele size - smallest observed allele size)^c
  that is applied to all observed allele sizes in order to minimize rounding error.

- **Fixed:** This information is given if you've specified a fixpoint for a particular locus (see 'Advanced settings'). It is not included in the summary table, but shows up in the text given for every locus, just above the SVG plots. This also tells you how far all values had to be shifted in order to agree with your specified fixpoint.

- **Rounded to:** In case of dinucleotide repeats, this tells you whether tandem rounded to even or odd numbers. Similar for tri-, tetra-, etc- repeats.

- **Error:** The average rounding error over all allele sizes of this locus. The error is always between 0 and the repeat size divided by two, and is highlighted in red, if it is above 0.1 times the repeat size. Thus, for dinucleotide repeats it is between 0 and 1 and highlighted when above 0.2. High error values indicate either problems with your data, or problems with the rounding method (which, again, is more likely when your data is bad). Check this value in any case, and consider it for publication!

- **Error outliers:** If you mark this checkbox, the 30 worst alleles are highlighted by asterisks in both the full range and detail plots, and listed in a table below both plots. Only alleles with rounding errors greater than 0.5 are given. Move the mouse cursor the a particular asterisk, and the name and row of the respective individual will appear in a tooltip. The table further includes information on original and transformed allele size, bin center, and rounding error. JavaScript must be enabled for this feature.

Advanced settings

This section explains possible arguments when using the source code version on the command line.

- **ruby tandem.rb -v**
  tandem tells you it's version number and exits.

- **ruby tandem.rb -verbose -i <input_file_name>**
  tandem runs as normal, but gives you detailed information about the progress of its exhaustive search.

- **ruby tandem.rb -linear -i <input_file_name>**
  tandem uses a linear function instead of the power function described above. This means that parameter c is fixed to 1.000. With this setting, the exhaustive search is roughly seven

times faster that with the power function, but results may not be as good. If you have a very large data set, it may be a good idea to use this function. You will be able to tell from the SVG plots whether tandem found a good solution or not. If it hasn't, you should run tandem again with standard settings.

This argument can be combined with -verbose.

- **ruby tandem.rb -neldermead -i <input_file_name>**
  tandem runs a heuristic search using the Nelder-Mead Downhill Simplex method to optimize parameters a, b, and c of the above power function. This search is much faster than the exhaustive search, however, the method is prone to getting trapped in local optima. If you have a very large data set and a small number of alleles, you may consider using this setting. However, you should carefulley inspect the SVG plots and rerun tandem with standard settings in case it couldn't find a good solution.
  This argument can be combined with -verbose.

- **ruby tandem.rb -i <input_file_name> -f <fixpoint_file_name>**
  If you know the actual absolute allele size of a specimen (*e.g.* because you've sequenced the locus for this specimen), you can specify this allele size in a seperate file, and tandem will adjust its transformation of allele sizes so that the specimen's observed allele size is rounded to the specified actual allele size. The file should be in the same folder as the input file, tandem.rb, and the folder tandem_resources, and it should look like shown in fig. 7. Per line, tandem expects (i) the name of the locus, exactly as stated in the input file, (ii) the observed allele size of the respective specimen, and (iii) the actual allele size as determined by sequencing. As in the input file, values should be separated by tabs (if you prepare the fixpoint_file in MS Excel, save as tab-delimited file). You can specify fixpoints for several loci in the same file, but you can only specify one fixpoint per locus.
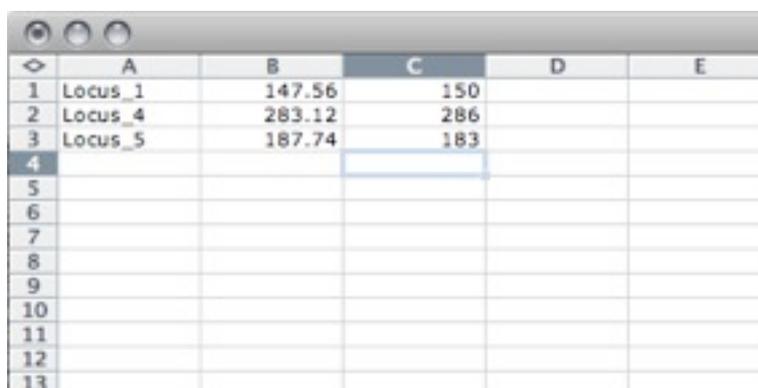
◇	A	B	C	D	E
1	Locus_1	147.56	150		
2	Locus_4	283.12	286		
3	Locus_5	187.74	183		
4					
5					
6					
7					
8					
9					
10					
11					
12					
13					

Fig. 7: *Fixpoint file with locus names in column A, observed allele sizes in column B, and actual allele size in column C. The file should be saved in tab-delimited format, in the folder of the input file.*

How to cite tandem

Matschiner M, Salzburger W (2009) TANDEM: integrating automated allele binning into genetics and genomics workflows. *Bioinformatics*, **25**(15), 1982-1983.

References

Amos W, Hoffman JI, Frodsham A *et al.* (2007) Automated binning of microsatellite alleles: problems and solutions. *Molecular Ecology Resources*, **7**, 10-14.

Beerli P (2006) Comparison of Bayesian and maximum-likelihood inference of population genetic parameters. *Bioinformatics*, **22**, 341-345.

Coombs JA, Letcher BH, Nislow KH (2007) Create 1.0 - Software to create and convert codominant molecular data. Available at http://lsc.usgs.gov/CAFL/Ecology/Ecology/html.

Dieringer D, Schlötterer, C (2003) Microsatellite Analyser (MSA): a platform independent analysis tool for large microsatellite data sets. *Molecular Ecology Notes*, **3**, 167-169.

Excoffier L, Laval G, Schneider S (2005) Arlequin ver. 3.0: an integrated software for population genetics data analysis. *Evolutionary Bioinformatics Online*, **1**, 47-50.

Falush D, Stephens M, Pritchard J (2003) Inference of population genetic structure using multilocus genotype data: linked loci and correlated allele frequencies. *Genetics*, **164**, 1567-1587.

Glaubitz, JC (2004) Convert: A user-friendly program to reformat diploid genotypic data for commonly used population genetic software packages. *Molecular Ecology Notes*, **4**, 309-310.

Hey J, Nielsen R (2004) Multilocus methods for estimating population sizes, migration rates and divergence time, with applications to the divergence of *Drosophila pseudoobscura* and *D. persimilis*. *Genetics*, **167**, 747-760.

Idury RM, Cardon LR (1997) A simple method for automated allele binning in microsatellite markers. *Genome Research*, **7**, 1104-1109.

Wenz HM, Robertson JM, Menchen S *et al.* (1998) High-precision genotyping by denaturing capillary electrophoresis. *Genome Research*, **8**, 69-80.

III   Phylogenetics

The rapid diversification of notothenioid fishes has been the subject of a large number of phylogenetic analyses. Based on morphological characters, Eastman (1993) considered the notothenioid family Bovichtidae (then including *Pseudaphritis*) to be ancestral to a clade combining Nototheniidae (then including *Eleginops*) and the four strictly Antarctic families Harpagiferidae, Artedidraconidae, Bathydraconidae, and Channichthyidae.

With the advent of PCR technology, notothenioids were among the first adaptive radiations to be investigated with molecular phylogenies (Bargelloni et al. 1994), which confirmed their assumed overall family-level relationships, and suggested that antifreeze glycoproteins evolved at a single point prior to the diversification of Antarctic notothenioids. The most ancestral lineage, Bovichtidae, was found paraphyletic with *Pseudaphritis* in a sister group position to all remaining notothenioid families (Ritchie et al. 1996), and *Eleginops* was recovered as ancestral to all other nototheniids and the strictly Antarctic families (Bargelloni et al. 2000). Thus, the monotypic families Pseudaphritidae and Eleginopidae were erected (Eastman, 2000), which had earlier been proposed on the basis of morphological data (Balushkin 1992, 2000). The taxonomic positions of Bovichtidae, Pseudaphritidae, and Eleginopidae have been corroborated by subsequent molecular analyses (Near et al 2004, Near & Cheng 2008), however, there is still disagreement regarding   the mono- or paraphyly of the families Nototheniidae and Bathy-draconidae (Bargelloni et al. 1997, Near et al. 2004, 2006, Near & Cheng 2008). Similarly, the sister group of notothenoids remains uncertain, but may be found among a group combining zoarcids, serranids, percids, scorpaenoids, trachinids and gasterosteoids (Dettaï & Lecointre 2004, 2005, Smith & Craig 2007).

Antifreeze glycoproteins have repeatedly been considered a key innovation for the adaptive radiation of the Antarctic clade (Bargelloni et al. 1994, 2000, Eastman 2000, 2005). It is assumed that antifreeze glycoproteins allowed survival of notothenioids at a time when the Antarctic water temperatures dropped below zero, which would have caused the extinction of a great part of the Antarctic ichthyofauna. Notothenioids could have taken over vacant resources, and proliferate as a response to these environmental changes (Near 2004, Eastman 2005). However, the timing of notothenioid diversification has long been unclear, and is notoriously difficult to estimate due to the paucity of the Antarctic fossil record. Estimates for the onset of the radiation of the Antarctic clade range between 7-15 and 24 Ma (Bargelloni et al. 1994, Near 2004) and are based on assumed molecular clock rates, or on a single questionable notothenioid fossil, †*Proeleginops grandeastmanorum* from the La Meseta Formation, Seymour Island. This fossil was originally described as a gadiform (Eastman & Grande 1991), but has later been considered to be a stem representative of the eleginopid lineage (Balushkin 1994). Thus, the time line of notothenioid diversification based on this fossil calibration remains questionable (Near 2004). However, a reliable estimate for the divergence of the Antarctic clade and the onset of the radiation would be important for a comparison with geological and climatic events of the Antarctic history, and to evaluate the role of antifreeze glycoproteins as a potential key innovation.

This question was addressed as part of this doctoral work. A large multi-marker phylogeny of acanthomorph fishes was established, including representatives of all notothenioid families, putative sister groups, as well as outgroups with rich fossil records. Using a relaxed molecular clock and 10 fossil and phylogeographic calibration points, the origin of the notothenioid Antarctic clade was identified near the Oligocene-Miocene boundary, at a time of cooling of Antarctic waters, which is consistent with the key innovation hypothesis for antifreeze glycoproteins, and highlights notothenioids as a prime model adaptive radiation. The dataset was subsequently extended to include over 100 species and fossil constraints for almost all divergence events, and a new model for the placement of prior distributions for fossil constraints was developed on the basis of probabilities of lineage nonpreservation. A reassessment of notothenoid divergence dates led to very similar results, further corroborating the key innovation hypothesis. In addition, the work that Rutschmann performed for her master thesis under my supervision extended the same data set to a population level phylogeny of notothenioid species, allowing a more focussed analysis of notothenioid divergence dates. Using isotopic signatures as proxies for trophic niche and habitat, this work further demonstrated parallel niche evolution in independent notothenioid clades, which is characteristic for adaptive radiation.

Three phylogenetic articles resulted from this doctoral work, of which two have been published in *PLoS ONE* and *Molecular Ecology*, and a third manuscript has been submitted to *Systematic Biology*:

3.1 Matschiner M, Hanel R, Salzburger W (2011) On the origin and trigger of the notothenioid adaptive radiation. *PLoS ONE* 6: e18911.

3.2 Matschiner M: Bayesian divergence priors based on probabilities of lineage non-preservation. Submitted to *Systematic Biology*.

3.3 Rutschmann S, Matschiner M, Damerau M, Muschick M, Lehmann MF, Hanel R, Salzburger W (2011) Parallel ecological diversification in Antarctic notothenioid fishes as evidence for adaptive radiation. *Molecular Ecology* 20: 4707-4721.

References

Balushkin AV (1992) Classification, phylogenetic, and origins of the families of the suborder Notothenioidei (Perciformes). *J Ichthyol* 32: 90–110.

Balushkin AV (1994) Fossil notothenioid, and not gadiform, fish *Proeleginops grandeastmanorum* gen. sp. nov. (Perciformes, Notothenioidei, Eleginopidae) from the late Eocene found in Seymour Island (Antarctica). *Voprosy Ikhtiologii* 34: 298–307.

Balushkin AV (2000) Morphology, classification, and evolution of Notothenioid fishes of the Southern Ocean (Notothenioidei, Perciformes). *J Ichthyol* 40: S74–S109.

Bargelloni L, Ritchie PA, Patarnello T, et al. (1994) Molecular evolution at subzero temperatures: mitochondrial and nuclear phylogenies of fishes from Antarctica (suborder Notothenioidei), and the evolution of antifreeze glycopeptides. *Mol Biol Evol* 11: 854–863.

Bargelloni L, Patarnello T, Ritchie PA, Battaglia B & Meyer A (1997) Molecular phylogeny and evolution of Notothenioid fish based on partial sequences of 12S and 16S ribosomal RNA mitochondrial genes. In: *Antarctic communities: species, structure, and survival* (Eds. Battaglia B, Valencia J, Walton DWH). Cambridge University Press., New York.

Bargelloni L, Marcato S, Zane L & Patarnello T (2000) Mitochondrial phylogeny of notothenioids: a molecular approach to Antarctic fish evolution and biogeography. *Syst Biol* 49: 114–129.

Dettaï A, Lecointre G (2004) In search of notothenioid (Teleostei) relatives. *Antarct Sci* 16: 71–85.

Dettaï A, Lecointre G (2005) Further support for the clades obtained by multiple molecular phylogenies in the acanthomorph bush. *C R Biol* 328: 674–689.

Eastman JT (1993) Antarctic Fish Biology: Evolution in a Unique Environment. Academic Press, Inc., San Diego, CA.

Eastman JT (2000) Antarctic notothenioid fishes as subjects for research in evolutionary biology. *Antarct Sci* 12: 276–287.

Eastman JT (2005) The nature of the diversity of Antarctic fishes. *Polar Biol* 28: 93–107.

Eastman JT & Grande L (1991) Late Eocene gadiform (Teleostei) skull from Seymour Island, Antarctic Peninsula. *Antarct Sci* 3: 87–95.

Near TJ (2004) Estimating divergence times of notothenioid fishes using a fossil-calibrated molecular clock. *Antarct Sci* 16: 37–44.

Near TJ, Cheng C-HC (2008) Phylogenetics of notothenioid fishes (Teleostei: Acanthomorpha): Inferences from mitochondrial and nuclear gene sequences. *Mol Phylogenet Evol* 47: 1–9.

Near TJ, Pesavento JJ, Cheng C-HC (2004) Phylogenetic investigations of Antarctic notothenioid fishes (Perciformes: Notothenioidei) using complete gene sequences of the mitochondrial encoded 16S rRNA. *Mol Phylogenet Evol* 32: 881–891.

Near TJ, Parker SK, Detrich HW III (2006) A genomic fossil reveals key steps in hemoglobin loss by the Antarctic icefishes. *Mol Biol Evol* 23: 2008–2016.

Ritchie PA, Bargelloni L, Meyer A, et al. (1996) Mitochondrial phylogeny of trematomid fishes (Nototheniidae, Perciformes) and the evolution of Antarctic fish. *Mol Phylogenet Evol* 5: 383–390.

Smith WL, Craig MT (2007) Casting the percomorph net widely: the importance of broad taxonomic sampling in the search for the placement of serranid and percid fishes. *Copeia* 2007: 35–55.

3.1 On the origin and trigger of the notothenioid adaptive radiation

Matschiner M, Hanel R, Salzburger W

PLoS one

On the Origin and Trigger of the Notothenioid Adaptive Radiation

Michael Matschiner[1], Reinhold Hanel[2], Walter Salzburger[1]*

1 Zoological Institute, University of Basel, Basel, Switzerland, 2 Institute of Fisheries Ecology, Johann Heinrich von Thünen-Institute, Federal Research Institute for Rural Areas, Forestry and Fisheries, Hamburg, Germany

Abstract

Adaptive radiation is usually triggered by ecological opportunity, arising through (*i*) the colonization of a new habitat by its progenitor; (*ii*) the extinction of competitors; or (*iii*) the emergence of an evolutionary key innovation in the ancestral lineage. Support for the key innovation hypothesis is scarce, however, even in textbook examples of adaptive radiation. Antifreeze glycoproteins (AFGPs) have been proposed as putative key innovation for the adaptive radiation of notothenioid fishes in the ice-cold waters of Antarctica. A crucial prerequisite for this assumption is the concurrence of the notothenioid radiation with the onset of Antarctic sea ice conditions. Here, we use a fossil-calibrated multi-marker phylogeny of nothothenioid and related acanthomorph fishes to date AFGP emergence and the notothenioid radiation. All time-constraints are cross-validated to assess their reliability resulting in six powerful calibration points. We find that the notothenioid radiation began near the Oligocene-Miocene transition, which coincides with the increasing presence of Antarctic sea ice. Divergence dates of notothenioids are thus consistent with the key innovation hypothesis of AFGP. Early notothenioid divergences are furthermore congruent with vicariant speciation and the breakup of Gondwana.

Introduction

Adaptive radiation - the evolution of ecological and phenotypic diversity within a rapidly multiplying lineage - has been implicated in the genesis of a great portion of the diversity of life [1,2]. According to Schluter [2], an adaptive radiation is characterized by rapid speciation, common ancestry, and a phenotype-environment correlation, whereby phenotypes must actually be beneficial in their respective environments. Adaptive radiation is often considered a consequence of ecological opportunity [1,2] arising through colonization of a new habitat with abundant niche-space, extinction of antagonists, and/or the origin of a key innovation [3]. All three settings induce the relaxation of selection pressure, which may promote diversification [3]. Key innovations can lead to ecological opportunity either by enabling the exploitation of new resources, or by boosting a clade's fitness relative to competing lineages. A third type of key innovation does not generate ecological opportunity, but directly enhances diversification rates by increasing the potential for reproductive isolation or ecological specialization, e.g. by decreasing dispersal distance and gene flow [4]. Corroboration of the key innovation hypothesis would, hence, involve the identification of (ecological) mechanisms linking a putative key innovation to increased speciation or decreased extinction rates, and comparative tests correlating it with inflating diversity [4].

So far, such tests have been applied to few key innovations only, and even the best examples of animal adaptive radiations provide only scanty evidence in support of the key innovation hypothesis.

Two of the most prominent examples of adaptive radiation, the Galapagos finches and Hawaiian honeycreepers, were, in fact, more likely triggered by the arrival of the ancestral species on competitor-free islands rather than by key innovations [5]. A number of key innovations have been proposed for the radiations of cichlid fishes in the Great Lakes of East Africa, including a highly variable pharyngeal jaw apparatus, egg-spots, and maternal mouth brooding behaviour [6,7]. However, based on a comparative analysis of successful and failed cichlid radiations, the role of all three traits as key innovations has been questioned [8]. Similarly, the acquisition of pharyngeal jaws provides a weak explanation for increased diversification rates in the radiation of labrid fishes [9]. A key innovation in Caribbean *Anolis* lizards, on the other hand, appears to pass both the ecological mechanism and comparative test: extended subdigital toe-pads enable *Anolis* to climb narrow twigs, leaves and grass blades. The resultant arborality distinguishes them from other iguanids. Toepads evolved at the base of the anole phylogeny, and also occur in the second-most species rich family of lizards, the Gekkonidae, thus linking its emergence with species richness [10].

Another vertebrate adaptive radiation, which has drawn increasing interest in recent years, has occurred on the isolated shelf areas surrounding the Antarctic continent in the perciform fish suborder Notothenioidei. A total of 132 notothenioid species are known to date, and new species are discovered at fast rates [11]. Nine species belong to three early diverging families (Bovichtidae, Pseudaphritidae, Eleginopidae) that occur almost exclusively outside Antarctic waters and are not usually considered

136

part of the radiation. The remaining 123 species in five families are often referred to as the "Antarctic Clade", which dominates the High Antarctic ichthyofauna in terms of species number (76.6%) and biomass (>90%) [11]. Notothenioids of the Antarctic Clade possess a wide range of adaptations to the extreme Antarctic environment, including antifreeze glycoproteins (AFGP) [12], retinal reorganization [13], and loss of heat shock response [14]. One of the five Antarctic notothenioid families (Channichthyidae) even lives without hemoglobin, which is unique among vertebrates [15]. Despite the loss of the swim bladder in their presumably benthos-dwelling ancestor, multiple notothenioid lineages have independently recolonized pelagic, semi-pelagic and cryopelagic habitats. Subsequent adaptations in ossification, scale mineralization, and lipid deposition led to a partial or full reacquisition of neutral buoyancy and a phenotype-environment correlation that is characteristic for adaptive radiations [2,11]. Another important phenotype-environment correlation exists between freezing avoidance and water temperature [16].

Antifreeze glycoproteins are present in almost all notothenioids of the Antarctic Clade, enabling them to cope with the subzero temperatures of Antarctic waters [17]. The widespread possession of AFGPs in the monophyletic Antarctic Clade, complete lack of AFGPs in non-Antarctic sister groups, and their highly conserved chemical structure [15] suggest that AFGPs evolved only once in notothenioids and that this occurred prior to the onset of diversification in the Antarctic Clade [12,17]. Therefore, it has been hypothesized that AFGPs represent a key innovation that allowed notothenioids to radiate at a time when Antarctic water temperatures dropped below zero, which presumably led to the extinction of a great part of the previous Antarctic shelf ichthyofauna [15]. Following Heard and Hauser [4] AFGPs would constitute a type I key innovation if the resulting fitness advantage enabled notothenioids to replace other clades, a type II key innovation if AFGPs allowed the invasion of previously unoccupied sea ice-associated habitats, or a combination of both. A crucial prerequisite for either hypothesis is the concurrence of the beginning of the notothenioid radiation and the onset of Antarctic sea ice conditions.

Cenozoic Antarctic water temperatures and the emergence of sea-ice in Antarctica can be inferred from deep sea isotope records and sediment analysis of drill cores [18–21]. The timing of the notothenioid radiation, on the other hand, is far less certain, which is in part due to the paucity of fossils in Antarctica. Existing molecular clock calibrations for notothenioids are based on few mitochondrial markers in combination with a single putative, but debated, eleginopid fossil [22], biogeographic patterns [23], or the presumed date of the perciform diversification [24]. Consequently, attempts to date the beginning of the notothenioid radiation have led to a wide range of contradicting results between 7 and 24 Ma [22,25].

Here we use a multi-marker (4599 bp, 6.53% missing data) phylogeny including representatives of all notothenioid families plus 69 non-notothenioid fishes with ten fossil and phylogeographic constraints to time-calibrate notothenioid divergences and AFGP evolution.

Results

Tree Topology

Partitioned Maximum Likelihood and Bayesian phylogenetic analyses of 83 acanthomorph taxa using GARLI-PART, RAxML, and BEAST (Fig. 1, Fig. S2) resulted in identical topologies with the exception of the position of *Antigonia capros*, which was placed as sister group to Lophiiformes and Tetraodontiformes in RAxML's

optimal tree. The phylogenetic placement of zeioids within Paracanthopterygii [26] and scarids within Labridae [27,28] was confirmed in all analyses. Gasterosteiformes and Zoarcidae appeared within Scorpaeniformes, thus rendering this order paraphyletic (albeit with low support values). Notothenioids were recovered as a sister group of a clade containing percids, trachinoids, and *Serranus atricauda*. The highly supported placement of the latter (Bayesian Posterior Probability (BPP) 1.0; Fig. S2, Table S1) is in accordance with previous phylogenetic hypotheses and suggests polyphyly of serranids, with representatives of the family in close phylogenetic affiliation with notothenioids, percids, and trachinoids [29,30]. Removal of *S. atricauda* from the data set affected the tree topology only in the weakly supported position of *Antigonia capros* (Table S1). The notothenioids were covered by 14 (phylogenetically) representative species. While this number might appear small in the context of notothenioid phylogenetics and divergence rate estimates (which was, notably, not the purpose of this study), it is absolutely balanced with respect to the timing of their radiation and their relative coverage in the total dataset. Our trees confirm the divergence of Bovichtidae prior to Pseudaphritidae, the monophyly of the Antarctic Clade, the paraphyly of the family Nototheniidae within the Antarctic Clade, and the interrelationships of derived notothenioid families (Fig. S2) [22,31].

Cross-Validation of Time Constraints

We first cross-validated the available 10 calibration points in order to test for their relative consistency. This step is important, as calibrations based on fossil and geological data show various degrees of uncertainty [32]. When estimated on the basis of all other constraints, five out of ten divergence dates were concordant with the respective fossil age assignments (Fig. 2). The split between gempylids and scombrids (node D; 79.2–18.4 Ma, 95% highest probability density (HPD)) seems to postdate respective fossil findings and suggest taxonomic or stratigraphic misinterpretations. The mean age estimate for the polymixiid lineage (99.2 Ma) falls into the Cenomanian, as does the oldest polymixiid fossil. Nevertheless, this constraint was excluded from further analyses, as nearly half of its HPD (133.1–70.5 Ma) postdates the Cenomanian. Age estimates for both cichlid and labrid divergences failed to match phylogeographic calibrations, which were therefore excluded. Estimated on the basis of nine constraints, the diversification of Percomorpha (203.3–135.0 Ma) seems to predate the earliest euteleost fossils (150.9 Ma) [33]. However, after exclusion of constraints A, C, D, E, and F, re-estimation resulted in younger percomorph divergence estimates (150.9–114.7 Ma), being congruent with the euteleost fossil record.

Notothenioid Divergence Dates

According to cross-validation results, we estimated divergence dates of notothenioids on the basis of six fossil constraints (nodes B, C, G, H, I, and J; Fig. 3, Table S2). Our results support a late Cretaceous origin of Bovichtidae (node U, mean 71.4 Ma, 95% HPD 89.4–54.4 Ma), early Paleocene divergence of Pseudaphritidae (node V, 63.0 Ma, 80.2–46.7 Ma), and a Mid-Eocene split between Eleginopidae and the Antarctic Clade (node W, 42.9 Ma, 56.9–29.8 Ma). Mid-Eocene origin of the eleginopid lineage is congruent with the age of the only fossil putatively assigned to Notothenioidei. *Proeleginops grandeastmanorum* from the La Meseta Formation on Seymour Island (dated to ~40 Ma) was first described as a gadiform fossil [34], and subsequently reinterpreted as an eleginopid [35]. While previous attempts to date the notothenioid radiation used this fossil as a single calibration point, our analysis deliberately excluded this constraint due to its debated
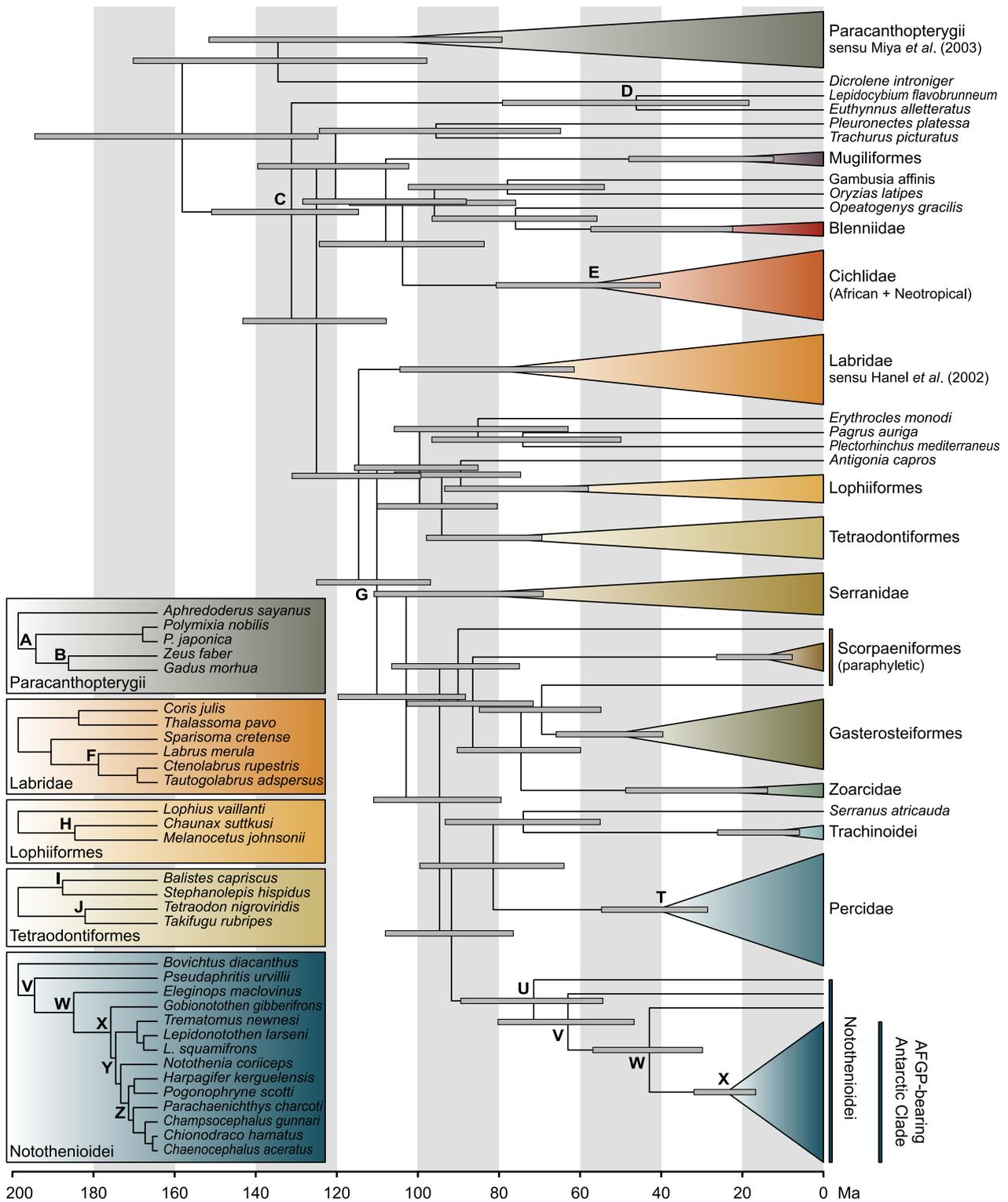
Figure 1. Time-calibrated phylogeny of acanthomorph fishes based on the concatenated dataset of two mitochondrial and four nuclear genes, and six calibration points (B, C, G, H, I, and J). All nodes used for constraint cross-validation are labelled with letters A–J, percid and notothenioid nodes are labelled with letters V–Z. Insets indicate nodes labels within Paracanthopterygii, Labridae, Lophiiformes, and Tetraodontiformes. Node bars show 95% HPD.
doi:10.1371/journal.pone.0018911.g001

138
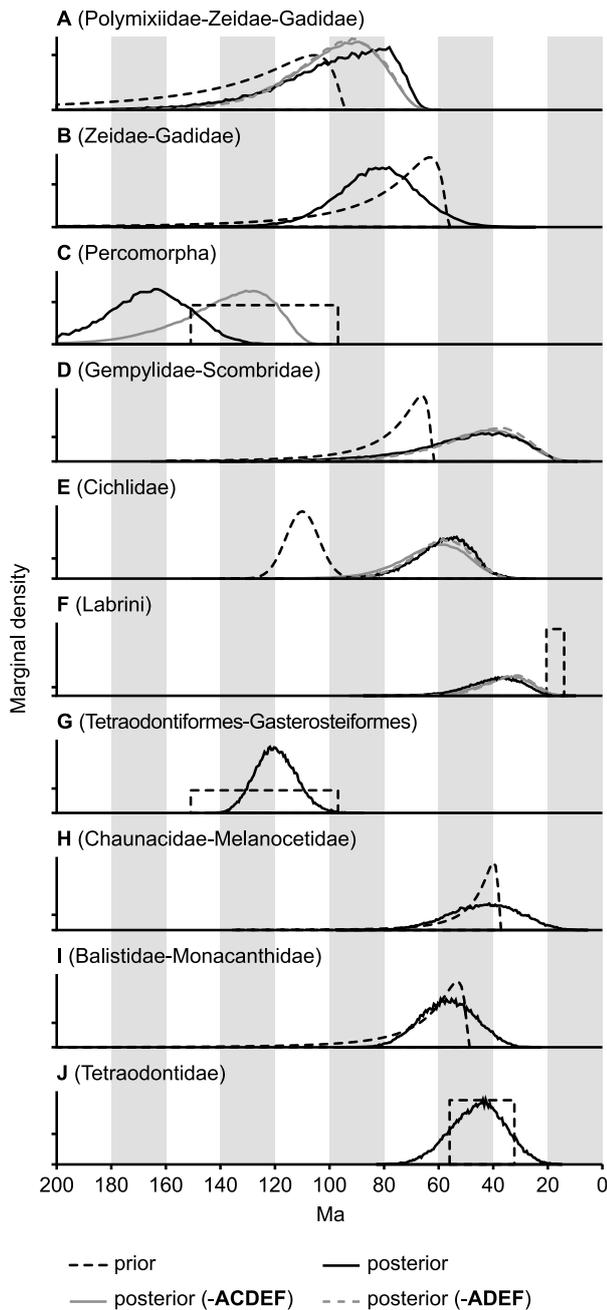
Figure 2. Cross validation of all fossil and phylogeographic constraints. The dashed black line indicates the prior, as specified in all BEAST runs including this constraint. The solid black line shows the marginal densities of BPPs for each node when its constraint is relaxed and its date estimated based on all other constraints. At this stage, only five out of ten (B, G, H, I, J) nodes showed a good fit between prior and posterior. A new BEAST analysis ('-ACDEF') was conducted, using only constraints B, G, H, I, and J. Results of this analysis (solid grey line) showed a good fit between prior and posterior for node C, therefore node C was reincluded in a final BEAST run ('-ADEF'; dashed grey line).
doi:10.1371/journal.pone.0018911.g002

taxonomic assignment [22]. Nevertheless, the concordance between divergence date estimates and fossil age corroborates the eleginopid interpretation of *P. grandeastmanorum*. Radiation of the Antarctic Clade began near the Oligocene-Miocene transition

with the divergence of *Gobionotothen* (node X, 23.9 Ma, 31.9–16.7 Ma) and was quickly followed by further diversification within the Antarctic Clade (node Y, 21.4 Ma, 28.2–15.3 Ma). Correspondingly, support values for early notothenioid divergences are low (node Y, Table S1), indicating rapid succession of speciation events. Excluding *G. gibberifrons* from the data set indicates that regardless of the exact topology of the Antarctic Clade, the radiation was underway in the early Miocene (23.0 Ma, 30.5–16.1 Ma). This is in agreement with previous age estimates (24.1±0.5 Ma) on the basis of the putative eleginopid fossil *P. grandeastmanorum* and a penalized likelihood molecular analysis approach [22]. Diversification of the four most derived notothenioid families Harpagiferidae, Artedidraconidae, Bathydraconidae, and Channichthyidae apparently began in the mid-Miocene (node Z, 14.7 Ma, 20.0–9.9 Ma). Close agreement of percid divergence dates (node T, 28.54–54.71 Ma) with biogeographical scenarios suggests that clades closely related to notothenioids were dated reliably (Text S1). All acanthomorph divergence date estimates are summarized in Table S2. Notothenioid divergence date estimates are relatively robust to the set of constraints used for our cross-validation (Table S3).

Discussion

Antifreeze Glycoproteins are a Key Innovation

Bayesian Inference of acanthomorph divergence dates shows that the adaptive radiation of Antarctic notothenioids, resulting in more than 120 morphologically highly diverse species that dominate Antarctic waters, began near the Oligocene-Miocene transition (mean 23.9 Ma, 95% HPD 31.9–16.7 Ma). While large-scale continental glaciation may not have been permanent before the middle Miocene climate transition (~14 Ma) [36,37], geological evidence supports temporal presence of Antarctic sea ice already 24 Ma: Deep-sea oxygen (δ^{18}O) isotopes (Fig. 3) provide a reliable record of relative temperature changes and demonstrate an overall cooling trend (~14°C) since the early Eocene [19]. Similarly, isotope levels of sedimentary alkenones reveal partial pressures of paleoatmospheric carbon dioxide (pCO$_2$; Fig. 3) and show a decrease from the middle to late Eocene that led to rapid expansions of large continental Antarctic ice sheets and widespread ice rafting as early as 34 Ma [18,38,39]. Numerical climate models with explicit, dynamical representations of sea ice show that moderate or full cenozoic glaciation of East Antarctica would have promoted extensive sea ice formation at least in cold austral summer orbits with low pCO$_2$ levels (560 ppmv) [40]. Direct evidence for continental glaciation and marine ice comes from cyclic glacimarine deposits in offshore drill cores, showing that glacial extensions well onto the continental shelf occurred repeatedly since the early Oligocene (Fig. 3) [20]. Furthermore, the long-term presence of local sea ice is suggested by findings of sea ice-dependent diatoms in lower Oligocene sediments [21]. Taken together, it is likely that Antarctic sea ice has existed with seasonal, orbital, and local constraints since the early Oligocene. Estimates for the onset of the Antarctic Circumpolar Current range widely [41], but its increasing strength presumably contributed to thermal isolation and cooling of Antarctic waters by up to 4°C during the Oligocene and Miocene [42]. Deep sea oxygen isotope records and glacimarine sediments further indicate a major period of global cooling and ice sheet expansion at the Oligocene-Miocene transition (Mi-1 event, 24.1–23.7 Ma, Fig. 3) [43]. This exactly coincides with our mean age estimate for the onset of the Antarctic notothenioid radiation (23.9 Ma), which is characterized by the presence of AFGPs. Based on the highly conserved chemical structure of AFGPs in
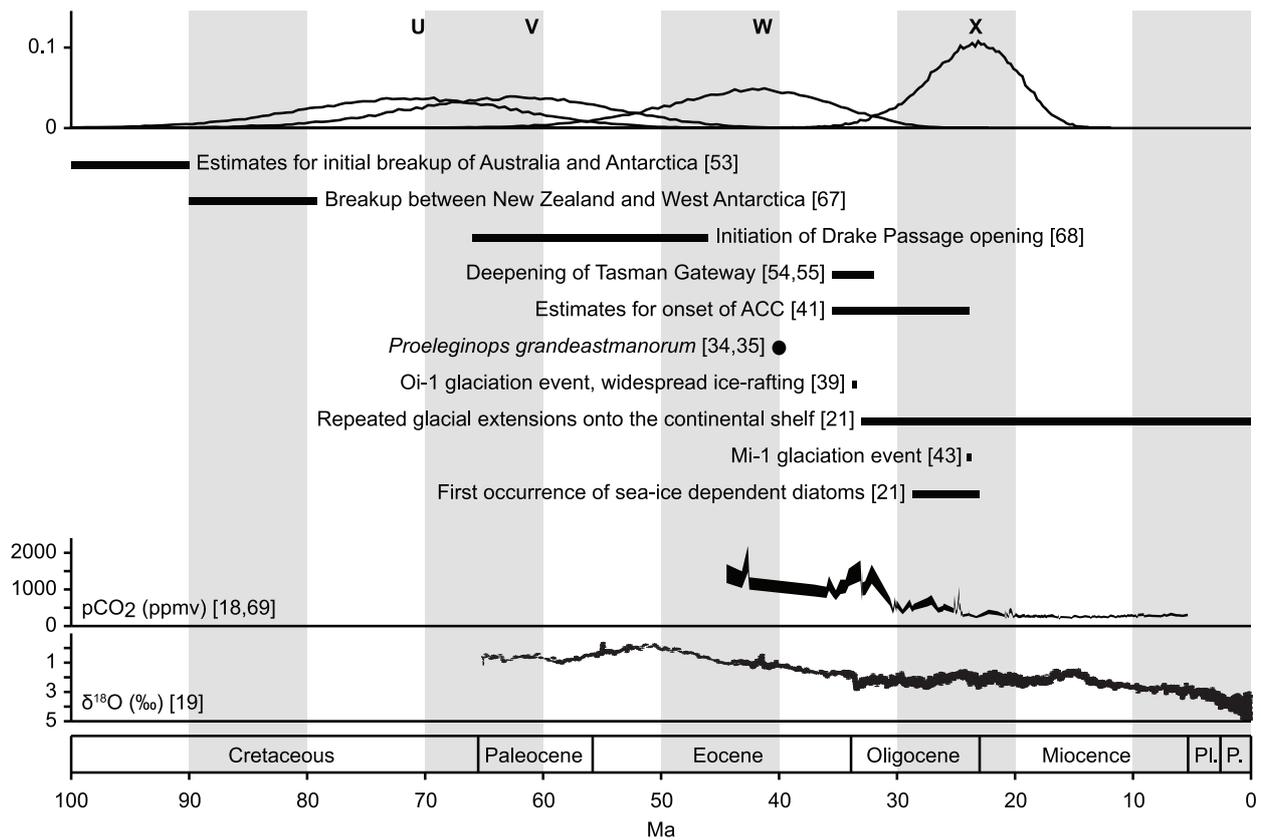
Figure 3. Comparison of notothenioid divergence dates and geological events. Phylogenetic nodes are labelled with letters U-X according to Fig. 1. The diversification of the Antarctic Clade (node X) coincides with an increase in frequency of glacial extensions well onto the shelf, the Mi-1 glacial event, the first occurences of sea-ice dependent diatoms in Antarctic waters, and a sharp decline of atmospheric CO₂ levels. Paleoc.:Paleocene, Oligoc.:Oligocene, Pl.:Pliocene, P.:Pleistocene.
doi:10.1371/journal.pone.0018911.g003

nearly all notothenioids of the Antarctic Clade [15], it is commonly assumed that AFGPs evolved only once, before the notothenioid radiation [12,17]. Members of the genus *Patagonotothen* apparently lack AFGPs [17], which is – however – likely due to secondary loss (as *Patagonotothen* is deeply nested within AFGP-bearing nototheniids, it would otherwise require at least 6 independent origins of notothenioid AFGPs [15,31]). Hence, the key innovation hypothesis of AFGP is consistent with our age estimate for the notothenioid radiation. Freezing avoidance could have allowed notothenioids to invade new, ice-associated niches, or to replace other clades subsequent to their extinction in a cooling environment. Without doubt, selection pressures must have been substantial, given that freezing avoidance is a matter of life and death in ice-laden habitats [44]. Emergence of AFGPs could have proceeded in a step-wise manner that began with accidental replication slippage in an intron of the ancestral trypsinogen gene [12]. Subsequent duplications of Thr-Ala-Ala tripeptides could have endowed some measures of freezing avoidance without immediate loss of trypsin activity [44].

The key innovation hypothesis of AFGP would further be corroborated if similar diversity was found in other clades that independently acquired freezing avoidance [4]. Outside notothenioids, near-identical AFGPs have convergently evolved in Arctic cod *Boreogadus saida* [45] and other taxa of the subfamily Gadinae [46,47]. Apparently, cod AFGPs share a common origin that dates back to the Miocene [48]. The subfamily consists of 23 species and

may thus be considered moderately species-rich. Type III antifreeze proteins (AFP) are found in zoarcids of both Antarctic and Arctic waters, and supposedly predate the bipolar distribution of zoarcids [47,49]. With over 200 species, the family Zoarcidae is indeed a highly diverse group [49], and surpasses even notothenioids in species richness. However, AFPs have been identified in comparatively few zoarcids to date [47], which may indicate secondary losses in many taxa. Whether AFP played a role in the zoarcid radiation remains to be elucidated. The distribution of type I AFPs over phylogenetically distant clupeids, osmerids, and cottids of the northern hemisphere provides a rare example of lateral gene flow in vertebrates [50] and indicates strong selection pressures. However, none of these AFP-bearing taxa have undergone substantial radiations. This could be due to external factors that mask the effect of AF(G)P emergence in non-notothenioid taxa. Most adaptive radiations occur in geographically confined areas, a potential prerequisite [51] that is satisfied for Antarctic continental shelves, but less so for Arctic habitats. In addition, dispersal of zoarcids to Antarctica in the Miocene [49], when the notothenioid diversification had already filled most available niche space, could have limited their radiation.

Phylogeography of Notothenioid Lineages

Estimates of early notothenioid divergence dates support most aspects of the phylogeographic scenario proposed by Balushkin [52]: The presence or even endemism of three out of four

bovichtid and pseudaphritid genera in Australia suggests occurrence of the presumably benthic notothenioid ancestor [15] on South Australian continental shelves in the late Cretaceous. Fragmentation of shelf areas between Australia and New Zealand ~70 Ma may have led to the separation of bovichtids from the pseudaphritid ancestor and to initial divergences within Bovichtidae. Extended pelagic larval durations could have contributed to long-ranged eastward dispersal of bovichtids with paleogene currents to South America and Tristan da Cunha [52]. The isolation of pseudaphritids in Southern Australia and divergence of the Antarctic lineage are presumably linked to the breakup of Australia and Antarctica. Separation between both landmasses started between 125 and 90 Ma [53], however, shallow water connections existed until ~33.5 Ma [54,55]. Vicariant speciation of benthic lineages could have occurred anytime between these dates, being supported by our age estimate of pseudaphritids (node V, 63.0 Ma, 80.3−46.7 Ma). Antarctic notothenioids then diversified into eleginopids and the ancestor of the Antarctic Clade in the Eocene (node W, 42.9 Ma, 56.9−29.8 Ma). Past presence of eleginopids in Antarctica is indicated by the fossil *P. grandeastmanorum*, presumably representing an early member of the lineage [35]. Finally, a drop in water temperatures and the increasing presence of sea ice in the Oligocene led to near-complete replacement of the Eocene Antarctic ichthyofauna [15], migration of eleginopids to South America, and adaptive radiation of the Antarctic Clade subsequent to AFGP emergence.

Materials and Methods

Phylogenetic Reconstruction

Four nuclear (myh6, Ptr, ENC1, tbr) and two mitochondrial (nd4, cyt *b*) markers were PCR amplified and sequenced for 14 notothenioid and 53 related acanthomorph fish species, and complemented with additional sequences from GenBank, Ensemble, and Genoscope to a total of 83 taxa (Text S2, Tables S4–S5). Phylogenetic analyses were conducted in GARLI-PART v0.97 [56] and RAxML v7.26 [57], and node support was assessed with nonparametric bootstraps. Detailed information on sample collection, marker selection, PCR amplification, and phylogenetic inferences are given in the Text S2, S3, S4, S5 and S6 and Fig. S1.

Time Constraints used for Dating

A total of eight fossil, and two phylogeographic constraints were chosen to time-calibrate acanthomorph divergences. Following Benton & Donoghue [58], we implemented time constraints for the origin of Tetraodontidae (node J; *Takifugu-Tetraodon* divergence; 56.0−32.25 Ma) and for the split between Gasterosteiformes and Tetraodontiformes (node G; 150.9−96.9 Ma). The minimum age for the Gasterosteiformes-Tetraodontiformes divergence is derived from the oldest known member of the tetraodontiform lineage, *Plectocretacicus clarae*, which also represents the oldest known percomorph. Also, the maximum constraint for divergence of gasterosteiform and tetraodontiform lineages is provided by the earliest euteleost record, *e.g. Tischlingerichthys viohli*. Therefore, we applied the same lower and upper bounds for the divergence of all Percomorpha (not including *Dicrolene introniger*, as the phylogenetic position of Ophidiiformes remains unclear [26]; node C). Uniform priors between minimum and maximum age of divergence were used for these constraints. In addition, we constrained five family divergences, which we expected to have relatively early fossil records in at least one of the descending lineages (Fig. S3). Thus, our analysis accounted for polymixiid fossils from the Cenomanian (node A; ≥93.5 Ma), a zeid fossil from the Thanetian (node B; ≥55.8 Ma), gempylid and scombrid

fossils from the Danian (node D; ≥61.7 Ma), a chaunacid fossil from the Bartonian (node H; ≥37.2 Ma), and a monacanthid fossil from the Ypresian (node I; ≥48.6 Ma). All used fossils are referenced in detail in Text S7. Lognormal priors were assigned to the above fossil constraints with hard lower bounds reflecting the age of the respective fossil, and soft upper bounds (Table S5, Fig. 2). Phylogeographic constraints for cichlid and labrid divergences were derived from the breakup of Gondwana, and from the closure of the connection between the Mediterranean and the Indian Ocean. We added the separation of Africa and South America as effective time constraint for the split between African and neotropical cichlids (node E), assuming vicariant divergence. Seafloor spreading in the South Atlantic started as early as 133 Myr ago [59] and a continuous North/South Atlantic Ocean presumably existed ~100 Myr ago [60], hence, we applied a normally distributed prior between 121.8 and 98.2 Ma (95% cumulative probability; mean: 110.0 Ma) to constrain cichlid divergence. The split between *Labrus* and both *Ctenolabrus* and *Tautogolabrus* (node F) represents the diversification of the labrid tribe Labrini [27]. Based on molecular clock calibrations and fossil evidence (Text S7) showing that Labrini were present in the Mediterranean 14.0 Ma, it has been suggested that the ancestor of Labrini migrated from the Indopacific into the Mediterranean prior to the closure of this seaway, 20.5−19.5 Ma [27]. Therefore, we constrained the diversification of Labrini with a uniform prior between 20.5 and 14.0 Ma.

Dating of Acanthomorph Divergences

In order to date notothenioid and non-notothenioid acanthomorph divergences, we generated time-calibrated phylogenies with BEAST v1.5.3 [61]. All BEAST runs were performed using mitochondrial and nuclear sequence alignments as separate partitions with unlinked substitution models. We employed a relaxed molecular clock model with branch rates drawn independently from a lognormal distribution [62], ten time constraints (Table S5), and the reconstructed birth-death process [63] as a tree prior (see Fig. S4 for substitution rates). The applicability of relaxed molecular clocks for cold-adapted organisms is discussed in Text S8. After optimization of operators according to preliminary run results, three different substitution models were implemented and evaluated. We included the codon position-based $HKY_{112}+CP_{112}+\Gamma_{112}$ model [64], in which all parameters are estimated independently for the first two and for the third codon positions. We also added the $GTR_{112}+CP_{112}+\Gamma_{112}$ model, using the same partitions. In a third set, we implemented $HKY+I+\Gamma$ for the first two mitochondrial codon positions, $TVM+\Gamma$ for the third mitochondrial codon position, $K80+I+\Gamma$ for the first nuclear codon position, and $GTR+\Gamma$ for the third nuclear codon position, as selected by BIC. For each of the three setting, we performed 20 independent analyses of 20 million generations each, discarding the first 2 million generations of every replicate as burnin. Replicate results were combined in LogCombiner v1.5.3 after removing the burnin. Convergence of run replicates was confirmed by effective sample sizes (ESS) >1200 for all parameters and by visual inspection of traces within and between replicates in Tracer v1.5. Substitution models were evaluated on the basis of Bayes Factors, again, as implemented in Tracer [65]. Bayes Factors provided 'very strong' [66] evidence that the substitution model combination selected by BIC was better-fitting than both the $HKY_{112}+CP_{112}+\Gamma_{112}$ (log 10 BF 350.6) and $GTR_{112}+CP_{112}+\Gamma_{112}$ (log 10 BF 276.0) models, and thus the BIC combination was used for all subsequent analyses. In order to assess the reliability of every individual time constraint, we conducted a cross-validation, whereby we relaxed constraints one

141

by one, and estimated divergence dates of relaxed constraints based on all other constraints (Fig. 2). BEAST runs were conducted as described above, but using 5 run replicates per cross-validation. We found good fit of posterior and prior distributions for constraints B, G, H, I, and J. Subsequently, 20 run replicates were performed with identical settings, but excluding the five unreliable constraints A, C, D, E, and F (run '-ACDEF' in Fig. 2). Posterior distributions of excluded constraints were again compared to their assumed prior distributions. After exclution of five constraints, node C (divergence of Percomorpha) provided adequate fit of posterior probability distribution to its suggested bounds [58], and was thus reincluded for yet another run with 20 independent replicates and unchanged settings (run '-ADEF' in Fig. 2 and Tables S1–S2). ESS values for this run were >900 for all parameters. As for GARLI-PART and RAxML analyses, the last run was repeated after removal of *Serranus atricauda* from the dataset ('-ADEF -*Serranus atricauda*' in Fig. 2 and Tables S1–S2), which had no impact on tree topology and little influence on node support (on average -0.19 BPP, Table S1) and divergence date estimates (average difference 0.57%, Table S2). All molecular data sets to date [22,23,31] failed to assign a reliable phylogenetic position to *G. gibberifrons* (node Y, BPP 0.64), thus indicating rapid divergence at the beginning of the notothenioid radiation. We also repeated the analysis without *G. gibberifrons* (8 replicates, unchanged settings) to obtain a minimum age estimate for the diversification of the Antarctic Clade that is robust to topological uncertainties. Maximum clade credibility trees were produced using TreeAnnotator v1.5.3.

Supporting Information

Figure S1 Partitioned ML consensus tree of full mitochondrial genomes (A), and best ML phylogenies of single mitochondrial markers (B-D), estimated with RAxML. Sequence data were taken from [3,4,6]. Markers ND4 (B) and cyt *b* (C) phylogenies show better agreement with the full mitogenome topology than other markers of comparable sequence length (D).
(EPS)

Figure S2 Partitioned ML phylogeny of 83 acanthomorph fishes, based on the concatenated dataset of two mitochondrial (ND4, cyt *b*) and four nuclear genes (myh6, Ptr, ENC1, tbr). Tree topology and branch lengths are as estimated using GARLI-PART, and near-identical topologies were recovered with RAxML and BEAST analyses. Filled circles indicate > 98% BS support (as calculated with GARLI-PART) and > 0.99 BPP (according to BEAST run '-ADEF'), white circles represent nodes with BS support > 80% and >0.90 BPP. Split circles indicate different levels of BS (left half) and BPP (right half) support. All node support values are summarized in Table S3. Nodes that were used for fossil and phylogeographic constraints are labelled with letters A–J, basal percid and notothenioid nodes are labelled with letters T–Z.
(EPS)

Figure S3 Partitioned BI phylogeny, based on the concatenated data set, reduced to family level. Node heights correspond to mean age estimates. The time scale is divided into phanerozoic stages (grey shades), and the presence of skeletal (black bars) or otolith fossils (dark grey bars) in a stage is plotted on top of family branches. Unless otherwise noted, all fossil information is taken from [22]. Fossils with questionable taxonomic or stratigraphic assignments are indicated by dashed bars. Numbers in brackets indicate the number of species included in this study and the total number of species per family [31]. [1] Fossil used to constrain node

D, however cross-validation results suggest unreliability of this constraint. As Gempylidae could be nested within Scombridae [25], fossils may have been misinterpreted. [2] [32,33] [3] According to Santini & Tyler [34], the oldest tetraodontid is *Archaeotetraodon winterbottomi* from the Rupelian. *Eotetraodon pygmaeus*, previously assigned to Tetraodontidae [22], has been moved to the family Triodontidae [4] [35,36].
(EPS)

Figure S4 Partitioned BI phylogeny of 83 acanthomorph fishes, based on the concatenated data set. Branch lengths are according to mean estimates of node ages (Table S4). Branch colors indicates substitution rates.
(EPS)

Table S1 Node support given as BS and BPP values for partitioned ML and BI phylogenetic reconstructions. Nodes marked with * were recovered in best ML tree topologies, but were not included in BS consensus trees. Nodes marked with. Nodes were labelled as in Fig. S2. BEAST analyses were based on six fossil constraints (run '-ADEF'). Exclusion of *Serranus atricauda* from the data set had little effect on node support.
(DOC)

Table S2 Divergence date estimates, estimated in BEAST on the basis of six reliable fossil calibrations (run '-ADEF'). For this analysis, time constraints were applied to nodes marked with *. Labels refer to nodes in Fig. S2. Exclusion of *Serranus atricauda* from the data set had negligible effects on age estimates. All dates are given in Ma.
(DOC)

Table S3 Estimates for the onset of the radiation of the AFGP-bearing Antarctic Clade (node X) when individual node constraints were removed during the constraint cross-validation. All dates are given in Ma.
(DOC)

Table S4 Genbank accession numbers for all sequences used for phylogenetic analyses. Sequences HM049934-HM050270 were produced as part of this study. * Nuclear *T. rubripes* and *T. nigroviridis* sequences were extracted from Ensembl (www.ensembl.org) and Genoscope (www.genoscope.cns.fr) genome browsers (Table S2).
(DOC)

Table S5 Ensembl and Genoscope identifiers of *Takifugu rubripes* and *Tetraodon nigroviridis* sequences. *T. rubripes* Ensembl identifiers were taken from [5], while *T. nigroviridis* Genoscope identifiers and sequences were found by BLAT-search against the *T. nigroviridis* genome, using the entire *T. rubripes* sequences as search templates.
(DOC)

Text S1
(DOC)

Text S2
(DOC)

Text S3
(DOC)

Text S4
(DOC)

Text S5
(DOC)

Text S6
(DOC)

142

Text S7
(DOC)

Text S8
(DOC)

Author Contributions

Conceived and designed the experiments: MM RH WS. Performed the experiments: MM. Analyzed the data: MM WS. Contributed reagents/materials/analysis tools: MM RH WS. Wrote the paper: MM RH WS.

References

1. Simpson GG (1953) The major features of evolution. New York: Colombia University Press. 434 p.
2. Schluter D (2000) The ecology of adaptive radiation. Oxford: Oxford University Press 288 p.
3. Yoder JB, Clancey E, Des Roches S, Eastman JM, Gentry L, et al. (2010) Ecological opportunity and the origin of adaptive radiations. J Evol Biol 23: 1581–1598.
4. Heard SB, Hauser DL (1995) Key evolutionary innovations and their ecological mechanisms. Hist Biol 10: 151–173.
5. Lack D (1947) Darwin's finches. Cambridge: Cambridge University Press. 211 p.
6. Salzburger W, Mack T, Verheyen E, Meyer A (2005) Out of Tanganyika: Genesis, explosive speciation, key-innovations and phylogeography of the haplochromine cichlid fishes. BMC Evol Biol 5: 17.
7. Salzburger W (2009) The interaction of sexually and naturally selected traits in the adaptive radiations of cichlid fishes. Molecular Ecology 18: 169–185.
8. Seehausen O (2006) African cichlid fish: a model system in adaptive radiation research. Proc R Soc B 273: 1987–1998.
9. Alfaro ME, Brock CD, Banbury BL, Wainwright PC (2009) Does evolutionary innovation in pharyngeal jaws lead to rapid lineage diversification in labrid fishes? BMC Evol Biol 9: 255.
10. Losos JB (2009) Lizards in an evolutionary tree: ecology and adaptive radiation of anoles. Berkeley: University of California Press. 528 p.
11. Eastman JT (2005) The nature of the diversity of Antarctic fishes. Polar Biol 28: 93–107.
12. Chen L, DeVries AL, Cheng C-HC (1997) Evolution of antifreeze glycoprotein gene from a trypsinogen gene in Antarctic notothenioid fish. Proc Natl Acad Sci U S A 94: 3811–3816.
13. Pointer MA, Cheng CH, Bowmaker JK, Parry JW, Soto N, et al. (2005) Adaptations to an extreme environment: retinal organisation and spectral properties of photoreceptors in Antarctic notothenioid fish. J Exp Biol 208: 2363–2376.
14. Hofmann GE, Lund SG, Place SP, Whitmer AC (2005) Some like it hot, some like it cold: the heat shock response is found in New Zealand but not Antarctic notothenioid fishes. J Exp Mar Biol Ecol 316: 79–89.
15. Eastman JT (1993) Antarctic fish biology: evolution in a unique environment. San Diego: Academic Press. 322 p.
16. Bilyk KT, DeVries AL (2010) Freezing avoidance of the Antarctic icefishes (Channichthyidae) across thermal gradients in the Southern Ocean. Polar Biol 33: 203–213.
17. Cheng C-HC, Chen L, Near TJ, Jin Y (2003) Functional antifreeze glycoprotein genes in temperate-water New Zealand nototheniid fish infer an Antarctic evolutionary origin. Mol Biol Evol 20: 1897–1908.
18. Pagani M, Zachos JC, Freeman KH, Tipple B, Bohaty S (2005) Marked decline in atmospheric carbon dioxide concentrations during the Paleogene. Science 309: 600–603.
19. Zachos JC, Dickens GR, Zeebe RE (2008) An early Cenozoic perspective on greenhouse warming and carbon-cycle dynamics. Nature, 451: 279–283.
20. Cape Roberts Science Team (2000) Studies from the Cape Roberts Project, Ross Sea, Antarctica. Initial report on CRP-3. Terra Antarctica, 7: 1–209.
21. Olney MP, Bohaty SM, Harwood DM, Scherer RP (2009) *Creania lacyae* gen. nov. et sp. nov. and *Synedropsis cheethamii* sp. nov., fossil indicators of Antarctic sea ice? Diatom Res 24: 357–375.
22. Near TJ (2004) Estimating divergence times of notothenioid fishes using a fossil-calibrated molecular clock. Antarct Sci 16: 37–44.
23. Bargelloni L, Marcato S, Zane L, Patarnello T (2004) Mitochondrial phylogeny of notothenioids: a molecular approach to Antarctic fish evolution and biogeography. Syst Biol 49: 114–129.
24. Chen W-J, Bonillo C, Lecointre G (1998) Phylogeny of the Channichthyidae (Notothenioidei, Teleostei) based on two mitochondrial genes. In: di Prisco G, Pisano E, Clarke A, eds. Fishes of Antarctica. A biological overview. Milano: Springer-Verlag Italia. pp 287–298.
25. Bargelloni L, Ritchie PA, Paternello T, Battaglia B, Lambert DM, et al. (1994) Molecular evolution at subzero temperatures: mitochondrial and nuclear phylogenies of fishes from Antarctica (suborder Notothenioidei), and the evolution of antifreeze glycopeptides. Mol Biol Evol 11: 854–863.
26. Miya M, Takeshima H, Endo H, Ishiguro NB, Inoue JG, et al. (2003) Major patterns of higher teleostean phylogenies: a new perspective based on 100 complete mitochondrial DNA sequences. Mol Phylogenet Evol 26: 121–138.
27. Hanel R, Westneat MW, Sturmbauer C (2002) Phylogenetic relationships, evolution of broodcare behavior, and geographic speciation in the wrasse tribe Labrini. J Mol Evol 55: 776–789.
28. Kazancioglu E, Near TJ, Hanel R, Wainwright PC (2009) Influence of feeding functional morphology and sexual selection on diversification rate in parrotfishes (Scaridae). Proc R Soc B 276: 3439–3446.
29. Smith WL, Craig MT (2007) Casting the percomorph net widely: the importance of broad taxonomic sampling in the search for the placement of serranid and percid fishes. Copeia 2007: 35–55.
30. Dettaï A, Lecointre G (2005) Further support for the clades obtained by multiple molecular phylogenies in the acanthomorph bush. CR Biol 328: 674–689.
31. Near TJ, Cheng C-HC (2008) Phylogenetics of notothenioid fishes (Teleostei: Acanthomorpha): Inferences from mitochondrial and nuclear gene sequences. Mol Phylogenet Evol 47: 832–840.
32. Near TJ, Bolnick DI, Wainwright PC (2005) Fossil calibrations and molecular divergence time estimates in centrarchid fishes (Teleostei: Centrarchidae). Evolution 59: 1768–1782.
33. Patterson C (1993) Osteichthyes: Teleostei. In: Benton MJ, ed. The fossil record 2. London: Chapman & Hall. pp 621–655.
34. Eastman JT, Grande L (1991) Late Eocene gadiform (Teleostei) skull from Seymour Island, Antarctic Peninsula. Antarct Sci 3: 87–95.
35. Balushkin AV (1994) Fossil notothenioid, and not gadiform, fish *Proeleginops grandeastmanorum* gen. nov. sp. nov. (Perciformes, Notothenioidei, Eleginopidae) from the late Eocene found in Seymour Island (Antarctica). J Ichthyol 34: 298–307.
36. Shevenell AE, Kennett JP, Lea DW (2004) Middle Miocene Southern Ocean cooling and Antarctic cryosphere expansion. Science 305: 1766–1770.
37. Shevenell AE, Kennett JP (2007) Cenozoic Antarctic cryosphere evolution: tales from deep-sea sedimentary records. Deep-Sea Res II 54: 2308–2324.
38. DeConto RM, Pollard D (2003) Rapid Cenozoic glaciation of Antarctica induced by declining atmospheric CO_2. Nature 421: 245–249.
39. Zachos JC, Quinn TM, Salamy KA (1996) High resolution (10^4 years) deep-sea foraminiferal stable isotope records of the Eocene-Oligocene climate transition. Paleoceanography 11: 251–266.
40. DeConto RM, Pollard D, Harwood D (2007) Sea ice feedback and Cenozoic evolution of Antarctic climate and ice sheets. Paleoceanography 22: PA3214.
41. Barker PF, Filippelli GM, Florindo F, Martin EE, Scher HD (2007) Onset and role of the Antarctic Circumpolar Current. Deep-Sea Res II 54: 2388–2398.
42. Nong GT, Najjar RG, Seidov D, Peterson WH (2000) Simulation of ocean temperature change due to the opening of Drake Passage. Geophys Res Lett 27: 2689–2692.
43. Naish TR, Woolfe KJ, Barret PJ, Wilson GC, Bohaty SM, et al. (2001) Orbitally induced oscillations in the East Antarctic ice sheet at the Oligocene/Miocene boundary. Nature 413: 719–723.
44. Cheng C-HC (1998) Origin and mechanism of evolution of antifreeze glycoproteins in polar fishes. In: di Prisco G, Pisano E, Clarke A, eds. Fishes of Antarctica. A biological overview. Milano: Springer-Verlag Italia. pp 311–328.
45. Chen L, DeVries AL, Cheng C-HC (1997) Convergent evolution of antifreeze glycoproteins in Antarctic notothenioid fish and Arctic cod. Proc Natl Acad Sci U S A 94: 3817–3822.
46. O'Grady SM, Schrag JD, Raymond JA, DeVries AL (1982) Comparison of antifreeze glycopeptides from arctic and antarctic fishes. J Exp Zool 224: 177–185.
47. Fletcher GL, Hew CL, Davies PL (2001) Antifreeze proteins of teleost fishes. Annu Rev Physiol 63: 359–390.
48. Bakke I, Johansen SD (2005) Molecular phylogenetics of Gadidae and related Gadiformes based on mitochondrial DNA sequences. Mar Biotechnol 7: 61–69.
49. Anderson ME (1994) Systematics and osteology of the Zoarcidae (Teleostei: Perciformes). Ichthyol Bull JLB Smith Inst Ichthyol 60: 1–120.
50. Graham LA, Lougheed SC, Ewart KV, Davies PL (2008) Lateral transfer of a Lectin-like antifreeze protein gene in fishes. PLoS ONE 3: e2616.
51. Salzburger W (2008) To be or not to be a hamlet pair in sympatry. Mol Ecol 17: 1397–1399.

143

52. Balushkin AV (2000) Morphology, classification, and evolution of notothenioid fishes of the Southern Ocean (Notothenioidei, Perciformes). J Ichthyol 40: S74–S109.

53. Anderson JB (1999) Antarctic Marine Geology Cambridge: Cambridge University Press. 289 p.

54. Kennett JP, Exon NF (2004) Paleoceanographic evolution of the Tasmanian Seaway and its climatic implications. In: Exon NF, Kennett JP, Malone MJ, eds. The Cenozoic Southern Ocean: Tectonics, sedimentation, and climate change between Australia and Antarctica. Geophysical Monograph. Washington, D.C.: American Geophysical Union. pp 345–367.

55. Stickley CE, Brinkhuis H, Schellenberg SA, Sluijs A, Roehl U, et al. (2004) Timing and nature of the deepening of the Tasmanian Gateway. Paleoceanography 19: PA4027.

56. Zwickl DJ (2006) Genetic algorithm approaches for the phylogenetic analysis of large biological sequence datasets under the maximum likelihood criterion. Ph.D. dissertation. Univ. of Texas at Austin.

57. Stamatakis A (2006) RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. Bioinformatics 22: 2688–2690.

58. Benton MJ, Donoghue P (2007) Paleontological evidence to date the tree of life. Mol Biol Evol 24: 26–53.

59. Storey BC (1995) The role of mantle plumes in continental breakup: case histories from Gondwanaland. Nature 377: 301–308.

60. Sereno PC, Wilson JA, Conrad JL (2004) New dinosaurs link southern landmasses in the Mid-Cretaceous. Proc R Soc B 271: 1325–1330.

61. Drummond AJ, Rambaut A (2007) BEAST: Bayesian evolutionary analysis by sampling trees. BMC Evol Biol 7: 214.

62. Drummond AJ, Ho SYW, Philips MJ, Rambaut A (2006) Relaxed phylogenetics and dating with confidence. PLoS Biol 4: e88.

63. Gernhard T (2008) The conditioned reconstructed process. J Theor Biol 253: 769–778.

64. Shapiro B, Rambaut A, Drummond AJ (2006) Choosing appropriate substitution models for the phylogenetic analysis of protein-coding sequences. Mol Biol Evol 23: 7–9.

65. Suchard MA, Weiss RE, Sinsheimer JS (2001) Bayesian selection of continuous-time Markov Chain evolutionary models. Mol Biol Evol 18: 1001–1013.

66. Kass RE, Raftery AE (1995) Bayes Factors. J Am Stat Assoc 90: 773–795.

67. Larter RD, Cunningham AP, Barker PF, Gohl K, Nitsche FO (2002) Tectonic evolution of the Pacific margin of Antarctica. 1. Late Cretaceous tectonic reconstructions. J Geophys Res 107: 2345.

68. Livermore R, Nankivell A, Eagles G, Morris P (2005) Paleogene opening of Drake Passage. Earth Planet Sc Lett 236: 459–470.

69. Henderiks J, Pagani M (2008) Coccolithophore cell size and the Paleogene decline in atmospheric CO_2. Earth Planet Sc Lett 269: 575–583.

144

Supporting Information

On the origin and trigger of the notothenioid adaptive radiation

Michael Matschiner[*]**, Reinhold Hanel**[†]**, and Walter Salzburger**[*]

Addresses:
[*]Zoological Institute, University of Basel, Vesalgasse 1, CH-4051 Basel, Switzerland
[†]Institute of Fisheries Ecology, Johann Heinrich von Thünen-Institute, Federal Research Institute for Rural Areas, Forestry and Fisheries, Palmaille 9, D-22767 Hamburg, Germany

Correspondence:
Walter Salzburger, Fax: +41 61 267 03 01, E-mail: walter.salzburger@unibas.ch

1 Figures

Fig. S1: Partitioned ML consensus tree of full mitochondrial genomes (A), and best ML phylogenies of single mitochondrial markers (B-D), estimated with RAxML. Sequence data were taken from Kawahara *et al.* (2008), Mabuchi *et al.* (2007), and Papetti *et al.* (2007). Markers ND4 (B) and cyt *b* (C) phylogenies show better agreement with the full mitogenome topology than other markers of comparable sequence length (D).

Fig. S2: Partitioned ML phylogeny of 83 acanthomorph fishes, based on the concatenated dataset of two mitochondrial (ND4, cyt *b*) and four nuclear genes (myh6, Ptr, ENC1, tbr). Tree topology and branch lengths are as estimated using GARLI-PART, and near-identical topologies were recovered with RAxML and BEAST analyses. Filled circles indicate >98% BS support (as calculated with GARLI-PART) and >0.99 BPP (according to BEAST run '-ADEF'), white circles represent nodes with BS support >80% and >0.90 BPP. Split circles indicate different levels of BS (left half) and BPP (right half) support. All node support values are summarized in Table S3. Nodes that were used for fossil and phylogeographic constraints are labelled with letters A–J, basal percid and notothenioid nodes are labelled with letters T–Z.

Fig. S3: Partitioned BI phylogeny, based on the concatenated data set, reduced to family level. Node heights correspond to mean age estimates. The time scale is divided into phanerozoic stages (grey shades), and the presence of skeletal (black bars) or otolith fossils (dark grey bars) in a stage is plotted on top of family branches. Unless otherwise noted, all fossil information is taken from Patterson (1993). Fossils with questionable taxonomic or stratigraphic assignments are indicated by dashed bars. Numbers in brackets indicate the number of species included in this study and the total number of species per family (Nelson 2006).
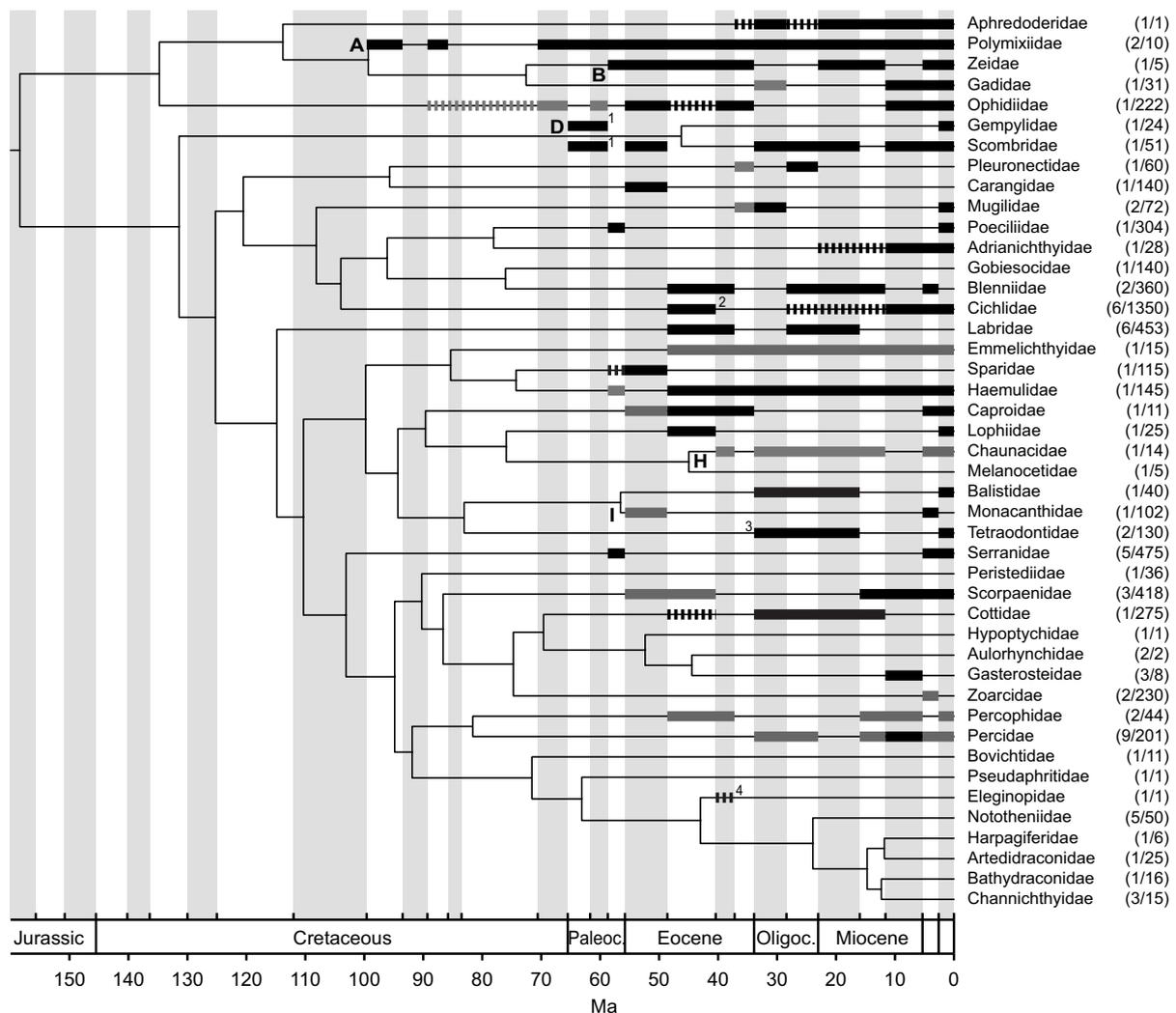
[1]Fossil used to constrain node D, however cross-validation results suggest unreliability of this constraint. As Gempylidae could be nested within Scombridae (Orrell *et al.* 2006), fossils may have been misinterpreted. [2]Murray (2001); Malabarba *et al.* (2010) [3]According to Santini & Tyler (2003), the oldest tetraodontid is *Archaeotetraodon winterbottomi* from the Rupelian. *Eotetraodon pygmaeus*, previously assigned to Tetraodontidae (Patterson 1993), has been moved to the family Triodontidae [4]Eastman & Grande (1991); Balushkin (1994).

Fig. S4: Partitioned BI phylogeny of 83 acanthomorph fishes, based on the concatenated data set. Branch lengths are according to mean estimates of node ages (Table S4). Branch colors indicates substitution rates.
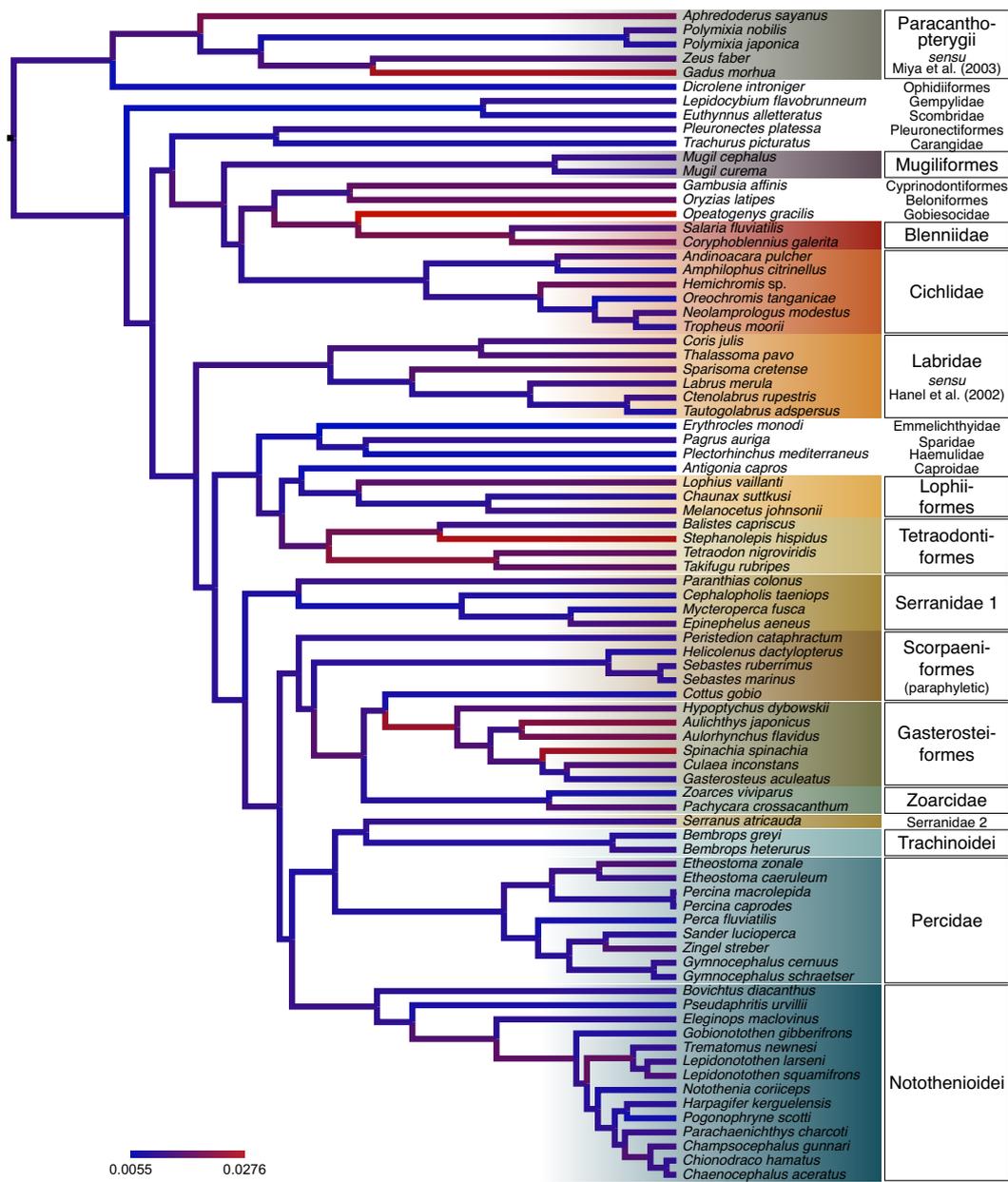
2   Tables

Tab.   S1: Node support given as BS and BPP values for partitioned ML and BI phylogenetic reconstructions. Nodes marked with * were recovered in best ML tree topologies, but were not included in BS consensus trees. Nodes were labelled as in SI Fig. S2. BEAST analyses were based on six fossil constraints (run '-ADEF'). Exclusion of *Serranus atricauda* from the data set had little effect on node support.

Node	full taxa set			without *Serranus atricauda*		
	GARLI-PART	RAxML	BEAST	GARLI-PART	RAxML	BEAST
A	62	38	0.94	56	38	0.95
B	93	86	1	86	85	1
C	100	100	0.92	100	100	91
D	100	100	1	100	100	1
E	100	100	1	100	100	1
F	100	100	1	100	100	1
G	75	79	0.98	77	67	0.98
H	100	100	1	100	100	1
I	100	100	1	100	100	1
J	100	100	1	100	100	1
T	100	100	1	100	100	1
U	100	100	1	100	100	1
V	86	83	1	84	83	1
W	100	100	1	100	100	1
X	100	100	1	100	100	1
Y	55	53	0.64	50	57	0.65
Z	98	97	1	97	98	1
1	100	100	0.99	100	100	0.99
2	100	100	1	100	100	1
3	100	100	0.45	100	100	0.47
4	50	61	0.84	49	63	0.86
5	57	78	0.89	64	74	0.89
6	92	94	1	91	94	1
7	97	99	0.99	98	99	0.99
8	100	100	1	100	100	1
9	*	21	0.60	*	16	0.60
10	48	47	0.89	57	45	0.85
11	65	90	1	75	89	1
12	85	94	1	90	94	1
13	100	100	1	100	100	1
14	100	100	1	100	100	1
15	100	100	1	100	100	1
16	94	98	1	99	98	1
17	100	100	1	100	100	1
18	92	98	1	95	96	1
19	100	100	1	100	100	1

Tab. S1 (continued)

Node	full taxa set			without *Serranus atricauda*		
	GARLI-PART	RAxML	BEAST	GARLI-PART	RAxML	BEAST
20	100	100	1	100	100	1
21	92	96	1	90	94	1
22	100	100	1	100	100	1
23	95	99	1	97	97	1
24	72	82	1	71	83	1
25	*	56	0.69	*	61	0.61
26	79	85	1	79	84	1
27	33	-	0.64	-	-	0.59
28	84	86	1	81	92	1
29	80	91	1	87	96	1
30	92	94	1	95	94	1
31	42	46	0.98	60	46	0.98
32	100	100	1	100	100	1
33	100	100	1	100	100	1
34	51	48	0.99	66	54	1
35	62	71	0.98	55	54	0.98
36	51	55	0.68	55	56	0.72
37	100	100	1	100	100	1
38	100	100	1	100	100	1
39	100	98	1	100	97	1
40	57	62	0.84	54	62	0.86
41	100	100	1	100	100	1
42	85	85	1	81	85	1
43	80	85	1	77	84	1
44	100	98	1	100	99	1
45	59	62	0.88	65	65	0.91
46	100	100	1	100	100	1
47	19	26	0.83	27	30	0.76
48	71	56	1	72	70	0.89
49	35	32	0.78	-	-	-
50	100	100	1	100	100	1
51	100	100	1	100	100	1
52	100	100	1	100	100	1
53	100	100	1	100	100	1
54	94	97	1	94	97	1
55	76	91	1	85	94	1
56	100	100	1	100	100	1
57	100	100	1	100	100	1
58	100	100	1	100	100	1
59	92	89	1	87	89	1
60	63	84	0.95	59	87	0.96
61	90	95	1	96	95	1
62	98	98	1	96	98	1

Tab. S1 (continued)

Node	full taxa set			without *Serranus atricauda*		
	GARLI-PART	RAxML	BEAST	GARLI-PART	RAxML	BEAST
63	100	100	1	100	100	1
64	100	100	1	100	100	1

Tab. S2: Divergence date estimates, estimated in BEAST on the basis of six reliable fossil calibrations (run '-ADEF'). For this analysis, time constraints were applied to nodes marked with *. Labels refer to nodes in SI Fig. S2. Exclusion of *Serranus atricauda* from the data set had negligible effects on age estimates. All dates are given in Ma.

Node	-ADEF			-ADEF -*Serranus atricauda*		
	95% HPD upper	Mean	95% HPD lower	95% HPD upper	Mean	95% HPD lower
A	133.05	99.16	70.51	130.78	98.12	69.53
B*	95.52	72.42	56.57	93.76	71.82	56.74
C*	150.89	131.18	114.66	150.89	131.20	114.62
D	79.15	46.14	18.36	82.56	47.34	18.64
E	80.71	59.70	40.24	81.24	59.36	39.26
F	52.40	34.89	19.31	52.12	34.79	19.04
G*	124.98	110.13	96.90	125.44	110.40	96.90
H*	56.32	44.89	37.54	56.36	44.86	37.54
I*	66.31	56.44	49.19	66.36	56.48	49.19
J*	53.53	42.96	32.26	53.53	42.98	32.25
T	54.71	41.15	28.54	55.08	41.36	28.83
U	89.44	71.44	54.41	89.45	71.72	54.41
V	80.26	63.01	46.70	80.21	63.09	46.52
W	56.86	42.93	29.81	57.17	43.09	29.99
X	31.94	23.88	16.72	31.84	24.06	16.80
Y	28.20	21.36	15.26	28.21	21.52	15.38
Z	20.02	14.71	9.93	20.07	14.86	10.04
root	194.47	158.13	124.65	193.97	157.61	124.51
1	151.51	113.60	79.27	150.85	112.86	78.96
2	20.77	11.90	4.54	20.62	11.93	4.78
3	170.20	134.51	97.80	168.79	133.79	96.93
4	143.15	125.02	107.81	143.31	124.97	107.84
5	139.52	120.30	102.26	139.41	120.09	102.18
6	124.30	95.53	64.82	124.39	95.59	64.83
7	128.41	107.94	88.07	129.05	107.83	88.05
8	47.99	29.03	12.29	49.42	29.41	12.67
9	124.33	103.79	83.66	125.24	103.63	83.65
10	116.87	95.95	75.88	117.13	95.64	75.04
11	102.38	77.94	54.02	102.31	77.30	52.52
12	96.51	75.92	55.83	96.80	75.71	55.95
13	57.36	39.22	22.39	57.23	39.33	22.25

Tab. S2 (continued)

Node	-ADEF			-ADEF -*Serranus atricauda*		
	95% HPD upper	Mean	95% HPD lower	95% HPD upper	Mean	95% HPD lower
14	45.29	28.36	13.12	44.49	28.10	12.95
15	47.18	32.45	19.47	46.58	32.25	19.13
16	29.43	19.44	10.51	29.23	19.32	10.51
17	15.90	9.65	4.21	16.02	9.64	4.26
18	131.03	114.62	99.28	131.34	114.80	99.30
19	104.13	82.56	61.47	103.96	82.38	60.26
20	68.95	46.61	25.72	68.01	46.27	25.26
21	84.66	63.05	42.06	84.97	63.03	41.70
22	19.80	11.77	4.73	19.75	11.66	4.88
23	115.64	99.59	85.17	116.31	99.71	85.26
24	105.84	85.20	62.96	106.78	85.55	63.98
25	96.57	74.13	49.94	97.46	74.45	50.17
26	110.04	94.14	80.43	109.91	94.21	79.87
27	105.80	89.46	74.66	106.04	89.57	74.57
28	93.42	75.80	57.98	93.44	75.91	57.92
29	97.90	82.92	69.41	97.64	82.84	69.04
30	119.69	102.91	88.28	120.01	102.84	88.15
31	110.83	90.16	69.12	111.64	90.26	69.35
32	74.98	51.24	29.00	74.65	51.33	29.39
33	41.51	25.22	11.09	42.11	25.44	11.24
34	110.90	94.68	79.48	111.17	94.43	79.34
35	106.47	90.10	74.97	106.57	89.84	74.74
36	102.68	86.48	71.55	102.64	86.30	71.30
37	26.35	16.21	7.71	26.59	16.32	7.79
38	6.66	3.83	1.55	6.68	3.89	1.50
39	90.32	74.58	59.89	90.14	74.44	59.82
40	84.84	69.47	54.87	84.73	69.28	54.89
41	65.91	52.29	39.60	65.43	52.13	39.64
42	56.82	44.37	32.54	56.48	44.17	32.61
43	49.42	36.90	24.57	49.55	36.79	24.81
44	43.38	31.93	21.39	43.25	31.84	20.98
45	37.72	26.13	15.43	37.24	25.94	15.12
46	48.77	30.37	13.78	48.95	30.21	13.70
47	108.03	91.72	76.49	108.54	91.90	76.68
48	99.51	81.45	63.96	100.36	81.79	64.05
49	93.26	74.02	55.05	-	-	-
50	26.12	15.24	5.89	25.94	15.20	6.17
51	41.90	29.67	18.36	42.37	29.94	18.75
52	28.56	18.41	9.17	28.66	18.61	9.56
53	1.42	0.77	0.25	1.45	0.79	0.25
54	45.16	33.00	21.77	45.36	33.09	21.83
55	36.14	25.56	15.82	36.19	25.49	15.59
56	26.24	16.87	8.51	26.03	16.73	8.16

Tab. S2 (continued)

Node	-ADEF			-ADEF -*Serranus atricauda*		
	95% HPD upper	Mean	95% HPD lower	95% HPD upper	Mean	95% HPD lower
57	9.34	5.38	2.13	9.37	5.42	2.18
58	15.16	10.33	6.06	15.28	10.49	6.13
59	11.02	6.93	3.37	11.03	7.00	3.37
60	25.08	18.74	12.94	25.16	18.89	13.13
61	17.34	11.78	6.65	17.34	11.91	6.63
62	17.10	12.28	7.77	17.24	12.44	7.91
63	9.35	6.20	3.35	9.43	6.32	3.45
64	3.96	2.36	0.97	4.13	2.43	1.00

Tab. S3: Estimates for the onset of the radiation of the AFGP-bearing Antarctic Clade (node X) when individual node constraints were removed during the constraint cross-validation. All dates are given in Ma.

Constraint Set	95% HPD upper	Mean	95% HPD lower
Full set	35.4	26.7	18.7
-A	34.3	25.9	17.8
-B	35.4	26.8	19.0
-C	39.0	29.3	20.0
-D	34.8	26.3	18.1
-E	31.3	23.4	16.4
-F	36.2	27.6	19.8
-G	34.6	26.2	18.5
-H	35.5	26.8	18.8
-I	36.0	27.0	19.0
-J	35.4	26.8	18.8
-ACDEF	34.2	25.0	16.8
-ADEF	31.9	23.9	16.7

Tab. S4: Genbank accession numbers for all sequences used for phylogenetic analyses. Sequences HM049934-HM050270 were produced as part of this study. *Nuclear *T. rubripes* and *T. nigroviridis* sequences were extracted from Ensembl (www.ensembl.org) and Genoscope (www.genoscope.cns.fr) genome browsers (SI Table S5).

Taxa	ND4	CytB	myh6	Ptr	ENC1	tbr1
Aequidens pulcher	HM050087	EF432944	HM050029	HM050147	HM049971	HM050208
Amphilophus citrinellus	HM050088	AB018985	HM050030	HM050148	HM049972	HM050209
Antigonia capros	NC_003191	NC_003191	EF536307	HM050149	HM049973	HM050210
Aphredoderus sayanus	NC_004372	NC_004372	EU001908	EU001962	EU002019	EU001990
Aulichthys japonicus	NC_011569	NC_011569	AB445150	AB445168	AB445222	AB445204
Aulorhynchus flavidus	NC_010268	NC_010268	AB445151	AB445169	AB445223	AB445205
Balistes capriscus	HM050089	EF392572	HM050031	HM050150	HM049974	HM050211
Bembrops greyi	HM050090	HM049934	HM050032	HM050151	HM049975	HM050212
Bembrops heterurus	HM050091	HM049935	HM050033	HM050152	HM049976	HM050213
Bovichtus diacanthus	HM050092	HM049936	HM050034	HM050153	HM049977	HM050214
Cephalopholis taeniops	HM050093	EF455991	HM050035	HM050154	HM049978	HM050215
Chaenocephalus aceratus	HM050094	HM049937	HM050036	HM050155		HM050216
Champsoceph. gunnari	HM050095	HM049938	HM050037	HM050156	HM049979	HM050217
Chaunax suttkusi	HM050096	HM049939	HM050038	HM050157	HM049980	HM050218
Chionodraco hamatus	HM050097	HM049940	HM050039	HM050158	HM049981	HM050219
Coris julis	HM050099	HM049942	HM050041	HM050160	HM049982	HM050221
Coryphoblenn. galerita	HM050098	HM049941	HM050040	HM050159		HM050220
Cottus gobio	HM050100	AY116366	HM050042			HM050222
Ctenolabrus rupestris	HM050101	HM049943	HM050043	HM050161	HM049983	HM050223
Culaea inconstans	NC_011577	NC_011577	AB445153	AB445171	AB445225	AB445207
Dicrolene introniger	HM050102	HM049944	HM050044	HM050162	HM049984	HM050224
Eleginops maclovinus	DQ526429	DQ526429	HM050045	HM050163	HM049985	HM050225
Epinephelus aeneus	HM050103	DQ197950	HM050046	HM050164	HM049986	HM050226
Erythrocles monodi	HM050104	EF456004	HM050047	HM050165	HM049987	HM050227
Etheostoma caeruleum	HM050105	DQ465142		HM050166	HM049988	HM050228
Etheostoma zonale	HM050106	AY964705		HM050167	HM049989	HM050229
Euthynnus alletteratus	NC_004530	EF439531	HM050048	HM050168	HM049990	HM050230
Gadus morhua	NC_002081	EU877717	EU001906	EU001960	EU002017	
Gambusia affinis	NC_004388	NC_004388	EU001907	EU001961	EU002018	EU001989
Gasterosteus aculeatus	AP002944	AP002944	AB445155	AB445173	AB445227	AB445209
Gobionoto. gibberifrons	HM050107	HM049945	HM050049	HM050169	HM049991	HM050231
Gymnocephalus cernuus	HM050108	AF045356	HM050050	HM050170	HM049992	HM050232
G. schraetser	HM050109	HM049946	HM050051	HM050171	HM049993	HM050233
Harpagifer kerguelensis	HM050110	HM049947	HM050052	HM050172	HM049994	HM050234
Helicol. dactylopterus	HM050111	EU492259	HM050053	HM050173	HM049995	
Hemichromis sp.	HM050112	HM049948		HM050174	HM049996	HM050235
Hypoptychus dybowskii	NC_004400	NC_004400	AB445149	AB445167	AB445221	AB445203
Labrus merula	HM050113	HM049949	HM050054	HM050175		HM050236
Lepidoc. flavobrunneum	HM050114	AM265576	HM050055	HM050176	HM049997	HM050237
Lepidonotothen larseni	HM050115	HM049950	HM050056	HM050177		HM050238

Tab. S4 (continued)

Taxa	ND4	CytB	myh6	Ptr	ENC1	tbr1
L. squamifrons	HM050116	HM049951	HM050057	HM050178	HM049998	HM050239
Lophius vaillanti	HM050117	HM049952	HM050058	HM050179	HM049999	HM050240
Melanocetus johnsonii	HM050118	HM049953	HM050059	HM050180		HM050241
Mugil cephalus	NC_003182	EU083840	HM050060		HM050000	HM050242
Mugil curema		EU715492	EU001913	EU001967	EU002023	EU001994
Mycteroperca fusca	HM050119	DQ197968	HM050061	HM050181	HM050001	HM050243
Neolamprol. modestus	HM050120	HM049954	HM050062	HM050182	HM050002	HM050244
Notothenia coriiceps	HM050121	HM049955	HM050063	HM050183	HM050003	HM050245
Opeatogenys gracilis	HM050122	HM049956	HM050064		HM050004	HM050246
Oreochromis tanganicae	HM050123	HM049957	HM050065	HM050184	HM050005	HM050247
Oryzias latipes	NC_004387	AB084730	EF032927	EF032953	EF032979	EF032966
Pachyc. crossacanthum	HM050124	HM049958	HM050066		HM050006	HM050248
Pagrus auriga	NC_005146	DQ197974	HM050067	HM050185	HM050007	HM050249
Parachaen. charcoti	HM050125	HM049959	HM050068	HM050186	HM050008	HM050250
Paranthias colonus	HM050126	HM049960		HM050187	HM050009	
Perca fluviatilis	HM050129	AY929376	HM050070	HM050189	HM050012	HM050253
Percina caprodes	HM050127	DQ493490			HM050010	HM050251
Percina macrolepida	NC_008111	DQ493495		HM050190	HM050013	HM050254
Perisled. cataphractum	HM050128	HM049961	HM050069	HM050188	HM050011	HM050252
Plectorh. mediterraneus	HM050130	DQ197979	HM050071	HM050191	HM050014	HM050255
Pleuronectes platessa	HM050131	EU224075	EU001930	HM050192		EU002008
Pogonophryne scotti	HM050132	HM049962	HM050072	HM050193		HM050256
Polymixia japonica	NC_002648	NC_002648	EU001926	EU001981	EU002037	
Polymixia nobilis	HM050133	DQ197980	HM050073	HM050194	HM050015	HM050257
Pseudaphritis urvillii	HM050134	HM049963	HM050074	HM050195	HM050016	HM050258
Salaria fluviatilis	HM050135	HM049964	HM050075	HM050196	HM050017	HM050259
Sander lucioperca	HM050136	HM049965	HM050076	HM050197	HM050018	HM050260
Sebastes marinus	HM050137	EF456022	HM050077		HM050019	HM050261
Sebastes ruberrimus	EU008930	AF031501	EU001929	EU001984	EU002040	EU002007
Serranus atricauda		EF439230	HM050078	HM050198	HM050020	HM050262
Sparisoma cretense	HM050138	HM049966	HM050079	HM050199		HM050263
Spinachia spinachia	NC_011582	NC_011582	AB445157	AB445175	AB445229	AB445211
Stephanolepis hispidus	HM050139	HM049967	HM050080	HM050200	HM050021	
Takifugu rubripes	NC_004299	NC_004299	*	*	*	*
Tautogolabrus adspersus	HM050140	HM049968	HM050081	HM050201	HM050022	HM050264
Tetraodon nigroviridis	NC_007176	AP006046	*	*	*	*
Thalassoma pavo	HM050141	DQ198011	HM050082	HM050202	HM050023	HM050265
Trachurus picturatus	HM050142	EF392634	HM050083	HM050203	HM050024	HM050266
Trematomus newnesi	HM050143	HM049969	HM050084	HM050204	HM050025	HM050267
Tropheus moorii	HM050144	AB018990		HM050205	HM050026	HM050268
Zeus faber	NC_003190	EU264027	EU001927	EU001982	EU002038	
Zingel streber	HM050145	HM049970	HM050085	HM050206	HM050027	HM050269
Zoarces viviparus	HM050146	EU492074	HM050086	HM050207	HM050028	HM050270

Tab. S5: Ensembl and Genoscope identifiers of *Takifugu rubripes* and *Tetraodon nigroviridis* sequences. *T. rubripes* Ensembl identifiers were taken from (Li *et al.* 2007), while *T. nigroviridis* Genoscope identifiers and sequences were found by BLAT-search against the *T. nigroviridis* genome, using the entire *T. rubripes* sequences as search templates.

Taxa	myh6	Ptr
T. rubripes	SINFRUE00000644156	SINFRUE00000786790
T. nigroviridis	GSTENT00008412001	GSTENT00035515001

Taxa	ENC1	tbr1
T. rubripes	SINFRUE00000681690	SINFRUE00000673034
T. nigroviridis	GSTENT00025143001	GSTENT00030575001

3 Text

Text S1: Percid divergence dates.
Our analysis recovered a percid clade consisting of the North American Etheostominae and a Eurasian clade combining Luciopercinae and Percinae. The basal position of *Perca fluviatilis* within the Eurasian clade and the inferred date (21.77-45.16 Ma) for its divergence is in concordance with the earliest *P. fluviatilis* fossils from the Miocene (26 Ma) (Carney & Dick 2000). In addition, our date estimate based on six constraints for the split between North American and Eurasian percids (node T, 28.54-54.71 Ma) agrees well with laurasian vicariance following the breakup of landbridges between North America and Europe 50-40 Ma (Milne & Abbott 2002). This shows that clades closely related to notothenioids were dated reliably. All acanthomorph divergence date estimates are summarized in Table S2.

Text S2: Sample collection.
Specimens of 14 notothenioid and 53 non-notothenioid acanthomorph fish species were acquired during field expeditions or from local dealers. We included putative sister groups of notothenioids (Dettaï & Lecointre 2004; Smith & Craig 2007), acanthopterygiid relatives, as well as paracanthopterygiid outgroups (Kawahara *et al.* 2008; Mabuchi *et al.* 2007). We deliberately constrained the number of notothenioid species to reduce sampling bias while at the same time including representatives of each family. Muscle or fin tissue samples were taken from all specimens and preserved in 95% ethanol. Genomic DNA was extracted by proteinase K digestion followed by sodium chloride extraction and ethanol precipitation.

Text S3: Marker selection.
Phylogenetic analyses were based on sequences of four nuclear and two mitochondrial genes. Out of ten nuclear markers developed by genome comparison strategy (Li *et al.* 2007), we chose myh6, Ptr, ENC1, and tbr1 in order to include slow, intermediate and fast-evolving genes. A similar approach

was taken for mitochondrial markers: preliminary phylogenetic analyses were performed with 53 published full mitochondrial genomes of acanthomorph fishes (Kawahara *et al.* 2008; Mabuchi *et al.* 2007), including three notothenioid species (Papetti *et al.* 2007). Hereby, species selection focused on clades G and H of (Kawahara *et al.* 2008), as percids had been suggested as a notothenioid sister group (Dettaï & Lecointre 2004). Mitogenomic sequences were aligned using MUSCLE v3.6 (Edgar 2004) and ProAlign 0.5a3 (Löytynoja & Milinkovitch 2003), and preliminary phylogenies were produced with with RAxML-VI-HPC (Stamatakis 2006b). Phylogenies of single mitochondrial genes were compared with the phylogeny of full mitogenomes (excluding dloop, ND6, and tRNAGlu sequences due to their poor phylogenetic performance (Kawahara *et al.* 2008)) to evaluate the suitability of every marker. We found ND4 (Miya *et al.* 2006) and cytochrome *b* (cyt *b*) sequences to reproduce the full mitogenome phylgeny better than other mitochondrial markers of comparable length, and thus included these markers in our analysis (SI Fig. S1).

Text S4: Primers, PCR, and sequencing.
Nuclear markers were PCR amplified using primer pairs myh6_F459 / myh6_R1325, ptr_F458 / ptr_R1248, enc1_F88 / enc1_R975, and tbr1_F86 / tbr1_R820 (Li *et al.* 2007), and the following cycling conditions: 94°C 2 min, [94°C 30 s, 51–60°C 30 s, 72°C 1 min] × 11–37 cycles, 72°C 7 min. Annealing temperatures were 51°C (myh6), 55°C (Ptr), 60–55°C (ENC1), and 57°C (tbr1), and the heating cycle was repeated 37 (myh6), 35 (Ptr), 46 (ENC1), or 32 (tbr1) times. ENC1 annealing temperature was reduced by 0.5°C per cycle over the first eleven cycles. The following, newly developed primers were used for amplification of mitochondrial genes: NotND4_F416 (CGN TGA GGD AAY CAR RCA GAA CG), NotND4_R1137 (TTD GGD AGD GGD GGD AGD GC), NotCytBf (GGC AAG CCT CCG AAA AAC CCA CCC), L14724t, H15915t (AAC CYY CGR TRT CCG GYT TAC AAG AC), and H15915n (AAC CTY CGG CCT CCG GTT TAC AAG AC). Of these, NotCytBf was designed to bind at position 4 of notothenioid cyt b sequences, since the binding site of traditional cyt *b* forward primers (Meyer *et al.* 1990) has been rearranged in notothenioids of the Antarctic Clade (Zhuang & Cheng 2010). L14724t, H15915t, and H15915n were modified from the commonly used cyt *b* primers L14724 and H15915 (Meyer *et al.* 1990) to account for variation found in teleost (Kawahara *et al.* 2008; Mabuchi *et al.* 2007), and notothenioid cyt *b* sequences (Papetti *et al.* 2007). NotND4_F416 and NotND4_R1137 design was based on mitogenomic alignments of acanthomorph fishes (Kawahara *et al.* 2008; Mabuchi *et al.* 2007). Cycling conditions for mitochondrial sequences were 95°C 3 min, [96°C 15 s, 54°C 30 s, 68°C 30 s] × 37 cycles, 72°C 7 min for ND4, and 94°C 2 min, [94°C 30 s, 67–63°C 30 s, 72°C 1 min] × 39 cycles, 72°C 7 min for cyt *b*, whereby, again the annealing temperature was decreased by 0.5°C per cycle over the first nine cycles. PCR products were purified with ExoSAP-IT (USB), and in some instances using GenElute PCR Clean-Up and Gel Extraction kits (Sigma-Aldrich). Cycle sequencing was performed using the BigDye Terminator v3.1 Cycle Sequencing kit (Applied Biosystems). Sequencing reactions contained 0.5 µM primer, 1.0 µl BigDye Terminator Reaction Mix (Applied Biosystems), and 1.0-3.0 µl purified DNA in a total volume of 8 µl. The following profile was used for cycle sequencing 94°C 1 min, [94°C 10 s, 52°C 20 s, 60°C 4 min] × 25 cycles. Sequence base calls were carried out with CodonCode Aligner 2.0.6 (CodonCode), and verified by eye.

Text S5: Sequence editing and alignment.

Forward and reverse sequences were assembled into contigs in CodonCode Aligner 2.0.6. The dataset was complemented with additional sequences from GenBank, Ensemble (www.ensembl.org), and Genoscope (www.genoscope.cns.fr) to a total of 83 acanthomorph fish species (Tables S1–S2). For every gene, sequences were aligned using MAFFT v6.717b (Katoh & Toh 2008). Alignments were trimmed to start and end with codon triplets, and uninformative insertions were removed. Alignment lengths were 627 bp (ND4), 1140 bp (cyt b), 783 bp (ENC1), 705 bp (myh6), 702 bp (Ptr), and 642 bp (tbr1). In addition to alignments for every single marker, we produced a full concatenated alignment ('full', 4599 bp, 6.53% missing data) as well as separate concatenations for mitochondrial ('mit123') and nuclear ('nuc123') gene sequences. Furthermore, alignments containing all first and second codon positions of mitochondrial ('mit12') and nuclear ('nuc12') sequences, and alignments containing only third codon positions ('mit3' and 'nuc3') were generated to allow codon position-based model selection and phylogenetic reconstruction.

Text S6: Model selection and phylogenetic reconstruction.

Likelihood scores were computed for 88 substitution models on the basis of Maximum Likelihood (ML) optimized phylogenies, as implemented in jModelTest v0.1.1 (Posada 2008; Guindon & Gascuel 2003). Best-fitting models of nucleotide substitution were selected for every alignment according to the Bayesian Information Criterion (BIC) (Schwarz 1978). Selected models were GTR+I+Γ (ND4 and mit123), TIM3+I+Γ (cyt *b* and ENC1),TPM2+I+Γ (myh6), HKY+I+Γ (Ptr, tbr1, and mit12) TVM+Γ (mit3), TPM1uf+I+Γ (nuc123), K80+I+Γ (nuc12), and GTR+Γ (nuc3). Maximum Likelihood phylogenetic inference was performed using a partition-enabled version of GARLI, GARLI-PART v0.97 (Zwickl 2006), as well as RAxML v7.26 (Stamatakis 2006b). For GARLI-PART analyses, alignments mit12, mit3, nuc12, and nuc3 were employed as a concatenated set and the four respective codon position-based models were implemented. Five paracanthopterygian species (*Polymixia japonica*, *Polymixia nobilis*, *Aphredoderus sayanus*, *Gadus morhua*, and *Zeus faber* (Miya *et al.* 2003)) were defined as outgroups and 10 independent run replicates were performed. Runs were set to terminate after a maximum of 5 million generations, or alternatively after 10000 generations without significant (p = 0.01) improvement of scoring topology. To assess node support, 100 nonparametric bootstrap (BS) replicates were run and summarized in a majority rule consensus tree in PAUP* v4.0a110 (Zwickl 2003). A combined bootstrap and ML search was conducted in RAxML, using 1000 rapid bootstrap inferences (Stamatakis *et al.* 2008). Substitution models selected by BIC were not available in RAxML, and the GTRCAT model (Stamatakis 2006a) was chosen instead for all partitions. Given the placement of the serranid species *Serranus atricauda* in a clade containing Percidae and Trachinoidei in both GARLI-PART and RAxML analyses (Fig. 1), all phylogenetic inferences were repeated excluding this species. Removal of *Serranus atricauda* from the dataset affected tree topology only in the position of *Antigonia capros* (now basal to all Lophiiformes and Tetraodontiformes), and on average improved BS support by 1.03% (GARLI-PART) and 0.13% (RAxML) per node (Table S3).

Text S7: Fossils used to constrain node ages.

We used eight fossil, and two phylogeographic constraints in order to estimate acanthomorph divergence dates. Here, we describe and reference each fossil constraint. Unless otherwise noted, all information is taken from Patterson (1993).

Node A (split between Polymixiidae and both Zeidae and Gadidae): This split is constrained by the earliest polymixiid fossils including *Berycopsis elegans* Dixon, 1850, and *Homonotichthys* spp. from the Lower Chalk of Kent and Sussex, England, UK, and *Omosoma tselfatensis* Gaudant, 1978, and *Omosomopsis simum* (Arambourg 1954) from Jebel Tselfat, Morocco, all skeletal and dated to the Cenomanian. We thus constrained node A with a wide lognormal prior (offset: 93.5 Ma, mean: 126.6 Ma, 95% cumulative prior probability (CPP): 328.6-98.2 Ma).

Node B (split between Zeidae and Gadidae): A putative minimum age for the divergence of Zeidae and Gadidae is given by an undescribed ?zeid from the Fur Formation, NW Jutland, Denmark, and by *Palaeocyttus princeps* Gaudant, 1978 from Laveiras, Portugal, both skeletal and dated to the Thanetian. However, the zeid assignment of *P. princeps* is questionable. A lognormal prior was applied (offset: 55.8 Ma, mean: 75.9 Ma, 95% CPP: 198.4–58.6 Ma).

Node C (divergence of Percomorpha): The minimum age derives from the oldest percomorph, *Plectocretacicus clarae* Sorbini, 1979, a skeletal fossil from the Lithographic Limestones of Hakel, Libanon, which has been dated to $\geq$ 96.9 Ma (Benton & Donoghue 2007). We assume that the diversification of Percomorpha postdates the earliest euteleost record, represented by *e.g. Tischlingerichthys viohli* Arratia, 1997, from the Solnhofen Limestone, Germany, that has an maximum age of 150.8 ± 0.1 Ma (Benton & Donoghue 2007). We thus applied a uniform prior between 150.9–96.9 Ma.

Node D (split between Gempylidae and Scombridae): The earliest gempylid record is provided by *Eutrichiurides orpiensis* Leriche, 1906, a skeletal fossil from the Montian Phosphates of Morocco, and an isolated teeth of *E. africanus* Dartevelle and Casier, 1949, from Landana, Angola. Both fossils have been dated to the Danian. The earliest scombrid fossils are *Landanichthys lusitanicus* Dartevelle and Casier, 1949, *L. moutai* Dartevelle and Casier, 1949, and *Sphyraenodus multidentatus* Dartevelle and Casier, 1949 of the scombrid tribe Scomberomorini, which are also from Landana, Angola, and are also assigned a Danian age. We applied a lognormal prior (offset: 61.7 Ma, mean 73.9 Ma, 95% CPP: 148.2–63.4 Ma). Note that the scombrid record traditionally includes Istiophoridae and Xiphiorhynchidae, and thus represents not necessarily a monophyletic group (Orrell *et al.* 2006).

Node E (split between African and Neotropical cichlids): Phylogeographic constraint, see Materials and Methods for details.

Node F (split between *Labrus* and both *Ctenolabrus* and *Tautogolabrus* within Labridae): Phylogeographic constraint, see Materials and Methods. Following Hanel *et al.* (2002), we assume that Labrini diversified subsequent to the closure of the seaway between the Indopacific and the Mediterranean. A minimum age for *Labrus* is provided by *Labrus agassizi* Heckel from St. Magarethen, Austria, which has been dated to the Upper Badenium (14.0 Ma). We applied a uniform prior between 20.5–14.0 Ma.

Node G (split between Gasterosteiformes and Tetraodontiformes): As the oldest percomorph,

Plectocretacicus clarae Sorbini, 1979, is considered a stem-Tetraodontiform, we applied the same prior as for node C (Divergence of Percomorpha). See above.

Node H (split between Chaunacidae and Melanocetidae): A minimum age for this split is provided by the earliest chaunacid, *Chaunax semiangulatus* Stinton, 1978, an otolith fossil from the Barton Formation, Hampshire, England, UK, that is dated to the Bartonian. We thus applied a lognormal prior (offset: 37.2 Ma, mean: 44.6 Ma, CPP: 89.7–38.2 Ma).

Node I (split between Balistidae and Monacanthidae): This split is constrained by the oldest monacanthid fossil, *Amanses sulcifer* Stinton, 1966, an otolith fossil from the London Clay Formation, England, UK, dated to the Ypresian. A lognormal prior was applied (offset: 48.6 Ma, mean: 60.8 Ma, CPP 135.1–50.3 Ma).

Node J (divergence of Tetraodontidae): The earliest tetraodontid is *Archaeotetraodon winterbottomi* Tyler and Bannikov, 1994, a skeletal fossil from the Pshekhsky Horizon, in the lower part of the Maikop Formation of the north Caucasus, Russia, which has been dated to ≥ 32.25 Ma (Benton & Donoghue 2007). Following Benton & Donoghue (2007), we assume that the divergence of tetraodontids postdates the earliest fossil record of potential sister clades, *e.g.* the balistid *Moclaybalistes danekrus* Santini and Tyler, 2002, from the Fur Formation NW Jutland, Denmark, that has been dated to the base of the Eocene. Therefore, we applied a uniform prior for the divergence of Tetraodontidae between 56.0–32.25 Ma.

Text S8: Applicability of relaxed molecular clocks for Notothenioidei.
Acanthomorph divergence dates were estimated using an uncorrelated lognormal relaxed molecular clock, as implemented in the software BEAST. While this approach allows for independent rate variation among branches, it does not account explicitly for parameters that have been shown to influence substitution rates, including temperature, metabolic rate, body size, and generation time (Estabrook *et al.* 2007; Nabholz *et al.* 2008). Thus, it could be argued that application of a naive uncorrelated lognormal relaxed clock introduces systematic bias to divergence date estimates of Antarctic ectotherms. As a result of particularly low substitution rates, short branches of Antarctic notothenioids could be misinterpreted as recent divergences. However, we see little evidence for unusually low substitution rates in Antarctic notothenioids. In fact, one of the lowest notothenioid substitution rates is inferred for *Pseudaphritis urvillii*, a basal riverine species occurring in southern Australia, rather than for high Antarctic species (SI Fig. S4). Similarly, ML branch lengths of Antarctic notothenioids are consistently longer, instead of shorter, than those of non-Antarctic notothenioids and closely related riverine percids from the northern hemisphere (Fig. 1). We thus conclude that systematic bias by temperature regime does not affect notothenioid divergence dates and that the use of an uncorrelated lognormal relaxed molecular clock appropriately accounts for occurring rate variation.

References

Arambourg C (1954) Les poissons crétacés du Jebel Tselfat (Maroc). *Notes Mém Serv Géol Maroc*, **118**, 1–188.

Balushkin AV (1994) Fossil notothenioid, and not gadiform, fish *Proeleginops grandeastmanorum* gen. sp. nov. (Perciformes, Notothenioidei, Eleginopidae) from the late Eocene found in Seymour Island (Antarctica). *Voprosy Ikhtiologii*, **34**, 298–307.

Benton M, Donoghue P (2007) Paleontological evidence to date the tree of life. *Mol Biol Evol*, **24**, 26–53.

Carney JP, Dick TA (2000) The historical ecology of yellow perch (*Perca flavenscens* [Mitchill]) and their parasites. *J Biogeogr*, **27**, 1337–1347.

Dettaï A, Lecointre G (2004) In search of notothenioid (Teleostei) relatives. *Antarct Sci*, **16**, 71–85.

Eastman JT, Grande L (1991) Late Eocene gadiform (Teleostei) skull from Seymour Island, Antarctic Peninsula. *Antarct Sci*, **3**, 87–95.

Edgar RC (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucl Acids Res*, **32**, 1792–1797.

Estabrook GF, Smith GR, Dowling TE (2007) Body mass and temperature influence rates of mitochondrial DNA evolution in North American cyprinid fish. *Evolution*, **61**, 1176–1187.

Guindon S, Gascuel O (2003) A simple, fast, and accurate algorithm to estimate large phylogenies by Maximum Likelihood. *Syst Biol*, **52**, 696–704.

Hanel R, Westneat MW, Sturmbauer C (2002) Phylogenetic relationships, evolution of broodcare behavior, and geographic speciation in the wrasse tribe Labrini. *J Mol Evol*, **55**, 776–789.

Katoh K, Toh H (2008) Recent developments in the MAFFT multiple sequence alignment program. *Brief Bioinform*, **9**, 286–298.

Kawahara R, Miya M, Mabuchi K *et al.* (2008) Interrelationships of the 11 gasterosteiform families (sticklebacks, pipefishes, and their relatives): A new perspective based on whole mitogenome sequences from 75 higher teleosts. *Mol Phylogenet Evol*, **46**, 224–236.

Li C, Ortí G, Zhang G, Lu G (2007) A practical approach to phylogenomics: the phylogeny of ray-finned fish (Actinopterygii) as a case study. *BMC Evol Biol*, **7**, 44.

Löytynoja A, Milinkovitch MC (2003) A hidden Markov model for progressive multiple alignment. *Bioinformatics*, **19**, 1505–1513.

Mabuchi K, Miya M, Azuma Y, Nishida M (2007) Independent evolution of the specialized pharyngeal jaw apparatus in cichlid and labrid fishes. *BMC Evol Biol*, **7**, 10–10.

Malabarba MC, Malabarba LR, Del Papa C (2010) *Gymnogeophagus eocenicus*, n. sp (Perciformes: Cichlidae), an Eocene cichlid from the Lumbrera formation in Argentina. *J Vert Paleont*, **30**, 341–350.

Meyer A, Kocher TD, Basasibwaki P, Wilson AC (1990) Monophyletic origin of Lake Victoria cichlid fishes suggested by mitochondrial DNA sequences. *Nature*, **347**, 550–553.

Milne RI, Abbott RJ (2002) The origin and evolution of Tertiary relict floras. In: *Advances in Botanical Research Vol. 38.* (eds. Callow JA). pp. 298–301. London: Academic Press.

Miya M, Takeshima H, Endo H *et al.* (2003) Major patterns of higher teleostean phylogenies: a new perspective based on 100 complete mitochondrial DNA sequences. *Mol Phylogenet Evol*, **26**, 121–138.

Miya M, Saitoh K, Wood R, Nishida M, Mayden RL (2006) New primers for amplifying and sequencing the mitochondrial ND4/ND5 gene region of the Cypriniformes (Actinopterygii: Ostariophysi). *Ichthyol Res*, **53**, 75–81.

Murray AM (2001) The oldest fossil cichlids (Teleostei: Perciformes): indication of a 45 million-year-old species flock. *Proc R Soc Lond B Biol Sci*, **268**, 679–684.

Nabholz B, Glemin S, Galtier N (2008) Strong variations of mitochondrial mutation rate across mammals - the longevity hypothesis. *Mol Biol Evol*, **25**, 120–130.

Nelson JS (2006) *Fishes of the World.* John Wiley & Sons, Inc., Hoboken, New Jersey.

Orrell TM, Collette BB, Johnson GD (2006) Molecular data support separate scombroid and xiphioid clades. *Bull Mar Sci*, **79**, 505–519.

Papetti C, Liò P, Rüber L, Patarnello T, Zardoya R (2007) Antarctic fish mitochondrial genomes lack ND6 gene. *J Mol Evol*, **65**, 519–528.

Patterson C (1993) Osteichthyes: Teleostei. In: *The fossil record 2*, pp. 621–656. Chapman & Hall, London, UK.

Posada D (2008) jModelTest: phylogenetic model averaging. *Mol Biol Evol*, **25**, 1253–1256.

Santini F, Tyler JC (2003) A phylogeny of the families of fossil and extant tetraodontiform fishes (Acanthomorpha, Tetraodontiformes), Upper Cretaceous to Recent. *Zool J Linn Soc*, **139**, 565–617.

Schwarz G (1978) Estimating the dimension of a model. *The Annals of Statistics*, **6**, 461–464.

Smith WL, Craig MT (2007) Casting the percomorph net widely: the importance of broad taxonomic sampling in the search for the placement of serranid and percid fishes. *Copeia*, **2007**, 35–55.

Stamatakis A (2006a) Phylogenetic models of rate heterogeneity: a high performance computing perspective. *Proceedings of IPDPS2006, High Performance Computational Biology Workshop, Rhodos, Greece.*

Stamatakis A (2006b) RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics*, **22**, 2688–2690.

Stamatakis A, Hoover P, Rougemont J (2008) A rapid bootstrap algorithm for the RAxML web servers. *Syst Biol*, **57**, 758–771.

Zhuang X, Cheng CHC (2010) ND6 gene "Lost" and found: evolution of mitochondrial gene rearrangement in Antarctic notothenioids. *Mol Biol Evol*, **27**, 1–13.

Zwickl D (2003) *PAUP*. Phylogenetic Analysis Using Parsimony (*and other Methods). Version 4.* Sinauer Associates, Sunderland, MA.

Zwickl D (2006) *Genetic algorithm approaches for the phylogenetic analysis of large biological sequence datasets under the maximum likelihood criterion.* Ph.D. dissertation. Univ. of Texas at Austin.

3.2 Bayesian divergence priors based on probabilities of lineage nonpreservation

Matschiner M

Submitted to *Systematic Biology.*

Version dated: December 14, 2011

Bayesian Divergence Priors Based on Probabilities of Lineage Nonpreservation

MICHAEL MATSCHINER[1]

[1]*Zoological Institute, University of Basel, Vesalgasse 1, CH-4051 Basel, Switzerland*

Corresponding author: Michael Matschiner, Zoological Institute, University of Basel, Vesalgasse 1, CH-4051 Basel, Switzerland; E-mail: michaelmatschiner@mac.com.

Abstract.— Divergence time estimation based on molecular phylogenies and the fossil record has provided insights into fundamental questions of evolutionary biology. However, in such analyses, fossils are frequently applied as minimum ages only. This practice has long been suspected to systematically overestimate divergence dates, and may thus confound conclusions drawn from these studies. Where applied, maximum age constraints for divergence dates are often defined subjectively, which directly affects age estimates. Here, a probabilistic model is developed on the basis of lineage preservation and diversification rates, which can both be estimated objectively from the fossil record and from birth-death models that account for extant species richness. The model is implemented in R package 'ageprior' and can be used to calculate Bayesian divergence prior distributions. The performance of these prior distributions is tested with simulated

phylogenies complemented with artificial fossil records. It is demonstrated that prior distributions calculated on the basis of this model lead to unbiased and robust age estimates, whereas uniform prior distributions substantially overestimate divergence dates. Finally I find that Marshall's bracketing method is extremely sensitive to departures from a strict molecular clock, and that it also tends to overestimated divergence dates. The proposed model is extended to account for a lag time in the preservation of young lineages and for rock outcrop bias, and is applied to a multi-marker phylogeny of teleost fishes. (Keywords: Bayesian divergence time estimation; prior distribution; relaxed molecular clock; fossil record; preservation rate; Teleostei; adaptive radiation; key innovation)

Fossil-constrained molecular divergence date estimation has recently been applied to several fundamental questions in evolutionary biology. Whether or not eutherian mammals have radiated as a response to the extinction of non-avian dinosaurs at the K-Pg (Cretaceous-Paleogene) boundary was investigated with a dated molecular supertree of nearly all extant mammals (Bininda-Emonds et al. 2007; Stadler 2011) and with a molecular supermatrix at the family level (Meredith et al. 2011), whereby the different approaches led to contradicting results. Time-calibrated molecular phylogenies have further been used to explain the biogeography of extant species (e.g. Azuma et al. 2008), to assess ecological limits on clade diversification (Rabosky 2009), and to reveal potential key innovations of adaptive radiations (Matschiner et al. 2011). These studies have in common that the oldest fossil records of included clades are identified and used as age constraints for the split between the given clade and its sister group. As it is increasingly accepted

among molecular biologists that even the oldest known fossil of a clade necessarily postdates the clade origin (Donoghue and Benton 2007; Brown et al. 2008), these are commonly used as minimum age constraints only, and must be complemented with maximum constraints for at least some divergence events in order to obtain a realistic time line of diversification. Unfortunately, maximum constraints cannot be read directly from the fossil record (Wilkinson et al. 2010), and can only rarely be derived from phylogeographic assumptions, such as an endemic diversification subsequent to the formation of the shared habitat (Salzburger et al. 2005). Due to the lack of rigid maximum age constraints, divergence date estimation is frequently performed with only one or very few maximum bounds near or at the root of the tree (Bininda-Emonds et al. 2007; Azuma et al. 2008; Beck 2008), despite a growing consensus that more constraints are generally better than few (Marshall 1990; Benton et al. 2009; Lukoschek et al. 2011). This practice is often regarded 'conservative', but has been suspected to systematically overestimate divergence dates (Pyron 2010; Hugall and Lee 2004), which can be explained by a simple consideration: Assume a phylogeny in which the root is dated correctly, and all other node ages are estimated with molecular data alone. In this case, there should be as many overestimates as underestimates, which means that age estimates would be free of bias and a good accuracy can be assumed. However the precision of node ages would be relatively poor and some of the underestimated dates could postdate the earliest known fossils of the respective clade. If these fossils were now added as strict minimum constraints, some of the underestimated dates would be corrected to agree with the fossil record, and other node ages would indirectly be pushed towards older ages. Thus, precision is gained at the cost of accuracy, and bias towards overestimation of divergence times is introduced. The degree of bias is likely to depend on branch rate variation, and on the number of fossils that are used as minimum constraints only.

Bayesian divergence date estimation, as implemented in the programs BEAST

(Drummond and Rambaut 2007), MCMCTree (Yang 2006), and MrBayes (Ronquist and Huelsenbeck 2003) permits the specification of exponential, lognormal, or gamma prior distributions for divergence dates, with so-called hard minimum, and soft maximum boundaries. These are considered to provide a wider appreciation of the fossil record than uniform priors with hard maximum bound, as they allow for a low probability that the clade origin is in fact older than the specified soft maximum (Ho 2007; Forest 2009; Benton et al. 2009; Wilkinson et al. 2010). However, the definition of both soft or hard maximum age constraints is often done subjectively to reflect the authors' belief that certain divergences are unlikely older than the specified date (Alfaro et al. 2007, 2009a). Several objective criteria for the definition of maximum age constraints have been proposed that are based on stratigraphic bounding, phylogenetic bracketing, or phylogenetic uncertainty. For stratigraphic bounding, the soft maximum is chosen as the age of the youngest deposit that could be expected to contain fossils of the respective clade, but does not. This approach has recently gained popularity (Ksepka et al. 2011), however, the decision whether or not a deposit can be anticipated to bear fossils of a clade is again subjective and can be difficult if the paleodistribution of the clade is uncertain. Phylogenetic bracketing uses the age of fossils of the nearest relatives (Reisz and Müller 2004; Müller and Reisz 2005), and has been criticized (Benton et al. 2009), as the age of these fossils could be very close to (or even younger than) the minimum constraint for a given divergence event, which would falsely suggest that the age of this divergence is known with high precision. Finally, when using phylogenetic uncertainty, the age of described fossils that may or may not be part of the clade, but predate the oldest certain record of the clade, is used as a soft maximum (Meredith et al. 2010). This approach is problematic regardless of whether the taxonomically uncertain fossils are in fact part of the clade or not. If they are, then the true clade age is older than the applied soft maximum, and if they are not, then there is no reason why they should constrain the age of the target clade.

These problems aside, even if appropriate hard minimum and soft maximum constraint are specified, and thus the offset and one additional parameter are fixed for parametric prior distributions, a second parameter remains to be chosen without objective criteria when lognormal or gamma distributions are used. Age estimates have been found highly sensitive to the choice of parameters for prior distributions (Inoue et al. 2010; Warnock et al. 2011), and thus objective specification of all parameters would be desirable.

By the use of a mathematical model of lineage preservation and diversification (Foote et al. 1999; Tavaré et al. 2002), not only hard minimum and soft maximum constraints, but the entire shape of divergence prior probability distributions could be defined in an objective manner (Benton and Donoghue 2007; Inoue et al. 2010). The model developed by Foote et al. (1999) determines the probability of nonpreservation of a clade for a given temporal gap between clade origin and the age of the first fossil, under the assumption of constant preservation and net diversification rates. Both parameters can be estimated objectively from the fossil record (Foote and Raup 1996; Foote 1997; Foote et al. 1999), and from birth-death models that take into account the extant species richness (Alfaro et al. 2009b). Here, the model of Foote et al. (1999) is extended to incorporate fossil age uncertainty, a lag time for the preservation of young lineages, and rock outcrop bias. The model is implemented in the new package 'ageprior' of the R programming environment (R Development Core Team 2011), and can be used to calculate parameters of divergence prior distributions based on probabilities of lineage nonpreservation. The performance of these prior distributions is tested with simulated phylogenies including an artificial fossil record, and the extended model is applied to a multi-marker phylogeny of teleost fishes.

Probability Distributions of Lineage Nonpreservation

The probability distribution of lineage nonpreservation is directly based on the preservation rate r, which implicitly incorporates preservation and recovery (Foote et al. 1999). Given lineage duration δ and preservation rate r, the probability for lineage nonpreservation is $P' = e^{-r\delta}$. By definition, a lineage is unpreserved between its origin at time t and the age of its oldest known fossil t_f, and so the probability of nonpreservation can be rewritten as a probability distribution for the time of lineage origin $P'_t = e^{-r(t-t_f)}$. For divergence dates of extant species, at least two lineages must have been present between the divergence event and the age of the oldest fossil in one of the descending lineages. Thus, the sum of lineage durations S is greater than or equal to $2 \times (t - t_f)$, and the combined probability of nonpreservation becomes $P'_t = e^{-rS}$ (Foote et al. 1999). The sum of lineage durations S depends not only on the time between divergence and the oldest fossil $t - t_f$, but also on the diversity D_t of each descending clade: $S_t = 2 \times \int_{t_f}^{t} D_t \, \mathrm{dt}$. Assuming an exponential diversification with constant speciation rate p and extinction rate q, the net diversification is $p - q$, and diversity D_t is equivalent to $D_t = e^{(p-q) \times (t-t_f)}$. This leads to

$$P'_t = e^{-\frac{2r}{p-q} \times (e^{(p-q) \times (t-t_f)} - 1)}, \ t \geq t_f. \tag{1}$$

P'_t is defined as $P'_t = 0$ for $t < t_f$. The normalized probability distribution is

$$P_t = \frac{P'_t}{\int_{t_f}^{\infty} P'_t}, \ t \geq t_f \tag{2}$$

that a given divergence event occurred at time t, which must be at least as old as the age of the earliest fossil t_f that can be assigned to one of the descending lineages. This prior distribution is based on the assumption that fossil age information is absolute and without

errors, which is rarely met in practice, and which I will relax in the following. Fossil age estimates derive from the age of the geological formation in which the respective fossil has been found. The age of a geological formation is usually defined as a time range between a younger date t_y and and older age t_o. Note that I will avoid using 'upper' and 'lower' in this study, because paleontologists and neontologist use these terms in exactly opposite ways (Inoue et al. 2010). Both dates represent estimates that have errors, which usually are comparatively small and without direction (Gradstein and Ogg 2009), and therefore neglected for the present purpose. In many cases, detailed information on the position of the fossil within the formation is not available, and the true age t_f of a given fossil can then be assumed to lie with uniform probability between the formations younger and older boundaries: $t_y \leq t_f \leq t_o$. If multiple representatives of a clade are preserved in different layers of the same geological formation, or in different formations with overlapping age estimates, the probability distribution for the age of the clade's oldest fossil is skewed toward older dates. However, the degree of skew is difficult to estimate as the occurrences of the clade's representatives are unlikely to be independent. Therefore, I here ignore the effect of multiple representatives. The divergence date probability distribution R_t is then proportional to

$$
\begin{aligned}
R'_t &= \int_{t_y}^{t} P_t \, \mathrm{dt_f} \text{ for } t_y \leq t \leq t_o; \\
&= \int_{t_y}^{t_o} P_t \, \mathrm{dt_f} \text{ for } t > t_o.
\end{aligned}
\tag{3}
$$

R'_t is defined as $R'_t = 0$ for $t < t_y$. The integral of R'_t is then $\int R'_t = t_o - t_y$, and the normalized divergence date probability distribution R_t becomes

$$
R_t = \frac{R'_t}{t_o - t_y}.
\tag{4}
$$

Probability distributions P_t and R_t are implemented in function 'find.prior' of R package 'ageprior', and calculated on the basis of parameters t_y, t_o, r, and $p - q$. P_t is used when a precise fossil age is known ($t_y = t_o$), otherwise R_t is applied. The distribution mean t_m is calculated as

$$t_m = \frac{t_o - t_y}{2} + \int t \times P_t \, \mathrm{dt}. \tag{5}$$

Using 10 steps per time unit (usually 1 Myr), distribution R_t (or P_t) is discretized in function 'find.prior' and approximated by discretized exponential, lognormal, and gamma distributions by minimization of root mean square deviation (RMSD). In each case, offset t_y and mean t_m are fixed to match those of R_t (or P_t). This fixation allows only one exponential distribution with $\lambda = 1/(t_m - t_y)$, however, lognormal and gamma distributions are defined by three parameters (including the offset), and thus allow multiple possibilities that match the given constraints of t_y and t_m. Therefore, the software exhaustively tests 130 values of σ between 0.01 and 1.30 for lognormal distributions, and 500 values for the gamma distributions' shape parameter between 0.01 and 5.00. These parameter boundaries were chosen according to initial tests in which optimal values were always found between these limits. Parameter μ of lognormal distributions, and the scale parameter of gamma distributions are calculated so that the mean of each distribution is equal to t_m. Finally, parameters of the overall best exponential, lognormal, or gamma distribution (Fig. 2A) are reported and can be used as divergence priors in molecular dating studies with BEAST (Drummond and Rambaut 2007), MCMCTree (Yang 2006), MrBayes 3.2 (Ronquist and Huelsenbeck 2003), or the forthcoming program RevBayes (Höhna 2011, priv. comm.)

Application to Simulated Phylogenies

Generation of datasets

The divergence prior distributions described above use fossil ages as hard constraints for minimum ages and soft maximum bounds are provided on the basis of preservation rate and net diversification rate. Using simulated phylogenies, the performance of these prior distributions was tested and compared to uniform prior distributions that use fossils as hard minimum constraints only. A constant birth-death model (Gernhard 2008) was used for tree generation. Simulations started with a single initial species and were conditioned to result in 20 extant taxa after 100 time units, each subdivided into 10 discrete time steps. The first speciation event was set to occur at the start of the simulation, and at each subsequent time step, branches originated or terminated with constant probability. In independent simulations, three different speciation rates p were applied ($p = 0.03$, 0.05, and 0.07 per time unit), and extinction rates q were set to result in net diversification rates $p - q$ of 0.02 per time unit ($q = 0.01$, 0.03, and 0.05 per time unit). The three combinations of speciation rate p and extinction rate q led to different values for the relative extinction (turnover) rate ϵ which determines the distribution of speciation times in a dated phylogeny, and thus the overall tree shape (Rabosky 2010). As the relative extinction rate is defined as $\epsilon = q/p$, my simulations tested $\epsilon = 0.33$ (low turnover), 0.6 (intermediate turnover), and 0.71 (high turnover). Dramatically different tree shapes corresponding to $\epsilon = 0.71$ and $\epsilon = 0.33$ are shown in Fig. 1.

The ancestral nucleotide sequence was produced at random, whereby 3000 nucleotides were drawn with probabilities 0.3 for A and T, and 0.2 for G and C, to result in GC content around 0.4, a common value in vertebrate genomes. Sequence evolution took place according to the general time reversible model (GTR). Relative individual substitution rates were taken from an empirical study (Matschiner et al. 2011), and set to AC = 0.040, AG = 0.159, AT = 0.060, CG = 0.019, CT = 0.196, and GT = 0.025. Sequences were partitioned into 3×1000 nucleotides, with different overall substitution

rates to simulate rate variation between codon positions. Relative overall rates between partitions were 0.20, 0.25, and 0.55, meaning that over half of all substitutions occurred in the third partition.

Finally, rate variation among branches was accounted for by drawing absolute branch rates from a lognormal distribution. This distribution directly influences the overall variation in the resulting sequences, and therefore affects phylogenetic inference. Lower rates result in a small number of phylogenetically informative sites, while high rates lead to greater amounts of homoplasies, impeding phylogenetic inference. Thus, a suitable scaling for this lognormal distribution was assessed by running sequence evolution simulations on a fixed pure birth phylogeny, with fixed relative branch rates, drawn from this distribution. In 15 independent simulations, a normal distribution with mean 0.0 and standard deviation 0.5 was log transformed and scaled with a factor increasing from 0.1 to 1.5, in steps of 0.1. After each sequence evolution simulation, the resulting 20 sequences were used for phylogenetic reconstruction with the software GARLI v.0.97 (Zwickl 2006), using 100 Bootstrap (BS) replicates. Hereby, the alignments were partitioned into 3×1000 nucleotides to account for the simulated rate variation among codon positions, and the GTR + Γ model of sequence evolution was applied for every partition, as suggested (in most cases) by the Bayesian Information Criterion (BIC; Schwarz 1978), implemented in the software jModelTest v.0.1.1 (Guindon and Gascuel 2003; Posada 2008). These simulations with increasing scale factor showed that the number of incorrectly identified nodes increases with scale factors > 0.3 per time unit, and that BS support for correct nodes decreases with increasing mean rate, suggesting homoplasies. According to these results, a factor of 0.3 was chosen to scale the lognormal distribution of rates for all subsequent simulations. With this scaling, the alignment contained 85.6% variant sites. Using 0.3 as a fixed scaling factor for the lognormal distribution of substitution rates, the effect of low, intermediate, and high rate variation was further explored by changing the

width of this distribution. To this end, the standard deviation was set to 0.25 (low rate variation), 0.5 (intermediate rate variation), and 1.0 (high rate variation) before log transformation. Lineages of simulated phylogenies were complemented with an artificial fossil record that was generated at random with low and high preservation rates ($r = 0.02$ and $r = 0.05$ per time unit) (see Fig. 1A,D,G,J). Hereby, the age and the taxonomic position of simulated fossils were known without error.

In summary, data sets included nine different combinations of speciation rates and branch rate variation, that were each tested with both low and high fossil preservation rate. For each setting, 30 independent simulation replicates were performed, resulting in a total of 540 data sets.

Divergence date estimation

Divergence dates were estimated with the software BEAST v.1.6.1 (Drummond and Rambaut 2007), using the birth-death tree prior (Gernhard 2008), and the relaxed uncorrelated lognormal molecular clock. The alignment was partitioned into simulated codon positions (3×1000 nucleotides), and the GTR $+\ \Gamma$ model of sequence evolution was applied to each partition. In order to save computation time, the tree topology was constrained to match the correct simulated tree, as this study focusses on the correctness of divergence age estimation, and not on topological uncertainties. The root age was fixed to 100.0 time units for all analyses. Whereas the use of error-free age constraints has long been criticized (Graur and Martin 2004), complete fixation of root ages is still frequently applied in molecular dating studies (Bininda-Emonds et al. 2007; Meredith et al. 2011). Here, the fixed root age constraint is used to provide a maximum bound for uniform prior distributions. It is equally applied when prior distributions were based on probabilities of lineage nonpreservation in order to allow the comparison of the performance of both distribution types. However, because BEAST does not allow complete fixation of node

ages, the root was constrained with a very narrow normal prior with mean 100.0, and standard deviation 0.1. The oldest fossil of each clade, including the stem group, was identified and used to constrain the age of this clade, either with uniform prior distributions between fossil age and root age, or with prior distributions based on probabilities of lineage nonpreservation. For calculation of the latter with function 'find.prior' of R package 'ageprior', the known values of preservation rate $r = 0.02$ or $r = 0.05$, and net diversification rate $p - q = 0.02$ were used. Accordingly, exponential prior distributions with offset t_f and mean $t_m = 14.59$ or $t_m = 7.52$ time units were used for all nodes when $r = 0.02$ and $r = 0.05$, respectively. These priors were applied even to clades without fossils, using $t_f = 0$ in these cases, as probabilities of nonpreservation apply equally to these lineages. For the setting with the greatest discrepancy between results obtained with uniform and exponential priors (high rate variation, high speciation; Fig. 1D–F,J–L), divergence date estimation was repeated with over- (200%) and underestimated (50%) values for r and $p - q$ in order to evaluate the robustness of the model.

To test the effect of fossil age uncertainty, all 540 data sets were reanalyzed after fossil ages were sorted into 10 bins that each covered 10 time units. For each clade's earliest fossil, the younger and older bin boundaries were used as t_y and t_o. Uniform priors were defined between its younger bin boundary t_y and the root age of 100 time units, and priors based on probabilities of lineage nonpreservation were again calculated with function 'find.prior', using t_y, t_o, and the known values for r and $p - q$. With these settings, the closest approximation to probability distribution R_t was provided by a gamma distribution with offset t_y, shape parameters $k = 2.37$ and $k = 3.10$, and scale parameters $\Theta = 8.26$ and $\Theta = 4.04$, when preservation rates were $r = 0.02$ and $r = 0.05$, respectively. Again, these distributions were applied to all nodes, including clades without fossils. All BEAST analyses were performed with 20 million generations of the Monte Carlo Markov Chain (MCMC), which was sampled every 2000 generations.

In addition to divergence date estimation with uniform priors and priors based on probabilities of lineage nonpreservation, all 540 data sets were analyzed with the method proposed by Marshall (2008). To this end, uncalibrated ultrametric trees were produced for all alignments, again by running the software BEAST for 20 million generations with fixation of the known tree topology. After discarding the first 2 million generations as burnin, mean heights were calculated for each node, and divided by the age of the earliest fossil constraining this node to obtain Marshall's (2008) empirical scaling factor s_i. The lineage with the highest empirical scaling factor was chosen as the calibration lineage, and its median age is derived from equation 14 of Marshall (2008), using $C = 0.5$, $\mathrm{FA_{cal}} = t_f$, and the total number of fossils per simulated data set as $n\overline{H}$. All other node age estimates are obtained by scaling of the ultrametric tree so that the node height of the calibration lineage matches its calculated median age.

Results

Convergence of BEAST runs was assessed with software TRACER (Rambaut and Drummond 2011), and confirmed by effective sample sizes (ESS) > 200 for nearly all parameters. The comparison of true and estimated node ages showed that uniform prior distributions between fossil age and root age tend to overestimate divergence dates (Fig. 1B, E, H, K, SI Fig. S1, S2), and that the mean overestimate can be as high as 30–40%. The degree of overestimation depends mainly on branch rate variation, but is exacerbated with higher speciation and preservation rates. The precision of estimation decreases with branch rate variation, but increases slightly with higher speciation rates. With uniform prior distributions, the RMSD between true and estimated node ages is between 2.28 and 11.70 time units, the true node age is contained within 62–82% of the 95% highest posterior density (HPD) intervals reported by BEAST, and the average width of 95% HPD intervals lies between 4.01 and 15.52 time units (SI Fig. S1, S2).

Analysis of the same data sets with divergence prior distributions based on probabilities of lineage nonpreservation yields age estimates that are free of bias in all tested settings. Estimated node ages are closer to true node ages than with uniform priors (RMSD 1.85–6.52 time units) and the true node age is more often contained within the reported 95% HPD interval (70–84%), even though the average widths of 95% HPD intervals are smaller (3.57–13.03 time units) (Figs. 1C, F, I, L, SI Figs. S3, S4). As with uniform priors, precision decreases with branch rate variation. The best results were obtained with low rate variation, high speciation rate, and high preservation rate (SI Fig. S4J). The robustness of prior distributions based on probabilities of lineage nonpreservation has been tested with false assumptions for preservation rate r and net diversification rate $p - q$ for all data sets with high branch rate variation and high speciation rate. When r and $p - q$ are over- (200%) or underestimated (50%), divergence date estimates are still better than those obtained with uniform priors in the same setting. The RMSD range between 3.81–6.15 time units (7.99–8.80 with uniform priors), 62–73% of 95% HPD intervals contain the true node age (66–68% with uniform priors), and the average widths of 95% HPD intervals range between 5.71 and 10.81 time units (12.32–12.44 with uniform priors) (SI Figs. S1L, S2L, S5, S6).

Results are generally similar when fossil ages are sorted into bins of 10 time units, and bin boundaries are used for dating instead of the exact fossil age. Using uniform priors between the younger bin boundary t_y and the root age, RMSD range between 2.24–11.22 time units, 62–81% of 95% HPD intervals contain the true node age, and the average widths of 95% HPD intervals lie between 4.39–16.00 time units. As with exactly known fossil ages, uniform priors show a tendency to overestimate divergence dates, and branch rate variation decreases the precision of estimates. With gamma distributions as divergence priors, RMSD are lower (2.09–6.70) and more 95% HPD intervals contain the true node age (70–82%) while being less wide (4.09–14.08) (SI Figs. S7–S10).

Divergence date estimates of Marshall's (2008) method cannot be directly compared to those obtained with Bayesian prior distributions, because Marshall's method does not allow specification of root age, which was known here, and used as a constraint in all other analyses. Nevertheless, this method appears to be highly sensitive to branch rate variation and results in mean overestimates of divergence date that are as high as 50–100% (SI Figs. S11, S12). Correspondingly, RMSD vary between 3.55 and 6.06 time units when branch rate variation is low, but are between 26.94 and 58.09 time units when branch rate variation is high. This demonstrates that the method relies heavily on an ultrametric tree with accurate relative branch lengths (Marshall 2008), for which a clock-like substitution rate would be necessary. With high branch rate variation, the calibration lineage is correctly identified in as few as 13–47% of the data sets when mean node heights of the ultrametric tree are used, or in 14–39% of the entire Bayesian samples of 9000 uncalibrated ultrametric trees per replicate. However, even with the settings chosen for low branch rate variation, an incorrect calibration lineage is identified in 20–70% of the data sets when using mean node heights, or in 45–62% of the Bayesian tree sample (SI Figs. S11, S12).

Model Extensions

Lag time for preservation of young lineages

Divergence prior distributions P_t and R_t assume a constant preservation rate per lineage between clade origin and the age of the clade's earliest fossil record. There are several reasons why this assumption is unlikely to be met in nature. The preservation rate depends on population size (Paul 2009) and geographic distribution, both of which may be small in recently diverged lineages (Tavaré et al. 2002; Donoghue and Benton 2007; Benton et al. 2009). The preservation rate further incorporates discovery and correct taxonomic

assignment of fossils. However, the oldest fossils of a clade may often lack the diagnostic crown group features, and are thus more prone to misidentification (Benton et al. 2009; Ksepka et al. 2011). On the other hand, the divergence of genetic markers may even predate the true speciation time, thus further increasing the time lag between genetic divergence and the first recognizable fossil occurrence (Brown et al. 2008). In order to account for the assumed lower preservation rate of young lineages, divergence prior distributions P_t and R_t are modified by a lag time parameter. Parameter lag represents the mean (in Myr) of an exponential function F for the impossibility of recognizable fossils in young lineages

$$F_{\Delta t} = e^{-\frac{1}{lag}\Delta t}, \Delta t > 0, \tag{6}$$

where Δt is the time since genetic divergence. Distribution P_t' is multiplied with $1 - F_{\Delta t}$ (here, $\Delta t = t - t_f$, $t > t_f$) prior to calculation of the normalized probability distributions P_t (Equation 2) and R_t (Equations 3,4). This leads to a shift of the means t_m of both distributions P_t and R_t towards older dates by the value of lag. With $lag > 0$, P_{t_f} becomes 0 and the shape of distribution P_t resembles that of lognormal or gamma distributions. The peak of distribution R_t is smoothed and predates t_o (Fig. 2B).

Application of a positive lag parameter is recommended even if no fossils are known to constrain a given divergence event, since the most likely divergence date for extant sister taxa (with or without a fossil record) is some time in the past, not the present (Weir and Schluter 2007; Bolnick and Near 2008). In part, this is due to the fact that speciation can only be recognized in retrospect and may collapse as long as diverging populations are connected by gene flow (Nosil et al. 2009). In addition, very young species are likely to lack diagnostic features, and may not be recognized as such (Purvis 2008).

Correction for rock outcrop bias

Preservation of a fossil is a complex process that includes fossilization, survival as a intact fossil in rock over millions of years, outcropping of deposits, discovery, identification, and description (Foote et al. 1999; Paul 2009). Lower probabilities for the fossilization of very young lineages (and for the correct identification of these) are accounted for by the *lag* parameter introduced above. Once fossils are embedded in rock, they are subject to geological activity, which may decrease fossil quality linearly with age (Benton et al. 2011), and thus reduce the probability of ancient fossils to be correctly identified. This bias is difficult to quantify and here considered small over the time spans for which preservation rates are usually calculated from the fossil record (e.g. the Cenozoic; Foote et al. 1999). On the other hand, the surface area of potentially fossil-bearing rock outcrop has long been known to be highly variable between stages (Raup 1972, 1976) and directly influences stage-specific preservation rates. Sedimentary rock outcrop bias has been quantified at global and regional scales, by counting the number of geological maps that include rock outcrop of a given stage (Smith and McGowan 2007; McGowan and Smith 2008), and by a direct reading of rock area mapped in the Geological Atlas of the World (Choubert and Faure-Muret 1976; Wall et al. 2009). Both types of quantifications distinguish between marine and terrestrial sediments. Corrections schemes for global and regional rock outcrop areas have been implemented in R package 'ageprior' and can be used to adjust divergence prior distributions P_t and R_t. The effect of correction for rock outcrop bias is demonstrated in Fig. 2: divergence date prior probability decreases faster in the Late Cretaceous than in the Early Cretaceous (Fig. 2D), because rock outcrop area of the latter (scaled by Myr) is about half that of the former (Fig. 2C). More specifically, calculation of distributions P_t and R_t starts at t_y using the preservation rate that applies for the stage of t_y. Going back in time, distributions are calculated anew with adjusted preservation rate, wherever the distributions cross a stage boundary with a change in rock outcrop area. Individual stage-specific distributions are truncated to stage boundaries, and combined with truncated

distributions of older stages, whereby the older distribution fragment is always scaled to match the height of the younger fragment at the stage boundary. Composite distributions are normalized, and mean ages t_m are adjusted and used to find the exponential, lognormal, or gamma distribution that best approximates the modified distributions P_t and R_t.

APPLICATION TO TELEOST FISHES

Multi-marker phylogeny

Prior distributions based on probabilities of nonpreservation were used to estimate divergence dates of teleost fishes. The multi-marker data set of Matschiner et al. (2011) was complemented with additional sequences to include members of 31 monophyletic groups that in combination represent nearly the entire extant diversity of acanthomorph fishes (see SI Table S1). Furthermore, six otomorph, protacanthopterygiid, and aulopiform representatives were added as outgroups (SI Table S2). Genbank, Ensemble, and Genoscope accession numbers are listed in SI Tables S3 and S4, and include seven syngnathoid sequences that were newly produced as part of this study (JN696468–JN696474). Sequence generation, alignment, and model selection followed protocols outlined elsewhere (Matschiner et al. 2011). Due to increased alignment size, the GTR + Γ model of sequence evolution was now selected by jModelTest (Posada 2008) for four partitions based on codon position and molecule type (Matschiner et al. 2011). The optimal tree topology was explored in two initial analyses with BEAST ('*BeastRun1*' and '*BeastRun2*'), using the same tree prior, clock models, and chain length as for simulated phylogenies, but running 10 replicates per analysis. Preliminary divergence prior distributions used in these exploratory analyses were calculated with function 'find.prior' of R package 'ageprior', assuming preservation rate $r = 0.01$ per Myr, net diversification rate

$p - q = 0.02$ per Myr, and a lag time parameter $lag = 2$ Myr (see SI Text S1). The resulting tree topology is shown in SI Fig. S13, and Bayesian Posterior Probability (BPP) support values for all nodes are listed in SI Table S5. In brief, tree topology agreed well with a recent taxonomic classification based on monophyletic groups (Wiley and Johnson 2010), and confirms most aspects of previous molecular phylogenies (discussed in detail in SI Text S2).

Divergence date estimation

The tree topology resulting from exploratory analyses *BeastRun1* and *BeastRun2* (SI Fig. S13) was fixed for the subsequent refinement of divergence date estimation with prior distributions based on probabilities of lineage nonpreservation. The fossil record of teleost fishes was analyzed rigorously, and the earliest known fossils could be identified for 77 of 106 divergence events (discussed in SI Text S2). In contrast to previous studies, neotropical cichlid fishes from the Lumbrera Formation, Argentina, were not accepted as the world's oldest cichlid records (Malabarba et al. 2006; Perez et al. 2010; Malabarba et al. 2010), and instead the age of †*Mahengechromis* spp. (Murray 2000) was used to constrain cichlid divergences. Also, the assignment of †*Archaeotetraodon winterbottomi* as a crown group tetraodontid (Benton and Donoghue 2007; Benton et al. 2009) is questioned on the basis of recent evidence confirming monophyly of the extinct genus †*Archaeotetraodon* (Carnevale and Tyler 2010) (see explanations in SI Text S2).

For each divergence event, the youngest (t_y) and oldest (t_o) possible age were determined for the earliest record of the two descending lineages. If no fossils were known, $t_y = t_o = 0$ were used. The most likely paleoenvironment (marine and/or freshwater) was identified based on sediment type of the formation bearing the clade's oldest record, and on the habitat of extant members of this clade (Nelson 2006). In order to investigate the effect of preservation rate on divergence dates, four different base preservation rates were

assumed to apply to Cenozoic marine teleosts: 0.03, 0.01, 0.003, and 0.001 per Myr. Stage-specific preservation rates in marine and freshwater habitats were derived from these base preservation rates with function 'adjust.pres.rate' of R package 'ageprior', using the appropriate correction for global rock outcrop bias (SI Text S2). The net diversification rate $p - q$ of acanthomorph fishes was estimated with the software MEDUSA (Harmon et al. 2008; Alfaro et al. 2009b) using pruned phylogenies and extant species richness of 31 monophyletic groups (Nelson 2006) (see SI Tables S1, Text S2). Note that this estimation of net diversification rates is to some extent circular, as it is based on time-constrained phylogenies, that were in turn inferred with divergence priors based on net diversification rate. Therefore, divergence date estimation with the software BEAST was performed iteratively (*BeastRun3a–d*, *BeastRun4a–d*, where *a–d* denote the different assumed base preservation rates), with net diversification rates based on phylogenies resulting from the previous round of inferences. It can be shown that estimates for divergence dates and net diversification rates converge quickly, regardless of the assumed base preservation rate (SI Table S6). Settings of BEAST analyses *BeastRun3a–d* and *BeastRun4a–d* were identical to those of *BeastRun1* and *BeastRun2*, except for divergence prior distributions. All divergence prior distributions of *BeastRun3a–d* and *BeastRun4a–d* were calculated with function 'find.prior' of R package 'ageprior', using stage-specific preservation rates, net diversification rates identified with MEDUSA, lag parameter $lag = 2$ Myr, and global corrections for rock outcrop bias (SI Text S2). The net diversification rates of basal outgroups were assumed to be equal to those of basal acanthomorph fishes.

Results

Analysis of diversification patterns with the software MEDUSA recognized a single major increase in speciation and extinction rates of acanthomorph fishes. In all used phylogenies (SI Table S6), this shift is identified in the stem leading to a clade combining

all Percomorpha except "Syngnathoidei", Dactylopteriformes, and Ophidiiformes (Fig. 3, node X; classification follows Wiley and Johnson 2010). This is largely congruent with the change of diversification patterns observed by Alfaro et al. (2009b) at the base of Percomorpha (there excluding Ophidiiformes and Scombriformes), with the only difference that Scombriformes were found basal to the faster-diverging clade in the phylogeny of Alfaro et al. (2009b). When phylogenetic trees of *BeastRun4a–d* were used, the estimated net diversification rates $p - q$ ranged between 0.044 and 0.045 per Myr in basal acanthomorph fishes, and between 0.056 and 0.073 in percomorph fishes descending from node X (Fig. 3). These values are comparable to net diversification rates estimated by Alfaro et al. (2009b) (basal acanthomorph fishes: 0.039, Percomorpha: 0.082 per Myr).

Relative extinction (turnover) rates ϵ appeared more variable and were estimated as 0.374–0.728 and 0.000–0.923, in basal acanthomorphs and descendants of node X, respectively (SI Table S6). Surprisingly, lower net diversification rates were estimated for generally younger phylogenies (*BeastRun4a,b*), presumably as a result of the dramatically different tree shapes indicated by relative extinction rates. These results suggest that the more loosely constrained phylogeny obtained with the lowest assumed base preservation rate $r = 0.001$ (*BeastRun4d*) provides a better fit to pure birth Yule trees, whereas more complex tree models are required to match patterns of more narrowly constrained phylogenies resulting from higher assumptions for the base preservation rate r.

The degree to which divergence date estimates depend on preservation rate was found to be highly variable. In the most extreme case, mean divergence dates estimated with $r = 0.03$ (*BeastRun4a*) and $r = 0.001$ (*BeastRun4d*) differed by as much as 34.87 Myr (48.26–83.19 Ma; common ancestor of Labrini and Julidini, SI Table S7). The age of stem group Tetraodontiformes, on the other hand, is estimated nearly independently of preservation rate (mean ages 98.66–99.56 Ma; node C in Fig. 3,4C, SI Table S7), and can apparently be determined with very high precision (95% HPD intervals range between

99.74–97.83 and 101.78–97.86 Ma). Thus, even despite the use of a lag parameter of 2 Myr, the estimated origin of stem Tetraodontiformes predates the minimum age of the earliest known tetraodontiform fossil †*Plectocretacicus clarae* Sorbini, 1979 by only 0.86–1.76 Myr. This suggests that age estimates for all Percomorpha, but especially those of the clade combining Tetraodontiformes, Nototheniiformes, "Labriformes" II, and other groups (see Fig. 3, node 44 in SI Fig. S13) are highly dependent on the correct taxonomic assignment of †*Plectocretacicus clarae*. This fossil species is known from five diminutive specimens from the Lithographic Limestone of Hakel, Lebanon (Patterson 1993; Tyler and Sorbini 1996), the age of which can be constrained as 99.6–97.8 Ma (Benton et al. 2009; see SI Text S2). It has frequently been cited as the oldest known percomorph (Patterson 1993; Benton et al. 2009) and has been used to constrain divergence dates in a large number of molecular dating studies (e.g. Azuma et al. 2008; Alfaro et al. 2009b; Santini et al. 2009; Miya et al. 2010; Matschiner et al. 2011). Its stem tetraodontiform assignment has been corroborated by an extensive phylogenetic analysis based on morphological characters of fossil and extant members of the clade (Santini and Tyler 2003). While the analysis of Santini and Tyler (2003) employed a questionable outgroup choice, reanalyses of their data set with multiple outgroups and both parsimony and Maximum Likelihood inference (SI Text S2) strongly supports the tetraodontiform position of †*Plectocretacicus clarae* and younger members of the fossil superfamily Plectocretacicoidea. In summary, †*Plectocretacicus clarae* seems to be one of the most reliably assigned Cretaceous acanthomorphs, that fossilized shortly after the origin of Tetraodontiformes, and thus provides a unique and well-constrained anchor point for time-calibrated percomorph phylogenies.

Divergence dates estimated for nototheniiform fishes are similar or slightly younger than those determined by Matschiner et al. (2011) (Fig. 4D–F). Depending on the assumed base preservation rate, mean estimates for the onset of the Antarctic notothenioid radiation range between 24.80 and 18.06 Ma, with 95% HPD intervals between 19.53–30.20

and 13.49–22.72 (Fig. 4F, SI Table S7). The genus-level preservation rate of Phanerozoic "Osteichthyes" has been quantified by Foote and Sepkoski Jr (1999) as 0.15 per 5.5 Myr, or 0.30 per 5.3 Myr, using two different time scales. Assuming that the average species richness of extant bony fish genera (6.27 species per genus; Nelson 2006) equals that of extinct genera, this would translate to Phaenerozoic species-level preservations rates of 0.0044–0.0092 per Myr. After correction for the lower global rock outcrop area per Myr in the Phanerozoic ($64\,700$ km^2/Myr) compared to the Cenozoic ($130\,500$ km^2/Myr; Wall et al. 2009), and assuming that the observed preservation rate refers mostly to marine genera, this suggests that the preservation rate of Cenozoic marine teleosts is on the order of 0.0089–0.0185 per Myr. If this is correct, then teleost divergence date estimates based on an assumed base preservation rate of 0.01 per Myr may be closest to the true ages. This would suggest onset of the Antarctic notothenioid radiation in the Early Miocene. It has been hypothesized that this radiation was triggered by the key innovation of antifreeze glycoproteins at a time of cooling of the Southern Ocean (Eastman 1993; Matschiner et al. 2011). With an assumed base preservation rate of 0.01, the mean age estimate for the radiation onset coincides with the Mi1a cold event (Fig. 5), during which Antarctic glaciation was even more extensive than today (Pekar and DeConto 2006). Thus, the key innovation hypothesis of antifreeze glycoproteins is supported by the results presented here (Matschiner et al. 2011).

Conclusions

Comparison of true and estimated node ages has demonstrated that Bayesian divergence prior distributions based on probabilities of lineage nonpreservation lead to unbiased and robust results, whereas the use of fossil ages as strict minimum constraints can severely overestimate divergence dates. The degree of overestimation mostly depends

on branch rate variation, but also increases with turnover and preservation rate. These results cast doubt on conclusions drawn from molecular dating studies, including a proposed Late Cretaceous radiation of eutherian mammals (Bininda-Emonds et al. 2007), as high branch rate variation and turnover seem to be common in nature (Tavaré et al. 2002; Drummond et al. 2006; Alfaro et al. 2009b; Stadler 2011; Morlon et al. 2011). In addition, the method proposed by Marshall (2008) is found to be extremely sensitive to departures from the strict molecular clock. However, even with very low branch rate variation, Marshall's method frequently failed to identify the correct calibration lineage, and thus resulted in divergence date overestimates. The same is true when the entire Bayesian posterior sample of ultrametric trees is used instead the single tree with mean node heights, and thus the Bayesian extension to Marshall's method, developed by Dornburg et al. (2011), may be equally sensitive to departures from a strict molecular clock.

I have demonstrated how a rigorous analysis of the fossil record, in combination with analyses of net diversification rate can be used to obtain unbiased divergence date estimates, when a preservation rate of the investigated clade is available. These have been quantified for groups as diverse as mammals (Foote et al. 1999), bivalves (Foote and Raup 1996), trilobites (Foote and Raup 1996), tetrapods (Friedman and Brazeau 2011), and fishes (Foote and Sepkoski Jr 1999), among others, and divergence prior distributions based on probabilities of nonpreservation can thus be readily applied to these groups. New preservation rates for additional groups can be calculated with the freqRat method of Foote and Raup (1996). Furthermore, the growing PaleoDB data base (www.paleodb.org) may become suitable for fast and flexible quantification of the preservation rate of a given taxonomic group.

The remarkable precision of divergence date estimates based on the probabilistic model presented here derives from the fact that most, if not all, nodes of a phylogeny can (and should) be constrained, as probabilities of nonpreservation apply to all lineages.

Rigorous analysis of the fossil is necessary to identify the earliest known fossil of every clade (Parham and Irmis 2008), which may be tedious, especially for extensive phylogenetic datasets. To facilitate and guide future molecular dating studies, a second paleontological data base is scheduled to launch in 2012 and will contain a collection of vetted fossil calibrations that are readily applied to divergence date estimation (Ksepka et al. 2011).

Given that fossil age uncertainty is implemented in the presented model, fossil cross-validation (Near et al. 2005; Marshall 2008; Matschiner et al. 2011) is of lesser importance when an extensive set of constraints is used. Wherever doubt remains concerning the earliest record of a clade, the fossil age uncertainty can be defined to encompass the maximum age of the oldest putative fossil (t_o), and the minimum age of the oldest fossil with certain taxonomic position (t_y). The posterior probability distribution for this divergence date will then provide a suggestion whether or not the older putative fossil had been correctly assigned. In this case, it may be preferable not to use a continuous uncertainty between t_o and t_y, but two separate age uncertainties between t_o and t_y of the older putative fossil, and t_o and t_y of the younger certain record. The combination of the two resulting divergence prior distributions would be a bimodal one, which is not yet available in Bayesian divergence date estimation software, but could be subject of future software developments.

Acknowledgements

*

References

Alfaro, M. E., C. D. Brock, B. L. Banbury, and P. C. Wainwright. 2009a. Does evolutionary innovation in pharyngeal jaws lead to rapid lineage diversification in labrid fishes? BMC Evol Biol 9.

Alfaro, M. E., F. Santini, and C. D. Brock. 2007. Do reefs drive diversification in marine teleosts? Evidence from the pufferfishes and their allies (order Tetraodontiformes). Evolution 61:2104–2126.

Alfaro, M. E., F. Santini, C. D. Brock, H. Alamillo, A. Dornburg, D. L. Rabosky, G. Carnevale, and L. J. Harmon. 2009b. Nine exceptional radiations plus high turnover explain species diversity in jawed vertebrates. P Natl Acad Sci USA 106:13410–13414.

Azuma, Y., Y. Kumazawa, M. Miya, K. Mabuchi, and M. Nishida. 2008. Mitogenomic evaluation of the historical biogeography of cichlids toward reliable dating of teleostean divergences. BMC Evol Biol 8:215.

Beck, R. M. D. 2008. A dated phylogeny of marsupials using a molecular supermatrix and multiple fossil constraints. Journal of Mammalogy 89:175–189.

Benton, M. and P. Donoghue. 2007. Paleontological evidence to date the tree of life. Mol Biol Evol 24:26–53.

Benton, M. J., P. C. J. Donoghue, and R. J. Asher. 2009. Calibrating and constraining molecular clocks. Pages 35–86 *in* The Timetree of Life (S. B. Hedges and S. Kumar, eds.). Oxford University Press, Oxford, UK.

Benton, M. J., M. A. Wills, and R. Hitchin. 2011. Quality of the fossil record through time. Nature 403:534–537.

Bininda-Emonds, O. R. P., M. Cardillo, K. E. Jones, R. D. E. MacPhee, R. M. D. Beck, R. Grenyer, S. A. Price, R. A. Vos, J. L. Gittleman, and A. Purvis. 2007. The delayed rise of present-day mammals. Nature 446:507–512.

Bolnick, D. I. and T. J. Near. 2008. Tempo of hybrid inviability in centrarchid fishes (Teleostei: Centrarchidae). Evolution 59:1754–1767.

Brown, J. W., J. S. Rest, J. García-Moreno, M. D. Sorenson, and D. P. Mindell. 2008. Strong mitochondrial DNA support for a Cretaceous origin of modern avian lineages. BMC Biology 6:6.

Carnevale, G. and J. C. Tyler. 2010. Review of the fossil pufferfish genus *Archaeotetraodon* (Teleostei, Tetraodontidae), with description of three new taxa from the Miocene of Italy. Geobios 43:283–304.

Choubert, G. and A. Faure-Muret. 1976. Geological Atlas of the World. 1:10,000,000. 22 Sheets with Explanations. UNESCO Commission for the Geologic Map of the World, Paris.

Donoghue, P. C. J. and M. J. Benton. 2007. Rocks and clocks: calibrating the Tree of Life using fossils and molecules. Trends Ecol Evol 22:424–431.

Dornburg, A., J. M. Beaulieu, J. C. Oliver, and T. J. Near. 2011. Integrating fossil preservation biases in the selection of calibrations for molecular divergence time estimation. Syst Biol 60:519–527.

Drummond, A. J., S. Y. W. Ho, M. J. Philips, and A. Rambaut. 2006. Relaxed phylogenetics and dating with confidence. PLoS Biol 4:e88.

Drummond, A. J. and A. Rambaut. 2007. BEAST: Bayesian evolutionary analysis by sampling trees. BMC Evol Biol 7:214.

Eastman, J. T. 1993. Antarctic Fish Biology: Evolution in a Unique Environment. Academic Press, Inc., San Diego, CA.

Foote, M. 1997. Estimating taxonomic durations and preservation probability. Paleobiology 23:278–300.

Foote, M., J. P. Hunter, C. M. Janis, and J. J. Sepkoski Jr. 1999. Evolutionary and preservational constraints on origins of biologic groups: divergence times of eutherian mammals. Science 283:1310–1314.

Foote, M. and D. M. Raup. 1996. Fossil preservation and the stratigraphic ranges of taxa. Paleobiology 22:121–140.

Foote, M. and J. J. Sepkoski Jr. 1999. Absolute measures of the completeness of the fossil record. Nature 398:415–417.

Forest, F. 2009. Calibrating the Tree of Life: fossils, molecules and evolutionary timescales. Annals of Botany 104:789–794.

Friedman, M. and M. D. Brazeau. 2011. Sequences, stratigraphy and scenarios: what can we say about the fossil record of the earliest tetrapods? Proc R Soc Lond B 278:432–439.

Gernhard, T. 2008. The conditioned reconstructed process. Journal of Theoretical Biology 253:769–778.

Gradstein, F. M. and J. G. Ogg. 2009. The geologic time scale. Pages 26–34 *in* The Timetree of Life. Oxford University Press, Oxford, UK.

Graur, D. and W. Martin. 2004. Reading the entrails of chickens: molecular timescales of evolution and the illusion of precision. Trends in Genetics 20:80–86.

Guindon, S. and O. Gascuel. 2003. A simple, fast, and accurate algorithm to estimate large phylogenies by Maximum Likelihood. Syst Biol 52:696–704.

Harmon, L. J., J. T. Weir, C. D. Brock, R. E. Glor, and W. Challenger. 2008. GEIGER: investigating evolutionary radiations. Bioinformatics 24:129–131.

Ho, S. Y. W. 2007. Calibrating molecular estimates of substitution rates and divergence times in birds. J Avian Biology 38:409–414.

Hugall, A. F. and M. S. Y. Lee. 2004. Molecular claims of Gondwanan age for Australian agamid lizards are untenable. Mol Biol Evol 21:2102–2110.

Inoue, J. G., P. C. J. Donoghue, and Z. Yang. 2010. The impact of the representation of fossil calibrations on Bayesian estimation of species divergence times. Syst Biol 59:74–89.

Ksepka, D. T., M. J. Benton, M. T. Carrano, M. A. Gandolfo, J. J. Head, E. J. Hermsen, W. G. Joyce, K. S. Lamm, J. S. L. Patane, M. J. Phillips, P. D. Polly, M. Van Tuinen, J. L. Ware, R. C. M. Warnock, and J. F. Parham. 2011. Synthesizing and databasing fossil calibrations: divergence dating and beyond. Biology Letters doi:10.1098/rsbl.2011.0356.

Lear, C. H., Y. Rosenthal, H. K. Coxall, and P. A. Wilson. 2004. Late Eocene to Early Miocene ice sheet dynamics and the global carbon cycle. Paleoceanography 19.

Lukoschek, V., J. S. Keogh, and J. C. Avise. 2011. Evaluating fossil calibrations for dating phylogenies in light of rates of molecular evolution: a comparison of three approaches. Syst Biol doi:10.1093/sysbio/syr075.

Malabarba, M. C., L. R. Malabarba, and C. Del Papa. 2010. *Gymnogeophagus eocenicus*, n. sp (Perciformes: Cichlidae), an Eocene cichlid from the Lumbrera formation in Argentina. J Vertebr Paleontol 30:341–350.

Malabarba, M. C., O. Zuleta, and C. Del Papa. 2006. *Proterocara argentina*, a new fossil cichlid from the Lumbrera Formation, Eocene of Argentina. J Vertebr Paleontol 26:267–275.

Marshall, C. R. 1990. The fossil record and estimating divergence times between lineages: maximum divergence times and the importance of reliable phylogenies. J Mol Evol 30:400–408.

Marshall, C. R. 2008. A simple method for bracketing absolute divergence times on molecular phylogenies using multiple fossil calibration points. Am Nat 171:726–742.

Matschiner, M., R. Hanel, and W. Salzburger. 2011. On the origin and trigger of the notothenioid adaptive radiation. PLoS ONE 6:e18911.

McGowan, A. J. and A. B. Smith. 2008. Are global Phanerozoic marine diversity curves truly global? A study of the relationship between regional rock records and global Phanerozoic marine diversity. Paleobiology 34:80–103.

Meredith, R. W., J. E. Janecka, J. Gatesy, O. A. Ryder, C. A. Fisher, E. C. Teeling, A. Goodbla, E. Eizirik, T. L. L. Simao, T. Stadler, D. L. Rabosky, R. L. Honeycutt, J. J. Flynn, C. M. Ingram, C. Steiner, T. L. Williams, T. J. Robinson, A. Burk-Herrick, M. Westerman, N. A. Ayoub, M. S. Springer, and W. J. Murphy. 2011. Impacts of the Cretaceous terrestrial revolution and KPg extinction on mammal diversification. Science doi:10.1126/science.1211028.

Meredith, R. W., M. A. Mendoza, K. K. Roberts, M. Westerman, and M. S. Springer.

2010. A phylogeny and timescale for the evolution of Pseudocheiridae (Marsupialia: Diprotodontia) in Australia and New Guinea. Journal of Mammalian Evolution 17:75–99.

Miya, M., T. W. Pietsch, J. W. Orr, R. J. Arnold, T. P. Satoh, A. M. Shedlock, H.-C. Ho, M. Shimazaki, M. Yabe, and M. Nishida. 2010. Evolutionary history of anglerfishes (Teleostei: Lophiiformes): a mitogenomic perspective. BMC Evol Biol 10:58.

Morlon, H., T. L. Parsons, and J. B. Plotkin. 2011. Reconciling molecular phylogenies with the fossil record. P Natl Acad Sci USA 108:16327–16332.

Müller, J. and R. R. Reisz. 2005. Four well-constrained calibration points from the vertebrate fossil record for molecular clock estimates. Bioessays 27:1069–1075.

Murray, A. M. 2000. Eocene cichlid fishes from Tanzania, East Africa. J Vertebr Paleontol 20:651–664.

Near, T. J., D. I. Bolnick, and P. C. Wainwright. 2005. Fossil calibrations and molecular divergence time estimates in centrarchid fishes (Teleostei: Centrarchidae). Evolution 59:1768–1782.

Nelson, J. S. 2006. Fishes of the World. John Wiley & Sons, Inc., Hoboken, New Jersey.

Nosil, P., L. J. Harmon, and O. Seehausen. 2009. Ecological explanations for (incomplete) speciation. Trends Ecol Evol Pages 1–12.

Parham, J. F. and R. B. Irmis. 2008. Caveats on the use of fossil calibrations for molecular dating: a comment on Near et al. Am Nat 171:132–136.

Patterson, C. 1993. An overview of the early fossil record of acanthomorphs. Bull Mar Sci 52:29–59.

Paul, C. R. C. 2009. The fidelity of the fossil record: the improbability of preservation. Palaeontology 52:485–489.

Pekar, S. F. and R. M. DeConto. 2006. High-resolution ice-volume estimates for the early Miocene: Evidence for a dynamic ice sheet in Antarctica. Palaeogeogr Palaeocl 231:101–109.

Perez, P. A., M. C. Malabarba, and C. Del Papa. 2010. A new genus and species of Heroini (Perciformes: Cichlidae) from the early Eocene of southern South America. Neotrop Ichthyol 8:631–642.

Posada, D. 2008. jModelTest: phylogenetic model averaging. Mol Biol Evol 25:1253–1256.

Purvis, A. 2008. Phylogenetic approaches to the study of extinction. Annu Rev Ecol Evol Syst 39:301–319.

Pyron, R. A. 2010. A Likelihood method for assessing molecular divergence time estimates and the placement of fossil calibrations. Syst Biol 59:185–194.

R Development Core Team. 2011. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing Vienna, Austria http://www.R-project.org.

Rabosky, D. L. 2009. Ecological limits on clade diversification in higher taxa. Am Nat 173:662–674.

Rabosky, D. L. 2010. Extinction rates should not be estimated from molecular phylogenies. Evolution 64:1816–1824.

Rambaut, A. and A. J. Drummond. 2011. Tracer 1.5. http://www.http://tree.bio.ed.ac.uk/software/tracer/.

Raup, D. M. 1972. Taxonomic diversity during the Phanerozoic. Science 177:1065–1071.

Raup, D. M. 1976. Species diversity in the Phanerozoic: an interpretation. Paleobiology 2:289–297.

Reisz, R. R. and J. Müller. 2004. Molecular timescales and the fossil record: a paleontological perspective. Trends in Genetics 20:237–241.

Ronquist, F. and J. P. Huelsenbeck. 2003. MrBayes 3: Bayesian phylogenetic inference under mixed models. Bioinformatics 19:1572–1574.

Salzburger, W., T. Mack, E. Verheyen, and A. Meyer. 2005. Out of Tanganyika: Genesis, explosive speciation, key-innovations and phylogeography of the haplochromine cichlid fishes. BMC Evol Biol 5:17.

Santini, F., L. J. Harmon, G. Carnevale, and M. E. Alfaro. 2009. Did genome duplication drive the origin of teleosts? A comparative study of diversification in ray-finned fishes. BMC Evol Biol 9:194.

Santini, F. and J. C. Tyler. 2003. A phylogeny of the families of fossil and extant tetraodontiform fishes (Acanthomorpha, Tetraodontiformes), Upper Cretaceous to Recent. Zool J Linn Soc 139:565–617.

Schwarz, G. 1978. Estimating the dimension of a model. The Annals of Statistics 6:461–464.

Smith, A. B. and A. J. McGowan. 2007. The shape of the Phanerozoic marine pelaeodiversity curve: how much can be predicted from the sedimentary rock record of Western Europe? Palaeontology 50:765–774.

Stadler, T. 2011. Mammalian phylogeny reveals recent diversification rate shifts. P Natl Acad Sci USA 108:6187–6192.

Tavaré, S., C. R. Marshall, O. Will, C. Soligo, and R. D. Martin. 2002. Using the fossil record to estimate the age of the last common ancestor of extant primates. Nature 416:726–729.

Tyler, J. C. and L. Sorbini. 1996. New superfamily and three new families of tetraodontiform fishes from the Upper Cretaceous: the earliest and most morphologically primitive plectognaths. Sm C Paleob 82:1–62.

Wall, P. D., L. C. Ivany, and B. H. Wilkinson. 2009. Revisiting Raup: exploring the influence of outcrop area on diversity in light of modern sample-standardization techniques. Paleobiology 35:146–167.

Warnock, R. C. M., Z. Yang, and P. C. J. Donoghue. 2011. Exploring uncertainty in the calibration of the molecular clock. Biology Letters .

Weir, J. T. and D. Schluter. 2007. The latitudinal gradient in recent speciation and extinction rates of birds and mammals. Science 315:1574–1576.

Wiley, E. O. and G. D. Johnson. 2010. A teleost classification based on monophyletic groups. Pages 123–182 *in* Origin and Phylogenetic Interrelationships of Teleosts. Verlag Dr. Friedrich Pfeil, München, Germany.

Wilkinson, R. D., M. E. Steiper, C. Soligo, R. D. Martin, Z. Yang, and S. Tavare. 2010. Dating primate divergences through an integrated analysis of palaeontological and molecular data. Syst Biol 60:16–31.

Yang, Z. 2006. Bayesian estimation of species divergence times under a molecular clock using multiple fossil calibrations with soft bounds. Mol Biol Evol 23:212–226.

Zwickl, D. J. 2006. Genetic algorithm approaches for the phylogenetic analysis of large

biological sequence datasets under the maximum likelihood criterion. Ph.D. thesis The University of Texas Austin, Texas.

Figures

Figure 1: Comparison of true node ages and estimated node ages. Frames A, D, G, and J each provide an example of a simulated phylogeny with low ($p = 0.03$, $q = 0.01$ per time unit) and high ($p = 0.07$, $q = 0.05$ per time unit) turnover rates, and its artificial fossil record based on low ($r = 0.02$) or high ($r = 0.05$) preservation rates. Fossils are indicated as circles placed on branches according to fossil age. Orange circles represent the oldest fossil of a given clade, that can be used to constrain the age of this clade. Each dot in frames B, C, E, F, H, I, K, and L marks a true node age in one out of 30 replicate phylogenies, and the percentage by which this age is misestimated when uniform or exponential prior distributions are used for fossil constraints. The solid black line shows a running window mean of node age estimate errors over ± 5 time units, and the solid grey lines indicate the standard deviation over the same running window. Numbers included in each individual frame report the RMSD of node age estimates, the proportion of estimated 95% HPD intervals that contain the true node age, and the mean width of 95% HPD intervals.

Figure 2: Exemplary divergence prior distributions based on lineage nonpreservation. Frame A shows probability distribution R_t (solid black line) when the clade's oldest fossil is between 20 and 50 Myr old ($t_y = 20, t_o = 50$), preservation rate $r = 0.003$, and net diversification rate $p - q = 0.03$ per Myr. The same probability distribution is shown in B, when a mean lag time of 20 Myr is assumed for the evolution of recognizable apomorphies ($lag = 20$). When a correction for relative global rock outcrop area (C) is applied (corr = "global"), the prior divergence probability is distributed as shown in D. Dashed lines indicate gamma distributions G_t that best approximate the theoretical probability distribution R_t.

Figure 3: Phylogeny of teleost fishes estimated with divergence prior distributions based on

lineage nonpreservation. Clades are collapsed to 37 groups that are assumed to be monophyletic (see SI Table S1) and colored according to extant species richness. With few exceptions, these 37 clades represent the entire extant acanthomorph diversity (SI Table S1). Outgroup lineages in grey were not used for the analysis of diversification rates. Branch lengths and node bars are according to the results of *BeastRun4b*, in which a base preservation rate of Cenozoic marine teleosts $r = 0.01$ per Myr was assumed. Nodes A–D are referred to in Fig. 4, and node X indicates the shift of diversification rates indentified with the software MEDUSA. The full phylogeny is shown in SI Fig. S13. [1](+ Synbranchiformes, Indostomatidae, Anabantiformes). [2](+ Pseudochromidae, Plesiopidae, Grammatidae, Opistognathidae, Notograptidae).

Figure 4: Posterior distributions for divergences of selected groups. The four distributions shown per group are according to *BeastRun4a–d*, whereby the youngest estimates always correspond to *BeastRun4a* ($r = 0.03$), and the oldest estimates result from *BeastRun4d* (which assumed a base preservation rate of Cenozoic marine teleost $r = 0.001$; SI Table S6). Nodes A–D are indicated in Fig. 4, and all nodes are shown in SI Fig. S13 (node labels in SI Fig. S13: A = 0, B = 7, C = 54, D = 82, E = 94, F = 97).

Figure 5: A) The posterior distribution for the divergence date of the notothenioid 'Antarctic Clade', assuming a base preservation rate of Cenozoic marine teleosts $r = 0.01$ per Myr. B) Benthic foraminiferal oxygen isotope records of ODP site 1090 in black (Pekar and DeConto 2006) and ODP site 1218 in grey (Lear et al. 2004). The Mi1a cold event 21.1 Ma is marked by high isotopic values $> 3.5‰$.

Figure 1: Comparison of true node ages and estimated node ages.
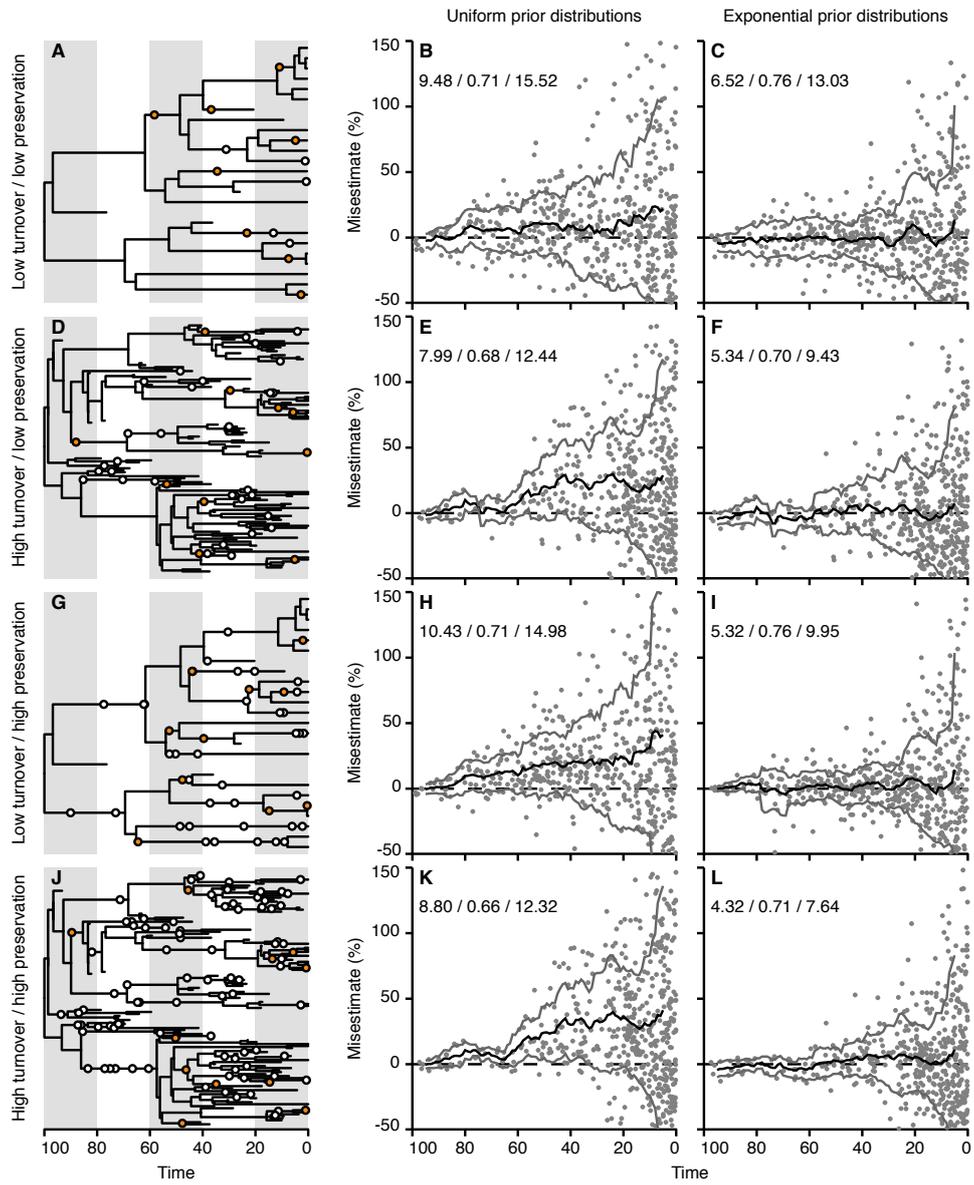
Figure 2: Exemplary prior divergence probability distributions based on lineage nonpreservation.
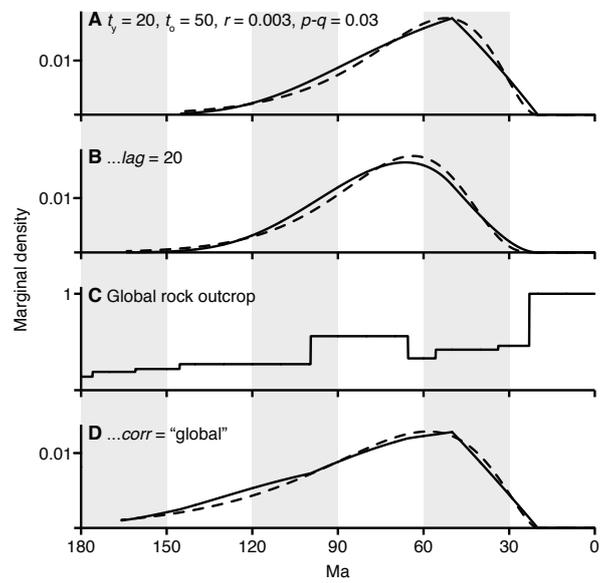
Figure 3: Phylogeny of teleost fishes estimated with divergence prior distributions based on lineage nonpreservation.
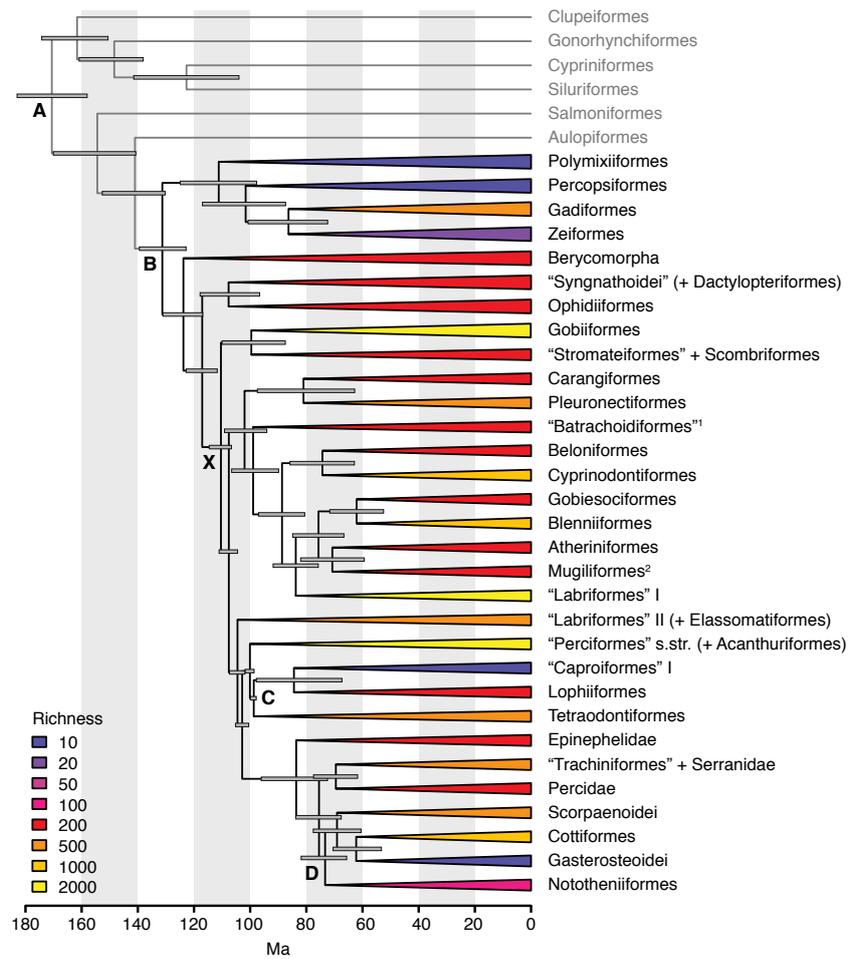
Figure 4: Posterior distributions for divergences of selected groups.
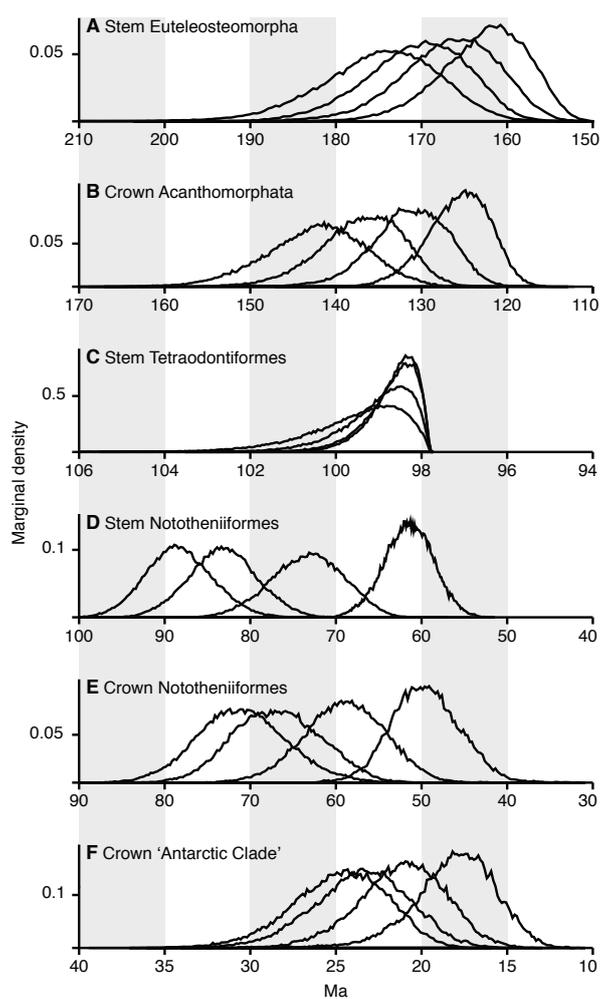
Figure 5: Posterior distribution for the divergence of the notothenioid 'Antarctic Clade' and benthic foraminiferal oxygen isotope records.
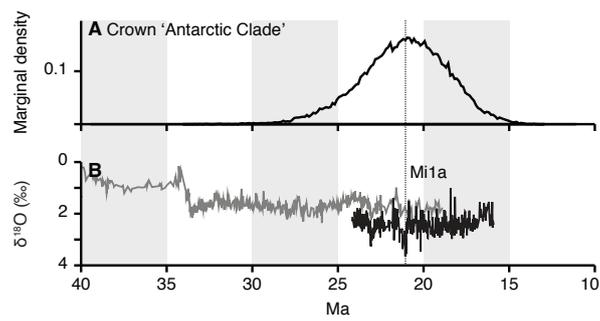
Supporting Information

Bayesian divergence priors based on probabilities of lineage nonpreservation

Michael Matschiner[*]

Addresses:
[*]Zoological Institute, University of Basel, Vesalgasse 1, CH-4051 Basel, Switzerland

Correspondence:
Michael Matschiner, Fax: +41 61 267 03 01, E-mail: michaelmatschiner@mac.com

1 Figures
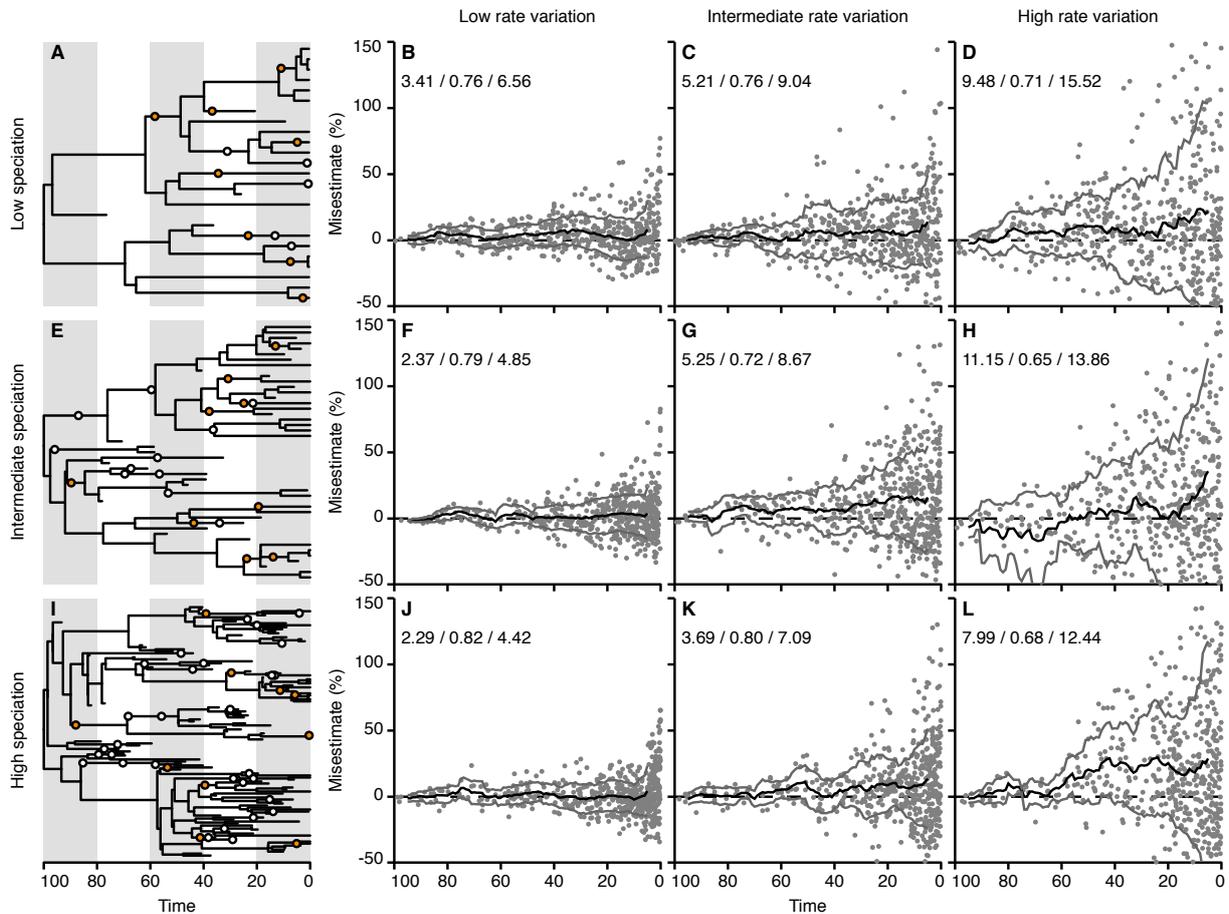
Fig. S1: Comparison of true node ages and estimated node ages in a total of 270 (9 × 30) simulated phylogenies with a root age of 100 time units and 20 extant taxa. Frames A, E, and I each provide an example of a simulated phylogeny with low, intermediate, and high speciation ($p = 0.03/0.05/0.07$ per time unit) and extinction rates ($q = 0.01/0.03/0.05$ per time unit; the net diversification rate was therefore $p - q = 0.02$ per time unit in all settings). For each phylogeny, a fossil record was simulated with preservation rate $r = 0.02$ per time unit. Examples of the artificial fossil record are indicated in frames A, E, and I as white and orange circles placed on branches according to fossil ages. Here, orange circles mark those fossils that are identified as the oldest record of a given extant clade, and can thus be used to constrain the age of this clade. For each setting of speciation and extinction rates, 90 (3 × 30) sequence alignments of the 20 extant taxa were produced with low, intermediate, and high substitution rate variation (standard deviation before log transformation: 0.25/0.5/1.0).

The simulated fossil record was used to estimates node ages with the software BEAST (Drummond & Rambaut 2007), whereby uniform prior distributions between the age of a clade's oldest

fossil and the root age were applied to clade ages. Each dot in frames B–D, F–H, and J–L represents a true node age (x-axis), and the percentage by which the node age estimate is misestimated when uniform prior distributions are used for dating (y-axis). The solid black line shows a running window mean of node age estimate errors over ± 5 time units, and the solid grey lines indicate the standard deviation over the same running window. Numbers included in each individual frame report the root mean square deviation of node age estimates, the proportion of estimated 95% HPD intervals that contain the true node age, and the mean width of 95% HPD intervals.

Fig. S2: As SI Fig. S1, but using preservation rate $r = 0.05$ for the simulated fossil record (illustrated in frames A, E, and I).
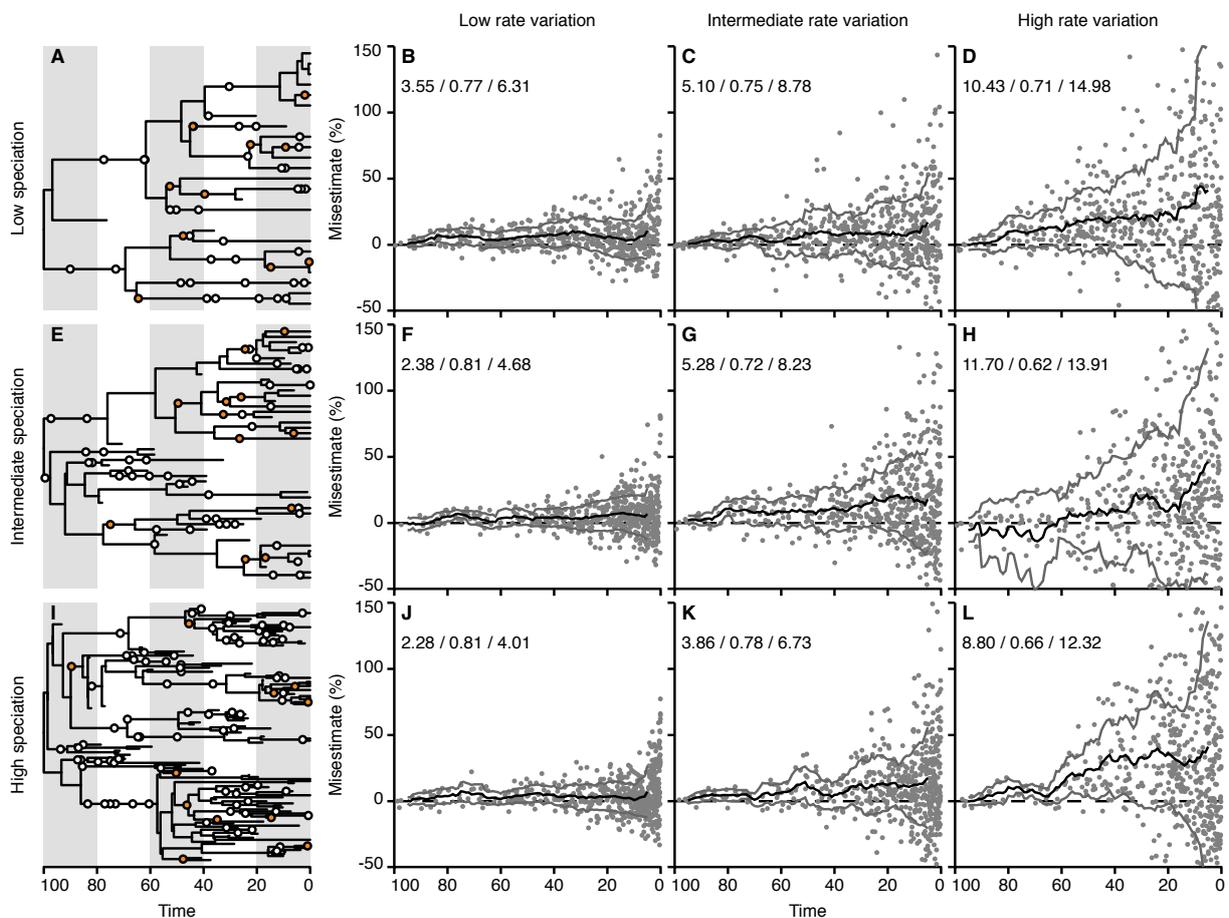
Fig. S3: As SI Fig. S1, but using exponential prior distributions for clade ages, calculated with R package 'ageprior' for the known preservation rate $r = 0.02$ and net diversification rate $p - q = 0.02$. Even the ages of clades without fossils were constrained, using fossil age $t_\mathrm{f} = 0$.
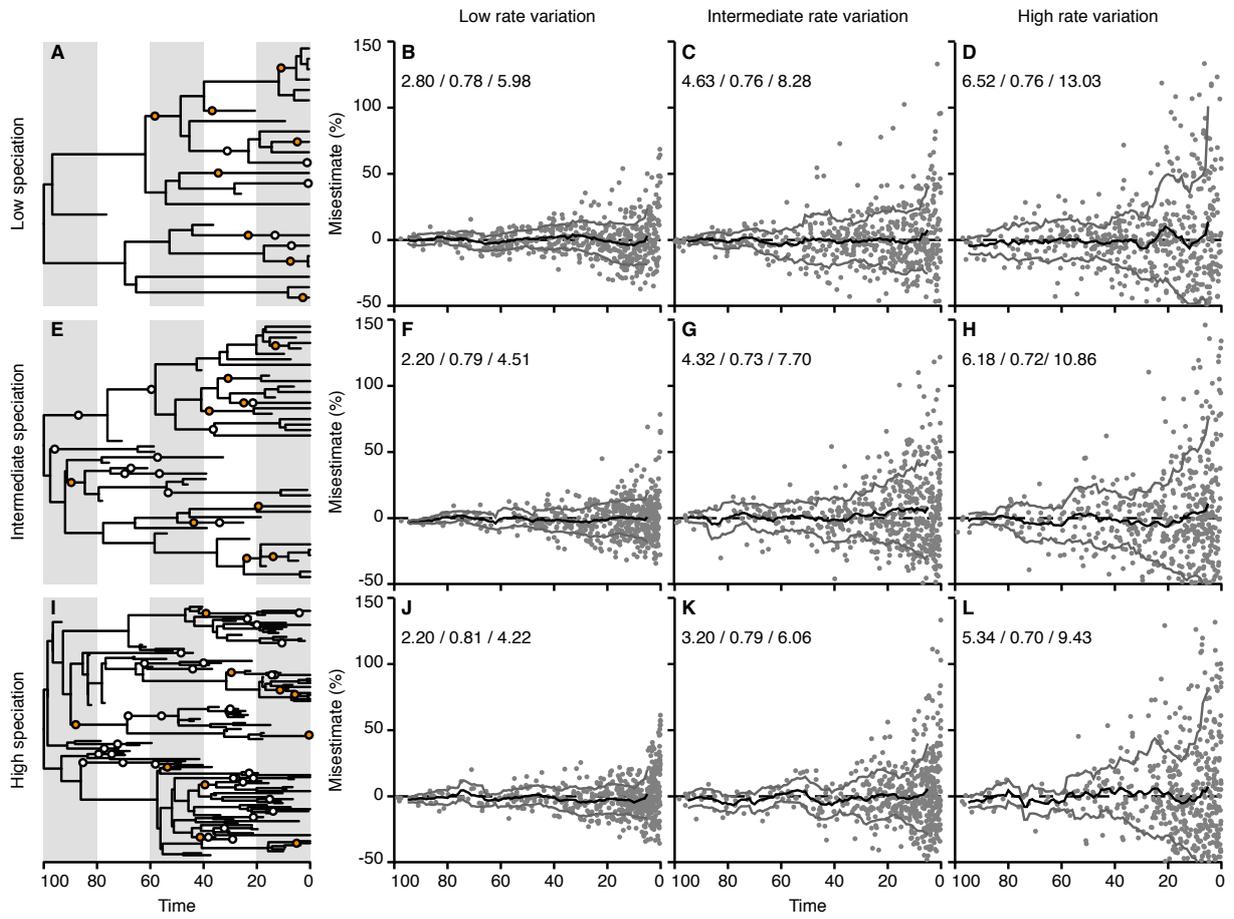
Fig. S4: As SI Fig. S3, but using preservation rate $r = 0.05$ for the simulated fossil record, and for calculation of exponential prior distributions with R package 'ageprior'.
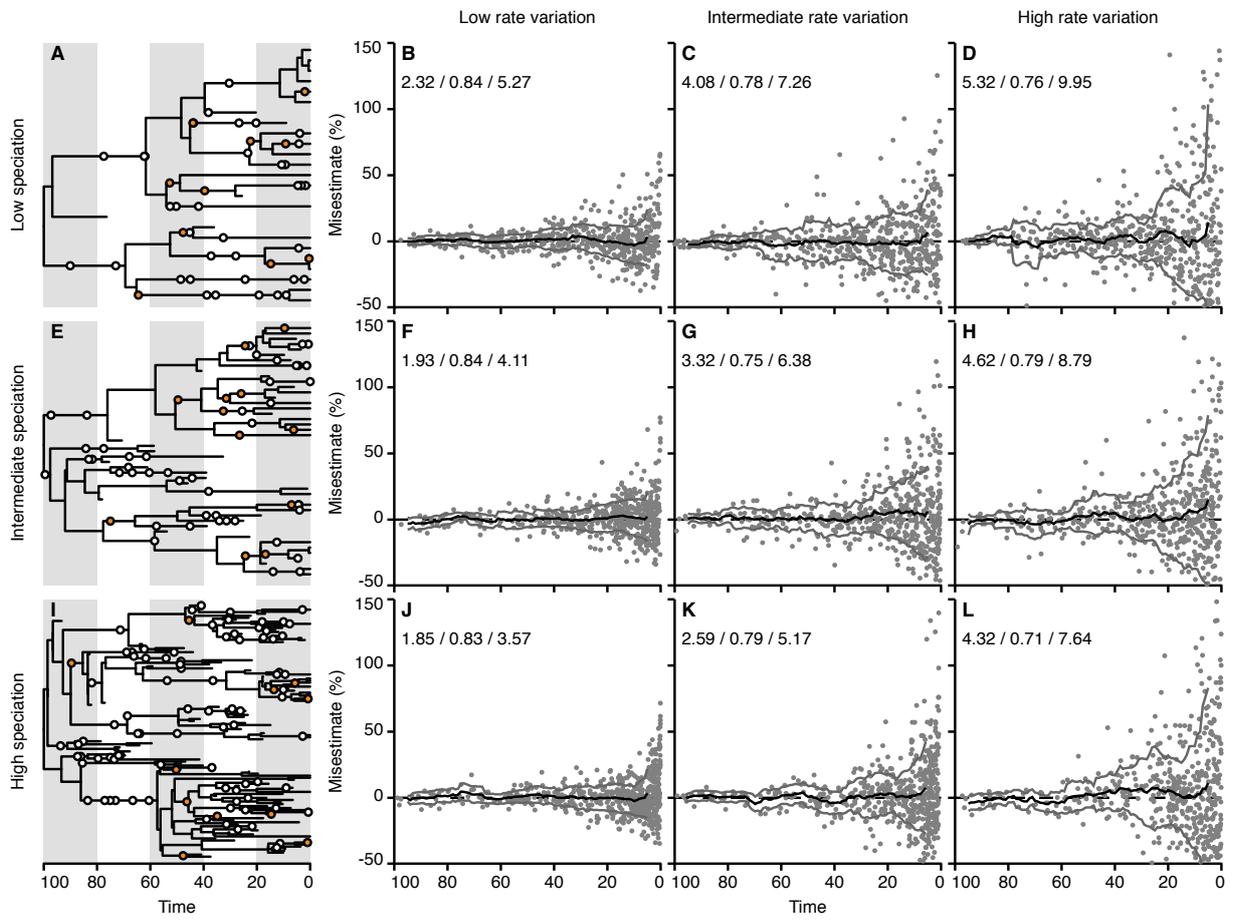
Fig. S5: As frame L of SI Fig. S3, but calculating exponential prior distributions with preservation rate r and net diversification rate $p - q$ that are over- or underestimated by a factor of 2.
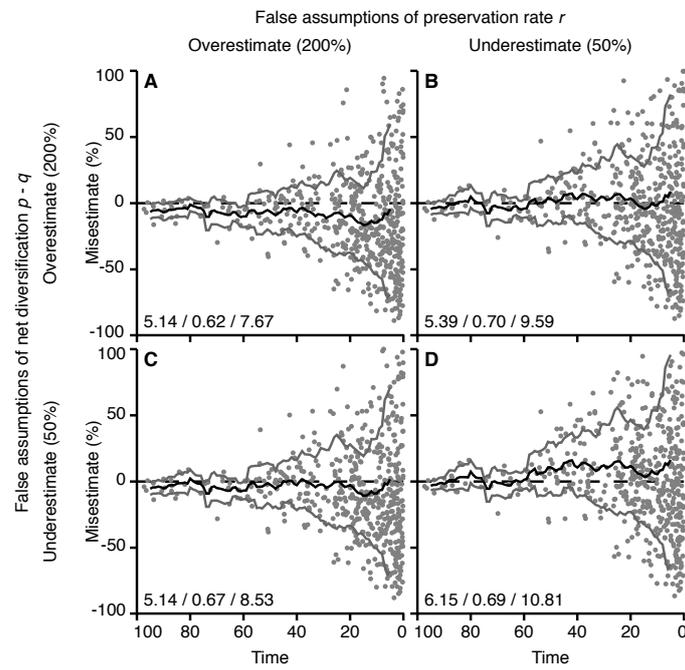
Fig. S6: As frame L of SI Fig. S4, but calculating exponential prior distributions with preservation rate r and net diversification rate $p - q$ that are over- or underestimated by a factor of 2.
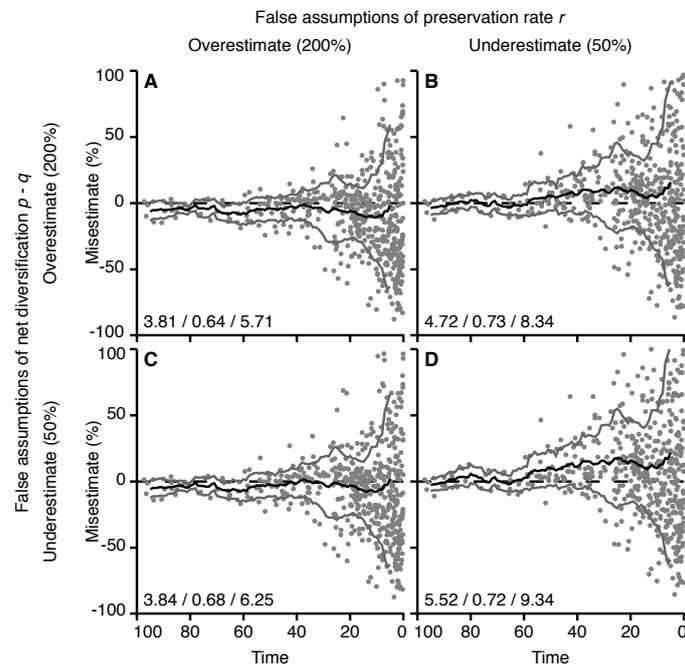
Fig. S7: As SI Fig. S1, but instead of precisely known fossil ages, only the interval of ten time units, into which the fossil age falls, is used for dating. Uniform priors between the younger end of this interval and the root age are applied.
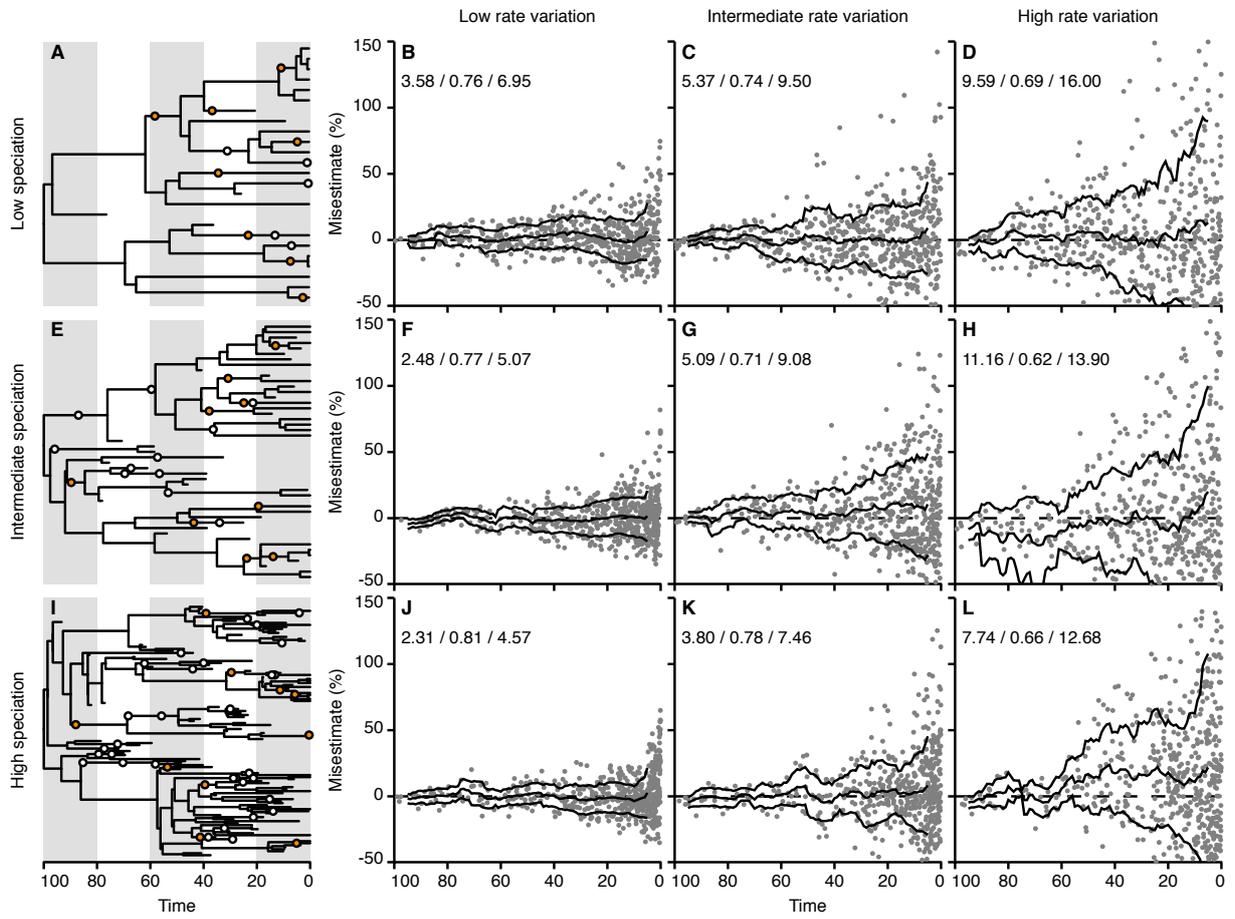
Fig. S8: As SI Fig. S7, but using preservation rate $r = 0.05$ for the simulated fossil record.
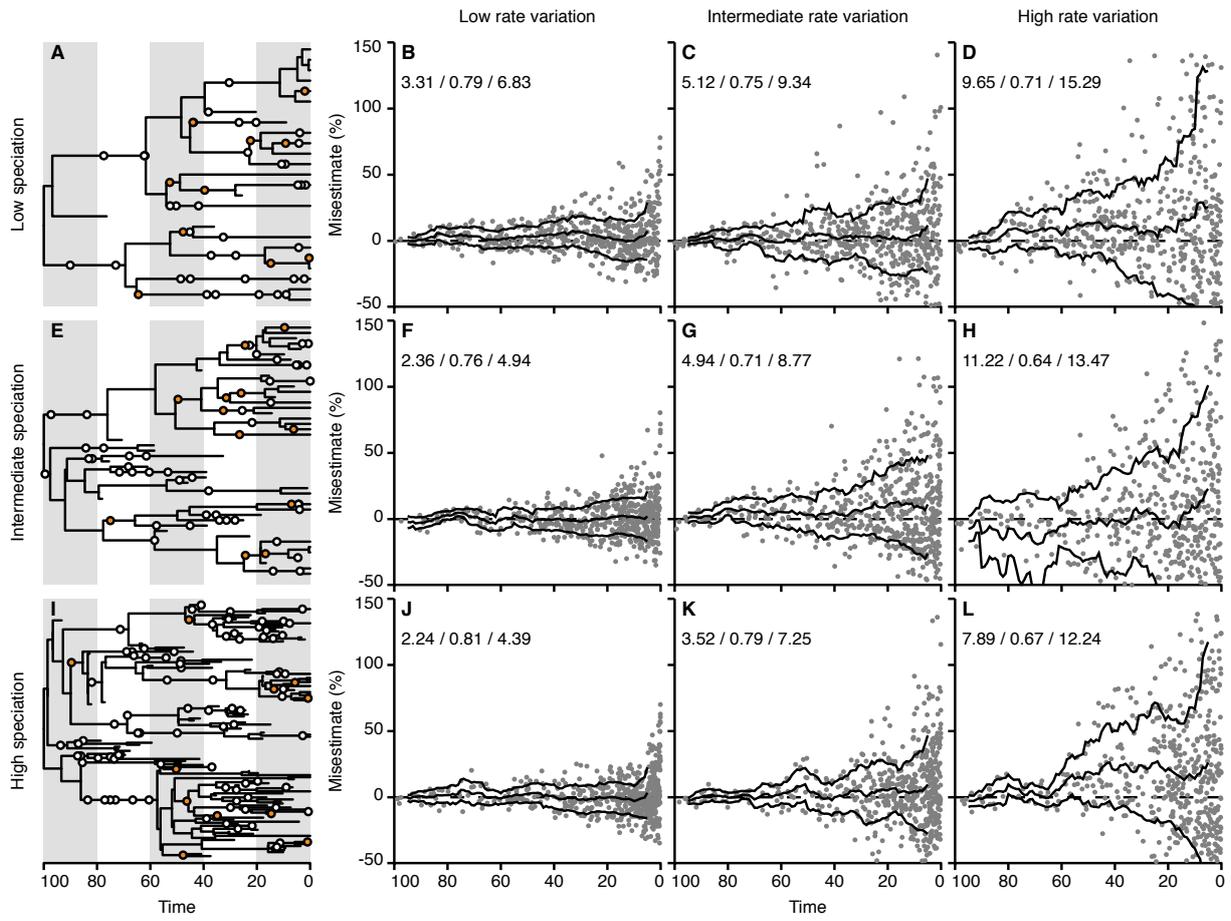
Fig. S9: As SI Fig. S7, but using gamma prior distributions for clade ages, calculated with R package 'ageprior' for the known preservation rate $r = 0.02$, net diversification rate $p - q = 0.02$, and a time interval $t_o - t_y = 10$ time units. Even the ages of clades without fossils were constrained, using fossil age $t_o = t_y = 0$.
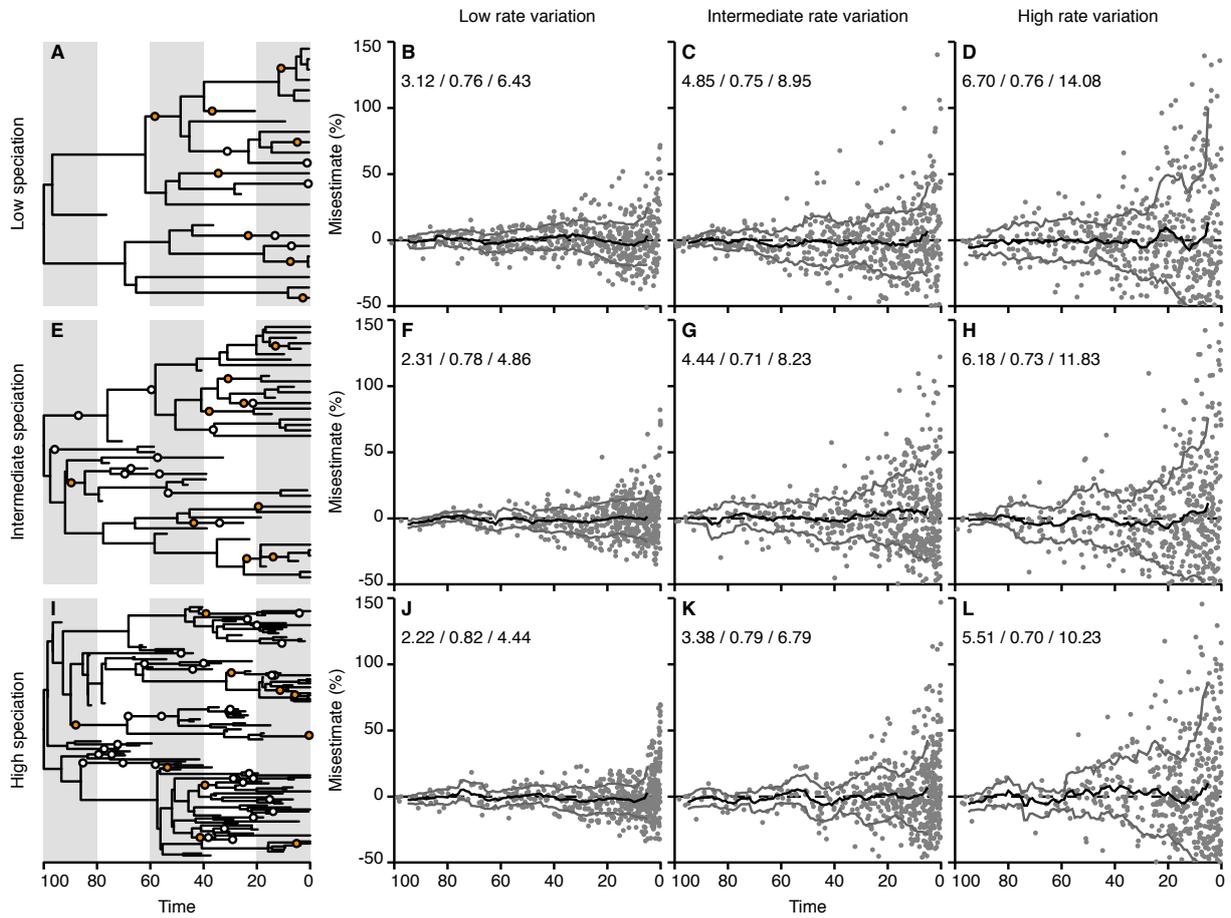
Fig. S10: As SI Fig. S9, but using preservation rate $r = 0.05$ for the simulated fossil record.
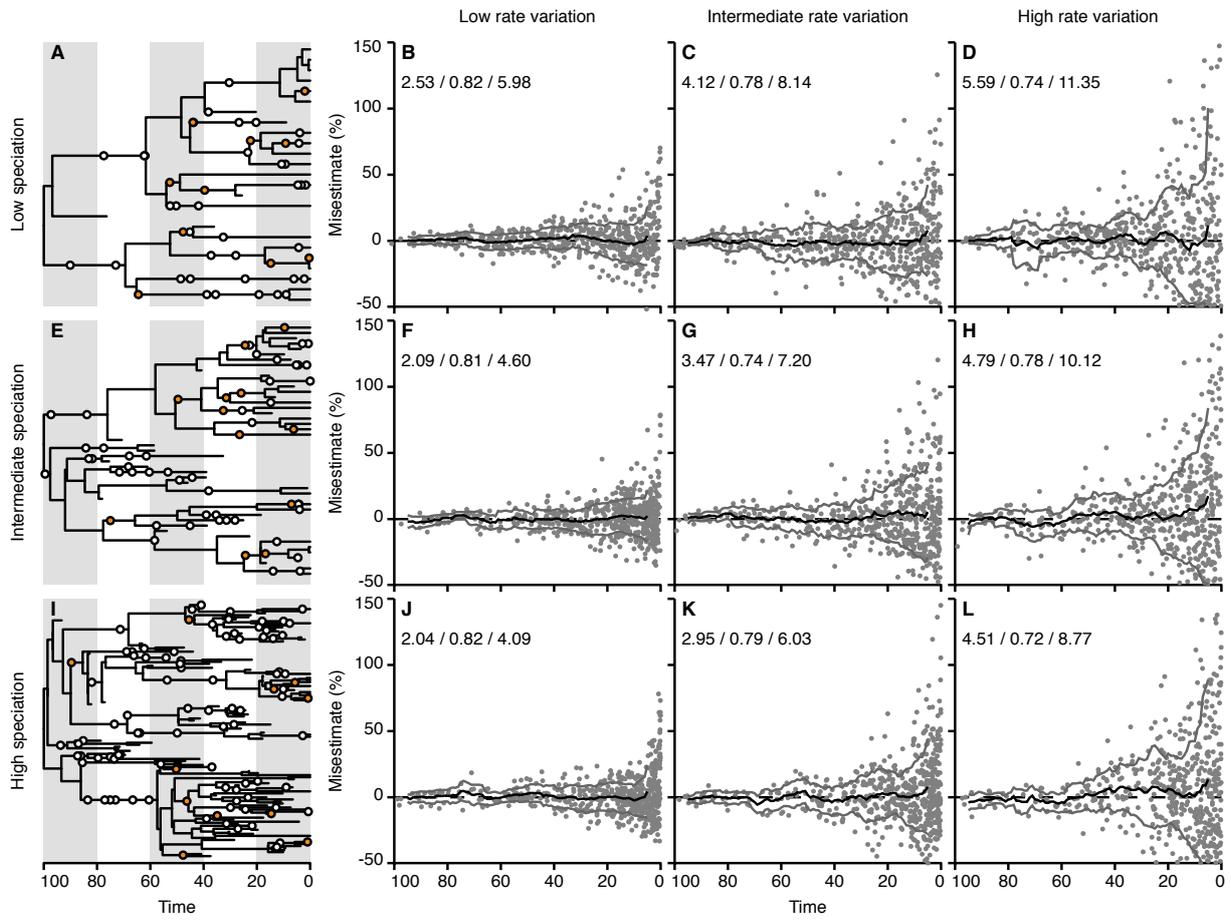
Fig. S11: Comparison of true node ages and age estimates when Marshall's (2008) method is used. Results are not directly comparable to those reported in SI Figs. S1–S10, because the known root age was specified in all other dating analyses, but cannot be used in combination with Marshall's method. Node ages of phylogenies, in which the calibration lineage was correctly identified with Marshall's (2008) empirical scaling factor s_i, are marked with dark grey dots, otherwise light grey dots are used. Numbers included in each individual frame report the root mean square deviation of node age estimates, the proportion of phylogenies, in which the calibration lineage is correctly identified with Marshall's empirical scaling factor s_i when Bayesian mean node heights are used, and the overall percentage of correctly identified calibration lineages when the entire Bayesian sample of node heights (from 9000 trees per replicate, after discarding the first 1000 trees as burnin) is used.
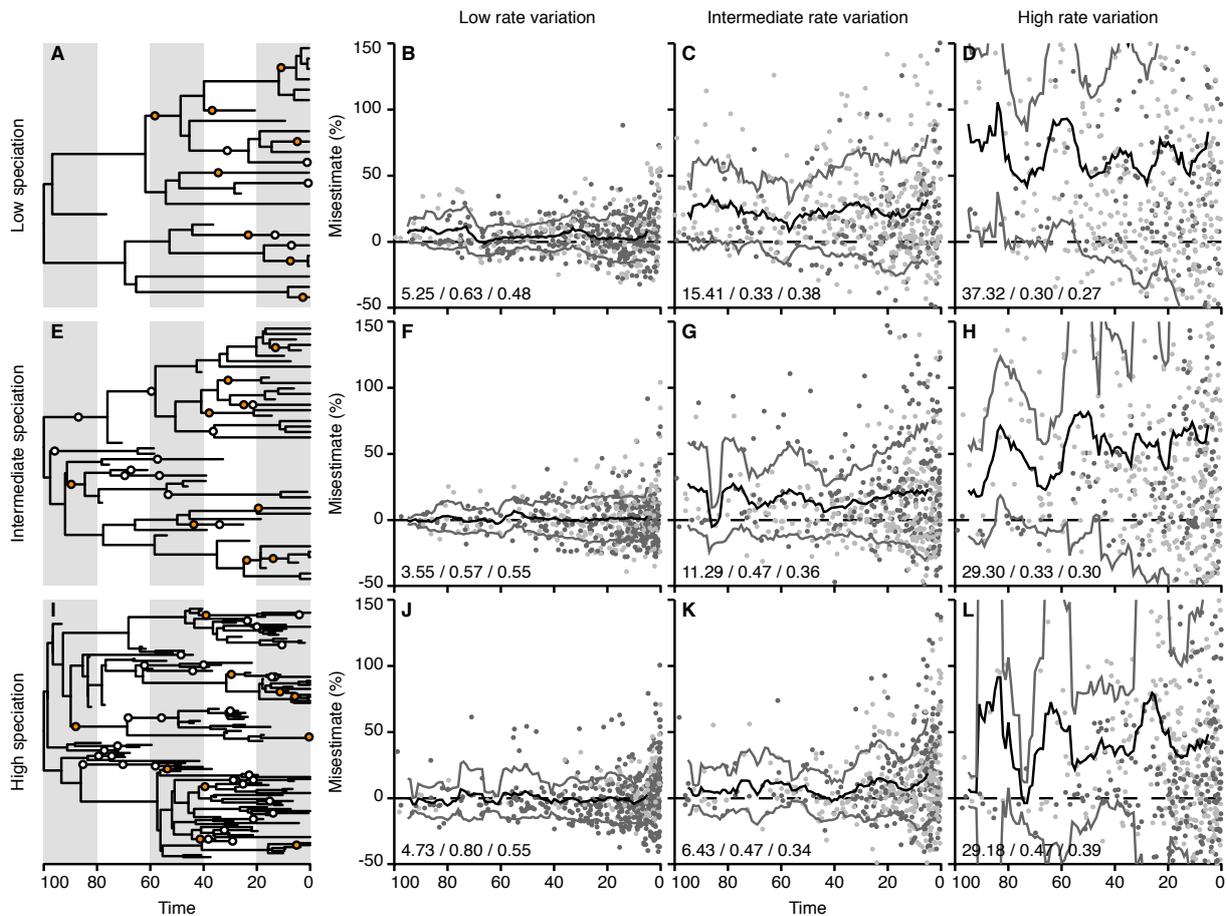
Fig. S12: As SI Fig. S11, but using preservation rate $r = 0.05$ for the simulated fossil record.
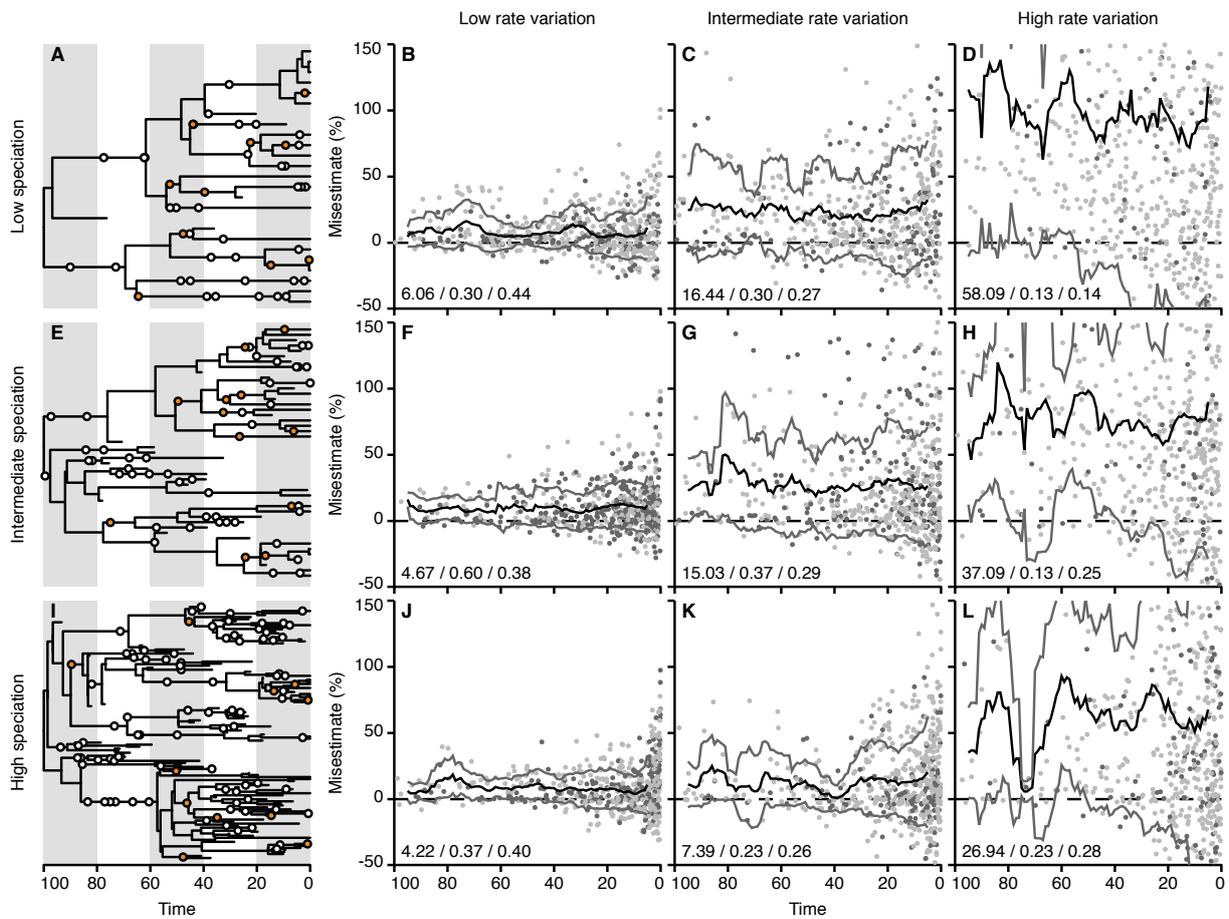
Fig. S13 (next page): Bayesian phylogeny of based on two mitochondrial and four nuclear markers. Branch lengths correspond to results of *BeastRun4b*, in which the base preservation rate of Cenozoic marine teleosts was assumed to be $r = 0.01$ per myr. Node labels are referred to in SI Tables S5–7, and SI Text S2. Outgroups not used for the MEDUSA analysis (nodes 1–6) are shown in grey. See SI Table S1 for a description of the monophyletic groups listed in boxes.
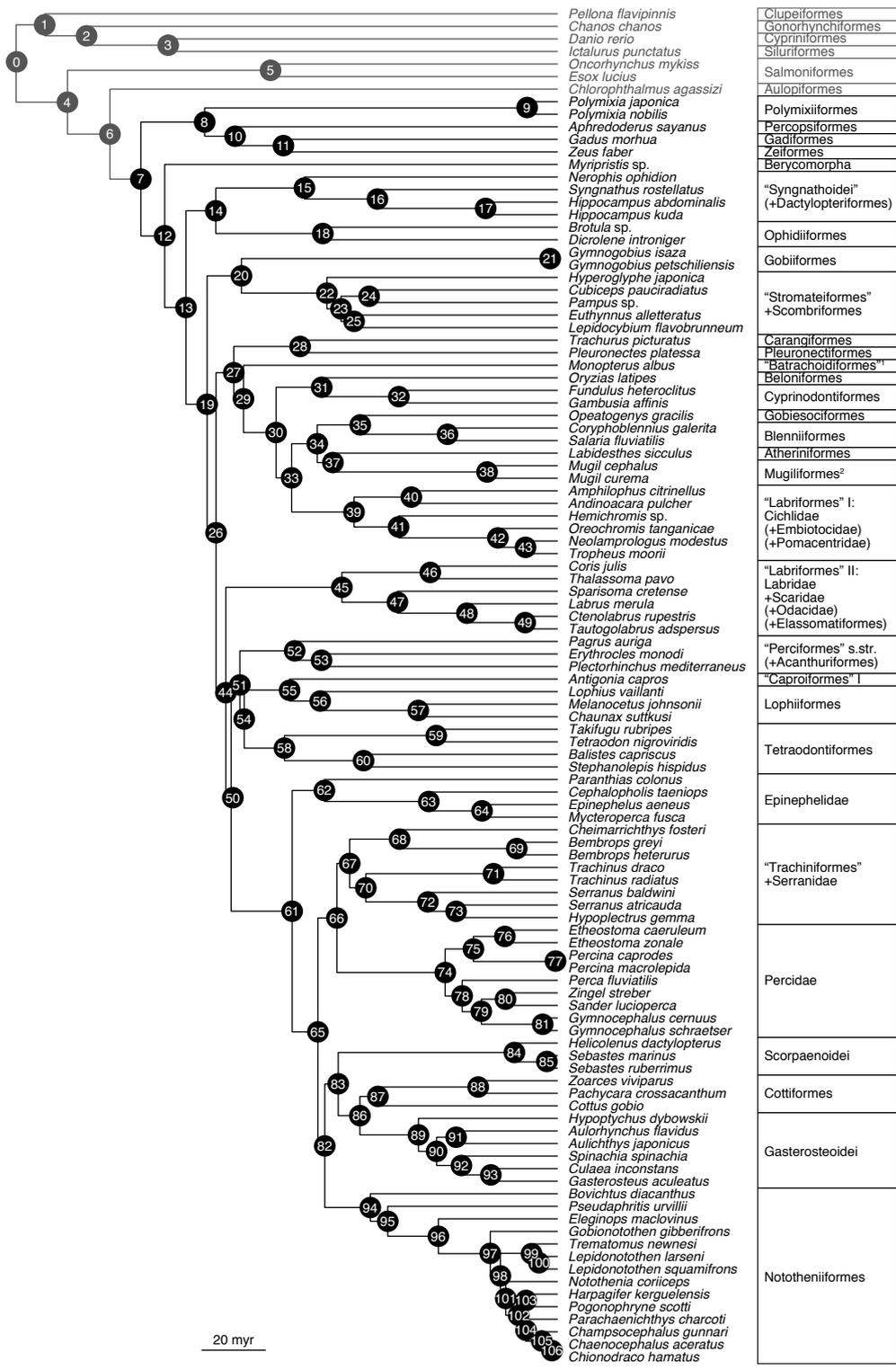
1(+ Synbranchiformes, Indostomatidae, Anabantiformes)

2(+ Pseudochromidae, Plesiopidae, Grammatidae, Opistognathidae, Notograptidae)

221

2 Tables

Tab. S1: Teleost groups used for the diversification analysis with the software MEDUSA, their species richness according to Nelson (2006), the species with which they were represented the molecular phylogenies, and the species IDs used in SI Table S3. Taxonomy follows Wiley & Johnson (2010), unless otherwise stated. All groups, as defined here, are assumed to be monophyletic. With the exception of Pholidichthyiformes *sensu* Springer and Johnson, 2004 (2 spp.), "Caproiformes" *sensu* Rosen, 1984 II (Caproinae, 1 sp.), Niphonidae *sensu* Smith and Craig, 2007 (1 sp.), and Icosteiformes Berg, 1937 (1 sp.), which are considered Acanthomorphata "incertia sedis", the acanthomorph species richness listed in Nelson (2006) is completely represented by the groups defined here (see footnotes).

Group	Richness	Species	ID
Polymixiiformes Lowe, 1838	10	*Polymixia japonica*	Poljap
		Polymixia nobilis	Polnob
Percopsiformes Berg, 1937	9	*Aphredoderus sayanus*	Aphsay
Gadiformes Goodrich, 1909	555	*Gadus morhua*	Gadmor
Zeiformes Regan, 1909	32	*Zeus faber*	Zeufab
Berycomorpha *sensu* Miya *et al.*, 2005	219	*Myripristis* sp.	Myrspc
"Syngnathoidei" *sensu* Kawahara *et al.*, 2008[1]	271	*Hippocampus abdominalis*	Hipabd
		Hippocampus kuda	Hipkud
		Syngnathus rostellatus	Synros
		Nerophis ophidion	Neroph
Ophidiiformes Berg 1937	337	*Dicrolene introniger*	Dicint
		Brotula sp.	Brospc
Gobiiformes Günther 1880	2211	*Gymnogobius petschiliensis*	Gympet
		Gymnogobius isaza	Gymisa
"Stromateiformes" Jordan, 1923 + Scombriformes *sensu* Johnson, 1986[2]	218	*Hyperoglyphe japonica*	Hypjap
		Cubiceps pauciradiatus	Cubpau
		Pampus sp.	Pamspc
		Lepidocybium flavobrunneum	Lepfla
		Euthynnus alletteratus	Eutall
Carangiformes Jordan, 1923	152	*Trachurus picturatus*	Trapic
Pleuronectiformes Bleeker, 1859	638	*Pleuronectes platessa*	Plepla
"Batrachoidiformes" Berg 1937[3]	301	*Monopterus albus*	Monalb
Beloniformes Berg, 1937	227	*Oryzias latipes*	Orylat
Cyprinodontiformes *sensu* Parenti, 1981	963	*Gambusia affinis*	Gamaff
		Fundulus heteroclitus	Funhet
Gobiesociformes Gill, 1872	334	*Opeatogenys gracilis*	Opegra
Blenniiformes *sensu* Springer, 1993	818	*Salaria fluviatilis*	Salflu
		Coryphoblennius galerita	Corgal
Atheriniformes Rosen, 1964	214	*Labidesthes sicculus*	Labsic
Mugiliformes Günther, 1880[4]	330	*Mugil cephalus*	Mugcep
		Mugil curema	Mugcur

Tab. S1 (continued)

Group	Richness	Species	ID
"Labriformes" *sensu* Kaufman and Liem, 1982 I[5]	1721	*Andinoacara pulcher*	Andpul
		Amphilophus citrinellus	Ampcit
		Hemichromis sp.	Hemspc
		Oreochromis tanganicae	Oretan
		Neolamprologus modestus	Neomod
		Tropheus moorii	Tromoo
"Labriformes" *sensu* Kaufman and Liem, 1982 II[6,7]	559	*Coris julis*	Corjul
		Thalassoma pavo	Thapav
		Sparisoma cretense	Spacre
		Labrus merula	Labmer
		Ctenolabrus rupestris	Cterup
		Tautogolabrus adspersus	Tauads
"Perciformes" *sensu* stricto[8,9]	2255	*Erythrocles monodi*	Erymon
		Pagrus auriga	Pagaur
		Plectorhinch. mediterraneus	Plemed
"Caproiformes" *sensu* Rosen, 1984 I[10]	10	*Antigonia capros*	Antcap
Lophiiformes Garman, 1899	313	*Lophius vaillanti*	Lopvai
		Chaunax suttkusi	Chasut
		Melanocetus johnsonii	Meljoh
Tetraodontiformes Berg, 1937	357	*Balistes capriscus*	Balcap
		Stephanolepis hispidus	Stehis
		Tetraodon nigroviridis	Tetnig
		Takifugu rubripes	Takrub
Epinephelidae *sensu* Smith and Craig, 2007	157	*Paranthias colonus*	Parcol
		Cephalopholis taeniops	Ceptae
		Mycteroperca fusca	Mycfus
		Epinephelus aeneus	Epiaen
"Trachiniformes" Bertin and Arambourg, 1958 + Serranidae *sensu* Smith and Craig, 2007[11,12]	532	*Trachinus radiatus*	Trarad
		Trachinus draco	Tradra
		Cheimarrichthys fosteri	Chefos
		Bembrops greyi	Bemgre
		Bembrops heterurus	Bemhet
		Serranus atricauda	Seratr
		Serranus baldwini	Serbal
		Hypoplectrus gemma	Hypgem
Percidae Cuvier, 1816	201	*Etheostoma zonale*	Ethzon
		Etheostoma caeruleum	Ethcae
		Percina macrolepida	Permac
		Percina caprodes	Percap
		Perca fluviatilis	Perflu
		Sander lucioperca	Sanluc
		Zingel streber	Zinstr

Tab. S1 (continued)

Group	Richness	Species	ID
Percidae Cuvier, 1816 (continued)		*Gymnocephalus cernuus*	Gymcer
		Gymnocephalus schraetser	Gymsch
Scorpaenoidei *sensu* Imamura and Shinohara, 1998	473	*Helicolenus dactylopterus*	Heldac
		Sebastes ruberrimus	Sebrub
		Sebastes marinus	Sebmar
Cottiformes *sensu* Wiley and Johnson, 2011[13]	1111	*Cottus gobio*	Cotgob
		Zoarces viviparus	Zoaviv
		Pachycara crossacanthum	Paccro
Gasterosteoidei *sensu* Britz and Johnson, 2002	14	*Hypoptychus dybowskii*	Hypdyb
		Aulichthys japonicus	Auljap
		Aulorhynchus flavidus	Aulfla
		Spinachia spinachia	Spispi
		Culaea inconstans	Culinc
		Gasterosteus aculeatus	Gasacu
Nototheniiformes Jordan, 1923	130[14]	*Bovichtus diacanthus*	Bovdia
		Pseudaphritis urvillii	Pseurv
		Eleginops maclovinus	Elemac
		Gobionotothen gibberifrons	Gobgib
		Trematomus newnesi	Trenew
		Lepidonotothen larseni	Leplar
		Lepidonotothen squamifrons	Lepsqu
		Notothenia coriiceps	Notcor
		Harpagifer kerguelensis	Harker
		Pogonophryne scotti	Pogsco
		Parachaenichthys charcoti	Parcha
		Champsocephalus gunnari	Chagun
		Chionodraco hamatus	Chiham
		Chaenocephalus aceratus	Chaace

[1](+ Dactylopteriformes; Kawahara *et al.* 2008).

[2]Supported by the molecular data presented here, "Stromateiformes" are paraphyletic, and include Scombriformes.

[3](+ Synbranchiformes *sensu* Gosline, 1983, Indostomidae, and Anabantiformes *sensu* Britz 1995; Kawahara *et al.* 2008).

[4](+ Pseudochromidae, Plesiopidae, Grammatidae, Opistognathidae, and Notograptidae; Setiamarga *et al.* 2008).

[5]"Labriformes" *sensu* Kaufman & Liem, 1982 I include Embiotocidae, Pomacentridae, and Cichlidae (Mabuchi *et al.* 2007).

[6]"Labriformes" *sensu* Kaufman & Liem, 1982 I comprise Labridae (including Scaridae and Odacidae; Hanel *et al.* 2002; Mabuchi *et al.* 2007).

[7](+ Elassomatiformes Johnson & Patterson, 1993; Setiamarga *et al.* 2008).

[8]Following Wiley & Johnson (2010), "Perciformes" *sensu* stricto are here considered to include

'the former Percoidei *sensu* Johnson (1984), except for members of that group that show affinities elsewhere'. Thus, the group is assumed to include the following families: Centropomidae, Ambassidae, Latidae, Moronidae, Percichthyidae, Perciliidae, Acropomatidae, Symphysanodontidae, Polyprionidae, Centrogeniidae, Ostracoberycidae, Callanthiidae, Dinopercidae, Banjosidae, Centrarchidae, Priacanthidae, Apogonidae, Epigonidae, Sillaginidae, Malacanthidae, Latilinae, Lactariidae, Dinolestidae, Scombropidae, Pomatomidae, Menidae, Leiognathidae, Bramidae, Caristiidae, Emmelichthyidae, Lutjanidae, Caesionidae, Lobotidae, Gerreidae, Haemulidae, Inermiidae, Neripteridae, Letherinidae, Sparidae, Centracanthidae, Polynemidae, Sciaenidae, Mullidae, Pempheridae, Glaucosomatidae, Leptobramidae, Bathyclupeidae, Monodactylidae, Toxotidae, Arripidae, Dichistiidae, Kyphosidae, Drepaneidae, Chaetodontidae, Pomacanthidae, Enoplosidae, Pentacerotidae, Nandidae, Polycentridae, Terapontidae, Kuhliidae, Oplegnathidae, Cirrhitidae, Chironemidae, Aplodactylidae, Cheilodactylidae, Latridae, Cepolidae.

[9](+ Acanthuriformes Jordan, 1923; Holcroft & Wiley 2008).

[11]"Trachiniformes" are here considered to include the families Chiasmodontidae, Champsodontidae, Trichodontidae, Pinguipedidae, Cheimarrhichthyidae, Trichonotidae, Creediidae , Percophidae, Leptoscopidae, Ammodytidae, Trachinidae, and Uranoscopidae.

[10](= Antigoniinae; Wiley & Johnson 2010; Nelson 2006).

[12]"Trachiniformes" are para- or polyphyletic by the inclusion of Serranidae *sensu* Smith and Craig, 2007, which is supported by the molecular phylogenies of Smith & Craig (2007) and Matschiner *et al.* (2011).

[13]Cottiformes *sensu* Wiley and Johnson, 2011 comprise the two suborders Cottoidei and Zoarcoidei.

[14](see Matschiner *et al.* 2011).

Tab. S2: Teleost outgroups included in phylogenies, but excluded from the diversification analysis with MEDUSA, the species with which they were represented in the molecular phylogenies, and the corresponding species IDs used in SI Table S3.

Group	Species	ID
Clupeiformes Goodrich, 1909	*Pellona flavipinnis*	Pelfla
Gonorynchiformes Regan, 1909	*Chanos chanos*	Chacha
Cypriniformes Goodrich, 1909	*Danio rerio*	Danrer
Siluriformes Hay, 1929	*Ictalurus punctatus*	Ictpun
Salmoniformes Greenwood *et al.*, 1966	*Oncorhynchus mykiss*	Oncmyk
	Esox lucius	Esoluc
Aulopiformes Rosen, 1973	*Chlorophthalmus agassizi*	Chlaga

Tab. S3: Genbank accession numbers for all sequences used for phylogenetic analyses.

ID	ND4	Cyt *b*	myh6	Ptr	ENC1	tbr1
Ampcit	HM050088	AB018985	HM050030	HM050148	HM049972	HM050209
Andpul	HM050087	EF432944	HM050029	HM050147	HM049971	HM050208
Antcap	NC_003191	NC_003191	EF536307	HM050149	HM049973	HM050210
Aphsay	NC_004372	NC_004372	EU001908	EU001962	EU002019	EU001990
Aulfla	NC_010268	NC_010268	AB445151	AB445169	AB445223	AB445205
Auljap	NC_011569	NC_011569	AB445150	AB445168	AB445222	AB445204
Balcap	HM050089	EF392572	HM050031	HM050150	HM049974	HM050211
Bemgre	HM050090	HM049934	HM050032	HM050151	HM049975	HM050212
Bemhet	HM050091	HM049935	HM050033	HM050152	HM049976	HM050213
Bovdia	HM050092	HM049936	HM050034	HM050153	HM049977	HM050214
Brospc		EF455989	EF032933	EF032959	EF032985	EF032972
Ceptae	HM050093	EF455991	HM050035	HM050154	HM049978	HM050215
Chaace	HM050094	HM049937	HM050036	HM050155		HM050216
Chacha	NC_004693	NC_004693	EU001904	EU001958	EU002015	EU001988
Chagun	HM050095	HM049938	HM050037	HM050156	HM049979	HM050217
Chasut	HM050096	HM049939	HM050038	HM050157	HM049980	HM050218
Chefos		AY722191				
Chiham	HM050097	HM049940	HM050039	HM050158	HM049981	HM050219
Chlaga	NC_003160	NC_003160	FJ918879		EU366600	
Corgal	HM050098	HM049941	HM050040	HM050159		HM050220
Corjul	HM050099	HM049942	HM050041	HM050160	HM049982	HM050221
Cotgob	HM050100	AY116366	HM050042			HM050222
Cterup	HM050101	HM049943	HM050043	HM050161	HM049983	HM050223
Cubpau	NC_013150	NC_013150				
Culinc	NC_011577	NC_011577	AB445153	AB445171	AB445225	AB445207
Danrer	NC_002333	NC_002333	EF032923	EF032949	EF032975	EF032962
Dicint	HM050102	HM049944	HM050044	HM050162	HM049984	HM050224
Elemac	DQ526429	DQ526429	HM050045	HM050163	HM049985	HM050225
Epiaen	HM050103	DQ197950	HM050046	HM050164	HM049986	HM050226
Erymon	HM050104	EF456004	HM050047	HM050165	HM049987	HM050227
Esoluc	NC_004593	NC_004593	EU001905	EU001959	EU002016	
Ethcae	HM050105	DQ465142		HM050166	HM049988	HM050228
Ethzon	HM050106	AY964705		HM050167	HM049989	HM050229
Eutall	NC_004530	EF439531	HM050048	HM050168	HM049990	HM050230
Funhet	NC_012312	NC_012312	EF032926	EF032952	EF032978	EF032965
Gadmor	NC_002081	EU877717	EU001906	EU001960	EU002017	
GamaffNC	_004388	NC_004388	EU001907	EU001961	EU002018	EU001989
Gasacu	AP002944	AP002944	AB445155	AB445173	AB445227	AB445209
Gobgib	HM050107	HM049945	HM050049	HM050169	HM049991	HM050231
Gymcer	HM050108	AF045356	HM050050	HM050170	HM049992	HM050232
Gymisa		AB560891	AB504103	AB504134		
Gympet	NC_008743	NC_008743				
Gymsch	HM050109	HM049946	HM050051	HM050171	HM049993	HM050233
Harker	HM050110	HM049947	HM050052	HM050172	HM049994	HM050234

Tab. S3 (continued)

ID	ND4	Cyt *b*	myh6	Ptr	ENC1	tbr1
Harker	HM050110	HM049947	HM050052	HM050172	HM049994	HM050234
Heldac	HM050111	EU492259	HM050053	HM050173	HM049995	
Hemspc	HM050112	HM049948		HM050174	HM049996	HM050235
Hipabd		JF273425	JN696469[1]	JN696472[1]		
Hipkud	NC_010272	NC_010272				
Hypdyb	NC_004400	NC_004400	AB445149	AB445167	AB445221	AB445203
Hypgem	NC_013832	NC_013832				
Hypjap	NC_013149	NC_013149				
Ictpun	NC_003489	NC_003489	EF032929	EF032955	EF032981	EF032968
Labmer	HM050113	HM049949	HM050054	HM050175		HM050236
Labsic		HQ691297	EU001919	EU001974		EU001999
Lepfla	HM050114	AM265576	HM050055	HM050176	HM049997	HM050237
Leplar	HM050115	HM049950	HM050056	HM050177		HM050238
Lepsqu	HM050116	HM049951	HM050057	HM050178	HM049998	HM050239
Lopvai	HM050117	HM049952	HM050058	HM050179	HM049999	HM050240
Meljoh	HM050118	HM049953	HM050059	HM050180		HM050241
Monalb	NC_003192	NC_003192	EU001928	EU001983	EU002039	EU002006
Mugcep	NC_003182	EU083840	HM050060		HM050000	HM050242
Mugcur		EU715492	EU001913	EU001967	EU002023	EU001994
Mycfus	HM050119	DQ197968	HM050061	HM050181	HM050001	HM050243
Myrspc	NC_003189	NC_003189	EU001916	EU001971	EU002026	EU001996
Neomod	HM050120	HM049954	HM050062	HM050182	HM050002	HM050244
Neroph	JN696471[1]	AF356043		JN696473[1]	JN696468[1]	
Notcor	HM050121	HM049955	HM050063	HM050183	HM050003	HM050245
Oncmyk	NC_001717	NC_001717	EF032924	EF032950	EF032976	EF032963
Opegra	HM050122	HM049956	HM050064		HM050004	HM050246
Oretan	HM050123	HM049957	HM050065	HM050184	HM050005	HM050247
Orylat	NC_004387	AB084730	EF032927	EF032953	EF032979	EF032966
Paccro	HM050124	HM049958	HM050066		HM050006	HM050248
Pagaur	NC_005146	DQ197974	HM050067	HM050185	HM050007	HM050249
Pamspc	NC_011707	NC_011707				
Parcha	HM050125	HM049959	HM050068	HM050186	HM050008	HM050250
Parcol	HM050126	HM049960		HM050187	HM050009	
Pelfla	NC_014268	NC_014268	EU001898	EU001953	EU002009	
Percap	HM050127	DQ493490		HM050010	HM050251	
Perflu	HM050129	AY929376	HM050070	HM050189	HM050012	HM050253
Permac	NC_008111	DQ493495		HM050190	HM050013	HM050254
Plemed	HM050130	DQ197979	HM050071	HM050191	HM050014	HM050255
Plepla	HM050131	EU224075	EU001930	HM050192		EU002008
Pogsco	HM050132	HM049962	HM050072	HM050193		HM050256
Poljap	NC_002648	NC_002648	EU001926	EU001981	EU002037	
Polnob	HM050133	DQ197980	HM050073	HM050194	HM050015	HM050257
Pseurv	HM050134	HM049963	HM050074	HM050195	HM050016	HM050258
Salflu	HM050135	HM049964	HM050075	HM050196	HM050017	HM050259

Tab. S3 (continued)

ID	ND4	Cyt *b*	myh6	Ptr	ENC1	tbr1
Salflu	HM050135	HM049964	HM050075	HM050196	HM050017	HM050259
Sanluc	HM050136	HM049965	HM050076	HM050197	HM050018	HM050260
Sebmar	HM050137	EF456022	HM050077		HM050019	HM050261
Sebrub	EU008930	AF031501	EU001929	EU001984	EU002040	EU002007
Seratr		EF439230	HM050078	HM050198	HM050020	HM050262
Serbal		AY321784	HQ731358			
Spacre	HM050138	HM049966	HM050079	HM050199		HM050263
Spispi	NC_011582	NC_011582	AB445157	AB445175	AB445229	AB445211
Stehis	HM050139	HM049967	HM050080	HM050200	HM050021	
Synros		AF356041	JN696470[1]	JN696474[1]		
Takrub	NC_004299	NC_004299	[2]	[2]	[2]	[2]
Tauads	HM050140	HM049968	HM050081	HM050201	HM050022	HM050264
Tetnig	NC_007176	AP006046	[2]	[2]	[2]	[2]
Thapav	HM050141	DQ198011	HM050082	HM050202	HM050023	HM050265
Tradra		EU036513				
Trapic	HM050142	EF392634	HM050083	HM050203	HM050024	HM050266
Trarad		DQ198015				
Trenew	HM050143	HM049969	HM050084	HM050204	HM050025	HM050267
Tromoo	HM050144	AB018990		HM050205	HM050026	HM050268
Zeufab	NC_003190	EU264027	EU001927	EU001982	EU002038	
Zinstr	HM050145	HM049970	HM050085	HM050206	HM050027	HM050269
Zoaviv	HM050146	EU492074	HM050086	HM050207	HM050028	HM050270

[1]Sequences produced as part of this study.

[2]Nuclear *T. rubripes* and *T. nigroviridis* sequences were extracted from Ensembl (www.ensembl.org) and Genoscope (www.genoscope.cns.fr) genome browsers (SI Table S4).

Tab. S4: Ensembl and Genoscope identifiers of *Takifugu rubripes* and *Tetraodon nigroviridis* sequences. *T. rubripes* Ensembl identifiers were taken from (Li *et al.* 2007), while *T. nigroviridis* Genoscope identifiers and sequences were found by BLAT-search against the *T. nigroviridis* genome, using the entire *T. rubripes* sequences as search templates.

Taxa	myh6	Ptr
T. rubripes	SINFRUE00000644156	SINFRUE00000786790
T. nigroviridis	GSTENT00008412001	GSTENT00035515001

Taxa	ENC1	tbr1
T. rubripes	SINFRUE00000681690	SINFRUE00000673034
T. nigroviridis	GSTENT00025143001	GSTENT00030575001

Tab. S5: BPP values for all nodes in two initial phylogenetic inferences (*BeastRun1* and *BeastRun2*; see SI Text S1), prior to fixation of the tree topology for subsequent divergence date estimation.

Node	BeastRun1	BeastRun2	Node	BeastRun1	BeastRun2	Node	BeastRun1	BeastRun2
1	1.00	[1]	37	0.47	0.48	73	0.58	0.63
2	1.00	[1]	38	1.00	[1]	74	1.00	[1]
3	1.00	[1]	39	1.00	[1]	75	1.00	1.00
4	1.00	[1]	40	1.00	1.00	76	1.00	1.00
5	1.00	[1]	41	1.00	1.00	77	1.00	1.00
6	1.00	[1]	42	1.00	1.00	78	1.00	1.00
7	0.94	[1]	43	1.00	1.00	79	0.99	0.99
8	0.74	[1]	44	1.00	1.00	80	1.00	1.00
9	1.00	1.00	45	0.99	[1]	81	1.00	1.00
10	0.41	0.56	46	1.00	1.00	82	0.73	0.73
11	1.00	1.00	47	1.00	1.00	83	1.00	1.00
12	0.71	[1]	48	1.00	1.00	84	1.00	[1]
13	0.63	1.00	49	1.00	1.00	85	1.00	1.00
14	[2]	0.56	50	0.69	0.71	86	1.00	1.00
15	0.84	[1]	51	1.00	1.00	87	[2]	[1]
16	1.00	1.00	52	0.99	[1]	88	1.00	1.00
17	1.00	1.00	53	0.88	0.91	89	1.00	[1]
18	1.00	[1]	54	1.00	0.96	90	1.00	1.00
19	1.00	1.00	55	0.93	0.98	91	1.00	0.99
20	0.90	0.95	56	1.00	[1]	92	1.00	1.00
21	1.00	[1]	57	1.00	1.00	93	1.00	1.00
22	1.00	[1]	58	1.00	[1]	94	1.00	[1]
23	[2]	0.36	59	1.00	1.00	95	1.00	1.00
24	[2]	0.32	60	1.00	1.00	96	1.00	1.00
25	0.34	0.35	61	0.99	1.00	97	1.00	1.00
26	0.84	0.90	62	0.87	[1]	98	0.85	0.87
27	0.51	0.69	63	1.00	1.00	99	1.00	1.00
28	1.00	1.00	64	1.00	1.00	100	1.00	0.99
29	0.51	0.65	65	0.98	1.00	101	0.85	0.83
30	1.00	1.00	66	0.99	1.00	102	1.00	1.00
31	1.00	1.00	67	0.71	[1]	103	1.00	1.00
32	1.00	[1]	68	0.95	0.99	104	1.00	1.00
33	0.99	1.00	69	1.00	1.00	105	1.00	1.00
34	1.00	1.00	70	0.63	0.84	106	1.00	1.00
35	1.00	1.00	71	1.00	1.00			
36	1.00	[1]	72	1.00	1.00			

[1]Node constrained as monophyletic.

[2]Node not recovered.

Tab. S6: Net diversification ($p - q$) and relative extinction rate (ϵ) per myr, as estimated with the software MEDUSA (Alfaro *et al.* 2009). Note that r designates the assumed base preservation rate per myr of Cenozoic marine teleosts, not the net diversification rate, as in Alfaro *et al.* (2009). Outgroups ancestral to node 7 were not used for diversification rate analysis with MEDUSA. In each case, the first iteration was based on a phylogeny with fixed topology (*BeastRun3a–d*; see SI Text S1), that was time-calibrated with the sofware BEAST (Drummond & Rambaut 2007), following protocols outlined elsewhere (Matschiner *et al.* 2011), and using divergence prior distributions for every node calculated with the R package 'ageprior'. Here, the age of each node's oldest fossil was used for parameters t_y and t_o, and the stage specific preservation rate r was derived with function 'adjust.pres.rate' of R package 'ageprior' (see SI Text S2). Furthermore, the lag time parameter was set to 2 myr, global marine or terrestrial corrections for outcrop bias were applied, and a universal net diversification rate $p - q$ of 0.02 per myr was assumed. Prior to the MEDUSA analysis, phylogenies were pruned to include only descendents of node 7, and all groups listed in SI Table S1 were reduced to single lineages. In the second iteration, divergence prior distributions for trees *BeastRun4a–d* were calculated as before, but using the net diversification rate estimates resulting from the first MEDUSA iteration with trees *BeastRun3a–d*. Again, phylogenies were pruned and reduced before the MEDUSA analysis. Root heights of the pruned phylogenies (equivalent to the age of node 7 in the full phylogenies), as well as net diversification ($p - q$) and relative extinction rate (ϵ) estimates converge rapidly, and are very similar between the first and the second iteration.

Tree (pruned)	r	Node 7 mean age (Ma)	Nodes 7–18 $p - q$	ϵ	Nodes 19–106 $p - q$	ϵ
BeastRun3a	0.03	125.32	0.0449	0.7293	0.0551	0.9343
BeastRun4a	0.03	125.69	0.0447	0.7278	0.0559	0.9228
BeastRun3b	0.01	130.71	0.0462	0.5573	0.0601	0.8437
BeastRun4b	0.01	131.28	0.0443	0.6382	0.0627	0.7805
BeastRun3c	0.003	136.70	0.0455	0.4599	0.0682	0.4849
BeastRun4c	0.003	137.09	0.0463	0.3787	0.0694	0.4143
BeastRun3d	0.001	142.49	0.0440	0.4228	0.0724	0.0001
BeastRun4d	0.001	142.35	0.0447	0.3737	0.0726	0.0001

Tab. S7: Mean age estimates and 95% HPD intervals for all divergence events with different assumptions for the base preservation rate r of Cenozoic marine teleost fishes. Node labels as in SI Fig. S13.

Node	*BeastRun4a* $r = 0.03$	*BeastRun4b* $r = 0.01$	*BeastRun4c* $r = 0.003$	*BeastRun4d* $r = 0.001$
0	165.95 (155.14-177.35)	170.62 (158.03-183.00)	175.05 (162.74-188.92)	179.85 (165.51-195.48)
1	159.35 (150.52-169.62)	161.62 (150.60-174.22)	163.96 (151.12-178.44)	166.01 (151.75-183.11)
2	146.60 (137.86-155.73)	148.40 (138.08-160.91)	149.74 (137.86-162.60)	151.12 (138.27-165.63)
3	119.86 (101.56-137.00)	122.68 (103.99-141.48)	124.73 (104.90-143.25)	125.49 (104.46-145.65)
4	147.64 (136.47-160.01)	154.43 (140.71-170.09)	160.39 (146.14-175.01)	166.71 (151.03-184.05)
5	87.57 (72.88-104.92)	90.41 (72.90-111.00)	93.29 (73.20-116.43)	95.40 (73.08-121.20)
6	135.12 (127.34-143.78)	141.03 (130.23-152.67)	146.57 (135.30-158.22)	152.36 (139.42-166.79)
7	125.69 (119.07-133.23)	131.28 (122.82-139.46)	137.09 (128.04-147.06)	142.35 (131.08-153.89)
8	105.28 (94.88-116.51)	111.19 (97.69-124.91)	118.80 (103.85-133.60)	123.85 (106.41-140.55)
9	7.01 (3.38-11.06)	9.53 (5.00-14.66)	11.19 (6.40-16.49)	12.20 (6.80-18.16)
10	94.93 (82.41-107.88)	101.61 (87.38-116.99)	109.77 (94.18-126.05)	114.49 (96.15-132.62)
11	81.19 (71.89-91.57)	86.36 (72.39-100.74)	94.23 (77.37-110.89)	97.96 (78.76-118.09)
12	118.76 (113.19-124.64)	123.78 (116.89-131.08)	129.08 (121.27-137.60)	133.57 (124.84-144.30)
13	112.76 (108.48-117.15)	117.08 (111.70-122.80)	121.93 (115.17-129.23)	125.85 (118.02-134.59)
14	95.98 (81.39-109.59)	107.68 (96.64-117.83)	114.05 (103.47-124.07)	118.41 (107.83-129.13)
15	67.33 (52.27-82.16)	79.73 (64.35-94.27)	87.55 (72.33-102.35)	92.31 (77.43-107.50)
16	46.33 (35.09-58.10)	56.72 (42.77-70.49)	63.41 (48.89-78.39)	67.66 (52.89-82.85)
17	18.98 (12.51-25.99)	22.62 (13.98-31.70)	25.73 (16.60-36.34)	27.16 (17.33-37.35)
18	66.38 (56.61-78.09)	73.93 (57.30-91.05)	80.40 (59.73-99.97)	85.15 (62.88-105.03)
19	107.89 (104.71-111.23)	110.41 (106.67-114.59)	113.64 (108.69-118.75)	116.63 (110.65-122.77)
20	92.74 (78.52-105.27)	99.59 (87.48-110.03)	103.89 (92.86-114.01)	106.93 (94.88-117.62)
21	1.98 (0.76-3.57)	2.28 (0.91-3.76)	2.41 (1.11-3.95)	2.53 (1.26-4.05)
22	69.90 (63.59-76.60)	72.65 (65.26-81.03)	74.35 (65.47-83.56)	75.71 (66.17-85.60)
23	66.38 (61.05-71.84)	68.21 (61.92-75.22)	69.37 (62.13-77.68)	70.37 (62.33-78.84)
24	58.33 (51.57-65.25)	59.35 (51.54-67.38)	60.06 (51.42-69.08)	60.80 (51.46-70.54)
25	63.15 (58.19-68.33)	64.08 (58.13-70.60)	64.69 (57.61-72.31)	65.28 (57.52-73.31)
26	105.74 (103.11-108.63)	107.60 (104.49-111.07)	110.29 (106.44-114.88)	113.09 (108.14-118.30)
27	89.61 (78.88-101.88)	102.01 (94.07-109.15)	106.61 (100.25-112.38)	109.67 (103.03-116.43)
28	70.05 (57.85-82.73)	81.01 (62.78-97.43)	87.74 (70.89-102.85)	91.01 (74.04-106.02)
29	86.69 (76.48-98.85)	98.95 (89.88-106.59)	103.97 (97.05-110.40)	107.28 (99.82-114.22)
30	77.69 (70.35-86.27)	88.66 (80.56-97.04)	95.17 (87.24-102.87)	98.81 (91.27-106.20)
31	65.91 (58.13-74.04)	74.28 (62.87-85.82)	80.32 (68.22-91.58)	83.65 (71.55-95.35)
32	43.59 (32.90-53.93)	50.00 (37.85-62.47)	53.86 (40.49-67.43)	56.63 (43.72-69.69)
33	73.65 (66.59-81.02)	83.76 (75.75-91.83)	90.52 (82.64-98.89)	94.20 (86.45-101.91)
34	66.39 (59.59-73.92)	75.69 (66.62-84.88)	82.74 (73.07-92.11)	86.68 (77.96-95.89)
35	55.16 (49.55-61.45)	62.11 (52.50-71.67)	68.54 (57.82-79.59)	72.37 (61.24-83.34)
36	26.22 (16.45-36.35)	34.63 (24.40-45.80)	40.95 (29.49-52.36)	44.65 (32.52-56.98)
37	61.93 (52.16-71.12)	70.75 (59.42-82.03)	77.93 (66.19-89.43)	81.66 (70.20-92.73)
38	16.95 (9.65-25.17)	22.16 (13.21-32.08)	25.72 (15.22-36.54)	28.10 (17.03-40.00)
39	60.13 (54.25-66.49)	64.11 (56.61-72.19)	67.62 (58.43-76.31)	69.19 (60.38-79.31)
40	45.29 (40.71-50.62)	46.04 (40.58-52.32)	46.82 (40.54-53.98)	47.22 (40.49-55.01)

Tab. S7 (continued)

Node	*BeastRun4a* $r = 0.03$	*BeastRun4b* $r = 0.01$	*BeastRun4c* $r = 0.003$	*BeastRun4d* $r = 0.001$
41	49.27 (45.81-53.72)	50.08 (45.82-55.51)	51.16 (45.86-57.77)	51.52 (45.98-58.64)
42	15.35 (10.36-21.09)	18.74 (11.96-26.55)	20.86 (13.41-29.33)	21.76 (14.06-29.77)
43	7.67 (3.83-11.42)	10.04 (5.64-15.13)	11.44 (6.40-17.09)	12.23 (7.18-18.19)
44	103.38 (101.06-105.69)	104.51 (101.88-107.37)	106.32 (102.99-109.84)	108.34 (104.30-112.71)
45	48.26 (35.97-61.62)	67.93 (50.99-85.79)	78.08 (62.25-93.44)	83.19 (68.38-97.67)
46	26.35 (16.00-36.82)	40.06 (25.52-55.43)	48.12 (33.10-64.91)	53.06 (36.08-69.32)
47	36.03 (25.58-46.78)	50.43 (35.07-66.30)	59.55 (43.28-75.85)	63.57 (47.59-81.23)
48	21.01 (14.51-27.96)	28.52 (18.93-38.54)	33.10 (21.90-44.50)	35.55 (24.65-46.88)
49	7.67 (4.24-11.74)	10.16 (5.62-15.27)	11.61 (6.73-17.10)	12.47 (7.25-18.11)
50	102.05 (100.13-104.11)	102.84 (100.53-105.20)	104.17 (101.42-107.20)	105.84 (102.37-109.59)
51	99.82 (98.57-101.36)	100.09 (98.60-101.79)	100.72 (98.76-102.79)	101.58 (99.03-104.30)
52	68.48 (54.81-86.00)	82.81 (62.71-98.22)	88.65 (72.04-99.91)	91.15 (76.52-101.85)
53	61.62 (45.95-81.04)	74.34 (54.08-92.25)	78.73 (60.53-94.19)	81.05 (63.91-95.60)
54	98.66 (97.83-99.74)	98.72 (97.83-99.94)	99.04 (97.83-100.64)	99.56 (97.86-101.78)
55	71.09 (55.43-93.16)	84.42 (67.32-97.79)	89.96 (77.69-98.70)	91.54 (80.87-99.85)
56	62.30 (49.89-78.21)	74.76 (58.46-90.25)	80.86 (67.43-92.43)	82.77 (70.24-93.70)
57	36.59 (22.88-50.88)	43.86 (27.93-59.90)	48.26 (33.75-64.34)	50.06 (35.32-65.73)
58	77.14 (61.62-92.24)	86.03 (74.84-95.65)	88.99 (79.94-97.16)	90.70 (82.60-98.11)
59	29.22 (17.96-40.68)	38.18 (23.78-52.22)	42.47 (29.60-56.19)	44.20 (30.40-58.32)
60	52.74 (39.89-65.62)	61.09 (46.33-76.38)	65.34 (50.66-79.08)	67.74 (52.53-82.05)
61	69.04 (62.02-76.60)	83.61 (72.47-96.06)	94.29 (85.43-102.92)	98.67 (91.16-105.43)
62	57.02 (39.32-69.80)	73.44 (56.58-89.79)	84.75 (70.59-98.09)	89.74 (76.69-101.27)
63	31.54 (22.94-41.25)	40.46 (27.85-53.14)	47.36 (33.33-62.28)	51.16 (35.75-66.89)
64	19.39 (14.15-25.06)	23.74 (15.39-32.71)	27.95 (17.53-38.50)	29.88 (18.89-41.52)
65	63.88 (58.76-69.02)	75.49 (67.58-83.77)	85.43 (77.69-92.92)	90.60 (83.16-97.74)
66	60.21 (55.80-64.82)	69.49 (61.74-77.44)	78.42 (69.97-86.66)	83.73 (75.35-92.05)
67	57.63 (53.49-61.90)	65.47 (57.87-73.25)	73.93 (65.33-82.94)	79.24 (70.35-88.01)
68	42.78 (28.87-54.67)	49.75 (35.97-63.78)	56.88 (41.73-72.09)	62.09 (46.86-76.70)
69	9.30 (4.74-14.78)	12.80 (6.87-19.67)	15.01 (8.62-22.37)	16.61 (9.51-24.45)
70	55.01 (51.56-58.67)	60.39 (53.06-68.17)	67.57 (57.90-77.52)	72.57 (61.80-82.88)
71	15.13 (7.59-23.64)	20.13 (11.52-30.17)	23.97 (14.04-34.35)	25.63 (15.03-36.62)
72	29.64 (19.39-40.18)	40.83 (29.44-51.93)	48.72 (35.72-60.91)	52.87 (40.47-65.86)
73	22.06 (13.47-31.61)	31.92 (21.48-42.82)	38.90 (26.44-51.09)	42.68 (29.97-55.44)
74	27.42 (20.98-34.83)	35.39 (28.02-43.19)	40.48 (32.11-48.47)	43.41 (34.40-52.19)
75	19.43 (13.20-26.36)	26.46 (19.18-34.08)	30.82 (22.79-38.95)	33.20 (24.61-42.13)
76	11.93 (6.91-17.37)	16.56 (10.24-23.16)	19.64 (12.66-26.63)	21.29 (13.53-29.19)
77	0.61 (0.23-1.07)	0.60 (0.22-1.02)	0.62 (0.26-1.01)	0.64 (0.27-1.07)
78	22.94 (16.96-29.11)	30.01 (22.88-37.38)	34.24 (26.32-42.27)	36.53 (27.80-44.98)
79	18.05 (12.63-23.54)	23.98 (17.47-30.47)	27.44 (20.16-34.68)	29.36 (22.07-37.27)
80	11.97 (7.35-16.83)	16.32 (10.26-22.27)	18.60 (12.31-25.44)	20.01 (13.31-27.43)
81	4.07 (1.92-6.52)	4.64 (2.42-7.24)	5.20 (2.79-7.82)	5.47 (2.95-8.17)
82	61.49 (55.69-66.96)	73.32 (65.65-81.84)	83.16 (75.46-90.89)	88.30 (80.61-95.77)
83	56.96 (50.76-63.66)	69.08 (60.53-77.54)	78.80 (70.54-86.98)	83.97 (75.25-92.19)

Tab. S7 (continued)

Node	BeastRun4a $r = 0.03$	BeastRun4b $r = 0.01$	BeastRun4c $r = 0.003$	BeastRun4d $r = 0.001$
84	11.74 (7.65-16.62)	13.55 (8.42-19.09)	14.94 (9.60-20.90)	15.93 (10.13-22.56)
85	3.13 (1.46-5.08)	3.33 (1.72-5.17)	3.52 (1.89-5.39)	3.72 (2.06-5.56)
86	49.62 (43.12-56.41)	62.25 (53.27-70.49)	71.35 (62.84-80.04)	76.84 (68.16-85.53)
87	44.79 (36.60-52.89)	56.58 (45.94-67.04)	65.22 (54.12-75.54)	70.28 (58.29-81.55)
88	16.64 (9.15-24.66)	24.97 (14.85-36.17)	29.76 (18.66-41.30)	32.82 (20.84-45.01)
89	34.09 (28.08-40.51)	43.79 (36.51-51.99)	51.16 (42.53-59.27)	55.24 (46.19-63.84)
90	29.06 (23.39-34.55)	38.08 (30.96-45.80)	44.84 (36.91-52.83)	48.47 (39.64-56.53)
91	23.63 (17.18-29.93)	31.94 (23.78-40.18)	38.31 (29.54-47.11)	41.76 (32.03-51.27)
92	23.11 (18.11-28.45)	30.25 (23.18-38.23)	35.92 (27.57-44.13)	39.06 (30.47-48.04)
93	17.41 (13.41-21.96)	20.97 (14.60-27.87)	23.83 (16.06-31.94)	25.54 (17.33-33.76)
94	49.52 (41.58-57.16)	58.97 (49.77-68.01)	66.85 (57.03-76.37)	70.92 (60.38-81.13)
95	44.74 (37.22-52.55)	53.48 (44.45-62.25)	60.96 (51.47-70.71)	64.59 (54.16-75.02)
96	31.56 (24.13-39.16)	37.52 (29.44-45.99)	42.61 (33.50-52.04)	44.95 (35.54-54.83)
97	18.06 (13.49-22.72)	21.15 (16.24-26.23)	23.53 (18.15-29.09)	24.80 (19.53-30.20)
98	15.36 (11.62-19.03)	17.96 (14.16-22.08)	19.78 (15.72-24.19)	20.92 (16.83-25.27)
99	7.47 (4.80-10.35)	8.22 (5.65-11.14)	8.66 (5.92-11.40)	9.15 (6.39-11.94)
100	5.22 (2.95-7.82)	5.82 (3.48-8.28)	6.11 (3.89-8.58)	6.47 (4.10-8.91)
101	14.16 (10.67-17.72)	16.28 (12.47-19.99)	17.74 (13.69-21.74)	18.67 (14.70-22.98)
102	10.87 (7.96-13.95)	12.20 (9.21-15.18)	12.96 (9.95-16.04)	13.57 (10.46-16.90)
103	8.67 (5.40-12.09)	9.89 (6.49-13.37)	10.63 (7.38-14.02)	11.13 (7.77-14.65)
104	8.91 (6.22-11.75)	10.00 (7.24-12.80)	10.56 (7.97-13.57)	11.11 (8.27-14.21)
105	4.38 (2.60-6.29)	4.97 (3.11-7.04)	5.37 (3.54-7.41)	5.65 (3.77-7.63)
106	1.54 (0.69-2.51)	1.73 (0.90-2.72)	1.90 (0.99-2.87)	1.96 (1.09-2.96)

3   Text

Text S1: Tree topology.

The application of prior constraints for divergence dates is most effective with a fixed tree topology. Therefore, the optimal tree topology was explored with two initial phylogenetic inferences with the software BEAST (Drummond & Rambaut 2007) ('*BeastRun1*' and '*BeastRun2*'), and the resulting tree topology was fixed for the final round of divergence date estimations. Monophyly of all groups defined in SI Table S1, except Cottiformes, was confirmed in the first phylogenetic analysis performed without monophyly constraints (*BeastRun1*). Cottiformes appeared paraphyletic, with *Cottus gobio* as the sister group of Gasterosteoidei, instead of Zoarcoidei. However, support for the clade combining Cottoidei and Gasterosteoidei was low (BPP 0.55), and following Wiley & Johnson (2010), Cottiformes were nevertheless assumed monophyletic for subsequent analyses. In a second phylogenetic analysis (*BeastRun2*), the groups listed in SI Table S1, as well as Otomorpha and Euteleostomorpha were constrained to be monophyletic, and their earliest fossil records (see SI Text S2) were used to calculate prior distributions for the stem group divergences with R package 'ageprior' (here using $r = 0.01$, $p - q = 0.02$, $lag = 2$, and the respective correction for rock outcrop bias). The resulting topology (SI Fig. S13) was used for divergence date estimation in all subsequent analyses ('*BeastRun3a-d*', '*BeastRun4a-d*').

Text S2: Node constraints.

The following describes the constraints chosen for every single node of the molecular phylogeny presented in Fig. S13. The known fossil record is discussed and arguments are given for the assumed paleoenvironments. These are given only as 'marine', 'freshwater', or 'marin/freshwater', and are used to apply the respective correction for rock outcrop bias. Nomenclature follows Wiley & Johnson (2010), and occurrence data is taken from Nelson (2006) unless otherwise stated. Also, the timescale of Gradstein & Ogg (2009), presented in detail in Ogg & Ogg (2008a,b,c), is used for ICS stage definitions, unless otherwise declared.

Parameters t_y and t_o were chosen to reflect the earliest and latest possible age for the oldest known fossil record per clade. Here, uncertainties in the age of stage boundaries were neglected, and otolith fossils were generally assumed to be potentially misinterpreted. The fossil record has previously been used to calculate preservation rates for Cenozoic and Late Cretaceous mammals (Foote *et al.* 1999), Devonian tetrapods (Friedman & Brazeau 2011), Upper Cambrian-Lower Ordovician trilobites, and Jurassic bivalves (Foote & Raup 1996), however, no preservation rate estimate exists for teleost fishes. Therefore, a range of base preservation rates between 0.03 and 0.001 (0.03/0.01/0.003/0.001 in *BeastRun3a/b/c/d* and *BeastRun4a/b/c/d*, respectively) per lineage per myr were here used for Cenozoic marine teleosts, thus assuming that the quality of their fossil record is similar to that of Cenozoic mammals (Foote *et al.* 1999), or lower by a factor of 3, 10, or 30. Generally lower base preservation rates were assumed for Cenozoic freshwater teleosts (0.0169/0.0056/0.0017/0.0006 per lineage per myr), that directly reflect the difference between global marine (8.55×10^6 km^2) and continental (4.82×10^6 km^2) rock outcrop area, in which fossils could potentially be discovered (Wall *et al.* 2009). Intermediate Cenozoic base preservation rates were assumed for groups that could have occupied both marine and freshwater habitats

234

(0.0235/0.0078/0.0023/0.0008 per lineage per myr). These base preservation rates were used to extrapolate the stage specific preservation rate at time t_y for every node with the function 'adjust.pres.rate' of R package 'ageprior'. As the base preservation rates were assumed to apply to Cenozoic teleosts, the parameters t_1 and t_2 (the younger and older end of the range for which the preservation rate has been estimated) of function adjust.pres.rate are set to 0 and 65.5 Ma, and the respective correction for rock outcrop bias ('global_marine', 'global_terrestrial', or 'global') is used.

Net diversification rates $p - q$ were calculated iteratively with the software MEDUSA (Alfaro *et al.* 2009), implemented in the R package 'geiger' (Harmon *et al.* 2008), on the basis of different phylogenies (*BeastRun3a/b/c/d, BeastRun4a/b/c/d*) resulting from the four assumed base preservation rates (SI Table S6). Results of the second MEDUSA iteration are included in the parameter specifications given below. The lag time parameter was set to 2 myr for all nodes, assuming that the mean duration for recognizable apomorphies is constant between clades and over time. Finally, rock outcrop bias was corrected according to the assumed paleoenvironment. Global rock outcrop corrections ('global_marine', 'global_terrestrial', or 'global') are used for all nodes, as regional correction schemes would require local preservation rate estimates.

Node: 0 (root)

Group 1: Otomorpha (= Otocephala, Ostarioclupeomorpha)
Earliest record: †*Tischlingerichthys viohli* Arratia, 1997 from the lithographic limestone of Solnhofen, Bayern, Germany (Tithonian: 150.8-149.9 Ma).
Paleoenvironment: Marine/freshwater.
Comments: Otomorpha combine Clupeiformes, Gonorynchiformes, Cypriniformes, Characiformes, Siluriformes, and Gymnotiformes (Wiley & Johnson 2010). This is largely supported by molecular data (Peng *et al.* 2006; Azuma *et al.* 2008; Ortí & Li 2009). The earliest representative of this group is the stem ostariophysan †*Tischlingerichthys viohli* Arratia, 1997 from the upper Solnhofen Limestone Formation, Bayern, Germany (Tithonian) (Davis & Fielitz 2010). Dating of the upper Solnhofen Limestone Formation is based on ammonite zonation (Benton & Donoghue 2007; Benton *et al.* 2009). Otomorpha comprise marine (e.g. Clupeiformes) as well as freshwater groups (*e.g.* Cypriniformes). The Solnhofen Limestone represents marine deposits. The early diversification could have occurred in both marine and freshwater environments.

Group 2: Euteleosteomorpha
Earliest record: †*Leptolepides sprattiformis* Blainville, 1818 from the lithographic limestone of Cerin, France (Kimmeridgian: 155.7–150.8 Ma).
Paleoenvironment: Marine/freshwater.
Comments: Euteleosteomorpha unite Protacanthopterygii and Neoteleostei, which receives strong support from molecular data (Azuma *et al.* 2008; Santini *et al.* 2009; Davis & Fielitz 2010). The Solnhofen limestone (Tithonian) yields several euteleost fossils, including †*Leptolepides sprattiformis* Blainville, 1818, †*L. haerteisi* Arratia, 1997, †*Orthogonikleithrus hoelli* Arratia, 1997, and †*O. leichi*

Arratia, 1987 (Arratia & Tischlinger 2010). Specimens of †*L. sprattiformis* have also been preserved in the lithographic limestone of Cerin, France, which has been assigned a Kimmeridgian age, and is thus older than the Solnhofen Limestone (Wenz *et al.* 1993).Like Otomorpha, Euteleosteomorpha comprise both predominantly marine (Acanthomorphata), and freshwater groups (Salmoniformes). While the earliest fossils are found in marine deposits, freshwater members of Euteleosteomorpha at the time of early diversification cannot be excluded.

Parameters: $t_\mathrm{y} = 150.8$, $t_\mathrm{o} = 155.7$, $r = 0.0084/0.0028/0.0008/0.0003$, $p - q = 0.0447/0.0443/0.0463/0.0447$, $lag = 2$, $corr =$ "global"

Node: 1

Group 1: Clupei (= Clupeomorpha Greenwood et al. 1966)
Earliest record: †*Ellimmichthys longicaudatus* Cope, 1886 from the Marfim Formation, Recôncavo Basin, Bahia, Brazil (Hauterivian–Barremian: 136.4–125.0 Ma).
Paleoenvironment: Marine/freshwater.
Comments: Clupeomorph fishes appear in the fossil record in the Early Cretaceous, with the genera †*Diplomystus* Cope, 1877, †*Ellimma* Jordan, 1910, †*Ellimmichthys* Jordan, 1919, †*Spratticeps* Patterson, 1970, †*Erichalcis* Forey, 1975, †*Eoknightia* Taverne, 1976, †*Histiurus* Costa, 1850, †*Nolfia* Taverne, 1976, †*Paraclupea* Du, 1950, †*Santanaclupea* Maisey, 1993, and †*Pseudoellimma* De Figueiredo, 2009. Of these, the oldest specimens may be †*Diplomystus kokuraensis* Uyeno, 1979 and †*Diplomystus primotinus* Uyeno, 1979 from the Wakino Formation, Kanmon group, Kyushu Island, Japan, or †*Ellimmichthys longicaudatus* Cope, 1886 from Bahia, northeastern Brazil. The two specimens of †*Diplomystus* have been discovered in the upper layer (fourth layer) of the Wakino subgroup (Uyeno & Yabumoto 1980), which has been claimed to be Neocomian (Berriasian-Hauterivian) by the original authors, but is more likely Barremian (Matsumoto *et al.* 1982) or Hauterivian–Barremian Kimura:1992vx. †*Ellimmichthys longicaudatus* has been reported from the Marfim Formation, Recncavo Basin, Bahia, Brazil, that is of Late Hauterivian-Early Barremian age (Maisey 2000). Both the Wakino and Marfim formations represent non-marine deposits (Kimura *et al.* 1992; Maisey 2000), whereas extant Clupei are predominantly marine. The earliest representatives could have occupied both marine and freshwater environments.

Group 2: Ostariophysi
Earliest record: †*Tischlingerichthys viohli* Arratia, 1997 from the lithographic limestone of Solnhofen, Bayern, Germany (Tithonian: 150.8–149.9 Ma).
Paleoenvironment: Marine/freshwater.
Comments: Ostariophysi unite Gonorynchiformes, Cypriniformes, Characiformes, Siluriformes, and Gymnotiformes. Appearance of this group is marked by the stem ostariophysan †*Tischlingerichthys viohli* (see above). Most extant Ostariophysi are freshwater fishes, with the exception of six out of 37 gonorynchiform fishes. The earliest record of Ostariophysi, however, is from the marine

deposits of the Solnhofen Limestone.

Parameters: $t_y = 149.9$, $t_o = 150.8$, $r = 0.0084/0.0028/0.0008/0.0003$, $p - q = 0.0447/0.0443/$ $0.0463/0.0447$, $lag = 2$, $corr = $ "global"

Node: 2

Group 1: Gonorynchiformes
Earliest record: †*Rubiesichthys gregalis* Wenz, 1984 from Montsec, Lérida, Spain, (Berriasian–Valanginian: 145.5–136.4 Ma).
Paleoenvironment: Freshwater.
Comments: †*Rubiesichthys gregalis* was first discovered in the Early Cretaceous outcrop in Montsec, Lérida, Spain (Berriasian–Valanginian), and subsequently also found to be abundant in the younger outcrop of Las Hoyas, Spain (Hauterivian–Barremian) (Poyato-Ariza 1996). It has been described as a chanid Wenz:1993tm and a later revision confirmed its similarity to gonorynchiform fishes traditionally considered chanids (Poyato-Ariza 1996). Most extant gonorynchiforms are freshwater fishes, as was †*Rubiesichthys gregalis* (De Gibert *et al.* 2000).

Group 2: Otophysa
Earliest record: †*Santanichthys diasii* Silva Santos, 1958 from the Santana Formation (Romualdo Member) of the Araripe Basin, northeastern Brazil (Aptian–Turonian: 125.0–89.3 Ma).
Paleoenvironment: Marine/freshwater.
Comments: Otophysa combine Cypriniformes, Characiformes, Siluriformes, and Gymnotiformes, a clade that is supported by molecular data (Peng *et al.* 2006; Santini *et al.* 2009). †*Santanichthys diasii* has been described as a stem characiform and the oldest otophysan (Filleul & Maisey 2004), however the age of the Santana Formation, and especially that of the Romualdo Member, remains uncertain, and can only be constrained to Aptian–Turonian on the basis of vertebrate palaeontology and palynological data (Martill 2007). Nearly all extant members of Otophysa are freshwater fishes, however many of the earliest representatives were marine, and †*Santanichthys diasii* has been regarded as a marine or brackish water fish (Filleul & Maisey 2004). Thus, both marine and freshwater otophysans could have existed at the time of †*Santanichthys diasii*.

Parameters: $t_y = 136.4$, $t_o = 145.5$, $r = 0.0103/0.0034/0.0010/0.0003$, $p - q = 0.0447/0.0443/$ $0.0463/0.0447$, $lag = 2$, $corr = $ "global"

Node: 3

Group 1: Cypriniformes
Earliest record: †*Molinichthys inopinatus* Gayet, 1982 from the El Molino Formation, Agua Clara,

Bolivia (Campanian–Danian: 83.5–61.7 Ma).

Paleoenvironment: Freshwater.

Comments: The taxonomic assignment of †*Molinichthys inopinatus* from the bolivian El Molino Formation (Late Campanian–Early Danian) (Gayet *et al.* 1991) as a cypriniform is based on a fragment of a pharyngeal bone and has long been debated (Fink *et al.* 1984), in part because the fossil has been discovered in Bolivia, but no extant cypriniforms are known from South America. Besides †*Molinichthys inopinatus*, the oldest known cypriniforms are catostomids of the genus †*Amyzon* from the Paskapoo Formation, Alberta, Canada (Tiffanian) (Liu & Chang 2009). Extant cypriniforms are almost exclusively freshwater fishes, the El Molino Formation contains marine and continental sequences (Gayet *et al.* 1991), and the Paskapoo Formation is composed of alluvial deposits (Burns *et al.* 2010). Thus a freshwater paleoenvironment is assumed for early cypriniforms.

Group 2: Characiphysae (= Characiphysi Fink and Fink, 1981)

Earliest record: †*Santanichthys diasii* Silva Santos, 1958 from the Santana Formation (Romualdo Member) of the Araripe Basin, northeastern Brazil (Aptian–Turonian: 125.0–89.3 Ma).

Paleoenvironment: Marine/freshwater.

Comments: Characiphysae combines the three orders Characiformes, Siluriformes, and Gymnotiformes. †*Santanichthys diasii* is the oldest known member of this group (see above). Whereas almost all extant members of Characiphysae are freshwater fishes, †*Santanichthys diasii* and other early representatives were likely marine or brackish waters species (see above).

Parameters: $t_y = 89.3$, $t_o = 125.0$, $r = 0.0213/0.0071/0.0021/0.0007$, $p - q = 0.0447/0.0443/0.0463/0.0447$, $lag = 2$, $corr =$ "global"

Node: 4

Group 1: Protacanthopterygii

Earliest record: †*Nybelinoides brevis* Traquair, 1911 from Bernissart, Belgium (Upper Barremian–lowermost Aptian: 127.24–124.0 Ma).

Paleoenvironment: Marine/freshwater.

Comments: Protacanthopterygii unite Salmoniformes and Argentiniformes, whereby Salmoniformes contain the suborders Esocoidei, Osmeroidei, and Salmonoidei (Azuma *et al.* 2008; Wiley & Johnson 2010; Davis & Fielitz 2010). Some molecular evidence however supports a sister clade relationship of Argentiniformes with Neoteleostei (Santini *et al.* 2009). The earliest argentiniform record is provided by †*Nybelinoides brevis* Traquair, 1911 from Bernissart, Belgium (Barremian–Aptian) (Taverne 1982; Santini *et al.* 2009). †*Estesesox foxi* Wilson et al., 1992 and †*Oldmanesox canadensis* Wilson et al., 1992, from the Judith River and Milk River Formations, Alberta, Canada (Campanian) are the oldest esocoid fossils, and †*Spaniodon* Pictet, 1850 from Sahel Alma, Lebanon (Late Santonian) likely represents the earliest preserved osmeroid (Taverne & Filleul 2003). The oldest undebated salmonoid fossil, †*Eosalmo driftwoodensis* Wilson, 1977 from the Driftwood Creek

Formation, British Columbia, Canada, is no older than the Middle Eocene (Wilson & Li 1999). Several cretaceous fossils with unclear taxonomic affinities have been assigned to the order Salmoniformes, to the point that the group has been considered a 'waste-basket' for fossil fishes (Gallo *et al.* 2009). These include †*Casieroides yamangiensis* Taverne, 1975, †*Chardonius longicaudatus* Taverne, 1975, and †*Pseudoleptolepis minor* Taverne, 1975 from the Loia Beds, Democratic Republic of the Congo (Albian) (Murray 2000b), †*Helgolandichthys schmidi* Taverne, 1981 from Töck, Helgoland, Germany (Aptian), †*Barcarenichthys joneti* Gayet, 1981 from Barcarena, Portugal (Cenomanian), †*Gaudryella gaudryi* Pictet and Humbert, 1866, †*Pseudoberyx longispina* Pictet and Humbert, 1866, †*Gharbouria libanica* Gayet, 1988, and †*Ginsburgia operta* Patterson, 1970 from Hakel, Lebanon (Cenomanian), †*Goudkoffia delicata* David, 1946, and †*Natlandia ornata* David, 1946 from Southern California (Maastrichtian). However, all cretaceous 'Salmoniformes' remain of questionable position (Gallo *et al.* 2009). Given that morphological (Wiley & Johnson 2010), and most molecular evidence (Azuma *et al.* 2008; Davis & Fielitz 2010) supports a protacanthopterygiid position of Argentiniformes, I here assume the argentiniform †*Nybelinoides brevis* to represent the earliest known protacanthopterygiid. The Bernissart beds have been dated as 127.24–124 Ma (Schnyder *et al.* 2009).

The two orders included in Protacanthopterygii represent opposing lifestyles, whereby Salmoniformes are anadromous, and extant Argentiniformes are strictly marine. The deposits of Bernissart, Belgium, are lacustrine (Schnyder *et al.* 2009), which suggests that at least some early argentiniforms, including †*Nybelinoides brevis* were freshwater fishes.

Group 2: Neoteleostei
Earliest record: †*Atolvorator longipectoralis* Gallo and Coelho, 2008 from the Coqueiro Seco Formation of the Sergipe-Alagoas Basin, northeastern Brazil (Barremian: 130.0–125.0 Ma).
Paleoenvironment: Marine/freshwater.
Comments: Neoteleostei comprise Acanthomorphata (= Acanthomorpha Rosen, 1973), Myctophiformes, Aulopiformes, and Stomiatiformes (= Stomiiformes Fink et Weitzmann 1982). Monophyly of Neoteleostei is supported by molecular data (Azuma *et al.* 2008; Davis & Fielitz 2010). Acanthomorphata are abundant in the fossil record of the Cenomanian (Patterson 1993a), and seem to be absent in Lower Cretaceous deposits. Only few recent discoveries suggest the possibility that acanthomorphs could have originated before the Cenomanian. Wilson and Murray Wilson & Murray (1996) described the sphenocephalid †*Xenyllion zonensis* from the Fish Scale Zone of Alberta, Canada, and found Sphenocephaliformes to be the sister group to all other paracanthopterygiids (Murray & Wilson 1999). The authors report the Fish Scale Zone within the Shaftesbury Formation to be of Albian–Cenomanian age. The same sphenocephlid genus also occurs in the Mowry Shale in Utah, which is also assigned a late Albian or early Cenomanian age (Stewart 1996). In both cases, acanthomorph fossils were found in the zone of †*Neogastroplites americanus* (Stewart 1996), that has been dated at 97.92–97.53 Ma (Scott *et al.* 2009), and thus falls into the early Cenomanian. González-Rodríguez & Fielitz (2008) describe the acanthomorph †*Muhichthys cordobai* from the Muhi Quarry, Hidalgo State, Mexico, which is also reported to be Albian-Cenomanian in age (Fielitz & González-Rodríguez 2010). Finally, Applegate et al. Applegate *et al.* (2000)

claimed the presence of Beryciformes in the Tlayúa Quarry, Puebla, Mexico, which is Albian in age. However, taxonomic relationships of these fossil remain to be confirmed (González-Rodríguez & Fielitz 2008). Otolith fossils of "genus *Acanthomorphorum*" *forcallensis* Nolf, 2004, "genus *Aulopiformorum*" *pseudocentrolophus* Nolf, 2004, and "genus *Protacanthopterygiorum*" *scalpellum* Nolf, 2004 from Maestrazgo, Spain, that predate the first skeletal acanthomorph fossils by $\tilde{2}5$ myr (Early Aptian) are here considered questionable. The earliest acanthomorph records are thus older than the first myctophiform and stomiiform fossils, the genera †*Sardinioides* von der Marck, 1858 and †*Telepholis* von der Marck, 1858 from Sendenhorst, Westphalia, Germany (Campanian) (Dietze 2009), but postdate the first aulopiform, †*Atolvorator longipectoralis* Gallo and Coelho, 2008 from the Coqueiro Seco Formation of the Sergipe-Alagoas Basin, northeastern Brazil (Barremian). Most extant neoteleosts, including the most ancestral groups Aulopiformes, Myctophiformes, Paracanthopterygii (with the exception of Percopsiformes), and Berycomorpha, are marine. However, the paleoenvironment of †*Atolvorator longipectoralis* has been described as a brackish lake with irregular inputs of marine water flow (de Figueiredo 2009), thus, early freshwater neoteleosts cannot be excluded.

Parameters: $t_\mathrm{y} = 124.0$, $t_\mathrm{o} = 127.24$, $r = 0.0103/0.0034/0.0010/0.0003$, $p - q = 0.0447/0.0443/0.0463/0.0447$, $lag = 2$, $corr =$ "global"

Node: 5

Group 1: Salmonoidei
Earliest record: †*Eosalmo driftwoodensis* Wilson, 1977 from the Driftwood Creek Formation, British Columbia, Canada (Middle Eocene: 48.6–40.4 Ma).
Paleoenvironment: Freshwater.
Comments: †*Eosalmo driftwoodensis* is the oldest undebated salmonoid fossil (see above). Extant salmonoids are anadromous and freshwater fishes, and the Driftwood Creek Formation is of evaporitic origin (Simandl & Hancock 1991). Anadromous life style is likely a derived feature of modern salmonoids, and the early history of salmonoids is thus assumed to be non-marine.

Group 2: Esocoidei
Earliest record: †*Estesesox foxi* Wilson et al., 1992 and †*Oldmanesox canadensis* Wilson et al., 1992, from the Judith River and Milk River Formations, Alberta, Canada (Campanian: 83.5–70.6 Ma).
Paleoenvironment: Freshwater.
Comments: Molecular data support Osmeroidei basal to a clade combining Salmonoidei and Esocoidei (López *et al.* 2004), thus the osmeroid fossil †*Spaniodon* from the Late Santonian is not used to constrain this node. All extant esocoids are freshwater fishes. The Milk River Formation represents the transition from an open shelf to a non-marine environment (McCrory & Walker 1986), and the Judith River sediments are primarily of fluvial origin (Dawson *et al.* 1994). Thus, a freshwater

paleoenvironment of esocoid fishes is assumed.

Parameters: $t_y = 70.6$, $t_o = 83.5$, $r = 0.0122/0.0041/0.0012/0.0004$, $p - q = 0.0447/0.0443/0.0463/0.0447$, $lag = 2$, $corr = $ "global_terrestrial"

Node: 6

Group 1: Aulopiformes
Earliest record: †*Atolvorator longipectoralis* Gallo and Coelho, 2008 from the Coqueiro Seco Formation of the Sergipe-Alagoas Basin, northeastern Brazil (Barremian: 130.0–125.0 Ma).
Paleoenvironment: Marine/freshwater.
Comments: See above.

Group 2: Ctenosquamata
Earliest record: †*Xenyllion zonensis* Wilson and Murray, 1996 from the Fish Scale Zone of Alberta, Canada (earliest Cenomanian: 97.92–97.53 Ma), or †*Muhichthys cordobai* from the Muhi Quarry, Hidalgo State, Mexico (Cenomanian: 99.6–93.5 Ma).
Paleoenvironment: Marine.
Comments: Ctenosquamata include the two subsections Myctophata (with only one order, Myctophiformes) and Acanthomorphata (= Acanthomorpha Rosen 1973). The earliest myctophiform record is provided by upper Cretaceous fossils from Sendenhorst, Westphalia, Germany (Campanian), that were originally described under the name †*Osmeroides monasteri* Agassiz, 1835 and were later assigned to the genus †*Sardionioides* (Prokofiev 2006; Dietze 2009). The acanthomorph record predates the earliest myctophiforms, the stem paracanthopterygiid †*Xenyllion zonensis* Wilson and Murray, 1996 is older than 97.53 Ma (earliest Cenomanian), but whether the acanthomorph record extends to the Albian remains questionable (see above). Here, an early Cenomanian age is assumed for the first acanthomorph record (99.6–97.53 Ma). Most extant members of Acanthomorphata and Myctophiformes are marine, including the most basal lineages Paracanthopterygii and Berycomorpha, and the earliest record †*Xenyllion zonensis* is also considered a marine fish (Murray & Wilson 1999). Thus the paleoenvironment of Ctenosquamata is assumed to be marine.

Parameters: $t_y = 125.0$, $t_o = 130.0$, $r = 0.0103/0.0034/0.0010/0.0003$, $p - q = 0.0447/0.0443/0.0463/0.0447$, $lag = 2$, $corr = $ "global"

Node: 7

Group 1: Paracanthopterygii *sensu* Miya *et al.* 2003
Earliest record: †*Xenyllion zonensis* Wilson and Murray, 1996 from the Fish Scale Zone of Alberta, Canada (earliest Cenomanian: 97.92–97.53 Ma).

Paleoenvironment: Marine.

Comments: Paracanthopterygii *sensu* Miya *et al.* (2003) combine Polymixiiformes, Percopsiformes, Gadiformes, and Zeiformes (excluding Caproidei). Monophyly of this group has been corroborated by further work based on molecular data (Azuma *et al.* 2008; Matschiner *et al.* 2011). Presumably polymixiiform fossils from the Cenomanian are often used to constrain divergence dates in molecular dating studies (Azuma *et al.* 2008; Miya *et al.* 2010; Matschiner *et al.* 2011). The Late Cretaceous polymixiid genera listed by Patterson (1993a) are †*Berycopsia* Radovcic, 1975, †*Berycopsis* Dixon, 1850, †*Dalmatichthys* Radovcic, 1975, †*Homonotichthys* Whitley, 1933, and †*Omosoma* Costa, 1857. The oldest of these are †*Berycopsis elegans* from the English Chalk of Sussex, United Kingdom, †*Homonotichthys* from the English Chalk of Malling and Brighton, United Kingdom and †*Omosoma simum* Arambourg, 1954 from Djebel Tselfat, Morocco. The English Chalk is thought to be of Cenomanian–Turonian age (Patterson 1993a), but †*Berycopsis elegans* Dixon, 1850 and †*Homonotichthys rotundus* Smith Woodward, 1902 were both recorded from the zone of †*Holaster subglobosus* of Kent and Sussex by Woodward (1902), which can be constrained as Middle to Upper Cenomanian (Jacobs *et al.* 2005). An additional species of †*Berycopsis* was described from Hajula, Lebanon, confirming that †*Berycopsis* was present in the Cenomanian (Bannikov & Bacchia 2005). The age of †*Omosoma simum* is somewhat less certain: While Djebel Tselfat is frequently cited as cenomanian, there are no absolute age assignments, nor has any microfauna been recorded. The Cenomanian age assignment is based on vertebrate assemblages (Arambourg 1954).

†*Berycopsis* was removed from Berycopsidae and placed in Polymixiidae by Patterson (1964). However, a recent systematic analysis recovers †*Berycopsis* in a basal position to a clade combining Polymixiiformes and Stephanoberyciformes (Dietze 2009). †*Homonotichthys* was initially placed in Berycidae by Woodward (1902), and later transferred to the beryciform family Holocentridae by Regan (1911). Patterson (1964) finally placed †*Homonotichthys* in the family Polymixiidae, due to similarities with the extant genus *Polymixia*, particulary in the structure of the neurocranium and hyoid arch. †*Omosoma* was transferred from Stromateidae to Polymixiidae by Regan (1911), based on its resemblances to †*Berycopsis*. †*Omosoma* was excluded from Dietzes (2009) systematic analysis because of uncertain character states, however †*Omosoma* is very similar to †*Berycopsis*, and may thus belong to the same ancestral position (Dietze 2009). Thus, the taxonomic assignment of both †*Berycopsis* and †*Omosoma* is questionable. †*Homonotichthys*, on the other hand (together with the Santonian †*Pycnosteroides* Hay, 1903) seems to be more closely related to Polymixia (Patterson 1964), and is thus here accepted as the oldest polymixiiform fossil.

The earliest record of Percopsiformes is provided by the stem group member †*Mcconichthys longipinnis* Grande, 1988 from the Tullock Formation, Montana (Murray 1996; Murray & Wilson 1999). The Tullock formation is frequently cited as early Paleocene, and therefore, a Danian age of †*Mcconichthys longipinnis* is assumed here.

A large gadiform otolith assemblage is known from the Paleocene deposits of Denmark (Schwarzhans 2003; Kriwet & Hecht 2008), but the only skeletal remains from the Paleocene are †"*Protacodus*" sp. from Greenland, considered to be the most ancient gadiform by Fedotov & Bannikov (1989). According to the description in Kriwet & Hecht (2008), otolith remains from Sundkrogen, Den-

mark, such as †*Coelorhynchus balticus* Koken, 1885 are older than †"*Protacodus*" (Selandian).The earliest zeiform record is †*Cretazeus rinaldii* Tyler et al., 2000 from the Calcare di Mellissano, Nardò, Italy (Upper Campanian–Lower Maastrichtian), according to Tyler & Santini (2005). The Middle Campanian-Upper Campanian boundary is 76.4 Ma, while the Lower Maastrichtian-Upper Maastrichtian boundary is 69.2 Ma (Ogg & Ogg 2008a).

In summary, the stem paracanthopterygiid genus †*Xenyllion* (see above) and †*Homonotichthys* may be the only known representatives of Paracanthopterygii from the Cenomanian. While †*Xenyllion* is known from the earliest Cenomanian, †*Homonotichthys* is of middle to upper Cenomanian age.

Most extant members of Paracanthopterygii (with the exception of percopsiforms) are marine, and the same is assumed for the earliest known members of this group (see above). Thus, the early diversification likely took place in a marine paleoenvironment.

Group 2: Acanthopterygii

Earliest record: †*Plesioberyx maximus* Gayet, 1980 and †*Plectocretacicus clarae* Sorbini, 1979 from the Lithografic Limestone of Hakel, Lebanon (Early Cenomanian: 99.1–97.8 Ma), or †*Lissoberyx anceps* Arambourg, 1954 from Djebel Tselfat, Morocco (Cenomanian: 99.6–93.5 Ma), or †*Aipichthyoides* Gayet, 1980 from the Bet Meir/Amminadav Formation, 'Ein Yabrud quarries, Israel.

Paleoenvironment: Marine/freshwater.

Comments: Based on morphological and molecular evidence (Miya *et al.* 2003, 2005; Mabuchi *et al.* 2007; Azuma *et al.* 2008; Kawahara *et al.* 2008; Wiley & Johnson 2010; Matschiner *et al.* 2011), Acanthopterygii are here considered to include the orders Berycomorpha *sensu* Miya *et al.* 2005, Ophidiiformes, Gobiiformes, "Stromateiforme", Scombriformes, Carangiformes, Pleuronectiformes, Batrachoidiformes, Synbranchiformes, Anabantiformes, Beloniformes, Cyprinodontiformes, Gobiesociformes, Blenniiformes, Atheriniformes, Mugiliformes, "Labriformes" *sensu* Kaufman & Liem 1982 , Elassomatiformes, "Perciformes" *sensu* stricto (see Wiley & Johnson 2010), "Caproiformes", Lophiiformes, Tetraodontiformes, "Trachiniformes", "Scorpaeniformes", Cottiformes, "Gasterosteiformes", Nototheniiformes, Dactylopteriformes, Pholidichthyiformes, Icosteiformes. The following families are considered Acanthopterygii *incertia sedis*: Pseudochromidae, Plesiopidae, Grammatidae, Opistognathidae, Notograptidae, Percidae.

Disregarding putative "berciform" fossils from the Albian Tlayúa Quarry, Puebla, Mexico, the appearance of acanthopterygiids in the fossil record is marked by various "berciforms", trachichthyoids, holocentroids, and a single putative tetraodontiform from the Cenomanian (Patterson 1993a). Additional Cretaceous tetraodontiforms have been described subsequently (Santini & Tyler 2003). Following Miya *et al.* (2005), Berycomorpha are here considered to include the reciprocally paraphyletic clades "Berciformes" and "Stephanoberciformes", whereby "Berciformes" encompasses Berycidae, Holocentridae, Trachichthyidae, Monocentridae, Anomalopidae, Anoplogastridae, and Diretmidae (Johnson & Patterson 1993). Thus, the large number of "berciform", trachichthyoid, and holocentroid fossils from Cenomanian deposits of Lebanon, Israel, Morocco, and the United Kingdom are strong evidence of the presence of Berycormorpha in the Cenomanian. While the

Moroccan Djebel Tsefat, as well as the Israelian Bet Meir/Amminadav Formation have been dated only to stage level (99.6–93.5 Ma), more precise age assignments are possible for †*Trachichthyoides* Woodward, 1902 from the English Chalk of Bromley, Kent, United Kingdom and the Lebanese localities of Namoura, Hajula, and Hakel, all of which bear "beryciform" fossils (Forey *et al.* 2003). †*Trachichthyoides* is from the Cenomanian zone of †*Holaster subglosus*, that can be constrained as middle to upper Cenomanian (see above), the Namoura deposits are placed in the middle portion of the Middle Cenomanian, and a later Lower Cenomanian age is assigned to the fish-bearing layers of both Hajula and Hakel, with Hakel being the older of the two deposits (Forey *et al.* 2003; Wippich & Lehmann 2004). The occurences of *Mantelliceras mantelli* and the foraminifer *Orbitulina concava* in the Lithographic Limestone of Hakel suggest an age between 99.1 and 97.8 Ma (Benton *et al.* 2009). "Beryciform" fossils from Hakel include †*Plesioberyx maximus* Gayet, 1980, †*Plesioberyx discoides* Gayet, 1980, †*Caproberyx pharsus* Patterson, 1967, and †*Stichopteryx lewisi* Davis, 1887 (Forey *et al.* 2003). Because †*Lissoberyx anceps* Arambourg, 1954 from Djebel Tselfat, Morocco, and †*Aipichthyoides* Gayet, 1980 from the Bet Meir/Amminadav Formation, 'Ein Yabrud quarries, Israel, could theoretically be older than the Hakel Limestone (while still being Cenomanian), the oldest berycomorph fossil is between 99.6 and 97.8 million years old.

The Lithographic Limestone of Hakel is also the type locality of the oldest tetraodontiform, †*Plectocretacicus clarae* Sorbini, 1979 that has been placed in the superfamily †Plectocretacicoidea, together with †*Protriacanthus gortanii* d'Erasmo, 1946 from Comen, northwestern Slovenia (Upper Cenomanian–Lower Turonian) and †*Cretatriacanthus guidottii* from Canale, Nardò, Italy (Upper Campanian–Lower Maastrichtian) (Tyler & Sorbini 1996). An extensive phylogenetic analysis of 56 extant and fossil tetraodontiforms (Santini & Tyler 2003) corroborated the close relationship of the three plectocretacicoids, and apparently suggested a stem-tetraodontiform position of †Plectocretacicoidea. However, only a single outgroup (either the zeid *Cyttus novazelandiae*, or the caproid *Antigonia capros*) was included in this analysis, which by definition renders the ingroup monophyletic, and thus does not in itself support the tetraodontiform assignment of plectocretacicoids. A reanalysis of the Santini & Tyler (2003) dataset, including *Cyttus novazelandiae* as an outgroup, but also *Antigonia capros* as part of the ingroup, shows that plectocretacicoids cluster with all other tetraodontiforms in 100% of the most parsimonious trees, whereas *Antigonia capros* is found at a more ancestral position (not shown). Given that caproids may indeed be part of the sister group of tetraodontiforms (which is suggested by the molecular data of this study), this does indeed support a stem-tetraodontiform position of plectocretacicoids.

Most of the extant members of acanthopterygii are marine. Also, most, but not all of the earliest acanthopterygiid fossils has been found in marine deposits. †*Spinocaudichthys oumtkoutensis* Filleul and Dutheil, 2001 has been described from freshwater deposits of the Kem Kem beds, Tafraout, Morocco, shows that freshwater acanthomorphs existed as early as the Cenomanian (Filleul & Dutheil 2001).

Parameters: $t_\mathrm{y} = 97.8$, $t_\mathrm{o} = 99.6$, $r = 0.0103/0.0034/0.0010/0.0003$, $p - q = 0.0447/0.0443/0.0463/0.0447$, $lag = 2$, $corr =$ "global"

Node: 8

Group 1: Polymixiiformes
Earliest record: †*Homonotichthys rotundus* Smith Woodward, 1902 from the English Chalk of Sussex, United Kingdom (Middle–Upper Cenomanian: 96.0–93.5 Ma)
Paleoenvironment: Marine.
Comments: See above for a discussion of the polymixiiform fossil record. An ancestral position of Polymixiiformes within Paracanthopterygii is supported by molecular data (Miya *et al.* 2003). The Middle to Upper Cenomanian dates to 96.0–93.5 Ma according to Ogg & Ogg (2008a). Extant Polymixiiformes are exclusively marine, as are the polymixiiform fossils from the Cenomanian (Filleul & Dutheil 2001).

Group 2: Percopsiformes, Gadiformes, Zeiformes
Earliest record: †*Cretazeus rinaldii* Tyler et al., 2000 from the Calcare di Mellissano, Nardò, Italy (Upper Campanian–Lower Maastrichtian: 76.4–69.2 Ma).
Paleoenvironment: Marine/freshwater.
Comments: †*Cretazeus rinaldii* is older than the earliest percopsiform, †*Mcconichthys longipinnis*, and the earliest gadiform record, †*Coelorhynchus balticus* (see above). Of the three orders, Percopsiformes appears to be the most basal lineage (Matschiner *et al.* 2011). Given that all percopsiforms are freshwater fishes, whereas both gadiforms and zeiforms are marine, the early history of the three order could have occurred in both marine and freshwater environments.

Parameters: $t_y = 93.5$, $t_o = 96.0$, $r = 0.0213/0.0071/0.0021/0.0007$, $p - q = 0.0447/0.0443/0.0463/0.0447$, $lag = 2$, $corr =$ "global"

Node: 9

Group 1: *Polymixia japonica*
Earliest record: Extant.
Paleoenvironment: Marine.
Comments: -

Group 2: *Polymixia nobilis*
Earliest record: Extant.
Paleoenvironment: Marine.
Comments: -

Parameters: $t_y = 0$, $t_o = 0$, $r = 0.0441/0.0147/0.0044/0.0015$, $p-q = 0.0447/0.0443/0.0463/0.0447$, $lag = 2$, $corr =$ "global_marine"

Node: 10

Group 1: Percopsiformes
Earliest record: †*Mcconichthys longipinnis* Grande, 1988 from the Tullock Formation, Montana (Danian: 65.5–61.7).
Paleoenvironment: Freshwater.
Comments: Sedimentation in the Tullock Formation occurred in ponds, floodplains, and river channels (Sheehan & Fastovsky 1992).

Group 2: Gadiformes, Zeiformes
Earliest record: †*Cretazeus rinaldii* Tyler *et al.* 2000 from the Calcare di Mellissano, Nardò, Italy (Upper Campanian–Lower Maastrichtian: 76.4–69.2 Ma).
Paleoenvironment: Marine.
Comments: Monophyly of a clade combining Gadiformes and Zeiformes is supported by molecular data (Miya *et al.* 2003, 2005). †*Cretazeus rinaldii* is older than the earliest gadiform, †*Coelorhynchus balticus* (see above). Nearly all extant gadiforms and zeiforms are marine, as was †*Cretazeus rinaldii* (Friedman 2010).

Parameters: $t_y = 69.2$, $t_o = 76.4$, $r = 0.0213/0.0071/0.0021/0.0007$, $p - q = 0.0447/0.0443/ 0.0463/0.0447$, $lag = 2$, $corr =$ "global"

Node: 11

Group 1: Gadiformes
Earliest record: †*Coelorhynchus balticus* Koken, 1885 from Sundkrogen, Denmark (Danian: 65.5–61.7)
Paleoenvironment: Marine.
Comments: See above.

Group 2: Zeiformes
Earliest record: †*Cretazeus rinaldii* Tyler *et al.*, 2000 from the Calcare di Mellissano, Nardò, Italy (Upper Campanian–Lower Maastrichtian: 76.4–69.2 Ma).
Paleoenvironment: Marine.
Comments: See above.

Parameters: $t_y = 69.2$, $t_o = 76.4$, $r = 0.0309/0.0103/0.0031/0.0010$, $p - q = 0.0447/0.0443/0.0463/ 0.0447$, $lag = 2$, $corr =$ "global_marine"

Node: 12

Group 1: Berycomorpha
Earliest record: †*Plesioberyx maximus* Gayet, 1980 from the Lithografic Limestone of Hakel, Lebanon (Early Cenomanian: 99.1–97.8 Ma), or †*Lissoberyx anceps* Arambourg, 1954 from Djebel Tselfat, Morocco (Cenomanian: 99.6–93.5 Ma), or †*Aipichthyoides* Gayet, 1980 from the Bet Meir/ Amminadav Formation, 'Ein Yabrud quarries, Israel.
Paleoenvironment: Marine.
Comments: An ancestral position of Berycomorpha within Acanthopterygii is supported by molecular data Miya *et al.* (2005). See above for a discussion of the "beryciform" fossil record.

Group 2: Percomorpha
Earliest record: †*Plectocretacicus clarae* Sorbini, 1979 from the Lithografic Limestone of Hakel, Lebanon (Early Cenomanian: 99.1–97.8 Ma).
Paleoenvironment: Marine.
Comments: Percomorpha are assumed to include all Acanthopterygii except Berycomorpha (Miya *et al.* 2003, 2005, 2010). See above for a list of Acanthopterygii, as defined here. †*Plectocretacicus clarae* is the oldest tetraodontiform fossil (Santini & Tyler 2003, see above).

Parameters: $t_y = 97.8$, $t_o = 99.6$, $r = 0.0309/0.0103/0.0031/0.0010$, $p-q = 0.0447/0.0443/0.0463/ 0.0447$, $lag = 2$, $corr = $ "global_marine"

Node: 13

Group 1: Ophidiiformes, "Syngnathoidei" (+ Dactylopteriformes)
Earliest record: †*Gasterorhampus zuppichinii* Sorbini, 1981 from the Calcare di Mellissano, Nardò, Italy, (Upper Campanian–Lower Maastrichtian: 76.4–69.2 Ma).
Paleoenvironment: Marine.
Comments: The monophyly of "Gastererosteiformes" has long been debated (Wilson & Orr 2011). As shown by Kawahara *et al.* (2008), the order consists of two monophyletic groups, the suborder Gasterosteioidei (comprising the families Hypoptychidae, Aulorhynchidae, and Gasteosteidae), and the family Indostomidae (that is recovered within Synbranchiformes, thus rendering this order paraphyletic), as well as the paraphyletic suborder "Syngnathoidei" into which Dactylopterifomes appear to be inserted (Smith & Wheeler 2004; Kawahara *et al.* 2008). A clade combining "Syngnathoidei" and Ophidiiformes is not supported by morphological analyses, and has received only little support from molecular data (reviewed in Kawahara *et al.* 2008). In Smith & Wheeler (2004), "Syngnathoidei" (represented by the genera *Pegasus* and *Aulostomus*) appeared polyphyletic, whereby *Pegasus* clustered with the ophidiiform *Chilara taylori* and the "trachinoid" *Champsodon* c.f. *atridorsalis*. *Aulostomus* on the other hand clustered with the dactylopteriform *Dactylopterus volitans*. In contrast, Kawahara *et al.* (2008) included representatives of all "syngnathoid" families, and found a monopyletic, albeit weekly supported, clade combining "Syngnathoidei" and Dactylopteriformes,

but no affinities with Ophidiiformes. Here a monophyletic clade comprising "Syngnathoidei" and Ophidiiformes is tentatively accepted on the basis of the sequence data presented in this study. Following Smith & Wheeler (2004), Kawahara *et al.* (2008), Setiamarga *et al.* (2008), and Wilson & Orr (2011), inclusion of Dactylopteriformes into "Syngnathoidei" is also assumed.

Ophidiiforms have an extensive fossil record based on otoliths, dating back to the Early Eocene (Rossi *et al.* 2000). Skeletal remains are scarce, but include specimens from Monte Bolca, Italy (Ypresian–Lutatian), and the ophidiid †*Eolamprogrammus senectus* Danil'chenko, 1968 from the Danata Formation, Turkmenistan. The fish-bearing layer of the Danata Formation has been considered synchronous with the Upper Thanetian anoxic event (Muzylev 1994).

The earliest "syngnathoid" record is provided by †*Gasterorhampus zuppichinii* Sorbini, 1981 from the Calcare di Mellissano, Nardò, Italy (Upper Campanian–Lower Maastrichtian; see above), which has been described as a macrorhamphosid, but has been argued to branch from the common stem of Macrorhamphosidae and Centriscidae (Orr 1995; Friedman 2009), or from the stem of a clade composed of Macrorhamphosidae, Centriscidae, and Syngnathidae, Solenostomidae, and Pegasidae (Natale 2008). Both studies agree that †*Gasterorhampus zuppichinii* is a member of "Syngnathoidei" (Orr 1995; Natale 2008). The dactyliform fossil record is poor, or even nonexisting (Imamura 2000), thus, †*Gasterorhampus zuppichinii* also provides the earliest record of the clade combining Ophidiiformes, "Syngnathoidei", and Dactylopteriformes. Nearly all extant members of this clade, and the earliest fossils, are marine.

Group 2: Percomorpha excluding Ophidiiformes, "Syngnathoidei", and Dactylopteriformes
Earliest record: †*Plectocretacicus clarae* Sorbini 1979 from the Lithografic Limestone of Hakel, Lebanon (Early Cenomanian: 99.1–97.8 Ma).
Paleoenvironment: Marine.
Comments: See above.

Parameters: $t_y = 97.8$, $t_o = 99.6$, $r = 0.0309/0.0103/0.0031/0.0010$, $p-q = 0.0447/0.0443/0.0463/0.0447$, $lag = 2$, $corr = $ "global_marine"

Node: 14

Group 1: "Syngnathoidei" (+ Dactylopteriformes)
Earliest record: †*Gasterorhampus zuppichinii* Sorbini, 1981 from the Calcare di Mellissano, Nardò, Italy, (Upper Campanian–Lower Maastrichtian: 76.4–69.2 Ma).
Paleoenvironment: Marine.
Comments: See above for a discussion of the "syngnathoid" and dactylopteriform fossil record.

Group 2: Ophidiiformes
Earliest record: †*Eolamprogrammus senectus* Danil'chenko, 1968 from the Danata Formation, Uylya-Kushlyuk, Turkmenistan (Upper Thanetian: 57.23–55.8 Ma).

Paleoenvironment: Marine.

Comments: See above for a brief description of the ophidiiform fossil record. The Upper Thanetian is here interpreted as the interval between the Th-4 boundary and the Thanetian–Ypresian boundary (57.23–55.8 Ma).

Parameters: $t_y = 69.2$, $t_o = 76.4$, $r = 0.0309/0.0103/0.0031/0.0010$, $p-q = 0.0447/0.0443/0.0463/0.0447$, $lag = 2$, $corr = $ "global_marine"

Node: 15

Group 1: Gastrophori *sensu* Duncker 1912, 1915
Earliest record: †*Hipposyngnathus neriticus* Jerzmañska, 1968 from the Menilite Formation (Jamna Dolma Member), outer Carpathian basin, Poland (earliest Rupelian: 33.1–32.5 Ma).
Paleoenvironment: Marine.
Comments: All four "syngnathoids" included in this study are members of the family Syngnathidae. Following Duncker (1912, 1915), Wilson & Orr (2011) accepted the partition of Syngnathidae into two natural groups, the abdominal-bearing Gastrophori and the tail-bearing Urophori. *Nerophis ophidion* is the only Gastrophori here included.

The oldest gastrophorine fossil is †*Hipposyngnathus neriticus* Jerzmañska, 1968 from the Menilite Formation, Poland (Wilson & Orr 2011). The Jalmna Dolma Member is synchronous with nannofossil zone NP22 (Kotlarczyk *et al.* 2006), which is earliest Rupelian in age (33.1–32.5 Ma) (Ogg & Ogg 2008c). †*Hipposyngnathus convexus* Danil'chenko, 1960 from the Maikop Group, Perekishkyul, northern Azerbaijan, has also been considered as an Oligocene gastrophorine by Wilson & Orr (2011), however, the age of the fish-bearing layers of the Maikop Group may have been misinterpreted and be Miocene rather than Oligocene (Popov *et al.* 2008).

Group 2: Urophori *sensu* Duncker 1912, 1915
Earliest record: †*Syngnathus incompletus* Cosmovici, 1887 from the Menilite Formation (IPM2 zone) outer Carpathian basin, Poland (early Rupelian: 33.1–30.0 Ma).
Paleoenvironment: Marine.
Comments: Urophori combine six syngnathid subfamilies, including the genera *Syngnathus* and *Hippocampus*. The earliest Urophori may be †*Syngnathus incompletus* from the Menilite Formation, Poland, which first occurs in the IPM2 zone, corresponding to nannofossil zones NP22 and NP23 (Kotlarczyk *et al.* 2006), and is thus early Rupelian in age (33.1–30.0 Ma) (Ogg & Ogg 2008c). †*Syngnathus incertus* Danil'chenko, 1960 from the Maikop Group is probably younger (see above).

Parameters: $t_y = 32.5$, $t_o = 33.1$, $r = 0.0225/0.0075/0.0023/0.0008$, $p-q = 0.0447/0.0443/0.0463/0.0447$, $lag = 2$, $corr = $ "global_marine"

Node: 16

Group 1: *Syngnathus* (+ *Enneacampus*, *Urocampus*, *Hippichthys*)
Earliest record: †*Syngnathus incompletus* Cosmovici, 1887 from the Menilite Formation (IPM2 zone) outer Carpathian basin, Poland (early Rupelian: 33.1–30.0 Ma).
Paleoenvironment: Marine.
Comments: Molecular data moderately supports a clade combining *Syngnathus*, *Enneacampus*, *Urocampus*, and *Hippichthys* (Wilson & Orr 2011). No fossils of the latter three genera are known (see above).

Group 2: *Hippocampus*
Earliest record: †*Hippocampus sarmaticus* Žalohar, Hitij & Kriznar, 2009 from the Coprolitic Horizon, Tunjice, Slovenia (middle Serravallian: 13.0–12.0 Ma).
Paleoenvironment: Marine.
Comments: †*Hippocampus sarmaticus* occurs in the second bed of the Coprolithic Horizon, and is thus the oldest member of the genus *Hippocampus* (Žalohar *et al.* 2009; Wilson & Orr 2011). The Coprolithic Horizon is dated to a lower Sarmatian stage (a Central Paratethys stage), which is synchronous with the middle Serravallian (Žalohar *et al.* 2009). The boundaries of Chron C5A (13.0–12.0 Ma) are here used to define middle Serravallian (Ogg & Ogg 2008b).

Parameters: $t_\mathrm{y} = 30.0$, $t_\mathrm{o} = 33.1$, $r = 0.0225/0.0075/0.0023/0.0008$, $p-q = 0.0447/0.0443/0.0463/0.0447$, $lag = 2$, $corr =$ "global_marine"

Node: 17

Group 1: *Hippocampus abdominalis* (+ *H. trimaculatus*, *H. camelopardalis*, *H. breviceps*, *H. coronatus*, *H. mohnikei*, *H. histrix*, *H. whitei*, *H. barbouri*, *H. subelongatus*, *H. comes*)
Earliest record: †*Hippocampus sarmaticus* Žalohar, Hitij & Kriznar, 2009 from the Coprolitic Horizon, Tunjice, Slovenia (middle Serravallian: 13.0–12.0 Ma).
Paleoenvironment: Marine.
Comments: The phylogeny of Wilson & Orr (2011) supports a clade of *H. abdominalis*, *H. comes*, and *H. barbouri*, while Casey *et al.* (2004) recover a group containing *H. abdominalis*, *H. trimaculatus*, *H. camelopardalis*, *H. breviceps*, *H. coronatus*, *H. mohnikei*, *H. histrix*, *H. whitei*, *H. barbouri*, *H. subelongatus*, and *H. comes* with high support. As †*Hippocampus sarmaticus* resembles *H. trimaculatus* (Žalohar *et al.* 2009), it is here used as the earliest record of this group.

Group 2: *Hippocampus kuda* (+ *H. ingens*, *H. capensis*, *H. algiricus*, *H. reidi*)
Earliest record: Extant.
Paleoenvironment: Marine.
Comments: The phylogeny of Wilson & Orr (2011) supports a clade of *H. kuda*, *H. ingens*, *H.*

reidi, *H. erectus*, and *H. zosterae*, while a group of *H. kuda*, *H. capensis*, *H. algiricus*, and *H. ingens* is highly supported in the analysis of Casey *et al.* (2004). No fossils are known of this group.

Parameters: $t_y = 12.0$, $t_o = 13.0$, $r = 0.0441/0.0147/0.0044/0.0015$, $p-q = 0.0447/0.0443/0.0463/$ 0.0447, $lag = 2$, $corr =$ "global_marine"

Node: 18

Group 1: Bythitoidei
Earliest record: †*Propteridium* Arambourg, 1967 from Istehbanat, Iran (Rupelian: 33.9–28.4 Ma).
Paleoenvironment: Marine.
Comments: The suborder Bythitoidei comprises the families Bythitidae and Aphyonidae. The earliest skeletal record is †*Propteridium* Arambourg, 1967 from Istehbanat, Iran (Cohen *et al.* 1990), which is from the Rupelian (Tyler 2000).

Group 2: "Ophidioidei"
Earliest record: †*Eolamprogrammus senectus* Danil'chenko, 1968 from the Danata Formation, Uylya-Kushlyuk, Turkmenistan (Upper Thanetian: 57.23–55.8 Ma).
Paleoenvironment: Marine.
Comments: According to Nielsen *et al.* (1999), the suborder "Ophidioidei" combines the families Carapodidae and Ophidiidae. It is considered problematic by Wiley & Johnson (2010), as no convincing synapomorphies have been proposed. According to Rossi *et al.* (2000), †*Eolamprogrammus senectus* could be assigned to the family Ophidiidae, and thus provides the earliest "ophidioid" record. Even if the suborder is paraphyletic, this fossil can be used to constrain the split between the ophidiid *Dicrolene* and the bythitid *Brotula*.

Parameters: $t_y = 55.8$, $t_o = 57.23$, $r = 0.0168/0.0056/0.0017/0.0006$, $p-q = 0.0447/0.0443/0.0463/$ 0.0447, $lag = 2$, $corr =$ "global_marine"

Node: 19

Group 1: Gobiiformes, "Stromateiformes", Scombriformes
Earliest record: †*Eutrichiurides opiensis* (Leriche 1906) Casier, 1944 from the Phosphates of Morocco (Danian–Lutetian: 65.5–40.4), †*Sphyraenodus multidentatus* Dartevelle & Casier, 1959 from the Landana Cliffs, Cabinda enclave, Angola (Danian–Lutetian: 65.5–40.4), or †*Ardiodus mariotti* White, 1931 from the Oldhaven Beds, Upnor, Kent, United Kingdom (Upper Thanetian, NP9, C25n: 57.2–56.6 Ma).
Paleoenvironment: Marine.

Comments: The clade combining "Stromateiformes" and Scombriformes is strongly supported by the molecular data presented here, and a close relationship of Gobiiformes with Scombriformes has been found previously (Mabuchi *et al.* 2007). The fossil record of Gobiiformes is scarce. Fossil remains of the extant genus *Gobiomorphus* have been described from the Bannockburn Formation, Central Otago, New Zealand (Early Miocene) (McDowall *et al.* 2006), which is predated by "*Gobius*" *gracilis* (Laube 1901) Obrhelová, 1961, skeletal fossils from the Early Oligocene of Bohemia (Böhme 2007; Gaudant 2009). The oldest "stromateiform" record is probably from the Ypresian or Upper Paleocene: According to Bannikov (1995), two "stromateioids" were reported by Bonde in the Mo-Clay (Fur/Ølst) Formation, Danmark. More recently, Sytchevskaya & Prokofiev (2005) claim the discovery of a centrolophid in the Danata Formation, Turkmenistan. The Mo-Clay (Fur/Ølst) Formation has been assigned an Ypresian age (Willumsen 2004), and the Danata Formation is Upper Thanetian in age (see above). Scombriformes *sensu* Johnson 1986 comprise the families Scombrolabracidae, Sphyraenidae, Gempylidae, Trichiuridae, Scombridae, Istiophoridae, and Xiphiidae. The monophyly of this group, as well as its interfamiliar relationships are controversial. On the basis of molecular data, Xiphiidae, Istiophoridae, and Sphyraenidae were found ancestral to a clade combining the remaining scombriform families and Sparidae, Moronidae, and Nemipteridae, albeit with low support (Orrell *et al.* 2006). Li *et al.* (2009) recovered Sphyraenidae as the sister group of Carangidae, and Xiphiidae combined with Menidae, again, with low support. On the other hand, Trichiuridae and Scombridae were found closely related to the perciform family Bramidae, and to the stromateiform family Nomeidae.

Regardless of the taxonomic placement of Xiphiidae, Istiophoridae, and Sphyraenidae, the earliest known scombriform record most likely dates from the Paleocene. Cretaceous fossils assigned to †*Blochiidae*, such as †*Cylindracanthus* Leidy, 1856 and †*Congorhynchus* Dartevelle & Casier, 1949 have probably been misinterpreted as scombriforms (Monsch 2004). Following Monsch (2004) and Fierstine (2006), the next-oldest scombriform fossils include †*Trichiurus gulincki* Casier, 1967 from the Tienen Formation, Dormaal, Belgium (MP7: earliest Ypresian) (Fairon-Demaret 2002; Ogg & Ogg 2008c), †*Scomberomorus* Lacepède, 1802, †*Auxides turkmenita* Danil'chenko, 1968, †*Hemingwaya sarissa* Sytchevskaya and Prokofiev, 2002, †*Eocoelopoma portentosa* Bannikov, 1985, and †*Palaeothunnus parvidentatus* Danil'chenko, 1968 from the Danata Formation, Turkmenistan (Upper Thanetian, see above), and †*Ardiodus mariotti* White, 1931 from the Oldhaven Beds, Upnor, Kent, United Kingdom (NP9, C25n: Upper Thanetian) (Ellison *et al.* 1994; Ogg & Ogg 2008c) and from the Phosphates of Morocco (Monsch 2004). Patterson (1993b) further lists †*Eutrichiurides opiensis* (Leriche 1906) Casier, 1944, described from the Phosphates of Morocco, and its congeneric †*Eutrichiurides africanus* Dartevelle and Casier, 1949, which was found in the Landana Cliffs of the Cabinda enclave, Angola. †*Eutrichiurides* has been placed in Gempylidae by Patterson (1993b), but Monsch (2004) agrees with Casier (1944) that it is a member of Trichiurinae instead. Furthermore, †*Sarda palaeocenica* Leriche, 1909, †*Sphyraenodus multidentatus* Dartevelle and Casier, 1959, †*Landanichthys lusitanicus* Dartevelle and Casier, 1949 and †*Landanichthys moutai* Dartevelle and Casier, 1949 are all known from the Landana Cliffs. This formation has been considered 'Montian' (Danian) in age Murray (2000b), but may be as young as Ypresian–Bartonian (Figueiredo *et al.* 2011). The age of the Moroccan Phosphates is similary unclear: The specimen of †*Ardiodus mariotti*

described by Monsch (2004) is reported from 'Morocco, Phosphates, Late Paleocene-Early Eocene: Thanetian-Ypresian', whereas specimens of †*Eutrichiurides opiensis* have been discovered in the 'Montian Phosphates, Morocco' (Patterson 1993b). However, a reference is not given for this claim, thus it is difficult to substantiate. In Morocco, vertebrate remains have been found in Phosphates of Khouribga, that are of Maastrichtian–Lower Lutetian age (Chernoff & Orris 2002). In the absence of a more precise dating of the Moroccan Phosphates and the Landana Cliffs, Danian–Lutetian age is assumed for both localities. Some evidence for Middle Paleocene presence of Scombriformes comes form otolith remains of 'genus Scombridarum' sp. from the Tashlik Formation, Luzanivka, Cherkassy Region, Ukraine (Selandian, NP5–6).

A lower limit for the oldest scombriform fossil is given by †*Ardiodus mariotti*: The Oldhaven Beds, near Upnor, Kent, have been assigned to the NP9 nannofossil zone and to chron C25n, which constrains its age at 57.2–56.6 Ma (Ogg & Ogg 2008c). Thus, the earliest scombriform record, and the earliest record of the clade combining Gobiiformes, "Stromateiformes", and Scombriformes is between 65.5 and 56.6 million years old.

Extant members of "Stromateiformes" and Scombriformes are almost exclusively marine, whereas about 10% of the gobiiform species live in freshwater habitats. The earliest record of the group comining Gobiiformes, "Stromateiformes", and Scombriformes is reported from the Moroccan Phosphates, the Landana Cliffs, Angola, and the Oldhaven Beds, United Kingdom, which are all marine deposits (Weir & Catt 1969; Murray 2000b). Thus this clades early history is assumed to be marine.

Group 2: Percomorpha excluding "Syngnathoidei", Dactylopteriformes, Ophidiiformes, Gobiiformes, "Stromateiformes", and Scombriformes
Earliest record: †*Plectocretacicus clarae* Sorbini, 1979 from the Lithografic Limestone of Hakel, Lebanon (Early Cenomanian: 99.1–97.8 Ma).
Paleoenvironment: Marine.
Comments: See above.

Parameters: $t_{\mathrm{y}} = 97.8$, $t_{\mathrm{o}} = 99.1$, $r = 0.0309/0.0103/0.0031/0.0010$, $p{-}q = 0.0559/0.0627/0.0694/0.0726$, $lag = 2$, $corr =$ "global_marine"

Node: 20

Group 1: Gobiiformes
Earliest record: "*Gobius*" *gracilis* (Laube 1901) Obrhelová, 1961 from Seifhennersdorf-Varnsdorf, Bohemia, Czech Republic (Rupelian: 33.9–30.7 Ma).
Paleoenvironment: Marine/freshwater.
Comments: At Seifhennersdorf, the top of the fossiliferous layer has been dated as 30.7 Ma, and the formation is assigned a Rupelian age (Gaudant 2009). *Gobius gracilis* is considered to be an eleotrid (Böhme 2007; Gaudant 2009). The earliest records of Gobiiformes are found in lacustrine sediments (Böhme 2007) whereas the great majority of extant gobiiforms are marine. Thus, both

types of paleoenvironments could have been used by early gobiiforms.

Group 2: "Stromateiformes", Scombriformes
Earliest record: †*Eutrichiurides opiensis* (Leriche 1906) Casier, 1944 from the Phosphates of Morocco (Danian–Lutetian: 65.5–40.4), †*Sphyraenodus multidentatus* Dartevelle and Casier, 1959 from the Landana Cliffs, Cabinda enclave, Angola (Danian–Lutetian: 65.5–40.4), or †*Ardiodus mariotti* White 1931 from the Oldhaven Beds, Upnor, Kent, United Kingdom (Upper Thanetian, NP9, C25n: 57.2–56.6 Ma).
Paleoenvironment: Marine.
Comments: See above.

Parameters: $t_y = 56.6$, $t_o = 65.5$, $r = 0.0125/0.0042/0.0013/0.0004$, $p-q = 0.0559/0.0627/0.0694/0.0726$, $lag = 2$, $corr =$ "global"

Node: 21

Group 1: *Gymnogobius isaza*
Earliest record: Extant.
Paleoenvironment: Freshwater.
Comments: *Gymnogobius isaza* is endemic to Lake Biwa, Japan.

Group 2: *Gymnogobius petschiliensis* (+ *G. urotaenia*)
Earliest record: Extant.
Paleoenvironment: Marine/freshwater.
Comments: *Gymnogobius petschiliensis* and *G. urotaenia* are the sister group of *G. isaza* (Sota *et al.* 2005). *Gymnogobius* has no fossil record. *Gymnogobius petschiliensis* occurs in the Yellow Sea, but also in brackish or freshwater habitats along the Japanes coast (Stevenson 2002).

Parameters: $t_y = 0$, $t_o = 0$, $r = 0.0380/0.0127/0.0038/0.0013$, $p - q = 0.0559/0.0627/0.0694/0.0726$, $lag = 2$, $corr =$ "global"

Node: 22

Group 1: Centrolophidae
Earliest record: †*Seriola paleocenica* Sytchevskaya and Prokofiev, 2005 from the Danata Formation, Uylya-Kushlyuk, Turkmenistan (Upper Thanetian: 57.23–55.8 Ma)
Paleoenvironment: Marine.
Comments: †*Seriola paleocenica* is mentioned in Bannikov (2009).

Group 2: Scombriformes, Nomeidae, Stromateidae (+ Ariommatidae)
Earliest record: †*Eutrichiurides opiensis* (Leriche 1906) Casier, 1944 from the Phosphates of Morocco (Danian–Lutetian: 65.5–40.4), †*Sphyraenodus multidentatus* Dartevelle and Casier, 1959 from the Landana Cliffs, Cabinda enclave, Angola (Danian–Lutetian: 65.5–40.4), or †*Ardiodus mariotti* White 1931 from the Oldhaven Beds, Upnor, Kent, United Kingdom (Upper Thanetian, NP9, C25n: 57.2–56.6 Ma).
Paleoenvironment: Marine.
Comments: See above for a discussion of the scombriform and "stromateiform" fossil record. Supported by morphological (Doiuchi *et al.* 2004) and molecular data (Doiuchi & Nakabo 2006), the "stromateiform" families Ariommatidae, Nomeidae, and Stromateidae form a sister group to Centrolophidae. The positions of the remaining two "stromateiform" families Amarsipidae (monotypic) and Tetragonuridae (three species) is uncertain (Doiuchi & Nakabo 2006). Scombriformes are the sister group to Nomeidae and Stromateidae according to molecular data presented here.

Parameters: $t_y = 56.6$, $t_o = 65.5$, $r = 0.0168/0.0056/0.0017/0.0006$, $p-q = 0.0559/0.0627/0.0694/0.0726$, $lag = 2$, $corr =$ "global_marine"

Node: 23

Group 1: Nomeidae, Stromateidae
Earliest record: Undescribed nomeid from the Mo-Clay (Fur/Ølst) Formation, Danmark (Ypresian: 55.8–48.6 Ma).
Paleoenvironment: Marine.
Comments: The position of Ariommatidae relative to Nomeidae, Stromateidae, and Scombriformes is not known. According to Bannikov (1995), one out of two probable "stromateoids" described by Bonde from the Mo-Clay (Fur/Ølst) Formation, Danmark, has been assigned to the Nomeidae.

Group 2: Scombriformes
Earliest record: †*Eutrichiurides opiensis* (Leriche 1906) Casier, 1944 from the Phosphates of Morocco (Danian–Lutetian: 65.5–40.4), †*Sphyraenodus multidentatus* Dartevelle and Casier, 1959 from the Landana Cliffs, Cabinda enclave, Angola (Danian–Lutetian: 65.5–40.4), or †*Ardiodus mariotti* White, 1931 from the Oldhaven Beds, Upnor, Kent, United Kingdom (Upper Thanetian, NP9, C25n: 57.2–56.6 Ma).
Paleoenvironment: Marine.
Comments: See above.

Parameters: $t_y = 56.6$, $t_o = 65.5$, $r = 0.0168/0.0056/0.0017/0.0006$, $p-q = 0.0559/0.0627/0.0694/0.0726$, $lag = 2$, $corr =$ "global_marine"

Node: 24

Group 1: Nomeidae
Earliest record: Undescribed nomeid from the Mo-Clay (Fur/Ølst) Formation, Danmark (Ypresian: 55.8–48.6 Ma).
Paleoenvironment: Marine.
Comments: See above.

Group 2: Stromateidae
Earliest record: †*Stromateus brailloni* NoIf 1975 (otolith remains) from the Sables dAuvers, Paris Basin, France (Lutetian–Bartonian, NP16: 42.4–39.7 Ma).
Paleoenvironment: Marine.
Comments: †*Stromateus brailloni* is listed in Patterson (1993b), and the Auvers Formation is there assigned a Lutetian age. According to Merle (2005), however, the Auvers Formation is nannofossil zone NP16 (Lutetian-Bartonian: 42.4–39.7 Ma) (Ogg & Ogg 2008c).

Parameters: $t_y = 48.6$, $t_o = 55.8$, $r = 0.0247/0.0082/0.0025/0.0008$, $p-q = 0.0559/0.0627/0.0694/0.0726$, $lag = 2$, $corr =$ "global_marine"

Node: 25

Group 1: Gempylidae (+ Scombrini, *Gasterochisma*)
Earliest record: †*Auxides turkmenita* Danil'chenko, 1968 from the Danata Formation, Uylya-Kushlyuk, Turkmenistan (Upper Thanetian: 57.23–55.8 Ma).
Paleoenvironment: Marine.
Comments: According to the cladogram of Block & Finnerty (1994) basd on molecular data, gempylids form a monophyletic clade with scombrins (genera Rastrelliger and Scomber) and the monotypic genus *Gasterochisma*. The oldest known member of this clade is †*Auxides turkmenita* from the Danata Formation, Turkmenistan. †*Auxides* is closely related to the extant genus *Scomber* (Monsch 2006).

Group 2: Thunnini (+ Sardini)
Earliest record: †*Palaeothunnus parvidentatus* Danil'chenko, 1968 from the Danata Formation, Uylya-Kushlyuk, Turkmenistan (Upper Thanetian: 57.23–55.8 Ma), or †*Sarda palaeocenica* Leriche, 1909 from the Landana Formation, Cabinda, Angola (Danian–Lutetian: 65.5–40.4 Ma).
Paleoenvironment: Marine.
Comments: Based on molecular data of Block & Finnerty (1994), Sardini (there represented by *Sarda*, and assuming monophyly of the tribe) are the sister group of Thunnini. †*Sarda palaeocenica* from the Landana Formation, Angola, is known from isolated teeth only (Patterson 1993b), and the Landana Formation has never been directly dated. †*Palaeothunnus parvidentatus* from the Danata

Formation, Turkmenistan, seems to be a basal member of the clade combining Thunnini and Sardini (Bannikov 2009).

Parameters: $t_y = 55.8$, $t_o = 65.5$, $r = 0.0168/0.0056/0.0017/0.0006$, $p-q = 0.0559/0.0627/0.0694/$ 0.0726, $lag = 2$, $corr = $ "global_marine"

Node: 26

Group 1: Carangiformes, Pleuronectiformes, "Batrachoidiformes" (+ Synbranchiformes, Indostomatidae, Anabantiformes), "Atherinomorphae", Gobiesociformes, Blenniiformes, Mugiliformes (+ Pseudochromidae, Plesiopidae, Grammatidae, Opistognathidae, Notograptidae), "Labriformes" I

Earliest record: †*Trachicaranx tersus* Danil'chenko, 1968 (and others) from the Danata Formation, Uylya-Kushlyuk, Turkmenistan (Upper Thanetian: 57.23–55.8 Ma), or *Cyprinodon* (?) *primulus* from the Maíz Gordo Formation, Argentina (Upper Thanetian, Riochican: 57.4–55.8 Ma).

Paleoenvironment: Marine/freshwater.

Comments: Close relationship of Carangiformes and Pleuronectiformes is strongly supported by molecular data (Miya *et al.* 2005; Setiamarga *et al.* 2008; Matschiner *et al.* 2011). Sister group relationship of Batrachoidiformes with a clade combining Atheriniformes, Beloniformes, Cyprinodontiformes, and Mugiliformes has been suggested by molecular data (Miya *et al.* 2005) and is confirmed here. "Batrachoidiformes" were found to be paraphyletic with Synbranchiformes and Indostomatidae nested inside of it, and Anabantiformes closely related to it (Miya *et al.* 2005; Mabuchi *et al.* 2007; Kawahara *et al.* 2008; Li *et al.* 2009). The superorder "Atherinomorphae" comprises Atheriniformes, Beloniformes, and Cyprinodontiformes, which is supported by morphological data (Wiley & Johnson 2010), but may be paraphyletic and include Gobiesociformes, Blenniiformes, Mugiliformes, and some "Labriformes", according to molecular data (Miya *et al.* 2005; Mabuchi *et al.* 2007; Matschiner *et al.* 2011). A close affiliation of Mugiliformes and Pseudochromidae is strongly supported (Mabuchi *et al.* 2007; Setiamarga *et al.* 2008), and the five "percoid" families Pseudochromidae, Grammatidae, Plesiopidae, Notograptidae, and Opistognathidae are considered to form a clade on the basis of morphology, with support from molecular data (Nelson 2006; Smith & Craig 2007). "Labriformes" *sensu* Kaufman & Liem 1982 comprise the four families Labridae (including Scaridae and Odacidae), Cichlidae, Embiotocidae, and Pomacentridae, which most likely represents a polyphyletic group (Mabuchi *et al.* 2007). Here, "Labriformes" I is considered to include Cichlidae, Embiotocidae, and Pomacentridae as a monophyletic group, whereas "Labriformes" II, composed of Labridae including Scaridae and Odacidae, is more closely related to Tetraodontiformes and allies. A clade combining "Atherinomorphae", Gobiesociformes, Blenniiformes, Mugiliformes, Pseudochromidae, Plesiopidae, Grammatidae, Opistognathidae, Notograptidae, and "Labriformes" I has long been hypothesized on the basis of morphology (Stiassny 1993), and has recently been corroborated by the molecular data of Li *et al.* (2009), who proposed the name Stiassnyformes for this clade. The sister group relationship of Carangiformes and Pleuronectiformes with "Batrachoid-

iformes" and Stiassnyformes must be considered tentative.

The earliest records of Carangiformes and Pleuronectiformes are provided by †*Trachicaranx tersus* Danil'chenko, 1968 from the Danata Formation, Turkmenistan (Upper Thanetian) and †*Heteronectes chaneti* Friedman, 2008 from Monte Bolca, Verona, Italy (Ypresian), respectively. The only skeletal remains of "Batrachoidiformes" is †*Halobatrachus didactylus* from Oran, Algeria (Miocene) (Carnevale 2004), which is predated by the otolith remains of †*"Batrachoididarum" trapezoidalis* Nolf, 1988 from Argile de Gan, Gan, France (Ypresian) (Patterson 1993b). No synbrachiform fossils are known (Patterson 1993b). Stiassnyformes are present in the Eocene deposits of Monte Bolca, Italy (Bellwood & Sorbini 1996), Mahenge, Tanzania (Murray 2000a), and the Lumbrera Formation, Argentina (Malabarba *et al.* 2010). The earliest record of Stiassnyformes may be provided by *Cyprinodon* (?) *primulus* Cockerell, 1936 from the Maíz Gordo Formation, Argentina (Cione & Báez 2007). The Maíz Gordo Formation has been assigned to the South American Land Mammal Age (SALMA) Riochican (Pascual *et al.* 1981), which is 57.4–55.8 Ma (Ogg & Ogg 2008c). Most extant members of this group are marine, however the earliest fossil record is provided by specimens from both marine (Bannikov 2000) and lacustrine taxa (Do Campo *et al.* 2007).

Group 2: "Labriformes" II, "Perciformes" *sensu* stricto (+ Acanthuriformes), "Caproiformes" I, Lophiiformes, Tetraodontiformes, Epinephelidae *sensu* Smith and Craig 2007, "Trachiniformes", Serranidae *sensu* Smith and Craig 2007, Percidae, Scorpaenoidei, Cottiformes, Gasterosteioidei, Nototheniiformes
Earliest record: †*Plectocretacicus clarae* Sorbini, 1979 from the Lithografic Limestone of Hakel, Lebanon (Early Cenomanian: 99.1–97.8 Ma).
Paleoenvironment: Marine.
Comments: See above. Acanthuriformes seem to be nested within "Perciformes" *sensu* stricto (Holcroft & Wiley 2008). "Caproiformes" are polyphyletic, and "Caproiformes" I is here considered to include Antigoniinae, which are assumed to be monophyletic. A single species remains in Caproiformes II (*Capros aper*), which is here ignored due to its uncertain taxonomic position. Nearly all extant members of this group are marine, and so is the ichthyofauna of Hakel, Lebanon.

Parameters: $t_y = 97.8$, $t_o = 99.1$, $r = 0.0213/0.0071/0.0021/0.0007$, $p-q = 0.0559/0.0627/0.0694/0.0726$, $lag = 2$, $corr =$ "global"

Node: 27

Group 1: Carangiformes, Pleuronectiformes
Earliest record: †*Trachicaranx tersus* Danil'chenko, 1968 (and others) from the Danata Formation, Uylya-Kushlyuk, Turkmenistan (Upper Thanetian: 57.23–55.8 Ma).
Paleoenvironment: Marine.
Comments: See above.

Group 2: "Batrachoidiformes" (+ Synbranchiformes, Indostomatidae, Anabantiformes), "Atherinomorphae", Gobiesociformes, Blenniiformes, Mugiliformes (+ Pseudochromidae, Plesiopidae, Grammatidae, Opistognathidae, Notograptidae), "Labriformes" I
Earliest record: *Cyprinodon* (?) *primulus* Cockerell, 1936 from the Maíz Gordo Formation, Argentina (Upper Thanetian, Riochican: 57.4–55.8 Ma).
Paleoenvironment: Marine/freshwater.
Comments: See above.

Parameters: $t_y = 55.8$, $t_o = 57.4$, $r = 0.0125/0.0042/0.0013/0.0004$, $p-q = 0.0559/0.0627/0.0694/0.0726$, $lag = 2$, $corr =$ "global"

Node: 28

Group 1: Carangiformes
Earliest record: †*Trachicaranx tersus* Danil'chenko, 1968 (and others) from the Danata Formation, Uylya-Kushlyuk, Turkmenistan (Upper Thanetian: 57.23–55.8 Ma).
Paleoenvironment: Marine.
Comments: See above.

Group 2: Pleuronectiformes
Earliest record: †*Heteronectes chaneti* Friedman, 2008 (and others) from Monte Bolca, Verona, Italy, (Ypresian, NP14, SB11: 49.4–49.1 Ma).
Paleoenvironment: Marine.
Comments: The age of the Monte Bolca outcrops is derived from the presence of *Discoaster sublodoensis*, which indicates nannoplancton zone NP14 (Monsch 2006). It can be further constrained to SB11 by larger foraminiferal assemblages of *Alveolina cremae*, *A. rugosa*, *A. distefanoi*, and *A. rutimeyeri* (Trevisani *et al.* 2005). The top of SB11 (49.1 Ma) is just above the base of NP14 (49.4 Ma), thus the age of the Monte Bolca outcrops can be dated to 49.4–49.1 Ma (Ogg & Ogg 2008c).

Parameters: $t_y = 55.8$, $t_o = 57.23$, $r = 0.0168/0.0056/0.0017/0.0006$, $p-q = 0.0559/0.0627/0.0694/0.0726$, $lag = 2$, $corr =$ "global_marine"

Node: 29

Group 1: "Batrachoidiformes" (+ Synbranchiformes, Indostomatidae, Anabantiformes)
Earliest record: †"*Batrachoididarum*" *trapezoidalis* Nolf, 1988 from Argile de Gan, Gan, France (Ypresian: 55.8–48.6 Ma).
Paleoenvironment: Marine/freshwater.
Comments: Only few "batrachoidiform" species, but most synbranchiform and anabantiform, and

all indostomatid fishes are confined to freshwater.

Group 2: "Atherinomorphae", Gobiesociformes, Blenniiformes, Mugiliformes (+ Pseudochromidae, Plesiopidae, Grammatidae, Opistognathidae, Notograptidae), "Labriformes" I
Earliest record: *Cyprinodon* (?) *primulus* Cockerell, 1936 from the Maíz Gordo Formation, Argentina (Upper Thanetian, Riochican: 57.4–55.8 Ma).
Paleoenvironment: Marine/freshwater.
Comments: See above.

Parameters: $t_y = 55.8$, $t_o = 57.4$, $r = 0.0125/0.0042/0.0013/0.0004$, $p-q = 0.0559/0.0627/0.0694/0.0726$, $lag = 2$, $corr =$ "global"

Node: 30

Group 1: Beloniformes, Cyprinodontiformes
Earliest record: *Cyprinodon* (?) *primulus* Cockerell, 1936 from the Maíz Gordo Formation, Argentina (Upper Thanetian, Riochican: 57.4–55.8 Ma).
Paleoenvironment: Marine/freshwater.
Comments: See above. About 50% of extant Beloniformes and almost all Cyprinodontiformes are freshwater fishes.

Group 2: Gobiesociformes, Blenniiformes, Atheriniformes, Mugiliformes (+ Pseudochromidae, Plesiopidae, Grammatidae, Opistognathidae, Notograptidae), "Labriformes" I
Earliest record: †*Palaeopomacentrus orphae* Bellwood and Sorbini, 1996 and †*Lorenzichthys olihan* Bellwood, 1999 from Monte Bolca, Verona, Italy (Ypresian, NP14, SB11: 49.4–49.1 Ma), or otoliths of the "genus *Atherinidarum*" from Argile de Gan, Gan, France (Ypresian: 55.8–48.6 Ma), and from the Cambay Shale Formation, Gujarat, India (Ypresian: 55.8–48.6 Ma).
Paleoenvironment: Marine/freshwater.
Comments: †*Palaeopomacentrus orphae* and †*Lorenzichthys olihan* Bellwood, 1999 have been described as pomacentrids, and are thus the earliest skeletal representative of "Labriformes" I (Cowman *et al.* 2009). They may be predated by otoliths of the "genus *Atherinidarum*" from Argile de Gan, France, and from the Cambay Shale Formation, India, which are both thought to be Ypresian in age (Nolf *et al.* 2006; Benton & Donoghue 2007; Benton *et al.* 2009). This group is composed of both marine and freshwater fishes.

Parameters: $t_y = 55.8$, $t_o = 57.4$, $r = 0.0125/0.0042/0.0013/0.0004$, $p-q = 0.0559/0.0627/0.0694/0.0726$, $lag = 2$, $corr =$ "global"

Node: 31

260

Group 1: Beloniformes
Earliest record: †*Rhamphexocoetus volans* Bannikov *et al.*, 1985 from Monte Bolca, Verona, Italy (Ypresian, NP14, SB11: 49.4–49.1 Ma).
Paleoenvironment: Marine/freshwater.
Comments: †*Rhamphexocoetus volans* is the earliest beloniform record according to Patterson (1993b).

Group 2: Cyprinodontiformes
Earliest record: *Cyprinodon* (?) *primulus* Cockerell, 1936 from the Maíz Gordo Formation, Argentina (Upper Thanetian, Riochican: 57.4–55.8 Ma).
Paleoenvironment: Freshwater.
Comments: See above.

Parameters: $t_y = 55.8$, $t_o = 57.4$, $r = 0.0125/0.0042/0.0013/0.0004$, $p-q = 0.0559/0.0627/0.0694/0.0726$, $lag = 2$, $corr =$ "global"

Node: 32

Group 1: Funduloidea
Earliest record: †*Tapatia occidentalis* Alvarez and Arriola-Longoria, 1972 from Santa Rosa, Jalisco, Mexico (Miocene: 23.03–5.332 Ma)
Paleoenvironment: Freshwater.
Comments: Within the suborder Cyprinodontoidei, the superfamily Funduloidea combines Profundulidae, Goodeidae, and Fundulidae (Nelson 2006). The earliest record of Funduloidea is provided by the goodeid †*Tapatia occidentalis* Alvarez and Arriola-Longoria, 1972 from Santa Rosa, Jalisco, Mexico (Miocene), which is listed in Patterson (1993b), and by specimens of *Fundulus* from the California, Nevada, and Montana (Miocene) (Cavender 1998).

Group 2: Poecilioidea (+ Cyprinodontoidea)
Earliest record: †'Poeciliidae indet.' from the Maíz Gordo Formation, Argentina (Upper Thanetian, Riochican: 57.4–55.8 Ma), or †*Prolebias delphinensis* Gaudant, 1989 from Montbrun-les-Bains, Drôme, France (Oligocene: 33.9–23.03 Ma).
Paleoenvironment: Freshwater.
Comments: The superfamily Poecilioidea also belongs to Cyprinodontoidei and comprises the families Anablepidae and Poeciliidae (Nelson 2006). It is considered the sister group to Cyprinodontoidea (Costa 1998). Undescribed poeciliid fossils are apparently present in the Argentinian Lumbrera (Alano Perez *et al.* 2010) and Maíz Gordo Formations (Cione 1986). Both formations are part of the Santa Barbara Group. Without additional information about these fossils, their taxonomic assignment as poeciliids may be questioned, especially since Cione & Báez (2007) do not

include poeciliids in their list of paleocene fishes from Argentina. According to López-Fernández & Albert (2011), the only confirmed fossils of poeciliids are from the Middle-Late Miocene Río Salí and San José Formations, Argentina. The Río Salí Formation is older than the San José Formation with an age around 14 Ma (Hernández *et al.* 2005). The European record of Cyprinodontidae extends to the Oligocene, with †*Prolebias delphinensis* Gaudant, 1989 from Montbrun-les-Bains, Drôme, France (Cavender 1998), and thus provides a minimum age of the clade combining Poecilioidea and Cyprinodontoidea, if the Argentinian poeciliid records should have been misidentified. Therefore, a wide range is used to reflect the uncertainty concerning the age of the earliest record of this group (57.4-23.03 Ma).

Parameters: $t_\mathrm{y} = 23.03$, $t_\mathrm{o} = 57.4$, $r = 0.0122/0.0041/0.0012/0.0004$, $p−q = 0.0559/0.0627/0.0694/0.0726$, $lag = 2$, $corr = $ "global_terrestrial"

Node: 33

Group 1: Gobiesociformes, Blenniiformes, Atheriniformes, Mugiliformes (+ Pseudochromidae, Plesiopidae, Grammatidae, Opistognathidae, Notograptidae)
Earliest record: Otoliths of the "genus *Atherinidarum*" from Argile de Gan, Gan, France (Ypresian: 55.8-48.6 Ma), and from the Cambay Shale Formation, Gujarat, India (Ypresian: 55.8–48.6 Ma), or skeletal remains of †*Palaeoatherina rhodanica* Gaudant, 1976 from Mormoiron, Vaucluse, France (Priabonian: 37.2–33.9 Ma), †*Palaeoatherina vardinis* Sauvage 1883, Bassin dAlés, France (Priabonian: 37.2–33.9 Ma), and †*Palaeoatherina formosa* Chedhomme and Gaudant, 1984, Orgnac-L'Aven, France (Priabonian: 37.2–33.9 Ma).
Paleoenvironment: Marine/freshwater.
Comments: The french formations of Mormoiron, Bassin dAlés, and Orgnac-L'Aven have all been considered Priabonian by Patterson (1993b). Thus, †*Palaeoatherina* provides a minimum age of this clade of 37.2–33.9 Ma if the otoliths of "genus *Atherinidarum*" should be misidentified (Patterson 1993b). Most extant members of this group are marine, however, different atheriniform families have entered freshwater environments.

Group 2: "Labriformes" I
Earliest record: †*Palaeopomacentrus orphae* Bellwood and Sorbini, 1996 and †*Lorenzichthys olihan* Bellwood, 1999 from Monte Bolca, Verona, Italy (Ypresian, NP14, SB11: 49.4–49.1 Ma).
Paleoenvironment: Marine/freshwater.
Comments: See above. Embiotocidae and Pomacentridae are exclusively marine, whereas cichlids are strictly confined to freshwater habitats.

Parameters: $t_\mathrm{y} = 49.1$, $t_\mathrm{o} = 55.8$, $r = 0.0160/0.0053/0.0016/0.0005$, $p−q = 0.0559/0.0627/0.0694/0.0726$, $lag = 2$, $corr = $ "global"

Node: 34

Group 1: Gobiesociformes, Blenniiformes
Earliest record: †?*Oncolepis isseli* Bassani, 1898 from Monte Bolca, Verona, Italy (Ypresian, NP14, SB11: 49.4–49.1 Ma).
Paleoenvironment: Marine.
Comments: While the genus level assignment of ?*Oncolepis isseli* may be questionable, it is frequently cited as a blenniid (Bellwood 1996), and is here accepted as such. Apparently, no gobiesociform fossils are known.

Group 2: Atheriniformes, Mugiliformes (+ Pseudochromidae, Plesiopidae, Grammatidae, Opistognathidae, Notograptidae)
Earliest record: Otoliths of the "genus *Atherinidarum*" from Argile de Gan, Gan, France (Ypresian: 55.8–48.6 Ma), and from the Cambay Shale Formation, Gujarat, India (Ypresian: 55.8–48.6 Ma), or skeletal remains of †*Palaeoatherina rhodanica* Gaudant, 1976 from Mormoiron, Vaucluse, France (Priabonian: 37.2–33.9 Ma), †*Palaeoatherina vardinis* Sauvage, 1883, Bassin dAlés, France (Priabonian: 37.2–33.9 Ma), and †*Palaeoatherina formosa* Chedhomme and Gaudant, 1984, Orgnac-L'Aven, France (Priabonian: 37.2–33.9 Ma).
Paleoenvironment: Marine/freshwater.
Comments: See above

Parameters: $t_y = 49.1$, $t_o = 55.8$, $r = 0.0160/0.0053/0.0016/0.0005$, $p-q = 0.0559/0.0627/0.0694/0.0726$, $lag = 2$, $corr = $ "global"

Node: 35

Group 1: Gobiesociformes
Earliest record: Extant.
Paleoenvironment: Marine.
Comments: -

Group 2: Blenniiformes
Earliest record: †?*Oncolepis isseli* Bassani, 1898 from Monte Bolca, Verona, Italy (Ypresian, NP14, SB11: 49.4–49.1 Ma).
Paleoenvironment: Marine.
Comments: See above.

Parameters: $t_y = 49.1$, $t_o = 49.4$, $r = 0.0247/0.0082/0.0025/0.0008$, $p-q = 0.0559/0.0627/0.0694/0.0726$, $lag = 2$, $corr = $ "global_marine"

Node: 36

Group 1: *Coryphoblennius*
Earliest record: Extant.
Paleoenvironment: Marine.
Comments: *Coryphoblennius* and *Salaria* both belong to the tribe Salariini (Nelson 2006), that was found paraphyletic by inclusion of *Blennius*, on the basis of molecular data (Almada *et al.* 2005). However, the position of *Blennius* as a sister to a group combining *Coryphoblennius* and *Lipophrys* was only weakly supported. Given that it contrasts with morphological data, *Blennius* is here not considered more closely related to *Coryphoblennius* than to *Salaria*, and thus, †*Blennius fossilis* Kramberger, 1891 from Dolje, Croatia (Badenian–Sarmatian) Patterson (1993b) is not used to constrain this node. No salariin fossils are known.

Group 2: *Salaria*
Earliest record: Extant.
Paleoenvironment: Marine.
Comments: -

Parameters: $t_y = 0$, $t_o = 0$, $r = 0.0441/0.0147/0.0044/0.0015$, $p - q = 0.0559/0.0627/0.0694/0.0726$, $lag = 2$, $corr = $ "global_marine"

Node: 37

Group 1: Atheriniformes
Earliest record: Otoliths of the "genus *Atherinidarum*" from Argile de Gan, Gan, France (Ypresian: 55.8–48.6 Ma), and from the Cambay Shale Formation, Gujarat, India (Ypresian: 55.8–48.6 Ma), or skeletal remains of †*Palaeoatherina rhodanica* Gaudant, 1976 from Mormoiron, Vaucluse, France (Priabonian: 37.2–33.9 Ma), †*Palaeoatherina vardinis* Sauvage, 1883, Bassin dAlés, France (Priabonian: 37.2–33.9 Ma), and †*Palaeoatherina formosa* Chedhomme and Gaudant, 1984, Orgnac-L'Aven, France (Priabonian: 37.2–33.9 Ma).
Paleoenvironment: Marine/freshwater.
Comments: See above.

Group 2: Mugiliformes (+ Pseudochromidae, Plesiopidae, Grammatidae, Opistognathidae, Notograptidae)
Earliest record: Otolith fossils of †*Haliophis colletti* Nolf and Lapierre, 1979 and †*"Opistognathidarium" bloti* Nolf and Lapierre, 1979 from the Calcaire Grossier and Auvers Formations, Paris Basin, France (Lutetian: 48.6–40.4 Ma), otolith remains of †*"Mugilidarum" debilis* Koken, 1888 from Jackson River, Mississippi (Priabonian: 37.2–33.9 Ma), or †*Mugil princeps* Agassiz, 1843 from the Menilite Beds, Poland and Ukraine (Rupelian: 33.9–28.4 Ma).

Paleoenvironment: Marine.

Comments: Otolith fossils of †*Haliophis colletti* Nolf and Lapierre, 1979 and †*"Opistognathidarium"* *bloti* Nolf and Lapierre, 1979 from the Calcaire Grossier and Auvers Formations, Paris Basin, France, support the presence of Pseudochromidae and Opistognathidae in the Lutetian. The earliest mugiliform fossils are otolith remains of †*"Mugilidarum" debilis* Koken, 1888 from Jackson River, Mississippi (Priabonian), and skeletal remains of †*Mugil princeps* Agassiz, 1843 from the Menilite Beds, Poland and Ukraine (Rupelian) (Patterson 1993b).

Parameters: $t_y = 33.9$, $t_o = 55.8$, $r = 0.0160/0.0053/0.0016/0.0005$, $p-q = 0.0559/0.0627/0.0694/0.0726$, $lag = 2$, $corr = $ "global"

Node: 38

Group 1: *Mugil cephalus*
Earliest record: Extant.
Paleoenvironment: Marine.
Comments: A number of skeletal fossils have been assigned to the genus *Mugil*, including †*M. princeps* Agassiz, 1843 from the Menilite Beds, Poland and Ukraine (Rupelian) (Patterson 1993b), †*M. minax* Bogatshov, 1933 from Pshekha River, Tsurevsky, Russia (Sarmatian), and †*M. acer* Switchenska, 1959 from Karpov Yar, Naslavcea, Moldowa (Sarmatian) (Bannikov 2009). However, their affinities with extant mugils are not clear, so these fossils are not used to constrain the split between *M. cephalus* and *M. curema*.

Group 2: *Mugil curema*
Earliest record: Extant.
Paleoenvironment: Marine.
Comments:

Parameters: $t_y = 0$, $t_o = 0$, $r = 0.0441/0.0147/0.0044/0.0015$, $p - q = 0.0559/0.0627/0.0694/0.0726$, $lag = 2$, $corr = $ "global_marine"

Node: 39

Group 1: Cichlinae *sensu* Sparks & Smith 2004
Earliest record: †*Gymnogeophagus eocenicus* Malabarba, Malabarba, and del Papa, 2010, †*Plesioheros chauliodus* Alano Perez, Malabarba, and del Papa, 2010, and †*Proterocara argentina* Malabarba, Zuleta, and del Papa, 2006 from the Lumbrera Formation, Argentina (Casamayoran: 45.4–39.9 Ma).
Paleoenvironment: Freshwater.

Comments: Cichlid fossils †*Gymnogeophagus eocenicus* Malabarba, Malabarba, and del Papa, 2010, †*Plesioheros chauliodus* Alano Perez, Malabarba, and del Papa, 2010, and †*Proterocara argentina* Malabarba, Zuleta, and del Papa, 2006 from the Lumbrera Formation have been claimed to be 48.6 Ma, from the Ypresian–Lutetian boundary (Alano Perez *et al.* 2010), however it is unclear how the date was obtained. The Lumbrera Formation has traditionally been assigned to the Casamayoran South American Land Mammal Age (SALMA), based on the evolutionary grade of fossil mammals (del Papa *et al.* 2010). However, it has been argued that the Lumbrera formation could be as young as Mustersan (Deraco *et al.* 2008). The upper Lumbrera Formation has been dated directly using U/Pb zircon determination, resulting in an estimate of 39.9 Ma, that serves as a minimum age estimate, given that the fossiliferous "Faja Verde" layers lie below the dated tuff layer (del Papa *et al.* 2010). Here, a Casamayoran age of the Lumbrera formation is accepted. A recent absolute dating of SALMA stages (Vucetich *et al.* 2007) indicates a Mid Lutetian to Bartonian age of the Casamayoran, defined by polarities C20–C18. This is equivalent to 45.4–38.0 Ma (Ogg & Ogg 2008c). Given the absolute U/Pb zircon determination, this range can be reduced to 45.4–39.9 Ma.

Group 2: Pseudocrenilabrinae
Earliest record: †*Mahengechromis* spp. Murray, 2000 from Mahenge, Singida, Tanzania (Lutetian: 46.3–45.7 Ma).
Paleoenvironment: Freshwater.
Comments: Pseudocrenilabrinae include all African cichlds, which were shown to be monophyletic, based on molecular data (Farias *et al.* 2000). †*Mahengechromis* spp. share two synapomorphies with all but the most basal members of Pseudocrenilabrinae, and can thus be assigned to this clade (Murray 2001b). Zircon dating has established a Lutetian age of the Mahenge site (46.3–45.7 Ma) (Murray 2001b).

Parameters: $t_y = 45.7$, $t_o = 46.3$, $r = 0.0070/0.0023/0.0007/0.0002$, $p−q = 0.0559/0.0627/0.0694/0.0726$, $lag = 2$, $corr = $ "global_terrestrial"

Node: 40

Group 1: Cichlasomatini
Earliest record: †*Tremembichthys garciae* Malabarba and Malabarba, 2008 from the Entre-Córregos Formation, Aiuruoca Tertiary Basin, Brazil, Eocene-Oligocene (35.0–30.0 Ma).
Paleoenvironment: Freshwater.
Comments: Cichlasomatini represent the sister clade to Heroini, which is supported by morphological and molecular data (Farias *et al.* 2000; López-Fernández *et al.* 2010). †*Tremembichthys garciae* has been assigned to the cichlid tribe Cichlasomatini with a phylogenetic analysis using 91 morphological characters (Malabarba & Malabarba 2008). The Entre-Córregos Formation has been dated to 35.0–30.0 Ma (Garcia *et al.* 2000).

Group 2: Heroini
Earliest record: †*Plesioheros chauliodus* Alano Perez, Malabarba, and del Papa, 2010 from the Lumbrera Formation, Argentina (Casamayoran: 45.4–39.9 Ma).
Paleoenvironment: Freshwater.
Comments: †*Plesioheros chauliodus* has been placed among Heroini with a phylogenetic analysis (Alano Perez *et al.* 2010). The age of the Lumbrera Formation is discussed above.

Parameters: $t_{\mathrm{y}} = 39.9$, $t_{\mathrm{o}} = 45.4$, $r = 0.0070/0.0023/0.0007/0.0002$, $p-q = 0.0559/0.0627/0.0694/0.0726$, $lag = 2$, $corr =$ "global_terrestrial"

Node: 41

Group 1: Hemichromini
Earliest record: †*Mahengechromis* spp. Murray, 2000 from Mahenge, Singida, Tanzania (Lutetian: 46.3–45.7 Ma).
Paleoenvironment: Freshwater.
Comments: Following Murray (2001a), †*Mahengechromis* spp. is here accepted as the sister group of Hemichromini.

Group 2: Haplotilapiini
Earliest record: †*Sarotherodon martyini* van Couvering, 1982 (= †*Oreochromis martyini*) from the Ngorora Formation, Lake Turkana, Kenya (Late Miocene: 12.0–9.3 Ma).
Paleoenvironment: Freshwater.
Comments: Haplotilapiini *sensu* Schwarzer *et al.*, 2009 comprise Etiini, Oreochromini, Boreotilapiini, and Austrotilapiini. The monophyly of this clade is strongly supported by molecular data (Schwarzer *et al.* 2009). The earliest fossils of Haplotilapiini are provided by †*Sarotherodon martyini* Van Couvering, 1982 from the Ngorora Formation, Lake Turkana, Kenya (Murray & Stewart 1999; Murray 2001a), that has been dated at 12.0–9.3 Ma (Murray & Stewart 1999).

Parameters: $t_{\mathrm{y}} = 45.7$, $t_{\mathrm{o}} = 46.3$, $r = 0.0070/0.0023/0.0007/0.0002$, $p-q = 0.0559/0.0627/0.0694/0.0726$, $lag = 2$, $corr =$ "global_terrestrial"

Node: 42

Group 1: Oreochromini
Earliest record: †*Sarotherodon martyini* van Couvering, 1982 (= †*Oreochromis martyini*) from the Ngorora Formation, Lake Turkana, Kenya (Late Miocene: 12.0–9.3 Ma).
Paleoenvironment: Freshwater.

Comments: †*Oreochromis lorenzoi* Carnevale, Sorbini, and Landini, 2003 from the Gessoso-Solfifera Formation, Italy (Messinian) is frequently considered the oldest member of the genus *Oreochromis*, however there is no reason to ignore †*Sarotherodon martyini*, that has been reassigned to the genus *Oreochromis*, and is older than †*Oreochromis lorenzoi* (Murray & Stewart 1999).

Group 2: Boreotilapiini, Austrotilapiini
Earliest record: Extant.
Paleoenvironment: Freshwater.
Comments: All cichilds of the East African Radiations, including Haplochromini, are grouped in Austrotilapiini (Schwarzer *et al.* 2009). Murray (2001a) lists two Oligocene fossils of 'fluviatile haplochromines': ?*Astatotilapia* from the Ad Darb Formation, Saudi Arabia (Lippitsch & Micklich 1998), and †*Macfadyena dabanensis* Van Couvering, 1982 from the Daban Formation, Somalia. Of these, the assignment of ?*Astatotilapia* cannot be upheld (Lippitsch & Micklich 1998), but it is "very probable" that the fossil is a haplochromine. †*Macfadyena dabanensis* has apparently been assigned only to the Pseudocrenilabrinae. The biogeography of East African cichlids supports a Tanganyikan origin of haplochromines (Salzburger *et al.* 2005), so the affiniity of ?*Astatotilapia* to Saudi Arabian haplochromines is questioned, and this fossil is not used to constrain the age of this group. No other fossils of Boreotilapiini or Austrotilapiini are known (Murray 2001a).

Parameters: $t_y = 9.3$, $t_o = 12.0$, $r = 0.0320/0.0107/0.0032/0.0011$, $p-q = 0.0559/0.0627/0.0694/0.0726$, $lag = 2$, $corr = $ "global_terrestrial"

Node: 43

Group 1: Lamprologini
Earliest record: Extant.
Paleoenvironment: Freshwater.
Comments: -

Group 2: Haplochromini (and other lineages)
Earliest record: Extant.
Paleoenvironment: Freshwater.
Comments: Based on molecular data, the 'modern haplochromines' *sensu* Salzburger *et al.*, 2005 include Tropheini as well as the Lake Malawi and Lake Victoria species flock, and are more closely related to Ectodini, Perissodini, and Limnochromini (and others) than to Lamprologini (Salzburger *et al.* 2005). See above for reasons why two putative 'haplochromine' fossils are not used here.

Parameters: $t_y = 0$, $t_o = 0$, $r = 0.0320/0.0107/0.0032/0.0011$, $p-q = 0.0559/0.0627/0.0694/0.0726$, $lag = 2$, $corr = $ "global_terrestrial"

Node: 44

Group 1: "Labriformes" II
Earliest record: †*Eocoris bloti* Bannikov and Sorbini, 1991, †*Phyllopharyngodon longipinnis* Bell-wood, 1990, and †*Bellwoodilabrus landinii* Bannikov and Carnevale, 2010 from Monte Bolca, Verona, Italy (Ypresian, NP14, SB11: 49.4–49.1 Ma).
Paleoenvironment: Marine.
Comments: "Labriformes" II is equivalent to Labridae, if the family is considered to include Scaridae and Odacidae (Patterson 1993b; Hanel *et al.* 2002; Cowman *et al.* 2009). Monophyly of "Labriformes" II, as well as its sister group relationship to a clade combining "Perciformes" *sensu* stricto, "Caproiformes" I, Lophiiformes, Tetraodontiformes is supported by molecular data (Mabuchi *et al.* 2007). The oldest fossils of this group are from Monte Bolca, Italy (Ypresian), including †*Eocoris bloti* Bannikov and Sorbini, 1991, †*Phyllopharyngodon longipinnis* Bellwood, 1990, and †*Bellwoodilabrus landinii* Bannikov and Carnevale, 2010 (Patterson 1993b; Cowman *et al.* 2009; Cowman & Bellwood 2011).

Group 2: "Perciformes" *sensu* stricto, "Caproiformes" I, Lophiiformes, Tetraodontiformes, Epinephelidae *sensu* Smith and Craig, 2007, "Trachiniformes", Serranidae *sensu* Smith and Craig, 2007, Percidae, Scorpaenoidei, Cottiformes, Gasterosteioidei, Nototheniiformes
Earliest record: †*Plectocretacicus clarae* Sorbini, 1979 from the Lithografic Limestone of Hakel, Lebanon (Early Cenomanian: 99.1–97.8 Ma).
Paleoenvironment: Marine.
Comments: See above.

Parameters: $t_y = 97.8$, $t_o = 99.1$, $r = 0.0309/0.0103/0.0031/0.0010$, $p-q = 0.0559/0.0627/0.0694/0.0726$, $lag = 2$, $corr = $ "global_marine"

Node: 45

Group 1: Julidines (+ Pseudolabrines, Labrichthyines, Novaculines, Pseudocheilines)
Earliest record: †*Julis sigismundi* Kner, 1862 from the Bay of Eisenstadt, Austria (Badenian: 13.7–12.7 Ma).
Paleoenvironment: Marine.
Comments: A clade combining Julidines, Pseudolabrines, Labrichthyines, Novaculines, and Pseudocheilines is supported by molecular data (Cowman *et al.* 2009). The earliest record of this clade may be provided by †*Julis sigismundi* Kner, 1862 from the Bay of Eisenstadt, Austria (Bellwood & Schultz 1991). While *Julis* is no longer a valid genus, †*Julis sigismundi* is here assumed to be a julidine, and therefore part of this group. The locality is Badenian in age (Bellwood & Schultz 1991), which lasted from 13.7 to 12.7 Ma (Harzhauser & Piller 2004).

Group 2: Labrines, Scarines (+ Cheilines)
Earliest record: †*Labrus agassizi* Münster, 1846, †*Labrus parvulus* Heckel, 1856, and †*Calotomus preisli* Bellwood and Schultz, 1991 from the Bay of Eisenstadt, Austria (Badenian: 13.7–12.7 Ma).
Paleoenvironment: Marine.
Comments: The molecular phylogeny of Cowman *et al.* (2009) strongly supports a clade combining Labrines, Scarines, and Cheilines, which is the sister group to Julidines, Pseudolabrines, Labrichthyines, Novaculines, and Pseudocheilines. †*Labrus agassizi*, and †*Labrus parvulus* likely are the oldest members of Scarines, and †*Calotomus preisli* has been considered the oldest scarine (Westneat & Alfaro 2005).

Parameters: $t_\mathrm{y} = 12.7$, $t_\mathrm{o} = 13.7$, $r = 0.0441/0.0147/0.0044/0.0015$, $p-q = 0.0559/0.0627/0.0694/0.0726$, $lag = 2$, $corr =$ "global_marine"

Node: 46

Group 1: *Coris julis*
Earliest record: Extant.
Paleoenvironment: Marine.
Comments: Both *Coris* and *Thalassoma* may be polyphyletic (Cowman *et al.* 2009). No fossils are known of either genus.

Group 2: *Thalassoma pavo*
Earliest record: Extant.
Paleoenvironment: Marine.
Comments: See above.

Parameters: $t_\mathrm{y} = 0$, $t_\mathrm{o} = 0$, $r = 0.0441/0.0147/0.0044/0.0015$, $p-q = 0.0559/0.0627/0.0694/0.0726$, $lag = 2$, $corr =$ "global_marine"

Node: 47

Group 1: Scarines (+ Cheilines)
Earliest record: †*Calotomus preisli* Bellwood and Schultz, 1991 from the Bay of Eisenstadt, Austria (Badenian: 13.7–12.7 Ma).
Paleoenvironment: Marine.
Comments: Scarines were found to group with Cheilines in the molecular phylogeny of Cowman *et al.* (2009). †*Calotomus preisli* has been considered the oldest scarine (Westneat & Alfaro 2005).

Group 2: Labrines

Earliest record: †*Labrus agassizi* Münster, 1846, and †*Labrus parvulus* Heckel, 1856 from the Bay of Eisenstadt, Austria (Badenian: 13.7–12.7 Ma).
Paleoenvironment: Marine.
Comments: See above.

Parameters: $t_y = 12.7$, $t_o = 13.7$, $r = 0.0441/0.0147/0.0044/0.0015$, $p-q = 0.0559/0.0627/0.0694/0.0726$, $lag = 2$, $corr = $ "global_marine"

Node: 48

Group 1: *Labrus*
Earliest record: †*Labrus agassizi* Münster, 1846, and †*Labrus parvulus* Heckel, 1856 from the Bay of Eisenstadt, Austria (Badenian: 13.7–12.7 Ma).
Paleoenvironment: Marine.
Comments: Scarines were found to group with Cheilines in the molecular phylogeny of Cowman *et al.* (2009). †*Calotomus preisli* has been considered the oldest scarine (Westneat & Alfaro 2005).

Group 2: *Ctenolabrus*, *Tautogolabrus* (+ *Symphodus*, *Centrolabrus*, *Tautoga*)
Earliest record: †*Crenilabrus woodwardi* Kramberger, 1891 from Dolje, Croatia (Badenian–Sarmatian: 13.7–11.6 Ma).
Paleoenvironment: Marine.
Comments: *Ctenolabrus* and *Tautogolabrus* were found to cluster with *Symphodus*, *Centrolabrus*, and *Tautoga* in Cowman *et al.* (2009). This group is represented in the fossil record by †*Crenilabrus woodwardi* from Dolje, Croatia (Badenian–Sarmatian: 13.7–11.6 Ma). *Crenilabrus* is synonymous with *Symphodus* (Hanel *et al.* 2002).

Parameters: $t_y = 12.7$, $t_o = 13.7$, $r = 0.0441/0.0147/0.0044/0.0015$, $p-q = 0.0559/0.0627/0.0694/0.0726$, $lag = 2$, $corr = $ "global_marine"

Node: 49

Group 1: *Ctenolabrus*
Earliest record: Extant.
Paleoenvironment: Marine.
Comments: *Ctenolabrus* and *Tautogolabrus* were recovered as sister groups in Cowman *et al.* (2009). No fossil record of either genus is known.

Group 2: *Tautogolabrus*
Earliest record: Extant.

Paleoenvironment: Marine.
Comments: See above.

Parameters: $t_y = 0$, $t_o = 0$, $r = 0.0441/0.0147/0.0044/0.0015$, $p - q = 0.0559/0.0627/0.0694/$
0.0726, $lag = 2$, $corr = $ "global_marine"

Node: 50

Group 1: "Perciformes" *sensu* stricto (+ Acanthuriformes), "Caproiformes" I, Lophiiformes, Tetra-odontiformes
Earliest record: †*Plectocretacicus clarae* Sorbini, 1979 from the Lithografic Limestone of Hakel, Lebanon (Early Cenomanian: 99.1–97.8 Ma).
Paleoenvironment: Marine.
Comments: The clade combining "Perciformes" *sensu* stricto, "Caproiformes" I, Lophiiformes, and Tetraodontiformes has previously been supported by molecular data (Mabuchi *et al.* 2007; Yamanoue *et al.* 2008; Kawahara *et al.* 2008; Matschiner *et al.* 2011). Acanthuriformes seem to be nested within "Perciformes" *sensu* stricto (Holcroft & Wiley 2008).

Group 2: Epinephelidae *sensu* Smith and Craig, 2007, "Trachiniformes", Serranidae *sensu* Smith and Craig, 2007, Percidae, Scorpaenoidei, Cottiformes, Gasterosteioidei, Nototheniiformes
Earliest record: †*Plesioserranus wemmeliensis* from the London Clay Formation of Kent, United Kingdom (Ypresian, MP8–9: 55.2–50.8 Ma).
Paleoenvironment: Marine.
Comments: The molecular phylogenies of Dettaï & Lecointre (2005) and Smith & Craig (2007) indicate support for this clade. Epinephelidae *sensu* Smith and Craig, 2007 comprise all epinpheline genera of Nelson (2006) except *Niphon*. The earliest record of Epinephelidae is †*Epinephelus casottii* Costa, 1858 from Retznei, Steiermark, Austria (Langhian: 15.97-13.65 Ma) (Schultz 2000).

"Trachiniformes" are here considered to combine the families Chiasmodontidae, Champsodontidae, Trichodontidae, Pinguipedidae, Cheimarrhichthyidae, Trichonotidae, Creediidae, Percophidae, Leptoscopidae, Ammodytidae, Trachinidae, and Uranoscopidae (Nelson 2006). Of these, the earliest skeletal fossils are provided by the trachinids †*Callipteryx speciosus* Agassiz, 1838 and †*Callipteryx recticaudatus* Agassiz, 1838 from Monte Bolca, Italy (Ypresian), followed by the chiasmodontid †*Pseudoscopelus grossheimi* Danil'chenko, 1960 from the lower Khadum Formation, Russia (Rupelian) (Patterson 1993a,b; Bellwood 1996). The early "trachiniform" fossil record is complemented with otolith remains of the percophid †*"Hemerocoetinarum" apertus* Schwarzhans, 1980 and the uranoscopid †*Uranoscopus ignavus* Schwarzhans 1980 from Waihao River, Canterbury, New Zealand (Lutetian) (Patterson 1993b), and of †*Ammodytes vasseuri* Nolf and Lapierre, 1977 from the Sables des Bois-Gouet, Loire-Atlantique, France (Lutetian) (Patterson 1993b). "Trachiniformes" are para- or polyphyletic by the inclusion of Serranidae *sensu* Smith and Craig, 2007. Note that Trachinidae were reduced to subfamily Trachininae and included in Smith & Craigs (2007) Serranidae. Whether

Trachinidae should be included in Serranidae or "Trachiniformes" is regarded irrelevant here, given that Serranidae appear nested in "Trachiniformes".

A large number of fossils has been assigned to Serranidae, or even the genus *Serranus* (Schultz 2000). The Cretaceous †*Eoserranus hislopi* Woodward, 1908 from the Lameta Formation, India (Maastrichtian) is considered the oldest serranid by Pondella II *et al.* (2003), but is classified as 'Percoidei *incertae sedis*' by Arratia *et al.* (2004). Similarly, †*Tretoperca vestita* Sytchevskaya, 1986 from a drill hole near Boltyshka village, Ukraine (Schultz 2000) (Late Paleocene–Early Eocene) (Skutschas & Gubin 2011), as well as other specimens assigned to serranids, such as †*Bilinia* Obrhelova, 1971, †*Blabe* White, 1936, and †*Kiinkerishia* Sytchevskaya, 1986 should be considered Percoidei *incertae sedis* (Prokofiev 2009) rather than serranids. Furthermore, Bannikov (2006) concludes that none of the fossils described as Serranidae, such as †*Serranus rugosus* Heckel, 1953 from Monte Bolca, belong to this family. This is corroborated by the comparative analysis of Schultz (2000), who found that the taxonomic assignment of nearly all fossil "Serraninae" is questionable. However, he reports that the denticulation of the edge of the preoperculum of the serranine †*Plesioserranus wemmeliensis* Storms 1897 resembles that of *Serranus*, which is here taken as evidence for their relatedness, and for the correct assignment of †*Plesioserranus wemmeliensis* in Serraninae (Schultz 2000). The earliest record of †*Plesioserranus wemmeliensis* is from the London Clay Formation of Kent, United Kingdom (Casier 1966), which is Ypresian and MP8–9 in age (Mlíkovsky 1996). The 'European Land Mammal Age' MP8–9 has been dated at 55.2–50.8 Ma (Ogg & Ogg 2008c).

Percidae appear comparatively late in the fossil record, with the earliest skeletal records provided by †*Perca fluviatilis* from La Montagne d'Andance, France, and from Murzak-Koba, Ukraine (Upper Miocene) (Lebedev 1952; Mein *et al.* 1983; Carney & Dick 2000). Otolith remains of Percidae belonging to †*Perca hassiaca* Weiler, 1961 and †*Perca* sp. 1 have been described in the Upper Rhine Graben and the Mainz Basin, Germany (Rupelian) (Martini & Reichenbacher 2007).

While Scorpaenoidei may be polyphyletic (Smith & Wheeler 2004), a clade combining the scorpaenoid families Scorpaenidae (with subfamilies Sebastinae, Setarchinae, Neosebastinae, Scorpaeninae, and Plectrogeninae, but excluding Apistinae, Tetraroginae and Synanceinae), Caracanthidae, Congiopodidae, and Bembridae is supported by their molecular data (their node 6, Fig. 3). The Oligocene †*Scorpaenoides popovicii* Priem, 1899 from Valea Caselor, Romania has been placed in Scorpaena by Woodward (1901), but is not a member of Scorpaeninae according to Schultz (1993). Due to its uncertain position, it can not be used to constrain the age of Smith & Wheeler's (2004) node 6. The earliest record of this clade may instead be represented by specimens of Scorpaena from the Mainz Basin and the Embayment of Leipzig, Germany, described by Weiler (1928; also see Schultz 1993). While the generic placement of these specimens is questioned by Schultz (1993), they are here considered to be representatives of the subfamily Scorpaeninae, and thus constrain the age of Smith & Wheeler's (2004) node 6. The age of the localities in the Mainz Basin and the Embayment of Leipzig has been given as Mid-Oligocene by Schultz (1993), and is reported to be Rupelian by Standke *et al.* (2005).

Cottoidei and Zoarcidei are often found closely related in molecular phylogenies, but also in association with Gasterosteoidei (Mabuchi *et al.* 2007; Kawahara *et al.* 2008; Matschiner *et al.* 2011). However, the relationships among these three suborders remain unclear in these studies.

Morphological evidence, on the other hand, supports a sister group relationship of Cottoidei and Zoarcidei (Wiley & Johnson 2010), and therefore, monophyly of Cottiformes, combining these two suborders, is here assumed. The earliest cottoid record is provided by †*Cottus cervicornis* Storms, 1984 from Argile de Boum, Belgium (Rupelian), whereas the zoarcoid fossil record is limited to Pliocene otolith remains of *Lycodes pacificus* Collett, 1879 (Patterson 1993b).

The polyphyly of Gasterosteiformes is discussed above. The earliest gasterosteiform record, †*Gasterorhamphosus zuppichinii* is closer to Syngnathidoidei than to Gasterosteoidei, and cannot be used to constrain gasterosteid divergences (Natale 2008). In the absence of aulorhynchid and hypoptychid fossils, the earliest record of Gasterosteoidei is provided by skeletal fossils of the extant *Gasterosteus aculeatus* Linné, 1758 from the Monterey Formation, California (Miocene) (Bell *et al.* 2009).

The fossil history of Nototheniiformes is limited to the single putative record of †*Proeleginops grandeastmanorum* from the La Meseta Formation, Seymour Island, Antarctic Peninsula. The fossil has originally been described as a gadiform (Eastman & Grande 1991), but was claimed to be a notothenioid (of the eleginopid familiy) by Balushkin (1994). The fossil has been used in a molecular dating study (Near 2004), however, its taxonomic assignment remains questionable. The type locality is reported as RV-8200, which is reported to be ˜40.0 myr old (Eastman & Grande 1991). However, according to Long (1992), RV-8200 corresponds to the lower section of 'Tertiary Eocene La Meseta' (Telm) 4, the age of which has recently been reevaluated and is now estimated between 52.5–51.0 Ma (Ivany *et al.* 2008). This age is older than the molecular date estimate for the split between Eleginopidae and the notothenioid families of the Antarctic Clade (42.9 Ma) (Matschiner *et al.* 2011). In their molecular analysis aimed at dating notothenioid divergences, Matschiner *et al.* (2011)) deliberately excluded †*Proeleginops grandeastmanorum* as a constraint due to its debated taxonomic assignment. The presumed fit of their results with the age of †*Proeleginops grandeastmanorum* (there assumed to be

textasciitilde 40.0 Ma) suggested correctness of the notothenioid interpretation, however, this does not hold if †*Proeleginops grandeastmanorum* is in fact 52.5–51.0 myr old. Thus, Nototheniiformes are here considered to have no fossil record.

In summary, the earliest record of the clade combining Epinephelidae *sensu* Smith and Craig, 2007, "Trachiniformes", Serranidae *sensu* Smith and Craig, 2007, Percidae, Scorpaenoidei, Cottiformes, Gasterosteioidei, and Nototheniiformes is the serranine †*Plesioserranus wemmeliensis* from the London Clay Formation of Kent, United Kingdom (Ypresian, MP8–9).

Parameters: $t_y = 97.8$, $t_o = 99.1$, $r = 0.0309/0.0103/0.0031/0.0010$, $p-q = 0.0559/0.0627/0.0694/0.0726$, $lag = 2$, $corr =$ "global_marine"

Node: 51

Group 1: Sparidae (+ Nemipteridae, Letherinidae, Centracanthidae, Lutjanidae), Emmelichthyidae, Haemulidae (+ Pomacanthidae)

Earliest record: †"genus *Haemulidarum*" *gullentopsi* Nolf, 1978, †"genus *Haemulidarum*" *maka-renkoi* Schwarzhans and Bratishko, 2011, and †"genus *Sparidarum*" *spatiatus* Schwarzhans and Bratishko, 2011 from the Tashlik Formation, Luzanivka, Cherkassy Region, Ukraine (Selandian, NP5–6: 60.5–58.3 Ma), †*Isacia remensis* Leriche, 1908 from from Sables de Chalons-sur-Vesle, Marne, France (Thanetian), and the Oldhaven Beds, United Kingdom (Upper Thanetian, NP9, C25n: 57.2–56.6 Ma), †*Sparus* sp. Arambourg 1952 from the Phosphates of Ouled Abdoun and Ganntour, Morocco (Thanetian: 58.7–55.8 Ma), or †*Sciaenurus bowerbanki* Agassiz, 1845 from the London Clay Formation, Bognor Regis, Sussex, United Kingdom (Ypresian, MP8–9: 55.2–50.8 Ma).

Paleoenvironment: Marine.

Comments: As defined by Wiley & Johnson (2010), "Perciformes" *sensu* stricto is restricted 'to the former Percoidei *sensu* Johnson (1984), except for members of that group that show affinities elsewhere'. Thus, the group is here considered to include the following families: Centropomidae, Ambassidae, Latidae, Moronidae, Percichthyidae, Perciliidae, Acropomatidae, Symphysanodontidae, Polyprionidae, Centrogeniidae, Ostracoberycidae, Callanthiidae, Dinopercidae, Banjosidae, Centrarchidae, Priacanthidae, Apogonidae, Epigonidae, Sillaginidae, Malacanthidae, Latilinae, Lactariidae, Dinolestidae, Scombropidae, Pomatomidae, Menidae, Leiognathidae, Bramidae, Caristiidae, Emmelichthyidae, Lutjanidae, Caesionidae, Lobotidae, Gerreidae, Haemulidae, Inermiidae, Nemipteridae, Letherinidae, Sparidae, Centracanthidae, Polynemidae, Sciaenidae, Mullidae, Pempheridae, Glaucosomatidae, Leptobramidae, Bathyclupeidae, Monodactylidae, Toxotidae, Arripidae, Dichistiidae, Kyphosidae, Drepaneidae, Chaetodontidae, Pomacanthidae, Enoplosidae, Pentacerotidae, Nandidae, Polycentridae, Terapontidae, Kuhliidae, Oplegnathidae, Cirrhitidae, Chironemidae, Aplodactylidae, Cheilodactylidae, Latridae, Cepolidae.

According to the molecular phylogeny of Holcroft & Wiley (2008), "Perciformes" *sensu* stricto may be paraphyletic, and include Acanthuriformes, "Caproiformes" I, Lophiiformes, and Tetraodontiformes. In the dataset presented here, "Perciformes" *sensu* stricto are represented by the sparid *Pagrus auriga*, the emmelichthyid *Erythrocles monodi*, and the haemulid *Plectorhinchus mediterraneus*. Sparidae, Emmelichthyidae, and Haemulidae, together with Lutjanidae and Pomacanthidae, form a well-supported monophyletic clade. In addition, Nemipteridae, Letherinidae, Centracanthidae, are considered to group with Sparidae (Nelson 2006). Whether the remaining families are closer to "Caproiformes" I, Lophiiformes, and Tetraodontiformes remains uncertain.

A number of fossil specimens from the Ypresian deposits of Monte Bolca have been described as haemulids or sparids (Patterson 1993a,b), however, the revision of Bannikov (2006) suggest that none of them was assigned to the correct family. He concludes that haemulids are not present in the Monte Bolca deposits, and that sparid assignments of Monte Bolca specimens can only be considered putative. Instead, at least four species of lutjanids are represented in the Bolca fauna (Bannikov 2006), that belong to the genera †*Ottaviania* (Agassiz 1839) Bannikov, 2006, †*Veranichthys* Bannikov, 2006, and †*Goujetia* Bannikov, 2006. The sparid †*Sciaenurus bowerbanki* Agassiz, 1845 is known from the London Clay Formation, Bognor Regis, Sussex, United Kingdom (Ypresian, MP8–9) (Mlíkovsky 1996).Some more questionable evidence exists for sparids and haemulids from the Paleocene: Isolated teeth from the Phosphates of Ouled Abdoun and Ganntour, Morocco (Thanetian), were assigned to †*Sparus* sp. Arambourg 1952, otolith remains from Sables de Chalons-sur-Vesle,

Marne, France (Thanetian), and the Oldhaven Beds, United Kingdom (Thanetian), were described as the haemulid †*Isacia remensis* Leriche, 1908. Otolith fossils of †"genus *Haemulidarum*" *gullentopsi* Nolf 1978 are known from the Landen Formation, Wansin, Belgium (Thanetian) (Patterson 1993b), and from the Tashlik Formation, Luzanivka, Cherkassy Region, Ukraine (Selandian, NP5–6) (Schwarzhans & Bratishko 2011). From the same locality, otoliths of †"genus *Haemulidarum*" *makarenkoi* Schwarzhans and Bratishko, 2011 and of †"genus *Sparidarum*" *spatiatus* Schwarzhans and Bratishko, 2011 were newly described (Schwarzhans & Bratishko 2011).

Thus, the oldest known representative of the group combining Sparidae, Emmelichthyidae, Haemulidae, Lutjanidae Pomacanthidae, Nemipteridae, Letherinidae, and Centracanthidae is Selandian–Ypresian in age (60.5–48.6 Ma), which is further constrained by the age of the London Clay Formation of Bognor Regis (Ypresian, MP8–9: 55.2–50.8 Ma).

Group 2: "Caproiformes" I, Lophiiformes, Tetraodontiformes
Earliest record: †*Plectocretacicus clarae* Sorbini, 1979 from the Lithografic Limestone of Hakel, Lebanon (Early Cenomanian: 99.1–97.8 Ma).
Paleoenvironment: Marine.
Comments: The clade combining "Caproiformes" I, Lophiiformes, and Tetraodontiformes is strongly supported by molecular data (Mabuchi *et al.* 2007; Yamanoue *et al.* 2008; Kawahara *et al.* 2008; Matschiner *et al.* 2011).

Parameters: $t_y = 97.8$, $t_o = 99.1$, $r = 0.0309/0.0103/0.0031/0.0010$, $p-q = 0.0559/0.0627/0.0694/0.0726$, $lag = 2$, $corr =$ "global_marine"

Node: 52

Group 1: Sparidae (+ Nemipteridae, Letherinidae, Centracanthidae, Lutjanidae)
Earliest record: †"genus *Sparidarum*" *spatiatus* Schwarzhans and Bratishko, 2011 from the Tashlik Formation, Luzanivka, Cherkassy Region, Ukraine (Selandian, NP5–6: 60.5–58.3 Ma), †*Sparus* sp. Arambourg, 1952 from the Phosphates of Ouled Abdoun and Ganntour, Morocco (Thanetian: 58.7–55.8 Ma), or †*Sciaenurus bowerbanki* Agassiz, 1845 from the London Clay Formation, Bognor Regis, Sussex, United Kingdom (Ypresian, MP8–9: 55.2–50.8 Ma).
Paleoenvironment: Marine.
Comments: See above for a discussion of the phylogenetic relationships within "Perciformes" *sensu* stricto, and for the sparid fossil record.

Group 2: Emmelichthyidae, Haemulidae (+ Pomacanthidae)
Earliest record: †"genus *Haemulidarum*" *gullentopsi* Nolf, 1978, and †"genus *Haemulidarum*" *makarenkoi* Schwarzhans and Bratishko, 2011 from the Tashlik Formation, Luzanivka, Cherkassy Region, Ukraine (Selandian, NP5–6: 60.5–58.3 Ma), or †*Parapristopoma prohumile* Arambourg, 1927 from the Oran-Ravin blanc Formation, Algeria (Messinian: 7.246–5.332 Ma).

Paleoenvironment: Marine.

Comments: Pomacanthidae form a well-supported sister group of Haemulidae, according to the molecular phylogeny of Holcroft & Wiley (2008). Disregarding presumed haemulids from Monte Bolca (Bannikov 2006), it seems that no skeletal remains of Haemulidae older than the Miocene are known. †*Parapristopoma prohumile* Arambourg 1927 has been found in the Oran-Ravin blanc Formation, Algeria (Messinian) (Gaudant 2008). Also, apparently no skeletal fossils of Emmelichthyidae or Pomacanthidae are known (Patterson 1993b). Thus, the earliest known fossils of this group can only be constrained to 61.7–5.332 Ma.

Parameters: $t_y = 50.8$, $t_o = 60.5$, $r = 0.0247/0.0082/0.0025/0.0008$, $p-q = 0.0559/0.0627/0.0694/0.0726$, $lag = 2$, $corr = $ "global_marine"

Node: 53

Group 1: Emmelichthyidae

Earliest record: †*Emmelichthys* sp. Nolf and Lapierre, 1977 from the Sables du Bois-Gouet, Loire-Atlantique, France (Lutetian: 48.6–40.4 Ma).

Paleoenvironment: Marine.

Comments: Apparently, only otolith fossils of emmelichthyids are known.Patterson (1993b) lists the otolith record †*Emmelichthys* sp. Nolf and Lapierre, 1977 from the Sables du Bois-Gouet, Loire-Atlantique, France (Lutetian).

Group 2: Haemulidae (+ Pomacanthidae)

Earliest record: †"genus *Haemulidarum*" *gullentopsi* Nolf, 1978, and †"genus *Haemulidarum*" *makarenkoi* Schwarzhans and Bratishko, 2011 from the Tashlik Formation, Luzanivka, Cherkassy Region, Ukraine (Selandian, NP5–6: 60.5–58.3 Ma), or †*Parapristopoma prohumile* Arambourg, 1927 from the Oran-Ravin blanc Formation, Algeria (Messinian: 7.246–5.332 Ma).

Paleoenvironment: Marine.

Comments: See above.

Parameters: $t_y = 5.332$, $t_o = 60.5$, $r = 0.0441/0.0147/0.0044/0.0015$, $p - q = 0.0559/0.0627/0.0694/0.0726$, $lag = 2$, $corr = $ "global_marine"

Node: 54

Group 1: "Caproiformes" I, Lophiiformes

Earliest record: †'*Lophius*' *brachysomus* Agassiz, 1835, †*Histionotophorus bassani* De Zigno, 1887, †*Orrichthys longimanus* Carnevale and Pietsch, 2010, and †*Antigonia veronensis* Sorbini and Bottura, 1988 from Monte Bolca, Verona, Italy (Ypresian, NP14, SB11: 49.4–49.1 Ma).

Paleoenvironment: Marine.

Comments: Sister group relationship of "Caproiformes" I and Lophiiformes is moderately supported by molecular data (Miya *et al.* 2010; Matschiner *et al.* 2011). Lophiiformes are generally rare in the fossil record, but present in the Monte Bolca deposits. Specimens from Monte Bolca have been assigned to the lophiid †'*Lophius' brachysomus* Agassiz, 1835, and the brachionichthyids †*Histionotophorus bassani* De Zigno, 1887 and †*Orrichthys longimanus* Carnevale and Pietsch, 2010 (Carnevale & Pietsch 2010). "Caproiformes" I are also represented in the Monte Bolca ichthyofauna with †*Antigonia veronensis* Sorbini and Bottura, 1988 (Tyler *et al.* 2003).

Group 2: Tetraodontiformes

Earliest record: †*Plectocretacicus clarae* Sorbini, 1979 from the Lithografic Limestone of Hakel, Lebanon (Early Cenomanian: 99.1–97.8 Ma).

Paleoenvironment: Marine.

Comments: Whereas a number of tetraodontid lineages have colonized freshwater habitats (Yamanoue *et al.* 2011), the early diversification probably took place in a marine environment (Santini & Tyler 2003).

Parameters: $t_\mathrm{y} = 97.8$, $t_\mathrm{o} = 99.1$, $r = 0.0309/0.0103/0.0031/0.0010$, $p-q = 0.0559/0.0627/0.0694/0.0726$, $lag = 2$, $corr =$ "global_marine"

Node: 55

Group 1: "Caproiformes" I

Earliest record: †*Antigonia veronensis* Sorbini and Bottura, 1988 from Monte Bolca, Verona, Italy (Ypresian, NP14, SB11: 49.4–49.1 Ma).

Paleoenvironment: Marine.

Comments: See above.

Group 2: Lophiiformes

Earliest record: †'*Lophius' brachysomus* Agassiz, 1835, †*Histionotophorus bassani* De Zigno, 1887, and †*Orrichthys longimanus* Carnevale and Pietsch, 2010 from Monte Bolca, Verona, Italy (Ypresian, NP14, SB11: 49.4–49.1 Ma).

Paleoenvironment: Marine.

Comments: See above.

Parameters: $t_\mathrm{y} = 49.1$, $t_\mathrm{o} = 49.4$, $r = 0.0247/0.0082/0.0025/0.0008$, $p-q = 0.0559/0.0627/0.0694/0.0726$, $lag = 2$, $corr =$ "global_marine"

Node: 56

Group 1: Lophioidei
Earliest record: †'*Lophius*' *brachysomus* Agassiz, 1835 from Monte Bolca, Verona, Italy (Ypresian, NP14, SB11: 49.4–49.1 Ma).
Paleoenvironment: Marine.
Comments: Lophioidei contain the single family Lophiidae. An ancestral position of Lophioidei compared to all other Lophiiformes is strongly supported by molecular data (Miya *et al.* 2010).

Group 2: Ceratioidei, Chaunacoidei (+ Antennarioidei, Ogcocephaloidei)
Earliest record: †*Histionotophorus bassani* De Zigno, 1887, and †*Orrichthys longimanus* Carnevale and Pietsch, 2010 from Monte Bolca, Verona, Italy (Ypresian, NP14, SB11: 49.4–49.1 Ma).
Paleoenvironment: Marine.
Comments: A clade combining Ceratioidei, Chaunacoidei, Antennarioidei, and Ogcocephaloidei is strongly supported by the molecular phylogeny of Miya *et al.* (2010). Antennarioidei include the family Brachionichthyidae, therefore the brachionichthyids †*Histionotophorus bassani* De Zigno, 1887 and †*Orrichthys longimanus* Carnevale and Pietsch, 2010 from Monte Bolca can be used to constrain the age of this clade.

Parameters: $t_{\mathrm{y}} = 49.1$, $t_{\mathrm{o}} = 49.4$, $r = 0.0247/0.0082/0.0025/0.0008$, $p-q = 0.0559/0.0627/0.0694/0.0726$, $lag = 2$, $corr =$ "global_marine"

Node: 57

Group 1: Ceratioidei
Earliest record: †*Chaenophryne* aff. *melanorhabdus* Regan and Trewavas, 1932, †*Leptacanthichthys* cf. *gracilispinis* Regan 1925, †*Oneirodes* sp. Lütken, 1871, †*Borophryne* cf. *apogon* Regan, 1925, and †*Linophryne* cf. *indica* Brauer, 1902 from the Puente Formation (Yorba Member), Los Angeles Basin, California (Mohnian: 8.6–7.6 Ma).
Paleoenvironment: Marine.
Comments: The earliest known ceratioid records are †*Chaenophryne* aff. *melanorhabdus* Regan and Trewavas, 1932, †*Leptacanthichthys* cf. *gracilispinis* Regan, 1925, †*Oneirodes* sp. Lütken, 1871, †*Borophryne* cf. *apogon* Regan, 1925, and †*Linophryne* cf. *indica* Brauer, 1902 from the Puente Formation, California (Mohnian) (Carnevale *et al.* 2008). The Yorba Member of the Californian Puente Formation has been reported to be Mohnian, and 8.6–7.6 myr old (Carnevale *et al.* 2008).

Group 2: Chaunacoidei
Earliest record: †*Chaunax semiangulatus* Stinton, 1978 from the Barton Formation, Hampshire, United Kingdom (Bartonian: 40.4–37.2 Ma).
Paleoenvironment: Marine.
Comments: Otolith remains of †*Chaunax semiangulatus* Stinton, 1978 from the Barton Formation,

United Kingdom (Patterson 1993b) have been found reliable according to the fossil cross-validation of Matschiner *et al.* (2011).

Parameters: $t_\mathrm{y} = 7.6$, $t_\mathrm{o} = 40.4$, $r = 0.0441/0.0147/0.0044/0.0015$, $p - q = 0.0559/0.0627/0.0694/0.0726$, $lag = 2$, $corr =$ "global_marine"

Node: 58

Group 1: Balistoidea
Earliest record: †*Eospinus daniltshenkoi* Tyler and Bannikov, 1992 from the Danata Formation, Uylya-Kushlyuk, Turkmenistan (Upper Thanetian: 57.23–55.8 Ma).
Paleoenvironment: Marine.
Comments: The superfamily Balastoidea combines Balistidae and Monacanthidae, which is strongly supported by morphology (Nelson 2006) and molecular phylogenies (Alfaro *et al.* 2007; Yamanoue *et al.* 2008). The extinct families †Moclaybalistidae and †Bolcabalistidae represent the stem group of Balistoidea. According to Bannikov & Tyler (2008), the oldest known members of these families are †*Moclaybalistes danekrus* Tyler and Santini, 2002 from the Mo-Clay (Fur/Ølst) Formation, Denmark, and †*Eospinus daniltshenkoi* Tyler and Bannikov, 1992 from the Danata Formation, Turkmenistan, which are both claimed to be Thanetian in age. Here, the Ypresian age estimate of Willumsen (2004) for the danish Mo-Clay Formation is accepted, and thus †*Eospinus daniltshenkoi* is used as the earliest known record of Balistoidea.

Group 2: Tetraodontidae (+ Diodontidae)
Earliest record: †*Heptadiodon echinus* Heckel, 1853 and †*Zignodon fornasieroae* Tyler and Santini, 2002 from Monte Bolca, Verona, Italy (Ypresian, NP14, SB11: 49.4–49.1 Ma).
Paleoenvironment: Marine.
Comments: Sister group relationship of Tetraodontidae and Diodontidae is unambiguously supported by morphological (Santini & Tyler 2003) and molecular data (Alfaro *et al.* 2007; Yamanoue *et al.* 2008, 2011), whereas the positions of Molidae, Triacanthidae, Triodontidae, Ostraciidae, and Triacanthodidae remain unresolved and could be ancestral to Balistoidea, Tetraodontidae, and Diodontidae. Of the tetraodontiform fossils included in the phylogenetic analysis of Santini & Tyler (2003), three Eocene taxa appear closely related to extant diodontids: †*Heptadiodon echinus* Heckel, 1853 and †*Zignodon fornasieroae* Tyler and Santini, 2002 from Monte Bolca, Italy (Ypresian), as well as †*Pshekhadiodon parini* Bannikov and Tyler, 1997 from Gorny Luch, Ukraine (Lutetian) (Tyler & Bannikov 2009).

Parameters: $t_\mathrm{y} = 48.6$, $t_\mathrm{o} = 55.8$, $r = 0.0247/0.0082/0.0025/0.0008$, $p - q = 0.0559/0.0627/0.0694/0.0726$, $lag = 2$, $corr =$ "global_marine"

Node: 59

Group 1: *Takifugu* (+ *Tetractenos*, *Torquigener*, *Marilyna*)
Earliest record: Extant.
Paleoenvironment: Marine.
Comments: According to the molecular phylogenies of Holcroft (2005) and Alfaro *et al.* (2007), the genus *Takifugu* forms a monophyletic clade with *Tetractenos*, *Torquigener*, and *Marilyna*. *Sphoeroides* appears closely related to *Takifugu* based on full mitochondrial genomes (Yamanoue *et al.* 2008), whereas *Sphoeroides* is recovered closer to *Tetraodon* in the phylogenies of Holcroft (2005) and Alfaro *et al.* (2007).

Group 2: *Tetraodon* (+ *Arothron*, *Canthigaster*, *Carinotetraodon*)
Earliest record: Extant.
Paleoenvironment: Marine/freshwater.
Comments: The clade combining the genera *Tetraodon*, *Arothron*, and *Canthigaster* is strongly supported by molecular data (Holcroft 2005; Alfaro *et al.* 2007; Yamanoue *et al.* 2008), and *Carinotetraodon* is considered closely related to *Canthigaster* (Carnevale & Tyler 2010). It has been recommended to use the fossil †*Archaeotetraodon winterbottomi* Tyler and Bannikov, 1994 from the Maikop Formation (Psekhsky Horizon), Russia (Oligocene), as an age constraint for the divergence of *Takifugu* and *Tetraodon* (Benton & Donoghue 2007; Benton *et al.* 2009), which has found frequent application in molecular dating studies (Alfaro *et al.* 2007; Azuma *et al.* 2008; Miya *et al.* 2010; Matschiner *et al.* 2011). This recommendation was based on the position of †*Archaeotetraodon winterbottomi* within crown group tetraodontids (there represented by the genera *Canthigaster*, *Sphoeroides*, and *Lagocephalus*) in the phylogenetic analysis of Santini & Tyler (2003), and the assumption that the *Takifugu–Tetraodon* split represents the origin of the tetraodontid crown group, as suggested by Holcroft (2005). However, more recent phylogenetic analyses suggested that the lineages *Lagocephalus*, *Sphoeroides*, and *Colomesus*, and thus the crown group, could be older than the *Takifugu–Tetraodon* divergence (Alfaro *et al.* 2007; Yamanoue *et al.* 2011). The fossil genus †*Archaeotetraodon* was represented in the analysis of Santini & Tyler (2003) by two species, †*A. winterbottomi*, and the younger †*A. jamestyleri* Bannikov, 1990 from the Kerch Peninsula, Ukraine. The genus was recovered as polyphyletic, whereby †*A. winterbottomi* was found in a sister group position to *Canthigaster rostrata* Bloch, 1786 with which it apparently shared two apomorphies that had convergently evolved in other clades. This close relationship suggested a crown group position of †*A. winterbottomi*, whereas †*A. jamestyleri* Bannikov, 1990 was found ancestral to all tetraodontids. The authors were hesitant to propose different generic names, because of 'the high degree of incompleteness of the fossils of the two species of this genus (especially of *A. winterbottomi*)'. Subsequent descriptions of the four new †*Archaeotetraodon* taxa, †*A. cerrinaferoni* Carnevale and Santini, 2006 from the Chelif Basin, Algeria, †*A. bannikovi* Carnevale and Tyler, 2010 from Piedmont, Italy, †*A. dicarloi* Carnevale and Tyler, 2010 from the Abruzzo Apennines, Italy, and †*A. zafaranai* Carnevale and Tyler, 2010 from Sicily, Italy, have led to a reanalysis of the †*Archaeotetraodon* affinities, and have corroborated the monophyly of the genus (Carnevale & Tyler

2010). The authors concluded that †*Archaeotetraodon* is unlikely to be closely related to any of the genera *Amblyrhynchotes*, *Auriglobus*, *Canthigaster*, *Carinotetraodon*, *Chelonodon*, *Contusus*, *Ephippion*, *Feroxodon*, *Javichthys*, *Lagocephalus*, *Marilyna*, *Monotrete*, *Omegophora*, *Polyspina*, *Reicheltia*, *Takifugu*, *Tetractenos*, *Tetraodon*, *Torquigener*, *Tylerius*, and *Xenopterus*, but that it shares similarities with *Arothron*, *Colomesus*, *Pelagocephalus*, and *Sphoeroides*. Thus, the position of †*Archaeotetraodon* within Tetraodontidae remains unresolved, and a stem group position, or a crown group position ancestral to the *Takifugu–Tetraodon* divergence cannot be excluded. The only fossils that unquestionably belong to crown group tetraodontids are skeletal remains of †*Sphoeroides hyperostosus* Tyler *et al.*, 1992, from the Yorktown Formation, North Carolina (Pliocene) (Carnevale & Tyler 2010), which are also not used to constrain the *Takifugu–Tetraodon* split, as *Sphoeroides* appears to be ancestral to it (Yamanoue *et al.* 2011). Tetraodontids have repeatedly colonized freshwater habitats (Yamanoue *et al.* 2011).

Parameters: $t_\mathrm{y} = 0$, $t_\mathrm{o} = 0$, $r = 0.0380/0.0127/0.0038/0.0013$, $p - q = 0.0559/0.0627/0.0694/0.0726$, $lag = 2$, $corr =$ "global"

Node: 60

Group 1: Balistidae
Earliest record: †*Gornylistes prodigiosus* Bannikov and Tyler, 2008 from Gorny Luch, Krasnodar Region, Ukraine (Bartonian: 40.4–37.2 Ma).
Paleoenvironment: Marine.
Comments: †*Gornylistes prodigiosus* Bannikov and Tyler, 2008 from Gorny Luch (Kuma Horizon), Ukraine (Bartonian) is the oldest member of Balistidae (Bannikov & Tyler 2008). The Kuma Horizon at Gorny Luch is characterized by foraminiferans *Turborotalia centralis*, *Globigerina praebulloides*, *Globigerina turkmenica*, and *Globanomalina micra* (Tyler & Bannikov 1992), which indicates a Bartonian age (Tyler & Bannikov 1997).

Group 2: Monacanthidae
Earliest record: †*Frigocanthus stroppanobili* Sorbini and Tyler, 2004 and †*Frigocanthus margaritatus* Sorbini and Tyler, 2004 from the Metauro River, Marche, Italy, and Crete, Greece (Pliocene: 5.332–2.588).
Paleoenvironment: Marine.
Comments: The earliest skeletal record of monacanthids are †*Frigocanthus stroppanobili* Sorbini and Tyler, 2004 and †*Frigocanthus margaritatus* Sorbini and Tyler, 2004 from the Metauro River, Italy, and from Crete, Greece (Pliocene) (Sorbini & Tyler 2004). Otolith fossils described as †*Amanses sulcifer* Stinton, 1966 from the London Clay Formation, Kent, United Kingdom (Ypresian), were found reliable according the fossil cross-validation of Matschiner *et al.* (2011), however the authors were unaware of the reanalysis of Schwarzhans (2003), which suggested that the otoliths represent a zeiform, instead of a tetraodontiform species. Thus, the results of Matschiner *et al.*

(2011) are here taken as evidence for a Paleocene-Eocene divergence of balistids and monacanthids rather than a corroboration of the †*Amanses sulcifer* assignment as a monacanthid.

Parameters: $t_y = 37.2$, $t_o = 40.4$, $r = 0.0247/0.0082/0.0025/0.0008$, $p-q = 0.0559/0.0627/0.0694/$ 0.0726, $lag = 2$, $corr =$ "global_marine"

Node: 61

Group 1: Epinephelidae *sensu* Smith and Craig, 2007
Earliest record: †*Epinephelus casottii* Costa, 1858 from Retznei, Steiermark, Austria (Langhian: 15.97–13.65 Ma).
Paleoenvironment: Marine.
Comments: See above.

Group 2: "Trachiniformes", Serranidae *sensu* Smith and Craig, 2007, Percidae, Scorpaenoidei, Cottiformes, Gasterosteioidei, Nototheniiformes
Earliest record: †*Plesioserranus wemmeliensis* from the London Clay Formation of Kent, United Kingdom (Ypresian, MP8–9: 55.2–50.8 Ma).
Paleoenvironment: Marine.
Comments: See above.

Parameters: $t_y = 50.8$, $t_o = 55.2$, $r = 0.0247/0.0082/0.0025/0.0008$, $p-q = 0.0559/0.0627/0.0694/$ 0.0726, $lag = 2$, $corr =$ "global_marine"

Node: 62

Group 1: *Paranthias*
Earliest record: Extant.
Paleoenvironment: Marine.
Comments: According to the molecular phylogenies of Smith & Craig (2007) and Craig & Hastings (2007), *Paranthias* appears to be nested within a polyphyletic *Cephalopholis*. No fossils are known of the genus.

Group 2: *Cephalopholis taeniops, Mycteroperca, Epinephelus* (+ *Hyporthodus, Alphestes, Dermatolepis, Triso*)
Earliest record: †*Epinephelus casottii* Costa, 1858 from Retznei, Steiermark, Austria (Langhian: 15.97–13.65 Ma).
Paleoenvironment: Marine.
Comments: Of the genus *Cephalopholis*, only *C. taeniops* is certainly included in this group, be-

cause other members of this genus could be closer to *Paranthias* (Smith & Craig 2007; Craig & Hastings 2007). A clade combining the genera *Epinephelus*, *Mycteroperca*, *Hyporthodus*, *Alphestes*, *Dermatolepis*, and *Triso* is strongly supported by molecular data (Craig & Hastings 2007). The earliest known record is †*Epinephelus casottii* (see above).

Parameters: $t_y = 13.65$, $t_o = 15.97$, $r = 0.0441/0.0147/0.0044/0.0015$, $p - q = 0.0559/0.0627/0.0694/0.0726$, $lag = 2$, $corr =$ "global_marine"

Node: 63

Group 1: *Cephalopholis taeniops*
Earliest record: Extant.
Paleoenvironment: Marine.
Comments: -

Group 2: *Epinephelus*, *Mycteroperca* (+ *Hyporthodus*, *Alphestes*, *Dermatolepis*, *Triso*)
Earliest record: †*Epinephelus casottii* Costa, 1858 from Retznei, Steiermark, Austria (Langhian: 15.97–13.65 Ma).
Paleoenvironment: Marine.
Comments: See above.

Parameters: $t_y = 13.65$, $t_o = 15.97$, $r = 0.0441/0.0147/0.0044/0.0015$, $p - q = 0.0559/0.0627/0.0694/ 0.0726$, $lag = 2$, $corr =$ "global_marine"

Node: 64

Group 1: *Epinephelus*
Earliest record: †*Epinephelus casottii* Costa 1858 from Retznei, Steiermark, Austria (Langhian: 15.97–13.65 Ma).
Paleoenvironment: Marine.
Comments: -

Group 2: *Mycteroperca*
Earliest record: Extant.
Paleoenvironment: Marine.
Comments: See above.

Parameters: $t_y = 13.65$, $t_o = 15.97$, $r = 0.0441/0.0147/0.0044/0.0015$, $p - q = 0.0559/0.0627/0.0694/0.0726$, $lag = 2$, $corr =$ "global_marine"

Node: 65

Group 1: "Trachiniformes", Serranidae *sensu* Smith and Craig, 2007, Percidae
Earliest record: †*Plesioserranus wemmeliensis* from the London Clay Formation of Kent, United Kingdom (Ypresian, MP8–9: 55.2–50.8 Ma).
Paleoenvironment: Marine/freshwater.
Comments: See above. Whereas "Trachiniformes" and Serranidae are exclusively marine, Percidae are confined to freshwater habitats.

Group 2: Scorpaenoidei, Cottiformes, Gasterosteioidei, Nototheniiformes
Earliest record: †*Scorpaena* sp. Linné, 1758 from the Mainz Basin and the Embayment of Leipzig, Germany (Rupelian: 33.9–28.4 Ma), or †*Cottus cervicornis* Storms, 1984 from Argile de Boum, Belgium (Rupelian: 33.9–28.4 Ma).
Paleoenvironment: Marine.
Comments: See above. Scorpaenoidei and Cottiformes are primarily marine, only few gasterosteioid and a single nototheniiform species have colonized freshwater habitats.

Parameters: $t_y = 50.8$, $t_o = 55.2$, $r = 0.0160/0.0053/0.0016/0.0005$, $p-q = 0.0559/0.0627/0.0694/0.0726$, $lag = 2$, $corr =$ "global"

Node: 66

Group 1: "Trachiniformes", Serranidae *sensu* Smith and Craig, 2007
Earliest record: †*Plesioserranus wemmeliensis* from the London Clay Formation of Kent, United Kingdom (Ypresian, MP8–9: 55.2–50.8 Ma).
Paleoenvironment: Marine.
Comments: See above.

Group 2: Percidae
Earliest record: †*Perca hassiaca* Weiler, 1961 and †*Perca* 'sp. 1' from the Upper Rhine Graben and the Mainz Basin, Germany (Rupelian: 33.9–28.4 Ma), or †*Perca fluviatilis* Linné, 1758 from La Montagne d'Andance, France, and from Murzak-Koba, Crimea, Ukraine (Tortonian–Messinian: 11.608–5.332 Ma).
Paleoenvironment: Freshwater.
Comments: Reports of Miocence fossils of †*Perca fluviatilis* are mentioned in Carney & Dick (2000), and their age is given as 26 Ma, which would actually be in the Oligocene. While the original studies of Lebedev (1952) and Mein *et al.* (1983) were unavailable, their titles show that the reported fossils date from the Upper Miocene (Tortonian–Messinian). This date agrees with the

earliest skeletal record listed in Patterson (1993b). The earliest known otolith remains of Percidae belong to †*Perca hassiaca* Weiler, 1961 and †*Perca* sp. 1 from the Upper Rhine Graben and the Mainz Basin, Germany (Rupelian) (Martini & Reichenbacher 2007).

Parameters: $t_y = 50.8$, $t_o = 55.2$, $r = 0.0160/0.0053/0.0016/0.0005$, $p-q = 0.0559/0.0627/0.0694/0.0726$, $lag = 2$, $corr =$ "global"

Node: 67

Group 1: "Trachiniformes" (excluding Trachininae and Chiasmodontidae)
Earliest record: †'*Hemerocoetinarum' apertus* Schwarzhans, 1980 and †*Uranoscopus ignavus* Schwarzhans, 1980 from Waihao River, Canterbury, New Zealand (Lutetian: 48.6–40.4 Ma), †*Ammodytes vasseuri* Nolf and Lapierre, 1977 from the Sables des Bois-Gouet, Loire-Atlantique, France (Lutetian: 48.6–40.4 Ma).
Paleoenvironment: Marine.
Comments: The affinities of Chiasmodontidae remain unclear, and may lie with "Stromateiformes" rather than with "Trachiniformes" (Dettaï & Lecointre 2005; Smith & Craig 2007), therefore the skeletal fossil of the chiasmodontid †*Pseudoscopelus grossheimi* Danil'chenko, 1960 from the lower Khadum Formation, Russia (Rupelian) (Patterson 1993b) is not used here. Disregarding Trachinidae and Chiasmodontidae, no skeletal fossils of "Trachiniformes" are known.

Group 2: Serranidae *sensu* Smith and Craig, 2007 (including Trachininae)
Earliest record: †*Plesioserranus wemmeliensis* from the London Clay Formation of Kent, United Kingdom (Ypresian, MP8–9: 55.2–50.8 Ma).
Paleoenvironment: Marine.
Comments: See above.

Parameters: $t_y = 50.8$, $t_o = 55.2$, $r = 0.0247/0.0082/0.0025/0.0008$, $p-q = 0.0559/0.0627/0.0694/0.0726$, $lag = 2$, $corr =$ "global_marine"

Node: 68

Group 1: Cheimarrhichthyidae (+ Pinguipedidae, Leptoscopidae, Ammodytidae, Uranoscopelidae)
Earliest record: †*Uranoscopus ignavus* Schwarzhans, 1980 from Waihao River, Canterbury, New Zealand (Lutetian: 48.6–40.4 Ma), †*Ammodytes vasseuri* Nolf and Lapierre, 1977 from the Sables des Bois-Gouet, Loire-Atlantique, France (Lutetian: 48.6–40.4 Ma).
Paleoenvironment: Marine.
Comments: Cheimarrhichthyidae is a monotypic familiy consisting of *Cheimarrichthys fosteri*. It

appears closely related to Pinguipedidae, Leptoscopidae, Ammodytidae, and Uranoscopelidae in the molecular phylogenies of Dettaï & Lecointre (2005), Smith & Craig (2007), and Li *et al.* (2009). All fossils listed are otolith remains, no skeletal fossils are known of this group (Patterson 1993b).

Group 2: Percophidae
Earliest record: †'*Hemerocoetinarum' apertus* Schwarzhans, 1980 from Waihao River, Canterbury, New Zealand (Lutetian: 48.6–40.4 Ma).
Paleoenvironment: Marine.
Comments: The relationships of percophids remain ambiguous (Smith & Craig 2007; Matschiner *et al.* 2011). Otolith remains have been described as †'*Hemerocoetinarum' apertus*. No skeletal fossils of percophids are known (Patterson 1993b). Given the general uncertainty of taxonomic assignments of otolith fossils, the age of the oldest fossil is here constrained only with 48.6–0 myr, allowing for the possibility that all "trachiniform" otoliths have been misidentified.

Parameters: $t_y = 0$, $t_o = 48.6$, $r = 0.0441/0.0147/0.0044/0.0015$, $p - q = 0.0559/0.0627/0.0694/0.0726$, $lag = 2$, $corr =$ "global_marine"

Node: 69

Group 1: *Bembrops greyi*
Earliest record: Extant.
Paleoenvironment: Marine.
Comments: -

Group 2: *Bembrops heterurus*
Earliest record: Extant.
Paleoenvironment: Marine.
Comments: -

Parameters: $t_y = 0$, $t_o = 0$, $r = 0.0441/0.0147/0.0044/0.0015$, $p - q = 0.0559/0.0627/0.0694/0.0726$, $lag = 2$, $corr =$ "global_marine"

Node: 70

Group 1: Trachinidae
Earliest record: †*Callipteryx speciosus* Agassiz, 1838 and †*Callipteryx recticaudatus* Agassiz, 1838 from Monte Bolca, Verona, Italy (Ypresian, NP14, SB11: 49.4–49.1 Ma).
Paleoenvironment: Marine.
Comments: See above.

Group 2: Serranidae *sensu* Smith and Craig, 2007 (excluding Trachinidae)
Earliest record: †*Plesioserranus wemmeliensis* from the London Clay Formation of Kent, United Kingdom (Ypresian, MP8–9: 55.2–50.8 Ma).
Paleoenvironment: Marine.
Comments: See above.

Parameters: $t_{\mathrm{y}} = 50.8$, $t_{\mathrm{o}} = 55.2$, $r = 0.0247/0.0082/0.0025/0.0008$, $p-q = 0.0559/0.0627/0.0694/0.0726$, $lag = 2$, $corr = $ "global_marine"

Node: 71

Group 1: *Trachinus draco*
Earliest record: Extant.
Paleoenvironment: Marine.
Comments: Whereas skeletal fossils have been assigned to the genus *Trachinus*, such as †*T. minutus* Jonet, 1958 from Froidefontaine, France (Oligocene) (Pharisat 1998), their affinities to *Trachinus draco* or *Trachinus radiatus* are not known, and they could be ancestral to both species. Thus, no fossils are known to constrain the *Trachinus draco*–*Trachinus radiatus* divergence.

Group 2: *Trachinus radiatus*
Earliest record: Extant.
Paleoenvironment: Marine.
Comments: See above.

Parameters: $t_{\mathrm{y}} = 0$, $t_{\mathrm{o}} = 0$, $r = 0.0441/0.0147/0.0044/0.0015$, $p-q = 0.0559/0.0627/0.0694/0.0726$, $lag = 2$, $corr = $ "global_marine"

Node: 72

Group 1: *Serranus baldwini*
Earliest record: Extant.
Paleoenvironment: Marine.
Comments: No fossils are known of *Serranus baldwini*.

Group 2: *Serranus atricauda*, *Hypoplectrus* (+ *Serranus tabacarius*, *Cratinus*, *Paralabrax*, *Plectranthias kelloggi*, *Diplectrum*)
Earliest record: Extant.
Paleoenvironment: Marine.

Comments: The molecular phylogenies of Pondella II *et al.* (2003) and Smith & Craig (2007) suggest a clade combining the genera *Hypoplectrus*, *Cratinus*, *Paralabrax*, *Plectranthias kelloggi*, and *Diplectrum*. The genera *Serranus* and *Plectranthias* appear paraphyletic (Smith & Craig 2007; this study), thus, only the species *S. atricauda*, *S. tabacarius*, and *P. kelloggi* are here considered to group with *Hypoplectrus*, *Cratinus*, *Paralabrax*, and *Diplectrum*. No fossils are known of this group.

Parameters: $t_\text{y} = 0$, $t_\text{o} = 0$, $r = 0.0441/0.0147/0.0044/0.0015$, $p{-}q = 0.0559/0.0627/0.0694/0.0726$, $lag = 2$, $corr =$ "global_marine"

Node: 73

Group 1: *Serranus atricauda*
Earliest record: Extant.
Paleoenvironment: Marine.
Comments: -

Group 2: *Hypoplectrus* (+ *Serranus tabacarius*, *Cratinus*, *Paralabrax*, *Plectranthias kelloggi*, *Diplectrum*)
Earliest record: Extant.
Paleoenvironment: Marine.
Comments: See above.

Parameters: $t_\text{y} = 0$, $t_\text{o} = 0$, $r = 0.0441/0.0147/0.0044/0.0015$, $p{-}q = 0.0559/0.0627/0.0694/0.0726$, $lag = 2$, $corr =$ "global_marine"

Node: 74

Group 1: Etheostomatinae
Earliest record: Extant.
Paleoenvironment: Freshwater.
Comments: Etheostomatinae include the four exclusively North American genera *Ammocrypta*, *Crystallaria*, *Percina*, and "*Etheostoma*", whereby "*Etheostoma*" may only be monophyletic after the exclusion of subgenus *Nothonotus* (Sloss *et al.* 2004). The earliest record of Etheostomatinae is from the Late Pleistocene, (0.1–0.0 Ma) and is here neglected (Cavender 1986).

Group 2: "Percinae", Luciopercinae
Earliest record: †*Perca hassiaca* Weiler, 1961 and †*Perca* 'sp. 1' from the Upper Rhine Graben and the Mainz Basin, Germany (Rupelian: 33.9–28.4 Ma), or †*Perca fluviatilis* Linné, 1758 from La Montagne d'Andance, France, and from Murzak-Koba, Crimea, Ukraine (Tortonian–Messinian:

11.608–5.332 Ma).

Paleoenvironment: Freshwater.

Comments: "Percinae" comprise the genera *Perca* and *Gymnocephalus* (Song *et al.* 1998) and could be paraphyletic (Sloss *et al.* 2004; Matschiner *et al.* 2011). Luciopercinae unite *Sander*, *Zingel*, and *Romanichthys* (Song *et al.* 1998), and appear to be a monophyletic group (Sloss *et al.* 2004; Matschiner *et al.* 2011).

Parameters: $t_\mathrm{y} = 5.332$, $t_\mathrm{o} = 33.9$, $r = 0.0320/0.0107/0.0032/0.0011$, $p{-}q = 0.0559/0.0627/0.0694/$ 0.0726, $lag = 2$, $corr = $ "global_terrestrial"

Node: 75

Group 1: *Etheostoma* (excluding *Nothonotus*)
Earliest record: Extant.
Paleoenvironment: Freshwater.
Comments: -

Group 2: *Percina*
Earliest record: Extant.
Paleoenvironment: Freshwater.
Comments:

Parameters: $t_\mathrm{y} = 0$, $t_\mathrm{o} = 0$, $r = 0.0320/0.0107/0.0032/0.0011$, $p{-}q = 0.0559/0.0627/0.0694/0.0726$, $lag = 2$, $corr = $ "global_terrestrial"

Node: 76

Group 1: *Etheostoma zonale*
Earliest record: Extant.
Paleoenvironment: Freshwater.
Comments: -

Group 2: *Etheostoma caeruleum*
Earliest record: Extant.
Paleoenvironment: Freshwater.
Comments: -

Parameters: $t_\mathrm{y} = 0$, $t_\mathrm{o} = 0$, $r = 0.0320/0.0107/0.0032/0.0011$, $p{-}q = 0.0559/0.0627/0.0694/0.0726$, $lag = 2$, $corr = $ "global_terrestrial"

Node: 77

Group 1: *Percina macrolepida*
Earliest record: Extant.
Paleoenvironment: Freshwater.
Comments: -

Group 2: *Percina caprodes*
Earliest record: Extant.
Paleoenvironment: Freshwater.
Comments: -

Parameters: $t_y = 0$, $t_o = 0$, $r = 0.0320/0.0107/0.0032/0.0011$, $p-q = 0.0559/0.0627/0.0694/0.0726$, $lag = 2$, $corr =$ "global_terrestrial"

Node: 78

Group 1: *Perca*
Earliest record: †*Perca hassiaca* Weiler, 1961 and †*Perca* 'sp. 1' from the Upper Rhine Graben and the Mainz Basin, Germany (Rupelian: 33.9–28.4 Ma), or †*Perca fluviatilis* Linné, 1758 from La Montagne d'Andance, France, and from Murzak-Koba, Crimea, Ukraine (Tortonian–Messinian: 11.608–5.332 Ma).
Paleoenvironment: Freshwater.
Comments: See above.

Group 2: *Gymnocephalus*, Luciopercinae
Earliest record: †*Sander teneri* Murray *et al.*, 2009 from Beaver Pond site, Ellesmere Island, Canada (Zanclean: 5.0–4.0 Ma).
Paleoenvironment: Freshwater.
Comments: The earliest known member of this group is represented by †*Sander teneri* from the Canadian Arctic. The Beaver Pond site of Ellesmere Island is supposed to be 5.0–4.0 myr old (Murray *et al.* 2009), which is here accepted.

Parameters: $t_y = 5.332$, $t_o = 33.9$, $r = 0.0320/0.0107/0.0032/0.0011$, $p-q = 0.0559/0.0627/0.0694/0.0726$, $lag = 2$, $corr =$ "global_terrestrial"

Node: 79

Group 1: Luciopercinae
Earliest record: †*Sander teneri* Murray *et al.*, 2009 from Beaver Pond site, Ellesmere Island, Canada (Zanclean: 5.0–4.0 Ma).
Paleoenvironment: Freshwater.
Comments: See above.

Group 2: *Gymnocephalus*
Earliest record: Extant.
Paleoenvironment: Freshwater.
Comments: -

Parameters: $t_{\mathrm{y}} = 4.0$, $t_{\mathrm{o}} = 5.0$, $r = 0.0320/0.0107/0.0032/0.0011$, $p - q = 0.0559/0.0627/0.0694/0.0726$, $lag = 2$, $corr = $ "global_terrestrial"

Node: 80

Group 1: *Zingel*
Earliest record: Extant.
Paleoenvironment: Freshwater.
Comments: -

Group 2: *Sander*
Earliest record: †*Sander teneri* Murray *et al.*, 2009 from Beaver Pond site, Ellesmere Island, Canada (Zanclean: 5.0–4.0 Ma).
Paleoenvironment: Freshwater.
Comments: -

Parameters: $t_{\mathrm{y}} = 4.0$, $t_{\mathrm{o}} = 5.0$, $r = 0.0320/0.0107/0.0032/0.0011$, $p - q = 0.0559/0.0627/0.0694/0.0726$, $lag = 2$, $corr = $ "global_terrestrial"

Node: 81

Group 1: *Gymnocephalus cernuus*
Earliest record: Extant.
Paleoenvironment: Freshwater.
Comments: See above.

Group 2: *Gymnocephalus schraetser*
Earliest record: Extant.
Paleoenvironment: Freshwater.
Comments: -

Parameters: $t_y = 0$, $t_o = 0$, $r = 0.0320/0.0107/0.0032/0.0011$, $p-q = 0.0559/0.0627/0.0694/0.0726$, $lag = 2$, $corr = $ "global_terrestrial"

Node: 82

Group 1: Scorpaenoidei, Cottiformes, Gasterosteioidei
Earliest record: †*Scorpaena* sp. Linné, 1758 from the Mainz Basin and the Embayment of Leipzig, Germany (Rupelian: 33.9–28.4 Ma), or †*Cottus cervicornis* Storms, 1984 from Argile de Boum, Belgium (Rupelian: 33.9–28.4 Ma).
Paleoenvironment: Marine.
Comments: See above.

Group 2: Nototheniiformes
Earliest record: Extant.
Paleoenvironment: Marine.
Comments: See above for reasons why †*Proeleginops grandeastmanorum* is not considered a nototheniiform fossil.

Parameters: $t_y = 28.4$, $t_o = 33.9$, $r = 0.0225/0.0075/0.0023/0.0008$, $p-q = 0.0559/0.0627/0.0694/0.0726$, $lag = 2$, $corr = $ "global_marine"

Node: 83

Group 1: Scorpaenoidei
Earliest record: †*Scorpaena* sp. Linné, 1758 from the Mainz Basin and the Embayment of Leipzig, Germany (Rupelian: 33.9–28.4 Ma).
Paleoenvironment: Marine.
Comments: See above.

Group 2: Cottiformes, Gasterosteioidei
Earliest record: †*Cottus cervicornis* Storms, 1984 from Argile de Boum, Belgium (Rupelian: 33.9–28.4 Ma).
Paleoenvironment: Marine.
Comments: Molecular phylogenies support the close association of Cottiformes and Gasterosteioidei

(Mabuchi *et al.* 2007; Kawahara *et al.* 2008). See above for a discussion of the cottiform and gasterosteioid fossil record.

Parameters: $t_y = 28.4$, $t_o = 33.9$, $r = 0.0225/0.0075/0.0023/0.0008$, $p-q = 0.0559/0.0627/0.0694/0.0726$, $lag = 2$, $corr =$ "global_marine"

Node: 84

Group 1: *Helicolenus*
Earliest record: Extant.
Paleoenvironment: Marine.
Comments: -

Group 2: *Sebastes*
Earliest record: †*Sebastes apostates* Jordan, 1920 and †*Sebastes ineziae* Jordan in Jordan and Gilbert, 1920 from the Monterey Formation (Lompoc Quarry), Santa Barbara County, California (Tortonian–Messinian, NPD7A: 7.7–6.8 Ma).
Paleoenvironment: Marine.
Comments: Within Sebastinae, the genera *Sebastes* and *Helicolenus* appear as sister groups in the phylogenetic analysis of Smith & Wheeler (2004). The earliest record of *Sebastes* may be provided by †*Sebastes apostates* Jordan, 1920, †*Sebastes ineziae* Jordan in Jordan and Gilbertm 1920, and others from the Lompoc Quarry of the Monterey Formation, California (Tortonian–Messinian) (Schultz 1993). The Lompoc Quarry has been correlated with diatom zone NPD7A (Barron *et al.* 2002), which is 7.7–6.8 myr old (Ogg & Ogg 2008b).

Parameters: $t_y = 6.8$, $t_o = 7.7$, $r = 0.0441/0.0147/0.0044/0.0015$, $p-q = 0.0559/0.0627/0.0694/0.0726$, $lag = 2$, $corr =$ "global_marine"

Node: 85

Group 1: *Sebastes ruberrimus*
Earliest record: Extant.
Paleoenvironment: Marine.
Comments: The relationships of extant species of Sebastes to their extinct congenerics are unclear.

Group 2: *Sebastes marinus*
Earliest record: Extant.
Paleoenvironment: Marine.
Comments:

Parameters: $t_y = 0$, $t_o = 0$, $r = 0.0441/0.0147/0.0044/0.0015$, $p-q = 0.0559/0.0627/0.0694/0.0726$, $lag = 2$, $corr =$ "global_marine"

Node: 86

Group 1: Cottiformes
Earliest record: †*Cottus cervicornis* Storms, 1984 from Argile de Boum, Belgium (Rupelian: 33.9–28.4 Ma).
Paleoenvironment: Marine.
Comments: See above. Cottiformes are marine with the exception of the species of Lake Baikal. Thus, the early diversification most probably took place in a marine environment.

Group 2: Gasterosteioidei
Earliest record: *Gasterosteus aculeatus* Linné, 1758 from the Monterey Formation, California (Serravalian: 13.3–13.0 Ma).
Paleoenvironment: Marine.
Comments: See above for a discussion of the gasterosteioid fossil record. The age of the *Gasterosteus aculeatus* fossil from the Monterey Formation is reported as 13.3–13.0 Ma. While *Gasterosteus aculeatus* is famous for its repeated colonization of novel freshwater habitats, the ancestral gasterosteioid lineages, and the earliest gasterosteoid fossil are marine (Bell 1977).

Parameters: $t_y = 28.4$, $t_o = 33.9$, $r = 0.0225/0.0075/0.0023/0.0008$, $p-q = 0.0559/0.0627/0.0694/0.0726$, $lag = 2$, $corr =$ "global_marine"

Node: 87

Group 1: Zoarcoidei
Earliest record: †*Ascoldia agnevica* Grechina, 1980, †*Ernogrammus litoralis* Grechina, 1980, †*Stichaeus brachigrammus* Nazarkin, 1998, †*Stichaeopsis sakhalinensis* Nazarkin, 1998, †*Nivchia makushoki* Nazarkin, 1998, †*Agnevichthys gretchinae* Nazarkin, 2002 and †*Palaeopholis laevis* Nazarkin, 2002 from the Agnev Formation, Sakhalin, Russia (Serravallian–Tortonian: 12.3–11.5 Ma).
Paleoenvironment: Marine.
Comments: Zoarcoidei comprise the families Zoarcidae, Bathymasteridae, Stichaeidae, Cryptacanthodidae, Pholididae, Anarhichadidae, Ptilichthyidae, and Scytalinidae. Nelson (2006) further includes the monotypic family Zaproridae in Zoarcoidei, however, the phylogenetic affinites of Zaproridae remain ambiguous (Mecklenburg 2003), and the fossil record of zaprorids is not considered here. Disregarding Zaproridae, the earliest zoarcoid fossils are provided by the stichaeids †*Ascoldia agnevica* Grechina, 1980, †*Ernogrammus litoralis* Grechina, 1980, †*Stichaeus brachigram-*

mus Nazarkin, 1998, †*Stichaeopsis sakhalinensis* Nazarkin, 1998, and †*Nivchia makushoki* Nazarkin, 1998 from the Agnev Formation, Sakhalin, Russia (Patterson 1993b; Nazarkin 1998), and by the pholids †*Agnevichthys gretchinae* Nazarkin, 2002 and †*Palaeopholis laevis* Nazarkin, 2002 from the same location (Nazarkin 2002). The Agnev Formation is reported to be 12.3–11.5 myr old (Nazarkin 2002), which is accepted here.

Group 2: Cottoidei
Earliest record: †*Cottus cervicornis* Storms, 1984 from Argile de Boum, Belgium (Rupelian: 33.9–28.4 Ma).
Paleoenvironment: Marine.
Comments: See above.

Parameters: $t_y = 28.4$, $t_o = 33.9$, $r = 0.0225/0.0075/0.0023/0.0008$, $p-q = 0.0559/0.0627/0.0694/0.0726$, $lag = 2$, $corr = $ "global_marine"

Node: 88

Group 1: Zoarcinae (+ Gymnelinae)
Earliest record: Extant.
Paleoenvironment: Marine.
Comments: The zoarcid subfamily Zoarcinae comprises four species of the genus *Zoarces* (Anderson & Fedorov 2004). On the basis of morphology, the subfamily Gymnelinae seems to be the the sister group of Zoarcinae (Anderson 1994). No fossils are known of Zoarcinae or Gymnelinae.

Group 2: Lycodinae
Earliest record: *Lycodes pacificus* Collett, 1879 from California (Pliocene: 5.332–2.588 Ma).
Paleoenvironment: Marine.
Comments: The zoarcid subfamily Lycodinae appears as the sister group to a clade combining Zoarcinae and Gymnelinae, according to morphological data (Anderson 1994). The fossil record of Zoarcidae is limited to otolith remains of the extant species *Lycodes pacificus* Collett, 1879 from California (Pliocene) (Patterson 1993b).

Parameters: $t_y = 0$, $t_o = 5.332$, $r = 0.0441/0.0147/0.0044/0.0015$, $p-q = 0.0559/0.0627/0.0694/0.0726$, $lag = 2$, $corr = $ "global_marine"

Node: 89

Group 1: Hypoptychidae
Earliest record: Extant.

Paleoenvironment: Marine.
Comments: The position of the family Hypoptychidae as the most ancestral gasterosteioid lineage is strongly supported by molecular data (Kawahara *et al.* 2008, 2009). No fossils are known of the family.

Group 2: Aulorhynchidae, Gasterosteidae
Earliest record: *Gasterosteus aculeatus* Linné, 1758 from the Monterey Formation, California (Serravalian: 13.3–13.0 Ma).
Paleoenvironment: Marine.
Comments: Sister group relationship of Aulorhynchidae and Gasterosteidae is strongly supported. See above.

Parameters: $t_\mathrm{y} = 13.0$, $t_\mathrm{o} = 13.3$, $r = 0.0441/0.0147/0.0044/0.0015$, $p-q = 0.0559/0.0627/0.0694/0.0726$, $lag = 2$, $corr = $ "global_marine"

Node: 90

Group 1: Aulorhynchidae
Earliest record: Extant.
Paleoenvironment: Marine.
Comments: Aulorhynchidae were found paraphyletic on the basis full mitochondrial genome data (Kawahara *et al.* 2008, 2009), which is not the case here. No aulorhynchid fossils are known.

Group 2: Gasterosteidae
Earliest record: *Gasterosteus aculeatus* Linné, 1758 from the Monterey Formation, California (Serravalian: 13.3–13.0 Ma).
Paleoenvironment: Marine.
Comments: See above.

Parameters: $t_\mathrm{y} = 13.0$, $t_\mathrm{o} = 13.3$, $r = 0.0441/0.0147/0.0044/0.0015$, $p-q = 0.0559/0.0627/0.0694/0.0726$, $lag = 2$, $corr = $ "global_marine"

Node: 91

Group 1: *Aulorhynchus*
Earliest record: Extant.
Paleoenvironment: Marine.
Comments: -

Group 2: *Aulichthys*
Earliest record: Extant.
Paleoenvironment: Marine.
Comments: -

Parameters: $t_y = 0$, $t_o = 0$, $r = 0.0441/0.0147/0.0044/0.0015$, $p-q = 0.0559/0.0627/0.0694/0.0726$, $lag = 2$, $corr =$ "global_marine"

Node: 92

Group 1: *Spinachia* (+ *Apeltes*)
Earliest record: Extant.
Paleoenvironment: Marine.
Comments: Sister group relationship between *Spinachia* and *Apeltes* is strongly supported by mitochondrial sequence data (Kawahara *et al.* 2009). No fossils are known of the genus *Spinachia*.

Group 2: *Gasterosteus*, *Culea* (+ *Pungitius*)
Earliest record: *Gasterosteus aculeatus* Linné, 1758 from the Monterey Formation, California (Serravalian: 13.3–13.0 Ma).
Paleoenvironment: Marine/freshwater.
Comments: The position of *Culea* relative to *Spinachia* and *Gasterosteus* is ambiguous (Kawahara *et al.* 2009; Matschiner *et al.* 2011). *Pungitius* was identified as the sister group of *Culea*, using molecular data (Kawahara *et al.* 2009).

Parameters: $t_y = 13.0$, $t_o = 13.3$, $r = 0.0380/0.0127/0.0038/0.0013$, $p-q = 0.0559/0.0627/0.0694/0.0726$, $lag = 2$, $corr =$ "global"

Node: 93

Group 1: *Culea* (+ *Pungitius*)
Earliest record: Extant.
Paleoenvironment: Freshwater.
Comments: No fossils of *Culea* or *Pungitius* are known.

Group 2: *Gasterosteus*
Earliest record: *Gasterosteus aculeatus* Linné, 1758 from the Monterey Formation, California (Serravalian: 13.3–13.0 Ma).
Paleoenvironment: Marine/freshwater.
Comments: See above.

Parameters: $t_y = 13.0$, $t_o = 13.3$, $r = 0.0380/0.0127/0.0038/0.0013$, $p-q = 0.0559/0.0627/0.0694/0.0726$, $lag = 2$, $corr = $ "global"

Node: 94

Group 1: Bovichtidae
Earliest record: Extant.
Paleoenvironment: Marine.
Comments: Bovichtidae are the most ancestral nototheniiform lineage, which is strongly supported by molecular and morphological data (Eastman 1993; Matschiner *et al.* 2011). No bovichtid fossils are known.

Group 2: Pseudaphritidae, Eleginopidae, 'Antarctic Clade'
Earliest record: Extant.
Paleoenvironment: Marine.
Comments: Monophyly of a group combining all nototheniiforms except Bovichtidae is strongly supported. The nototheniiform 'Antarctic Clade' comprises the families "Nototheniidae", Harpagiferidae, Artedidraconidae, Bathydraconidae, and Channichthyidae. The relationships of nototheniiform families are well investigated (Near & Cheng 2008; Matschiner *et al.* 2011; Rutschmann *et al.* 2011)). "Nototheniidae" have been shown to be paraphyletic, and the same has been suggested for Bathydraconidae, but needs to be confirmed with more extensive datasets (Rutschmann *et al.* 2011). Disregarding †*Proeleginops grandeastmanorum* (see above), no nototheniiform fossils are known. The paucity of the nototheniiform fossil record reflects the very small number of Antarctic rock outcrops, and is uninformative regarding the age of Nototheniiformes. As their preservation probability is assumed to be near zero, nototheniiform divergences are not constrained. Besides *Pseudaphritis urvillii*, all nototheniiforms are marine.

Parameters: -

Node: 95

Group 1: Pseudaphritidae
Earliest record: Extant.
Paleoenvironment: Freshwater.
Comments: -

Group 2: Eleginopidae, 'Antarctic Clade'
Earliest record: Extant.

Paleoenvironment: Marine.
Comments: -

Parameters: -

Node: 96

Group 1: Eleginopidae
Earliest record: Extant.
Paleoenvironment: Marine.
Comments: See above for reasons why †*Proeleginops grandeastmanorum* is not included here.

Group 2: 'Antarctic Clade'
Earliest record: Extant.
Paleoenvironment: Marine.
Comments: -

Parameters: -

Node: 97

Group 1: *Gobionotothen*
Earliest record: Extant.
Paleoenvironment: Marine.
Comments: The position of *Gobionotothen* within nototheniiformes of the 'Antarctic Clade' remains unresolved and disagrees between molecular phylogenies (Sanchez *et al.* 2007; Near & Cheng 2008; Matschiner *et al.* 2011; Rutschmann *et al.* 2011).

Group 2: *Trematomus*, *Lepidnotothen* (+ *Patagonotothen*), *Notothenia*, Harpagiferidae, Artedidraconidae, Bathydraconidae, Channichthyidae
Earliest record: Extant.
Paleoenvironment: Marine.
Comments: *Patagonotothen* is the sister group of *Lepidonotothen* (Sanchez *et al.* 2007; Near & Cheng 2008). The phylogenetic position of other "nototheniid" genera are here considered uncertain (Rutschmann *et al.* 2011).

Parameters: -

Node: 98

Group 1: *Trematomus*, *Lepidonotothen* (+ *Patagonotothen*)
Earliest record: Extant.
Paleoenvironment: Marine.
Comments: See above.

Group 2: *Notothenia*, Harpagiferidae, Artedidraconidae, Bathydraconidae, Channichthyidae
Earliest record: Extant.
Paleoenvironment: Marine.
Comments: -

Parameters: -

Node: 99

Group 1: *Trematomus*
Earliest record: Extant.
Paleoenvironment: Marine.
Comments: See above.

Group 2: *Lepidonotothen* (+ *Patagonotothen*)
Earliest record: Extant.
Paleoenvironment: Marine.
Comments: -

Parameters: -

Node: 100

Group 1: *Lepidonotothen larseni* (+ *L. nudifrons*)
Earliest record: Extant.
Paleoenvironment: Marine.
Comments: *Lepidonotothen nudifrons* is the sister group of *L. larseni* (Near & Cheng 2008; Rutschmann *et al.* 2011).

Group 2: *Lepidonotothen squamifrons*
Earliest record: Extant.
Paleoenvironment: Marine.

Comments: -

Parameters: -

Node: 101

Group 1: *Notothenia*
Earliest record: Extant.
Paleoenvironment: Marine.
Comments: -

Group 2: Harpagiferidae, Artedidraconidae, Bathydraconidae, Channichthyidae
Earliest record: Extant.
Paleoenvironment: Marine.
Comments: -

Parameters: -

Node: 102

Group 1: Harpagiferidae, Artedidraconidae
Earliest record: Extant.
Paleoenvironment: Marine.
Comments: -

Group 2: Bathydraconidae, Channichthyidae
Earliest record: Extant.
Paleoenvironment: Marine.
Comments: -

Parameters: -

Node: 103

Group 1: Harpagiferidae
Earliest record: Extant.
Paleoenvironment: Marine.

Comments: -

Group 2: Artedidraconidae
Earliest record: Extant.
Paleoenvironment: Marine.
Comments: -

Parameters: -

Node: 104

Group 1: Bathydraconidae
Earliest record: Extant.
Paleoenvironment: Marine.
Comments: -

Group 2: Channichthyidae
Earliest record: Extant.
Paleoenvironment: Marine.
Comments: -

Parameters: -

Node: 105

Group 1: *Champsocephalus*
Earliest record: Extant.
Paleoenvironment: Marine.
Comments: *Champsocephalus* has been found ancestral to all other channichthyids, based on molecular data (Near *et al.* 2003).

Group 2: Channichthyidae (except *Champsocephalus*)
Earliest record: Extant.
Paleoenvironment: Marine.
Comments: -

Parameters: -

Node: 106

Group 1: *Chaenocephalus*
Earliest record: Extant.
Paleoenvironment: Marine.
Comments: -

Group 2: *Chionodraco* (+ *Chaenodraco*, *Cryodraco*, *Chionobathyscus*)
Earliest record: Extant.
Paleoenvironment: Marine.
Comments: A clade combining *Chionodraco*, *Chaenodraco*, *Cryodraco*, and *Chionobathyscus* has been found in the molecular phylogeny of Near *et al.* (2003).

Parameters: -

References

Alano Perez P, Malabarba MC, Del Papa C (2010) A new genus and species of Heroini (Perciformes: Cichlidae) from the early Eocene of southern South America. *Neotrop Ichthyol*, **8**, 631–642.

Alfaro ME, Santini F, Brock CD (2007) Do reefs drive diversification in marine teleosts? Evidence from the pufferfishes and their allies (order Tetraodontiformes). *Evolution*, **61**, 2104–2126.

Alfaro ME, Santini F, Brock CD *et al.* (2009) Nine exceptional radiations plus high turnover explain species diversity in jawed vertebrates. *P Natl Acad Sci USA*, **106**, 13410–13414.

Almada F, Almada VC, Guillemaud T, Wirtz P (2005) Phylogenetic relationships of the northeastern Atlantic and Mediterranean blenniids. *Biol J Linnean Soc*, **86**, 283–295.

Anderson EM (1994) Systematics and osteology of the Zoarcidae (Teleostei: Perciformes). *Ichthyol Bull JLB Smith Inst Ichthyol*, **60**, 1–120.

Anderson ME, Fedorov VV (2004) Family Zoarcidae Swainson 1839 – Eelpouts. *Calif Acad Sci Annotated Checklists of Fishes*, **34**, 1–58.

Applegate SP, González-Rodríguez K, Alvarado-Ortega J (2000) Fish fauna of the Tlayúa quarries. In: *Paleontological highlights of the Mixteca Poblana in Central Mexico. Field Guide* (eds. Espinosa-Arrubarrena L, Montellano-Ballesteros M, Applegate SP), pp. 95–105. 60th Annual Meeting of the Soceity of Vertebrate Paleontology, Univ. Autón. Edo. Hidalgo, Spec. Pub.

Arambourg C (1954) Les poissons crétacés du Jebel Tselfat (Maroc). *Notes Mém Serv Géol Maroc*, **118**, 1–188.

Arratia G, Tischlinger H (2010) The first record of Late Jurassic crossognathiform fishes from Europe and their phylogenetic importance for teleostean phylogeny. *Foss Rec*, **13**, 317–341.

Arratia G, López-Arbarello A, Prasad GVR, Parmar V, Kriwet J (2004) Late Cretaceous-Paleocene percomorphs (Teleostei) from India – Early radiation of Perciformes. In: *Recent Advances in the Origin and Early Radiation of Vertebrates* (eds. Arratia G, Wilson MVH, Cloutier R), pp. 635–663. Verlag Dr. Friedrich Pfeil, München, Germany.

Azuma Y, Kumazawa Y, Miya M, Mabuchi K, Nishida M (2008) Mitogenomic evaluation of the historical biogeography of cichlids toward reliable dating of teleostean divergences. *BMC Evol Biol*, **8**, 215.

Balushkin AV (1994) Fossil notothenioid, and not gadiform, fish *Proeleginops grandeastmanorum* gen. sp. nov. (Perciformes, Notothenioidei, Eleginopidae) from the late Eocene found in Seymour Island (Antarctica). *Voprosy Ikhtiologii*, **34**, 298–307.

Bannikov AF (1995) Morphology and phylogeny of fossil stromateoid fishes (Perciformes). *Geobios*, **19**, 177–181.

Bannikov AF (2000) New data on Late Paleocene marine percoids (Perciformes) of Turkmenistan. *J Ichthyol*, **40**, 564–570.

Bannikov AF (2006) Fishes from the Eocene of Bolca, northern Italy, previously classified in the Sparidae, Serranidae, and Haemulidae. *Geodiversitas*, **28**, 249–275.

Bannikov AF (2009) On early Sarmatian fishes from the Eastern Paratethys. *Paleontol J*, **43**, 569–573.

Bannikov AF, Bacchia F (2005) New species of the Cenomanian Eurypterygii (Pisces, Teleostei) from Lebanon. *Paleontol J*, **39**, 514–522.

Bannikov AF, Tyler JC (2008) A new genus and species of triggerfish from the Middle Eocene of the Northern Caucasus, the earliest member of the Balistidae (Tetraodontiformes). *Paleontol J*, **42**, 615–620.

Barron JA, Lyle M, Koizumi I (2002) Late Miocene and Early Pliocene biosiliceous sedimentation along the California Margin. *Rev Mex Cienc Geol*, **19**, 161–169.

Bell MA (1977) A Late Miocene marine threespine stickleback, *Gasterosteus aculeatus aculeatus*, and its zoogeographic and evolutionary significance. *Copeia*, **1977**, 277–282.

Bell MA, Stewart JD, Park PJ (2009) The world's oldest fossil threespine stickleback fish. *Copeia*, **2009**, 256–265.

Bellwood DR (1996) The Eocene fishes of Monte Bolca: the earliest coral reef fish assemblage. *Coral Reefs*, **15**, 11–19.

Bellwood DR, Schultz O (1991) A review of the fossil record of the parrotfishes (Labroidei: Scaridae) with a description of a new *Calotomus* species from the Middle Miocene (Badenian) of Austria. *Annalen des naturhistorischen Museums in Wien*, **92**, A55–71.

Bellwood DR, Sorbini L (1996) A review of the fossil record of the Pomacentridae (Teleostei: Labroidei) with a description of a new genus and species from the Eocene of Monte Bolca, Italy. *Zool J Linn Soc*, **117**, 159–174.

Benton M, Donoghue P (2007) Paleontological evidence to date the tree of life. *Mol Biol Evol*, **24**, 26–53.

Benton MJ, Donoghue PCJ, Asher RJ (2009) Calibrating and constraining molecular clocks. In: *The Timetree of Life* (eds. Hedges SB, Kumar S), pp. 35–86. Oxford University Press, Oxford, UK.

Block BA, Finnerty JR (1994) Endothermy in fishes: a phylogenetic analysis of constraints, predispositions, and selection pressures. *Environ Biol Fish*, **40**, 283–302.

Böhme M (2007) Revision of the cyprinids from the Early Oligocene of the Ceske Stredohori Mountains, and the phylogenetic relationships of *Protothymallus* Laube 1901 (Teleostei, Cyprinidae, Gobioninae). *Acta Musei Nationalis Pragae: Series B - Historia Naturalis*, **63**, 175–194.

Burns ER, Bentley LR, Hayashi M *et al.* (2010) Hydrogeological implications of paleo-fluvial architecture for the Paskapoo Formation, SW Alberta, Canada: a stochastic analysis. *Hydrogeol J*, **18**, 1375–1390.

Carnevale G (2004) The first fossil ribbonfish (Teleostei, Lampridiformes, Trachipteridae). *Geol Mag*, **141**, 573–582.

Carnevale G, Pietsch TW (2010) Eocene handfishes from Monte Bolca, with description of a new genus and species, and a phylogeny of the family Brachionichthyidae (Teleostei: Lophiiformes). *Zool J Linn Soc*, **160**, 621–647.

Carnevale G, Tyler JC (2010) Review of the fossil pufferfish genus *Archaeotetraodon* (Teleostei, Tetraodontidae), with description of three new taxa from the Miocene of Italy. *Geobios*, **43**, 283–304.

Carnevale G, Pietsch TW, Takeuchi GT, Huddleston RW (2008) Fossil ceratioid anglerfishes (Teleostei: Lophiiformes) from the Miocene of the Los Angeles Basin, California. *J Paleont*, **82**, 996–1008.

Carney JP, Dick TA (2000) The historical ecology of yellow perch (*Perca flavenscens* [Mitchill]) and their parasites. *J Biogeogr*, **27**, 1337–1347.

Casey SP, Hall HJ, Stanley HF, Vincent ACJ (2004) The origin and evolution of seahorses (genus *Hippocampus*): a phylogenetic study using the cytochrome *b* gene of mitochondrial DNA. *Mol Phylogenet Evol*, **30**, 261–272.

Casier E (1944) Contributions à l'étude des poissons fossiles de la Belgique. V. – Les genres *Trichiurides* Winkler (s.str.) et *Eutrichiurides* nov., leurs affinités respectives. *Bulletin du Musée Royal d'Histoire Naturelle de Belgique*, **20**, 1–9.

Casier E (1966) *Faune Ichthyologique du London Clay.* British Museum (Natural History), London.

Cavender TM (1986) Review of the fossil history of North American freshwater fishes. In: *The Zoogeography of North American Freshwater Fishes* (eds. Hocutt CH, Wiley EO), pp. 699–724. John Wiley & Sons, New York.

Cavender TM (1998) Development of the North American Tertiary freshwater fish fauna with a look at parallel trends found in the European record. *Ital J Zool*, **65**, S149–161.

Chernoff CB, Orris GJ (2002) *Data set of world phosphate mines, deposits, and occurrences – Part A. Geologic data.* U.S. Geological Survey.

Cione AL (1986) Los perces continentales del Cenozoico de Argentina. *Actas del Congreso Argentino de Paleontologia y Biostratigrafia*, **4**, 101–106.

Cione AL, Báez AM (2007) Peces continentales y anfibios cenozoicos de Argentina: los últimos cincuenta años. *Ameghiniana*, **Publicación Especial 11**, 195–220.

Cohen DM, Rosenblatt RH, Moser HG (1990) Biology and description of a bythithid fish from deep-sea thermal vents in the tropical eastern Pacific. *Deep-Sea Res*, **37**, 267–283.

Costa WJEM (1998) Phylogeny and classification of the Cyprinodontiformes (Euteleostei: Atherinomorpha): a reappraisal. In: *Phylogeny and Classification of Neotropical Fishes* (eds. Malabarba LR, Reis RE, Vari RP, Lucena ZMS, Lucena CAS). EDIPUCRS, Porto Alegre, Brasil.

Cowman PF, Bellwood DR (2011) Coral reefs as drivers of cladogenesis: expanding coral reefs, cryptic extinction events, and the development of biodiversity hotspots. *J Evol Biol*, pp. no–no.

Cowman PF, Bellwood DR, Herwerden Lv (2009) Dating the evolutionary origins of wrasse lineages (Labridae) and the rise of trophic novelty on coral reefs. *Mol Phylogenet Evol*, **52**, 621–631.

Craig MT, Hastings PA (2007) A molecular phylogeny of the groupers of the subfamily Epinephelinae (Serranidae) with a revised classification of the Epinephelini. *Ichthyol Res*, **54**, 1–17.

Davis MP, Fielitz C (2010) Estimating divergence times of lizardfishes and their allies (Euteleostei: Aulopiformes) and the timing of deep-sea adaptations. *Mol Phylogenet Evol*, **57**, 1194–1208.

Dawson FM, Evans CG, Marsh R, Richardson R (1994) Uppermost Cretaceous and Tertiary strata of the Western Canada Sedimentary Basin. In: *Geological Atlas of the Western Canada Sedimentary Basin* (eds. Mossop GD, Shetsen I). Canadian Society of Petroleum Geologists and Alberta Research Council, Calgary.

De Gibert JM, Fregenal-Martinez MA, Buatois LA, Mángano MG (2000) Trace fossils and their palaeoecological significance in Lower Cretaceous lacustrine conservation deposits, El Montsec, Spain. *Palaeogeogr Palaeocl*, **156**, 89–101.

Deraco MV, Powell JE, Lopez G (2008) Primer leontínido (Mammalia, Notoungulata) de la Formación Lumbrera (Subgrupo Santa Bárbara, Grupo Salta-Paleógeno) del noroeste argentino. *Ameghiniana*, **45**, 1–13.

Dettaï A, Lecointre G (2005) Further support for the clades obtained by multiple molecular phylogenies in the acanthomorph bush. *C R Biol*, **328**, 674–689.

Dietze K (2009) Morphology and phylogenetic relationships of certain neoteleostean fishes from the Upper Cretaceous of Sendenhorst, Germany. *Cretaceous Res*, **30**, 559–574.

Do Campo M, del Papa C, Jiménez-Millán J, Nieto F (2007) Clay mineral assemblages and analcime formation in a Palaeogene fluvial-lacustrine sequence (Maíz Gordo Formation Palaeogen) from northwestern Argentina. *Sediment Geol*, **201**, 56–74.

Doiuchi R, Nakabo T (2006) Molecular phylogeny of the stromateoid fishes (Teleostei: Perciformes) inferred from mitochondrial DNA sequences and compared with morphology-based hypotheses. *Mol Phylogenet Evol*, **39**, 111–123.

Doiuchi R, Sato T, Nakabo T (2004) Phylogenetic relationships of the stromateoid fishes (Perciformes). *Ichthyol Res*, **51**.

Drummond AJ, Rambaut A (2007) BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evol Biol*, **7**, 214.

Duncker G (1912) Die Gattungen der Syngnathiden. *Mitteilungen aus dem Naturhistorischen Museum in Hamburg*, **29**, 219–240.

Duncker G (1915) Revision der Syngnathidae. *Mitteilungen aus dem Hamburgischen Zoologischen Museum und Institut*, **32**, 9–120.

Eastman JT (1993) *Antarctic Fish Biology: Evolution in a Unique Environment*. Academic Press, Inc., San Diego, CA.

Eastman JT, Grande L (1991) Late Eocene gadiform (Teleostei) skull from Seymour Island, Antarctic Peninsula. *Antarct Sci*, **3**, 87–95.

Ellison RA, Knox RWO, Jolley DW, King C (1994) A revision of the lithostratigraphical classification of the early Palaeogene strata of the London Basin and East Anglia. *P Geologist Assoc*, **105**, 187–197.

Fairon-Demaret M (2002) Fruits and seeds from the Tienen Formation at Dormaal, Palaeocene-Eocene transition in eastern Belgium. *Rev Paleobot Palyno*.

Farias IP, Ortí G, Meyer A (2000) Total evidence: molecules, morphology, and the phylogenetics of cichlid fishes. *J Exp Zool*, **288**, 76–92.

Fedotov VF, Bannikov AF (1989) On the phylogenetic relationships of fossil Gadidae. In: *Papers on the systematics of gadiform fishes* (ed. Cohen DM). Science Series. Natural History Museum of Los Angeles County, Los Angeles, CA.

Fielitz C, González-Rodríguez KA (2010) A new species of *Enchodus* (Aulopiformes: Enchodontidae) from the Cretaceous (Albian to Cenomanian) of Zimapán, Hidalgo, México. *J Vertebr Paleontol*, **30**, 1343–1351.

Fierstine HL (2006) Fossil history of billfishes (Xiphioidei). *B Mar Sci*, **79**, 433–453.

Figueiredo CA, Binga L, Castelhano J, Vining BA (2011) Exploring for gas: the future for Angola. In: *Petroleum Geology: From Mature Basins to New Frontierss – Proceedings of the 7th Petroleum Geology Conference. Geological Society, London.*, pp. 1043–1058.

de Figueiredo FJ (2009) A new clupeiform fish from the Lower Cretaceous (Barremian) of Sergipe-Alagoas Basin, Northeastern Brazil. *J Vertebr Paleontol*, **29**, 993–1005.

Filleul A, Dutheil DB (2001) *Spinocaudichthys oumtkoutensis*, a freshwater acanthomorph from the Cenomanian of Morocco. *J Vertebr Paleontol*, **21**, 774–780.

Filleul A, Maisey JG (2004) Redescription of *Santanichthys diasii* (Otophysi, Characiformes) from the Albian of the Santana Formation and comments on its implications for otophysan relationships. *Am Mus Novit*, **3455**, 21 pp.

Fink SV, Greenwood PH, Fink WL (1984) A critique of recent work on fossil ostariophysan fishes. *Copeia*, **1984**, 1033–1041.

Foote M, Raup DM (1996) Fossil preservation and the stratigraphic ranges of taxa. *Paleobiology*, **22**, 121–140.

Foote M, Hunter JP, Janis CM, Sepkoski Jr JJ (1999) Evolutionary and preservational constraints on origins of biologic groups: divergence times of eutherian mammals. *Science*, **283**, 1310–1314.

Forey PL, Yi L, Patterson C, Davies CE (2003) Fossil fishes from the Cenomanian (Upper Cretaceous) of Namoura, Lebanon. *J Syst Palaeontol*, **1**, 227–330.

Friedman M (2009) Ecomorphological selectivity among marine teleost fishes during the end-Cretaceous extinction. *P Natl Acad Sci USA*, **106**, 5218–5223.

Friedman M (2010) Explosive morphological diversification of spiny-finned teleost fishes in the aftermath of the end-Cretaceous extinction. *Proc R Soc Lond B*, **277**, 1675–1683.

Friedman M, Brazeau MD (2011) Sequences, stratigraphy and scenarios: what can we say about the fossil record of the earliest tetrapods? *Proc R Soc Lond B*, **278**, 432–439.

Gallo V, de Figueiredo FJ, Azevedo SA (2009) *Santanasalmo elegans* gen. et sp. nov., a basal euteleostean fish from the Lower Cretaceous of the Araripe Basin, northeastern Brazil. *Cretaceous Res*, **30**, 1357–1366.

Garcia MJ, Santos M, Hasui Y (2000) Palinologia da parte aflorante da Formação Entre-Córregos, Bacia de Aiuruoca, Terciário do Estado de Minas Gerais, Brasil. *Revista Universidade de Guarulhos – Geociências*, **5**, 259.

Gaudant J (2008) Paléobiodiversité et paléoenvironnements: l'exemple des gisements de poissons téléostéens du Messinien préévaporitique d'Oran et du bassin du Chélif (Algérie). *Geodiversitas*, **30**, 141–163.

Gaudant J (2009) Note complémentaire sur l'ichthyofaune oligocène de Seifhennersdorf (Saxe, Allemagne) et de Varnsdorf, Kundratice, Lbín, Skalice, Knížecí, etc. (Bohême, République tchèque). *Annalen des Naturhistorischen Museums Wien*, **111A**, 281–312.

Gayet M, Marshall LG, Sempéré T (1991) The Mesozoic and Paleocene vertebrates of Bolivia and their stratigraphic context : a review. In: *Fosiles y Facies de Bolivia – Vol. I Vertebrados* (ed. Suarez-Soruco R), pp. 393–433. Revista Technica de YPFB, Santa Cruz, Bolivia.

González-Rodríguez K, Fielitz C (2008) A new species of acanthomorph fish form the Upper Cretaceous Muhi Quarry, Hidalgo, Central Mexico. In: *Mesozoic Fishes 4 – Homology and Phylogeny*

(eds. Arratia G, Schultze HP, Wilson MVH), pp. 399–411. Verlag Dr. Friedrich Pfeil, München, Germany.

Gradstein FM, Ogg JG (2009) The geologic time scale. In: *The Timetree of Life*, pp. 26–34. Oxford University Press, Oxford, UK.

Hanel R, Westneat MW, Sturmbauer C (2002) Phylogenetic relationships, evolution of broodcare behavior, and geographic speciation in the wrasse tribe Labrini. *J Mol Evol*, **55**, 776–789.

Harmon LJ, Weir JT, Brock CD, Glor RE, Challenger W (2008) GEIGER: investigating evolutionary radiations. *Bioinformatics*, **24**, 129–131.

Harzhauser M, Piller WE (2004) Integrated stratigraphy of the Sarmatian (Upper Middle Miocene) in the western Central Paratethys. *Stratigraphy*, **1**, 65–86.

Hernández RM, Jordan TE, Dalenz Farjat A *et al.* (2005) Age, distribution, tectonics, and eustatic controls of the Paranense and Caribbean marine transgressions in southern Bolivia and Argentina. *J S Am Earth Sci*, **19**, 495–512.

Holcroft NI (2005) A molecular analysis of the interrelationships of tetraodontiform fishes (Acanthomorpha: Tetraodontiformes). *Mol Phylogenet Evol*, **34**, 525–544.

Holcroft NI, Wiley EO (2008) Acanthuroid relationships revisited: a new nuclear-based analysis that incorporates tetraodontiform representatives. *Ichthyol Res*, **55**, 274–283.

Imamura H (2000) An alternative hypothesis on the phylogenetic position of the family Dactylopteridae (Pisces: Teleostei), with a proposed new classification. *Ichthyol Res*, **47**, 203–222.

Ivany LC, Lohmann KC, Hasiuk F *et al.* (2008) Eocene climate record of a high southern latitude continental shelf: Seymour Island, Antarctica. *Geol Soc Am Bull*, **120**, 659–678.

Jacobs LL, Ferguson K, Polcyn MJ, Rennison C (2005) Cretaceous δ^{13}C stratigraphy and the age of dolichosaurs and early mosasaurs. *Neth J Geosci*, **84**, 257–268.

Johnson GD, Patterson C (1993) Percomorph phylogeny: a survey of Acanthomorphs and a new proposal. *B Mar Sci*, **52**, 554–626.

Kaufman LS, Liem KF (1982) Fishes of the suborder Labroidei (Pisces: Perciformes): phylogeny, ecology, and evolutionary significance. *Breviora*, **472**, 1–19.

Kawahara R, Miya M, Mabuchi K *et al.* (2008) Interrelationships of the 11 gasterosteiform families (sticklebacks, pipefishes, and their relatives): A new perspective based on whole mitogenome sequences from 75 higher teleosts. *Mol Phylogenet Evol*, **46**, 224–236.

Kawahara R, Miya M, Mabuchi K, Near TJ, Nishida M (2009) Stickleback phylogenies resolved: Evidence from mitochondrial genomes and 11 nuclear genes. *Mol Phylogenet Evol*, **50**, 401–404.

Kimura T, Ohana T, Naito G (1992) *Cupressionocladus* sp., newly found from the Lower Cretaceous Wakino Formation, West Japan. *Bulletin of the Kitakyushu Museum of Natural History*, **11**, 79–86.

Kotlarczyk J, Jerzmanska A, Swidnicka E (2006) A framework of ichthyofaunal ecostratigraphy of the Oligocene-Early Miocene strata of the Polish outer Carpathian basin. *Ann Soc Geol Pol*, **76**, 1–111.

Kriwet J, Hecht T (2008) A review of early gadiform evolution and diversification: first record of a rattail fish skull (Gadiformes, Macrouridae) from the Eocene of Antarctica, with otoliths preserved in situ. *Naturwissenschaften*, **95**, 899–907.

Lebedev VD (1952) Fishes from alate-paleolithic settlement at Murzak-Koba in Crimea. *Bulletin of the Society of Natural History, Moscow*, **51**, 46–51.

Li B, Dettaï A, Cruaud C *et al.* (2009) RNF213, a new nuclear marker for acanthomorph phylogeny. *Mol Phylogenet Evol*, **50**, 345–363.

Li C, Ortí G, Zhang G, Lu G (2007) A practical approach to phylogenomics: the phylogeny of ray-finned fish (Actinopterygii) as a case study. *BMC Evol Biol*, **7**, 44.

Lippitsch E, Micklich N (1998) Cichlid fish biodiversity in an Oligocene lake. *Ital J Zool*, **65**, 185–188.

Liu J, Chang MM (2009) A new Eocene catostomid (Teleostei: Cypriniformes) from northeastern China and early divergence of Catostomidae. *Sci China Ser D-Earth Sci*, **52**, 189–202.

Long DJ (1992) Sharks from the La Meseta Formation (Eocene), Seymour Island, Antarctic Peninsula. *J Vertebr Paleontol*, **12**, 11–32.

López JA, Chen WJ, Ortí G (2004) Esociform phylogeny. *Copeia*, **2004**, 449–464.

López-Fernández H, Albert JS (2011) Paleogene radiations. In: *Historical Biogeography of Neotropical Freshwater Fishes* (eds. Albert JS, Reis RE), pp. 105–118. University of California Press, London, United Kingdom.

López-Fernández H, Winemiller KO, Honeycutt RL (2010) Multilocus phylogeny and rapid radiations in Neotropical cichlid fishes (Perciformes: Cichlidae: Cichlinae). *Mol Phylogenet Evol*, **55**, 1070–1086.

Mabuchi K, Miya M, Azuma Y, Nishida M (2007) Independent evolution of the specialized pharyngeal jaw apparatus in cichlid and labrid fishes. *BMC Evol Biol*, **7**, 10–10.

Maisey JG (2000) Continental break up and the distribution of fishes of Western Gondwana during the Early Cretaceous. *Cretaceous Res*, **21**, 281–314.

Malabarba MC, Malabarba LR (2008) A new cichlid *Tremembichthys garcia* (Actinopterygii, Perciformes) from the Eocene-Oligocene of Eastern Brazil. *Rev Bras Paleontol*, **11**, 59–68.

Malabarba MC, Malabarba LR, Del Papa C (2010) *Gymnogeophagus eocenicus*, n. sp (Perciformes: Cichlidae), an Eocene cichlid from the Lumbrera formation in Argentina. *J Vertebr Paleontol*, **30**, 341–350.

Marshall CR (2008) A simple method for bracketing absolute divergence times on molecular phylogenies using multiple fossil calibration points. *Am Nat*, **171**, 726–742.

Martill DM (2007) The age of the Cretaceous Santana Formation fossil Konservat Lagerstätte of north-east Brazil: a historical review and an appraisal of the biochronostratigraphic utility of its palaeobiota. *Cretaceous Res*, **28**, 895–920.

Martini E, Reichenbacher B (2007) Nannoplankton und Fisch-Otolithen in den Mittleren Pechelbronn-Schichten (Unter-Oligozän, Oberrheingraben/Mainzer Becken). *Geologische Abhandlungen Hessen*, pp. 1–40.

Matschiner M, Hanel R, Salzburger W (2011) On the origin and trigger of the notothenioid adaptive radiation. *PLoS ONE*, **6**, e18911.

Matsumoto T, Obata I, Tashiro M *et al.* (1982) Correlation of marine and non-marine formation in the Cretaceous of Japan. *Fossils*, **31**, 1–26.

McCrory VLC, Walker RG (1986) A storm and tidally-influenced prograding shoreline-Upper Cretaceous Milk River Formation of Southern Alberta, Canada. *Sedimentology*, **33**, 47–60.

McDowall RM, Kennedy EM, Lindqvist JK *et al.* (2006) Probable *Gobiomorphus* fossils from the Miocene and Pleistocene of New Zealand (Teleostei: Eleotridae). *J Roy Soc New Zeal*, **36**, 97–109.

Mecklenburg CW (2003) Family Zaproridae Jordan & Evermann 1898 – Prowfishes. *Calif Acad Sci Annotated Checklists of Fishes*, **13**, 1–3.

Mein P, Meon H, Romaggi JP, Samuel E (1983) La vie ein Ardeche au Miocene superieur d'apres les document troues dans las carriere de la Montagne d'Andance. *Nouvelles Archives du Musée d'Histoire Naturelle de Lyon*, **Suppl. 21**, 37–44.

Merle D (2005) *Nucellopsis* (Gastropoda, Muricidae), a new genus from the paralic domain of the European Palaeogene. *C R Palevol*, **4**, 177–189.

Miya M, Takeshima H, Endo H *et al.* (2003) Major patterns of higher teleostean phylogenies: a new perspective based on 100 complete mitochondrial DNA sequences. *Mol Phylogenet Evol*, **26**, 121–138.

Miya M, Satoh TP, Nishida M (2005) The phylogenetic position of toadfishes (order Batrachoidiformes) in the higher ray-finned fish as inferred from partitioned Bayesian analysis of 102 whole mitochondrial genome sequences. *Biol J Linnean Soc*, **85**, 289–306.

Miya M, Pietsch TW, Orr JW *et al.* (2010) Evolutionary history of anglerfishes (Teleostei: Lophiiformes): a mitogenomic perspective. *BMC Evol Biol*, **10**, 58.

Mlíkovsky J (1996) Tertiary avian localities of the United Kingdom. *Acta U Carol Geol*, **39**, 759–771.

Monsch KA (2004) Revision of the scombroid fishes from the Cenozoic of England. *T Roy Soc Edin Earth*, **95**, 445–489.

Monsch KA (2006) A revision of scombrid fishes (Scombroidei, Perciformes) from the Middle Eocene of Monte Bolca, Italy. *Palaeontology*, **49**, 873–888.

Murray AM (1996) A new Paleocene genus and species of percopsid, *Massamorichthys wilsoni* (Paracanthopterygii) from Joffre Bridge, Alberta, Canada. *J Vertebr Paleontol*, **16**, 642–652.

Murray AM (2000a) Eocene cichlid fishes from Tanzania, East Africa. *J Vertebr Paleontol*, **20**, 651–664.

Murray AM (2000b) The Palaeozoic, Mesozoic and Early Cenozoic fishes of Africa. *Fish Fish*, **1**, 111–145.

Murray AM (2001a) The fossil record and biogeography of the Cichlidae (Actinopterygii: Labroidei). *Biol J Linnean Soc*, **74**, 517–532.

Murray AM (2001b) The oldest fossil cichlids (Teleostei: Perciformes): indication of a 45 million-year-old species flock. *Proc R Soc Lond B*, **268**, 679–684.

Murray AM, Stewart KM (1999) A new species of tilapiine cichlid from the Pliocene, Middle Awash, Ethiopia. *J Vertebr Paleontol*, **19**, 293–301.

Murray AM, Wilson MVH (1999) Contributions of fossils to the phylogenetic relationships of the percopsiform fishes (Teleostei: Paracanthopterygii): order restored. In: *Mesozoic Fishes 2 – Systematics and Fossil Record* (eds. Arratia G, Schultze HP), pp. 397–411. Verlag Dr. Friedrich Pfeil, München, Germany.

Murray AM, Cumbaa SL, Harington CR, Smith GR, Rybczynski N (2009) Early Pliocene fish remains from Arctic Canada support a pre-Pleistocene dispersal of percids (Teleostei: Perciformes). *Can J Earth Sci*, **46**, 557–570.

Muzylev NG (1994) Anoxic events of the Paleocene-Middle Eocene. In: *Ecosystem restructuring and the evolution of the biosphere* (eds. Rozanov AY, Semikhatov MA), pp. 160–166. Nedra, Moscow, Russia.

Natale M (2008) *Il gasterosteiforme cretaceo* Gasterorhamphosus zuppichinii *Sorbini, 1981: analisi morfoanatomica e considerazioni sistematiche.* Ph.D. thesis, Universita' di Pisa, Pisa.

Nazarkin MV (1998) New stichaeid fishes (Stichaeidae, Perciformes) from Miocene of Sakhalin. *J Ichthyol*, **38**, 293–306.

Nazarkin MV (2002) Gunnels (Perciformes, Pholidae) from the Miocene of Sakhaline Island. *J Ichthyol*, **42**, 279–288.

Near TJ (2004) Estimating divergence times of notothenioid fishes using a fossil-calibrated molecular clock. *Antarct Sci*, **16**, 37–44.

Near TJ, Cheng CHC (2008) Phylogenetics of notothenioid fishes (Teleostei: Acanthomorpha): Inferences from mitochondrial and nuclear gene sequences. *Mol Phylogenet Evol*, **47**, 1–9.

Near TJ, Pesavento JJ, Cheng CHC (2003) Mitochondrial DNA, morphology, and the phylogenetic relationships of Antarctic icefishes (Notothenioidei: Channichthyidae). *Mol Phylogenet Evol*, **28**, 87–98.

Nelson JS (2006) *Fishes of the World*. John Wiley & Sons, Inc., Hoboken, New Jersey.

Nielsen JG, Cohen DM, Markle DF, Robins CR (1999) Ophidiiform fishes of the world (order Ophidiiformes). An annotated and illustrated catalogue of pearlfishes, cusk-eels, brotulas and other ophidiiform fishes known to date. *FAO Fisheries Synopsis*, **125**, 1–179.

Nolf D, Rana RS, Singh H (2006) Fish otoliths from the Ypresian (early Eocene) of Vastan, Gujarat, India. *Bulletin de l'Institut Royal des Sciences Naturelles de Belgique, Sciences de la Terre*, **76**, 105–118.

Ogg JG, Ogg G (2008a) Late Cretaceous (65-100 Ma time-slice). International Commission on Stratigraphy. https://engineering.purdue.edu/Stratigraphy/charts/Timeslices/3_Late_Cret.pdf.

Ogg JG, Ogg G (2008b) Neogene-Late Oligocene (0-33 Ma time-slice). International Commission on Stratigraphy. https://engineering.purdue.edu/Stratigraphy/charts/Timeslices/1_Neogene.pdf.

Ogg JG, Ogg G (2008c) Paleogene (33-66 Ma time-slice). International Commission on Stratigraphy. https://engineering.purdue.edu/Stratigraphy/charts/Timeslices/2_Paleogene.pdf.

Orr JW (1995) *Phylogenetic relationships of gasterosteiform fishes (Teleostei: Acanthomorpha)*. Ph.D. thesis, University of Washington.

Orrell TM, Collette BB, Johnson GD (2006) Molecular data support separate scombroid and xiphioid clades. *B Mar Sci*, **79**, 505–519.

Ortí G, Li C (2009) Phylogeny and classification. In: *Reproductive Biology and Phylogeny of Fish (Agnatha and Osteichthyes)* (ed. Jamieson BMG), pp. 1–24. Science Publishers, Enfield, New Hampshire.

del Papa C, Kirschbaum A, Powell J *et al.* (2010) Sedimentological, geochemical and paleontological insights applied to continental omission surfaces: A new approach for reconstructing an Eocene foreland basin in NW Argentina. *J S Am Earth Sci*, **29**, 327–345.

Pascual R, Bond M, Vucetich MG (1981) El Subgrupo Santa Bárbara (Grupo Salta) y sus vertebrados. Cronología, paleoambientes y paleobiogeografía. In: *Actas del VIII Congreso Geológico Argentino*, pp. 743–758. San Luis, Argentina.

Patterson C (1964) A review of Mesozoic acanthopterygian fishes, with special reference to those of the English Chalk. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, **247**, 213–482.

Patterson C (1993a) An overview of the early fossil record of acanthomorphs. *Bull Mar Sci*, **52**, 29–59.

Patterson C (1993b) Osteichthyes: Teleostei. In: *The fossil record 2*, pp. 621–656. Chapman & Hall, London, UK.

Peng Z, He S, Wang J, Wang W, Diogo R (2006) Mitochondrial molecular clocks and the origin of the major otocephalan clades (Pisces: Teleostei): a new insight. *Gene*, **370**, 113–124.

Pharisat A (1998) Marine fish biodiversity in the Oligocene of Froidefontaine (Belfort Territory, France). *Ital J Zool*, **65**, 189–191.

Pondella II DJ, Craig MT, Franck JPC (2003) The phylogeny of Paralabrax (Perciformes: Serranidae) and allied taxa inferred from partial 16S and 12S mitochondrial ribosomal DNA sequences. *Mol Phylogenet Evol*, **29**, 176–184.

Popov SV, Sychevskaya EK, Akhmet'ev MA, Zaporozhets NI, Golovina LA (2008) Stratigraphy of the Maikop Group and Pteropoda Beds in northern Azerbaijan. *Stratigr Geol Correl*, **16**, 664–677.

Poyato-Ariza FJ (1996) A revision of Rubiesichthys gregalis WENZ 1984 (Ostariophysi, Gonorynchiformes), from the Early Cretaceous of Spain. In: *Mesozoic Fishes – Systematics and Paleoecology* (eds. Arratia G, Viohl G), pp. 319–328. Verlag Dr. Friedrich Pfeil, München, Germany.

Prokofiev AM (2006) Fossil myctophoid fishes (Myctophiformes: Myctophoidei) from Russia and adjacent regions. *J Ichthyol*, **46**, S38–S83.

Prokofiev AM (2009) Systematics of Oligocene percoids classified as "*Serranus budensis*", with the description of new taxa (in Russian). *Current Problems of Modern Science*, **2**, 199–222.

Regan CT (1911) The anatomy and classification of the Teleostean fishes of the orders Berycomorphi and Xenoberyces. *Ann Mag Nat Hist*, **8**, 1–9.

Rossi CMR, Gosztonyi AE, Cozzuol MA (2000) A Miocene cusk-eel (Ophidiiformes: Ophidiidae) from the Península Valdés, Argentina. *J Vertebr Paleontol*, **20**, 645–650.

Rutschmann S, Matschiner M, Damerau M *et al.* (2011) Parallel ecological diversification in Antarctic notothenioid fishes as evidence for adaptive radiation. *Mol Ecol*.

Salzburger W, Mack T, Verheyen E, Meyer A (2005) Out of Tanganyika: Genesis, explosive speciation, key-innovations and phylogeography of the haplochromine cichlid fishes. *BMC Evol Biol*, **5**, 17.

Sanchez S, Dettaï A, Bonillo C *et al.* (2007) Molecular and morphological phylogenies of the Antarctic teleostean family Nototheniidae, with emphasis on the Trematominae. *Polar Biol*, **30**, 155–166.

Santini F, Tyler JC (2003) A phylogeny of the families of fossil and extant tetraodontiform fishes (Acanthomorpha, Tetraodontiformes), Upper Cretaceous to Recent. *Zool J Linn Soc*, **139**, 565–617.

Santini F, Harmon LJ, Carnevale G, Alfaro ME (2009) Did genome duplication drive the origin of teleosts? A comparative study of diversification in ray-finned fishes. *BMC Evol Biol*, **9**, 194.

Schnyder J, Dejax J, Keppens E *et al.* (2009) An Early Cretaceous lacustrine record: Organic matter and organic carbon isotopes at Bernissart (Mons Basin, Belgium). *Palaeogeogr Palaeocl*, **281**, 79–91.

Schultz O (1993) Der Nachweis von *Scorpaena* s.s. (Pisces, Teleostei) im Badenien von St. Margarethen, Burgenland, Österreich. *Ann Naturhist Mus Wien*, **95**, 127–177.

Schultz O (2000) Ein Zackenbarsch (*Epinephelus*, Serranidae, Pisces) aus dem Mittel-Miozän von Retznei, Steiermark. *Joannea Geol Paläont*, **2**, 5–56.

Schwarzer J, Misof B, Tautz D, Schliewen UK (2009) The root of the East African cichlid radiations. *BMC Evol Biol*, **9**, 186.

Schwarzhans W (2003) Fish otoliths from the Paleocene of Denmark. *Geol Surv Den Greenl*, **2**, 1–94.

Schwarzhans W, Bratishko A (2011) The otoliths from the middle Paleocene of Luzanivka (Cherkasy district, Ukraine). *N Jb Geol Pal A*, **261**, 83–110.

Scott RW, Oboh-Ikuenobe FE, Benson Jr DG, Holbrook JM (2009) Numerical age calibration of the Albian/Cenomanian boundary. *Stratigraphy*, **6**, 17–32.

Setiamarga DHE, Miya M, Yamanoue Y *et al.* (2008) Interrelationships of Atherinomorpha (medakas, flyingfishes, killifishes, silversides, and their relatives): the first evidence based on whole mitogenome sequences. *Mol Phylogenet Evol*, **49**, 598–605.

Sheehan PM, Fastovsky DE (1992) Major extinctions of land-dwelling vertebrates at the Cretaceous-Tertiary boundary, eastern Montana. *Geology*, **20**, 556–560.

Simandl GJ, Hancock KD (1991) Geology of dolomite-hosted magnesite deposits of the Brisco and Driftwood Creek Areas, British Columbia. *Geological Fieldwork*, **1992**, 461–478.

Skutschas P, Gubin Y (2011) A new salamander from the late Paleocene–early Eocene of Ukraine. *Acta Palaeontol Pol*, p. in press.

Sloss BL, Billington N, Burr BM (2004) A molecular phylogeny of the Percidae (Teleostei, Perciformes) based on mitochondrial DNA sequence. *Mol Phylogenet Evol*, **32**, 545–562.

Smith WL, Craig MT (2007) Casting the percomorph net widely: the importance of broad taxonomic sampling in the search for the placement of serranid and percid fishes. *Copeia*, **2007**, 35–55.

Smith WL, Wheeler WC (2004) Polyphyly of the mail-cheeked fishes (Teleostei: Scorpaeniformes): evidence from mitochondrial and nuclear sequence data. *Mol Phylogenet Evol*, **32**, 627–646.

Song CB, Near TJ, Page LM (1998) Phylogenetic relations among percid fishes as inferred from mitochondrial cytochrome *b* DNA sequence data. *Mol Phylogenet Evol*, **10**, 343–353.

Sorbini C, Tyler JC (2004) Review of the fossil file fishes of the family Monacathidae (Tetraodontiformes), Pliocene and Pleistocene of Europe, with a new genus, *Frigocanthus*, and two new species related to the recent *Aluterus*. *Bollettino del Museo Civico di Storia Naturale di Verona, (Geologia, Paleontologia e Preistoria)*, **28**, 41–76.

Sota T, Mukai T, Shinozaki T, Sato H, Yodoe K (2005) Genetic differentiation of the gobies *Gymnogobius castaneus* and *G. taranetzi* in the region surrounding the sea of Japan as inferred from a mitochondrial gene genealogy. *Zool Sci*, **22**, 87–93.

Sparks JS, Smith WL (2004) Phylogeny and biogeography of cichlid fishes (Teleostei: Perciformes: Cichlidae). *Cladistics*, **20**, 501–517.

Standke G, Blumenstengel H, von Bülow W (2005) Das Tertiär Ostdeutschlands in der Stratigraphischen Tabelle von Deutschland 2002. *Newsl Stratigr*, **41**, 323–338.

Stevenson DE (2002) Systematics and distribution of fishes of the Asian goby genera *Chaenogobius* and *Gymnogobius* (Osteichthyes: Perciformes: Gobiidae), with the description of a new species. *Species Diversity*, **7**, 151–312.

Stewart JD (1996) Cretaceous acanthomorphs of North America. In: *Mesozoic Fishes – Systematics and Paleoecology* (eds. Arratia G, Viohl G), pp. 383–394. Verlag Dr. Friedrich Pfeil, München, Germany.

Stiassny MLJ (1993) What are grey mullets? *B Mar Sci*, **52**, 197–219.

Sytchevskaya EK, Prokofiev AM (2005) The first occurrence of fishes of the family Centrolophidae (Perciformes, Sromateoidei) in the Late Paleocene of Turkmenistan. *J Ichthyol*, **45**, 134–137.

Taverne L (1982) Sur *Pattersonella formosa* (Traquair R.H. 1911) et *Nybelinoides brevis* (Traquair R.H. 1911), téléostéens salmoniformes argentinoides du Wealdien inférieur de Bernissart, Belgique, précédemment attribués au genre *Leptolepis* Agassiz L. 1832. *Bulletin de l'Institut Royal des Sciences Naturelles de Belgique, Sciences de la Terre*, **54**, 1–27.

Taverne L, Filleul A (2003) Osteology and relationships of the genus *Spaniodon* (Teleostei, Salmoniformes) from the Santonian (Upper Cretaceous) of Lebanon. *Palaeontology*, **46**, 927–944.

Trevisani E, Papazzoni CA, Ragazzi E, Roghi G (2005) Early Eocene amber from the "Pesciara di Bolca" (Lessini Mountains, Northern Italy). *Palaeogeogr Palaeocl*, **223**, 260–274.

Tyler JC (2000) *Arambourgthurus*, a new genus of hypurostegic surgeonfish (Acanthuridae) from the Oligocene of Iran, with a phylogeny of the Nasinae. *Geodiversitas*, **22**, 525–537.

Tyler JC, Bannikov AF (1992) New genus of primitive ocean sunfish with separate premaxillae from the Eocene of Southwest Russia (Molidae, Tetraodontiformes). *Copeia*, **1992**, 1014–1023.

Tyler JC, Bannikov AF (1997) Relationships of the fossil and recent genera of rabbitfishes (Acanthuroidei: Siganidae). *Sm C Paleob*, **84**, 1–44.

Tyler JC, Bannikov AF (2009) Phylogenetic implications of the some cranial features of the porcupine pufferfish *Pshekhadiodon* (Tetraodontiformes, Diodontidae) from the Eocene of the Northern Caucasus. *J Ichthyol*, **49**, 703–709.

Tyler JC, Santini F (2005) A phylogeny of the fossil and extant zeiform-like fishes, Upper Cretaceous to Recent, with comments on the putative zeomorph clade (Acanthomorpha). *Zool Scripta*, **34**, 157–175.

Tyler JC, Sorbini L (1996) New superfamily and three new families of tetraodontiform fishes from the Upper Cretaceous: the earliest and most morphologically primitive plectognaths. *Sm C Paleob*, **82**, 1–62.

Tyler JC, O'Toole B, Winterbottom R (2003) Phylogeny of the genera and families of zeiform fishes, with comments on their relationships with tetraodontiforms and caproids. *Sm C Zool*, **618**, 1–120.

Uyeno T, Yabumoto Y (1980) Early Cretaceous freshwater fishes from Northern Kyushu, Japan: II. Restoration of two species of the clupeoid fish genus *Diplomystus*. *Bulletin of the Kitakyushu Museum of Natural History*, **2**, 25–31.

Vucetich MG, Reguero MA, Bond M *et al.* (2007) Mamíferos continentales del Paleógeno argentino:las investigaciones de los últimos cincuenta años. *Ameghiniana*, **Publicación Especial 11**, 239–255.

Wall PD, Ivany LC, Wilkinson BH (2009) Revisiting Raup: exploring the influence of outcrop area on diversity in light of modern sample-standardization techniques. *Paleobiology*, **35**, 146–167.

Weiler W (1928) Beiträge zur Kenntnis der tertiären Fische des Mainzer Beckens II. 3. Teil. Die Fische des Septarientones. *Abh Hess Geol Landesanstalt*, **8**, 63.

Weir AH, Catt JA (1969) The mineralogy of palaeogene sediments in Northeast Kent (Great Britain). *Sediment Geol*, **3**, 17–33.

Wenz S, Berner P, Bourseau JP *et al.* (1993) L'ichthyofaune des calcaires lithographiques du Kimméridgien supérieur de Cerin (Ain, France). *Geobios*, **27**, 61–70.

Westneat MW, Alfaro ME (2005) Phylogenetic relationships and evolutionary history of the reef fish family Labridae. *Mol Phylogenet Evol*, **36**, 370–390.

Wiley EO, Johnson GD (2010) A teleost classification based on monophyletic groups. In: *Origin and Phylogenetic Interrelationships of Teleosts*, pp. 123–182. Verlag Dr. Friedrich Pfeil, München, Germany.

Willumsen PS (2004) Palynology of the Lower Eocene deposits of northwest Jutland, Denmark. *B Geol Soc Denmark*, **52**, 141–157.

Wilson AB, Orr JW (2011) The evolutionary origins of Syngnathidae: pipefishes and seahorses. *J Fish Biol*, **78**, 1603–1623.

Wilson MVH, Li GQ (1999) Osteology and systematic position of the Eocene salmonid †*Eosalmo driftwoodensis* Wilson from western North America. *Zool J Linn Soc*, **125**, 279–311.

Wilson MVH, Murray AM (1996) Early Cenomanian acanthomorph teleost in the Cretaceous Fish Scale Zone, Albian/Cenomanian boundary, Alberta, Canada. In: *Mesozoic Fishes – Systematics and Paleoecology* (eds. Arratia G, Viohl G), pp. 369–382. Verlag Dr. Friedrich Pfeil, München, Germany.

Wippich MGE, Lehmann J (2004) *Allocrioceras* from the Cenomanian (Mid-Cretaceous) of the Lebanon and its bearing on the paleobiological interpretation of heteromorphic ammonites. *Palaeontology*, **47**, 1093–1107.

Woodward AS (1901) *Catalogue of the Fossil Fishes in the British Museum (Natural History), Cromwell Road, S.W*, vol. 4. British Museum, Natural History, London.

Woodward AS (1902) *The fossil fish of the English Chalk.* Mon Palaeont Soc, London, UK.

Yamanoue Y, Miya M, Matsuura K *et al.* (2008) A new perspective on phylogeny and evolution of tetraodontiform fishes (Pisces: Acanthopterygii) based on whole mitochondrial genome sequences: Basal ecological diversification? *BMC Evol Biol*, **8**, 212.

Yamanoue Y, Miya M, Doi H *et al.* (2011) Multiple invasions into freshwater by pufferfishes (Teleostei: Tetraodontidae): A mitogenomic perspective. *PLoS ONE*, **6**, e17410.

Žalohar J, Hitij T, Križnar M (2009) Two new species of seahorses (Syngnathidae, *Hippocampus*) from the Middle Miocene (Sarmatian) Coprolitic Horizon in Tunjice Hills, Slovenia: The oldest fossil record of seahorses. *Annales de Paléontologie*, **95**, 71–96.

Package 'ageprior'

October 19, 2011

Type Package

Title Prior distributions for molecular dating

Version 0.9.2

Date 2011-10-10

Author Michael Matschiner <michaelmatschiner@mac.com>

Maintainer Michael Matschiner <michaelmatschiner@mac.com>

Description Calculation of the parameters of Bayesian prior distributions for divergence dates based on fossil calibrations

License GPL (>= 2)

LazyLoad yes

R topics documented:

1

---

adjust.pres.rate   *Preservation rates with rock outcrop bias*

Description

Adjusts the preservation rate over a given period to the rock outcrop bias

Usage

```
adjust.pres.rate(r,ty,t1,t2,corr,plot)
```

Arguments

r	Preservation rate.
ty	The younger age boundary of the geological formation in which the oldest fossil constraining the node has been found (in Ma). This will be used as the prior offset.
t1	Younger end of the range for which the preservation rate has been calculated.
t2	Older end of the range for which the preservation rate has been calculated.
corr	The applied correction for outcrop bias. Options include: "global_marine" (Wall et al. 2009), "global_terrestrial" (Wall et al. 2009), "global" (Wall et al. 2009), "NorthAmerica_marine" (Wall et al. 2009), "NorthAmerica_terrestrial" (Wall et al. 2009), "NorthAmerica" (Wall et al. 2009), "WesternEurope_marine" (Smith & McGowan 2007), "WesternEurope_terrestrial" (Smith & McGowan 2007), "WesternEurope" (Smith & McGowan 2007), "Australia_marine" (McGowan & Smith 2008), "Australia_terrestrial" (McGowan & Smith 2008), and "Australia" (McGowan & Smith 2008). When using the adjusted preservation rate in function find.prior, the same correction must be applied there.
plot	This optional parameter allows to suppress graphical output with plot=FALSE. The default setting is plot=TRUE.

Details

The function find.prior allows for adjustment of prior distributions to bias in the fossil record that results from the heterogeneous availability of sedimentary rock outcrops per interval. If a correction is applied in find.prior, the preservation rate should also be adjusted to any rock outcrop bias at the time of the clade-constraining fossil. For example, if a preservation rate of 0.03 has been calculated for Cenozoic mammals, then this specifies the average rate between 65.5-0 Ma. If the oldest fossil of a given mammal clade dates to the Eocene (56-34 Ma), then find.prior would calculate an lognormal, gamma, or exponential prior distribution with offset ty=34 Ma, based on a parameter combination including the parameters r (preservation rate) and p_q (p-q, the net diversification rate). If however, the preservation rate at ty=34 Ma can be expected to be higher or lower than the average preservation rate (because higher- or lower-than- average amounts of rock outcrops are known of this age), then the preservation rate used to define the resulting prior distribution should be adjusted. Say, there

is twice as much rock outcrop of an age of ty=34 Ma compared to the average outcrop amount for the whole Cenozoic, then the preservation rate used to define any prior with offset ty=34 should be twice the overall Cenozoic preservation rate. In the example given above, it should be set to 0.06. The function adjust.pres.rate facilitates this calculation. Given parameters r (preservation rate - 0.03 in the above example), ty (prior offset / youngest possible age of the clade's oldest fossil, 34 Ma in the example), t1 (younger end of the range for which the preservation rate has been calculated - 0 Ma in the example of Cenozoic mammals, as the Cenozoic extends to the present), t2 (older end of the range for which the preservation rate has been calculated - 65.5 Ma, the begin of the Cenozoic in the example), and corr (the applied correction for outcrop bias, in the example, it could be "global_terrestrial"), the function adjust.pres.rate calculates a corrected preservation rate that should then be used as a parameter in the function `find.prior`.

Author(s)

Michael Matschiner

Maintainer: Michael Matschiner <michaelmatschiner@mac.com>

Source

Matschiner M (2011) in prep.

References

McGowan AJ, Smith AB (2008) *Paleobiology* 34:80-103.
Smith AB, McGowan AJ (2007) *Palaeontology* 50:765-774.
Wall PD, Ivany LC, Wilkinson BH (2009) *Paleobiology* 35:146-167.

See Also

`find.prior`

Examples

```
adjust.pres.rate(r=0.02,ty=34,t1=0,t2=65.5,corr="global_terrestrial")
```

---

ageprior	*Prior distributions for molecular dating*

Description

Calculation of the parameters of Bayesian prior distributions for divergence dates based on fossil calibrations. Necessary parameters are preservation rate, net diversification rate, and the age of the oldest fossil of the investigated clade. The preservation rate can be calculated from the fossil record (see Foote 1997, Foote et al. 1999), and should be adjusted with `adjust.pres.rate` if a correction for rock outcrop bias (option "corr" of function `find.prior`) is used. The net diversification rate can be estimated from the number of extant taxa and a rough age estimate for the clade (assuming an exponential diversification model with or without saturation) using `net.div.rate`, or with more sophisticated models implemented in MEDUSA (Alfaro et al. 2009; part of the R package Geiger). Once these parameter estimates are at hand, they can be used to calculate an age prior distribution with `find.prior`.

Details

Package:	ageprior
Type:	Package
Version:	0.92
Date:	2011-10-10
License:	GPL >= 2
LazyLoad:	yes

Author(s)

Michael Matschiner

Maintainer: Michael Matschiner <michaelmatschiner@mac.com>

Source

Matschiner M (2011) in prep.

References

Alfaro ME, Santini F, Brock CD, et al. (2009) *Proc Natl Acad Sci* 106:13410-13414.
Foote M (1997) *Paleobiology* 23:278-300.
Foote M, Hunter JP, Janis CM, Sepkoski Jr J (1999) *Science* 283:1310-1314.

See Also

find.prior adjust.pres.rate net.div.rate

Australia   *Australian rock outcrop bias*

Description

Quantification of the Australian rock outcrop of the Phanerozoic

Usage

```
data(Australia_marine)
data(Australia_terrestrial)
data(Australia)
```

Format

Data frames with 72 values for preservation rate bias (relr) and the younger end of the time interval (py) for which this preservation rate has been calculated.

Details

Australian marine and terrestrial rock outcrops of the Phanerozoic have been quantified by Mc-
Gowan & Smith (2008) as the number of geological maps of Australia containing this type of
rock outcrop of a given time interval (using bins with an average size of 7.5 myr). This quantifi-
cation is here translated into maps per myr and is scaled so that the maximum number of maps
per myr equals 1. This data serves as a proxy for preservation rate bias. It is used by functions ad-
just.pres.rate and find.prior when called with corr="Australia_marine", corr="Australia_terrestrial",
or corr="Australia" (which combines both the marine and terrestrial rock outcrop).

Author(s)

Michael Matschiner

Maintainer: Michael Matschiner <michaelmatschiner@mac.com>

Source

Matschiner M (2011) in prep.

References

McGowan AJ, Smith AB (2008) *Paleobiology* 34:80-103.

---

find.prior	*Prior distributions for molecular dating*

Description

Calculation of the parameters of Bayesian prior distributions for divergence dates based on fossil
calibrations

Usage

```
find.prior(id,ty,to,r,p_q,lag,corr,plot)
```

Arguments

`id`	A name for the node. If this matches the name of a taxonomic group defined in BEAUTi, then the output of this script can be integrated directly into the BEAST XML file to constrain the tmrca of this group.
`ty`	The younger age boundary of the geological formation in which the oldest fossil constraining the node has been found (in Ma). This will be used as the prior offset.
`to`	The older age boundary of the geological formation in which the oldest fossil constraining the node has been found (in Ma).
`r`	Preservation rate (see Foote 1997, Foote et al. 1999). If a correction for outcrop bias is applied (see below), the preservation rate may need to be adjusted with `adjust.pres.rate` before running find.prior.

p_q	An estimate for the net diversification rate p-q (origination rate - extinction rate) of the investigated clade. This can be calculated from the number of extant species, an age estimate, and with assumption of an exponential diversification model (with or without saturation) using the function `net.div.rate`.
lag	An optional lag time parameter to account for the facts that directly after speciation fossil preservation is less likely due to potentially smaller population sizes and geographic ranges, and for the decreased probability of correct taxonomic assignment of fossils that preserved shortly after the origin of a given clade. The lag parameter specifies the mean of an exponential distribution for the delay between clade origin and the point in time, from which on the above effects become negligible with respect to the preservation probability. These effects can be ignored completely by specifying lag=0 (the default).
corr	Allows an optional correction for bias in the fossil record that results from the heterogeneous availability of sedimentary rock outcrops per interval. Options include: "global_marine" (Wall et al. 2009), "global_terrestrial" (Wall et al. 2009), "global" (Wall et al. 2009), "NorthAmerica_marine" (Wall et al. 2009), "NorthAmerica_terrestrial" (Wall et al. 2009), "NorthAmerica" (Wall et al. 2009), "WesternEurope_marine" (Smith & McGowan 2007), "WesternEurope_terrestrial" (Smith & McGowan 2007), "WesternEurope" (Smith & McGowan 2007), "Australia_marine" (McGowan & Smith 2008), "Australia_terrestrial" (McGowan & Smith 2008), "Australia" (McGowan & Smith 2008), and "none". Unless this is set to "none", care should be taken that the preservation rate is adjusted to rock outcrop bias at the youngest possible age of the oldest fossil of a given clade. This adjustment is performed by function `adjust.pres.rate`.
plot	This optional parameter allows to suppress graphical output with plot=FALSE. The default setting is to plot the calculated probability distribution for linage nonpreservation, and the best log-normal, gamma, or exponential distribution approximation.

Details

This function finds the best log-normal, gamma, or exponential approximation to probabilistic bounds on lineage nonpreservation according to Foote et al. (1999). Offset and mean of the log-normal, gamma, or exponential distribution are fixed to match those of the theoretical distribution Rt (Matschiner 2011), and so is the offset. As a measure of fit between distributions, the root mean square deviation is calculated to decide which type of distribution provides the best approximation. For log-normal distributions parameter sigma is chosen to minimize the root mean square deviation, and the same is performed for the shape parameter of gamma distributions. The three types of distributions are readily implemented into divergence date estimation with the software BEAST (Drummond & Rambaut 2007), and output is written that can be inserted into BEAST XML files to apply the chosen prior distributions.

Author(s)

Michael Matschiner

Maintainer: Michael Matschiner <michaelmatschiner@mac.com>

Source

Matschiner M (2011) in prep.

References

Drummond AJ, Rambaut A (2007) *BMC Evol Biol* 7:214.
Foote M (1997) *Paleobiology* 23:278-300.
Foote M, Hunter JP, Janis CM, Sepkoski Jr J (1999) *Science* 283:1310-1314.
McGowan AJ, Smith AB (2008) *Paleobiology* 34:80-103.
Smith AB, McGowan AJ (2007) *Palaeontology* 50:765-774.
Wall PD, Ivany LC, Wilkinson BH (2009) *Paleobiology* 35:146-167.

See Also

adjust.pres.rate net.div.rate

Examples

```
find.prior(id="NodeA",ty=10,to=20,r=0.02,p_q=0.02,corr="global_marine")
find.prior(id="NodeB",ty=65.5,to=70.6,r=0.02,p_q=0.01,lag=2)
find.prior(id="NodeC",ty=145.5,to=150.8,r=0.02,p_q=0.04)
```

---

global	*Global rock outcrop bias*

Description

Quantification of the global rock outcrop of the Phanerozoic

Usage

```
data(global_marine)
data(global_terrestrial)
```

Format

Data frames with 23 values for preservation rate bias (relr) and the younger end of the time interval (py) for which this preservation rate has been calculated.

Details

Global marine and terrestrial rock outcrops of the Phanerozoic have been quantified by Wall et al. (2009) by combination of data on global outcrop areas from the UNESCO *Geological Atlas of the World* (Choubert & Faure-Muret 1976) with volumetric data on global sediment compositions (Ronov 1980, 1994). Bin size for the 22 used phanerozoic time intervals varies between 6 (Early Triassic) and 54 (Cambrian) myr, with an average length of 24.4 myr. The 23rd value results from an extrapolation of Miocene rock outcrop to the present. Rock outcrop quantification is here translated into area per myr and is scaled so that the maximum area per myr equals 1. This data serves as a proxy for preservation rate bias. It is used by functions adjust.pres.rate and find.prior when called with corr="global_marine", corr="global_terrestrial", or corr="global" (which combines both the marine and terrestrial rock outcrop).

Author(s)

Michael Matschiner

Maintainer: Michael Matschiner <michaelmatschiner@mac.com>

Source

Matschiner M (2011) in prep.

References

Choubert G, Faure-Muret A (1976) *Geological atlas of the world. 1:10,000,000. 22 sheets with explanations*. UNESCO Commission for the Geological Map of the World, Paris.

Ronov AB (1980) The earth's sedimentary shell: quantitative patterns of its structures, compositions and evolution (the 20th V. I. Vernadskiy Lecture). pp. 1-80 in Yaroshevskiy AA (ed.) *The earth's sedimentary shell*. Nauka, Moscow. American Geological Institute Reprint Series 5:1-73.

Ronov AB (1994) Phanerozoic transgressions and regressions on the continents: a quantitative approach based on areas flooded by the sea and areas of marine and continental deposition. *American Journal of Science* 294:777-801.

Wall PD, Ivany LC, Wilkinson BH (2009) *Paleobiology* 35:146-167.

---

net.div.rate	*Exponential and saturated net diversification rate*

Description

Calculates net diversification rates assuming a model of exponential diversification, with or without saturation

Usage

```
net.div.rate(n,dt,sat)
```

Arguments

n   Extant species diversity of the investigated clade.

dt   Time interval over which the extant species diversity is believed to have accumulated (= the clade age).

sat   Saturation (optional). If set to TRUE, then it is assumed that half the current species richness existed after half the clade's age, and that diversity slowly saturated after this. This is similar to logistic growth. By default, sat is set to FALSE.

Details

The function net.div.rate calculated the net diversification rate for a clade, given n (the number of extant species in this clade), dt (the time since clade origin), and assuming an exponential diversification model. Optionally, it can be assumed that half the extant species richness was present after half the time, and that since that time, diversity has saturated. In this case, the parameter sat should be set to TRUE. Estimation of the net diversification is problematic, because it requires an estimate for the age of the investigated clade, which one typically tries to find with

this type of analysis. Thus, the age estimate influences the net diversification rate estimate, which in turn again influences the age estimate. I recommend to estimate the net diversification rate before and after the molecular dating analysis, and repeat the analysis if the a posteriori net diversification rate is too different from the a priori assumptions. It may be very useful to run an analysis with the software MEDUSA (Alfaro et al. 2009), or the forthcoming FOSSILMEDUSA (www.webpages.uidaho.edu/~lukeh/software/index.html), to gain insights into the net diversification rate dynamics of the investigated clade. Experience has shown that molecular dating results are relatively robust to the chosen net diversification rate (higher estimates of net diversification lead to older age estimates). It may be wise to try a range of possible net diversification rates.

Author(s)

Michael Matschiner

Maintainer: Michael Matschiner <michaelmatschiner@mac.com>

Source

Matschiner M (2011) in prep.

References

Alfaro ME, Santini F, Brock CD, et al. (2009) *Proc Natl Acad Sci* 106:13410-13414.

See Also

find.prior

Examples

```
net.div.rate(n=4500,dt=180,sat=TRUE)
```

NorthAmerica	*North American rock outcrop bias*

Description

Quantification of the North American rock outcrop of the Phanerozoic

Usage

```
data(NorthAmerica_marine)
data(NorthAmerica_terrestrial)
data(NorthAmerica)
```

Format

Data frames with 23 values for preservation rate bias (relr) and the younger end of the time interval (py) for which this preservation rate has been calculated.

Details

North American marine and terrestrial rock outcrops of the Phanerozoic have been quantified by Wall et al. (2009) by combination of data on North American outcrop areas from the UNESCO *Geological Atlas of the World* (Choubert & Faure-Muret 1976) with volumetric data on North American sediment compositions (Ronov 1980, 1994). Bin size for the 22 used phanerozoic time intervals varies between 6 (Early Triassic) and 54 (Cambrian) myr, with an average length of 24.4 myr. The 23rd value results from an extrapolation of Miocene rock outcrop to the present. Rock outcrop quantification is here translated into area per myr and is scaled so that the maximum area per myr equals 1. This data serves as a proxy for preservation rate bias. It is used by functions adjust.pres.rate and find.prior when called with corr="NorthAmerica_marine", corr="NorthAmerica_terrestrial", or corr="NorthAmerica" (which combines both the marine and terrestrial rock outcrop).

Author(s)

Michael Matschiner

Maintainer: Michael Matschiner <michaelmatschiner@mac.com>

Source

Matschiner M (2011) in prep.

References

Choubert G, Faure-Muret A (1976) *Geological atlas of the world. 1:10,000,000. 22 sheets with explanations*. UNESCO Commission for the Geological Map of the World, Paris.

Ronov AB (1980) The earth's sedimentary shell: quantitative patterns of its structures, compositions and evolution (the 20th V. I. Vernadskiy Lecture). pp. 1-80 in Yaroshevskiy AA (ed.) *The earth's sedimentary shell*. Nauka, Moscow. American Geological Institute Reprint Series 5:1-73.

Ronov AB (1994) Phanerozoic transgressions and regressions on the continents: a quantitative approach based on areas flooded by the sea and areas of marine and continental deposition. *American Journal of Science* 294:777-801.

Wall PD, Ivany LC, Wilkinson BH (2009) *Paleobiology* 35:146-167.

WesternEurope   *Western European rock outcrop bias*

Description

Quantification of the Western European rock outcrop of the Phanerozoic

Usage

```
data(WesternEurope_marine)
data(WesternEurope_terrestrial)
data(WesternEurope)
```

Format

A data frame with 72 values for preservation rate bias (relr) and the younger end of the time interval (py) for which this preservation rate has been calculated.

Details

Western European marine and terrestrial rock outcrops of the Phanerozoic have been quantified by Smith & McGowan (2007) and McGowan & Smith (2008) as the number of geological maps of France, Spain, England, and Wales containing this type of rock outcrop of a given time interval (using bins with an average size of 7.5 myr). This quantification is here translated into maps per myr and is scaled so that the maximum number of maps per myr equals 1. This data serves as a proxy for preservation rate bias. It is used by functions adjust.pres.rate and find.prior when called with corr="WesternEurope_marine", corr="WesternEurope_terrestrial", or corr="WesternEurope" (which combines both the marine and terrestrial rock outcrop).

Author(s)

Michael Matschiner

Maintainer: Michael Matschiner <michaelmatschiner@mac.com>

Source

Matschiner M (2011) in prep.

References

McGowan AJ, Smith AB (2008) *Paleobiology* 34:80-103.
Smith AB, McGowan AJ (2007) *Palaeontology* 50:765-774.

Index

12

3.3 Parallel ecological diversification in Antarctic notothenioid fishes as evidence for adaptive radiation

Rutschmann S, Matschiner M, Damerau M, Muschick M, Lehmann MF, Hanel R, Salzburger W

MOLECULAR ECOLOGY

MOLECULAR ECOLOGY

Parallel ecological diversification in Antarctic notothenioid fishes as evidence for adaptive radiation

SEREINA RUTSCHMANN,* MICHAEL MATSCHINER,* MALTE DAMERAU,† MORITZ MUSCHICK,* MORITZ F. LEHMANN,‡ REINHOLD HANEL† and WALTER SALZBURGER*

*Zoological Institute, University of Basel, Vesalgasse 1, CH-4051 Basel, Switzerland, †Institute of Fisheries Ecology, Johann Heinrich von Thünen-Institute, Federal Research Institute for Rural Areas, Forestry and Fisheries, Palmaille 9, D-22767 Hamburg, Germany, ‡Institute of Environmental Geosciences, University of Basel, Bernoullistrasse 30, CH-4056 Basel, Switzerland

Abstract

Antarctic notothenioid fishes represent a rare example of a marine species flock. They evolved special adaptations to the extreme environment of the Southern Ocean including antifreeze glycoproteins. Although lacking a swim bladder, notothenioids have diversified from their benthic ancestor into a wide array of water column niches, such as epibenthic, semipelagic, cryopelagic and pelagic habitats. Applying stable carbon (C) and nitrogen (N) isotope analyses to gain information on feeding ecology and foraging habitats, we tested whether ecological diversification along the benthic–pelagic axis followed a single directional trend in notothenioids, or whether it evolved independently in several lineages. Population samples of 25 different notothenioid species were collected around the Antarctic Peninsula, the South Orkneys and the South Sandwich Islands. The C and N stable isotope signatures span a broad range (mean $\delta^{13}C$ and $\delta^{15}N$ values between −25.4‰ and −21.9‰ and between 8.5‰ and 13.8‰, respectively), and pairwise niche overlap between four notothenioid families was highly significant. Analysis of isotopic disparity-through-time on the basis of Bayesian inference and maximum-likelihood phylogenies, performed on a concatenated mitochondrial (cyt b) and nuclear gene (*myh6*, *Ptr* and *tbr1*) data set (3148 bp), showed that ecological diversification into overlapping feeding niches has occurred multiple times in parallel in different notothenioid families. This convergent diversification in habitat and trophic ecology is a sign of interspecific competition and characteristic for adaptive radiations.

Keywords: disparity-through-time, marine speciation, niche overlap, pelagization, phylogeny, stable nitrogen and carbon isotopes

Received 1 February 2011; revision received 7 July 2011; accepted 15 July 2011

Introduction

Adaptive radiation, the evolution of ecological and phenotypic diversity within a rapidly multiplying lineage, is thought to be responsible for a great portion of the diversity of life (Simpson 1953; Schluter 2000). The most famous examples of adaptive radiations are the Darwin's finches on Galápagos, the Caribbean *Anolis* lizards and the East African cichlid fishes. One of the key

features of an adaptive radiation is the correlation between the morphologically diverse phenotypes of the 'participating' species and the various habitats that these occupy (Schluter 2000). While it is conceivable how such 'adaptive disparity' is fulfilled by the paradigmatic Darwin's finches, anoles and cichlids with their characteristic adaptations in beaks, limbs and trophic structures, respectively, the inference of phenotype-environment correlation remains a challenge in other cases of adaptive radiation (Schluter 2000; Gavrilets & Losos 2009).

In fishes, most studies on adaptive radiation focus on freshwater systems, with the cichlid species flocks of

Correspondence: Walter Salzburger, Fax: +41 61 267 03 01; E-mail: walter.salzburger@unibas.ch and Reinhold Hanel, Fax: +49 40 38 90 52 61; E-mail: reinhold.hanel@vti.bund.de

the East African Great Lakes being the prime examples (Salzburger 2008, 2009). The Antarctic notothenioids represent a marine species flock that evolved under extreme environmental conditions (Eastman & Clarke 1998; Eastman 2000). The perciform suborder Notothenioidei diversified into at least 130 species in eight families, encompassing over 100 Antarctic species (Eastman 2005; Eakin *et al.* 2009). Three ancestral families, Bovichtidae, Pseudaphritidae and Eleginopidae, comprise eleven primarily non-Antarctic species, distributed around southern South America, the Falkland Islands, southern New Zealand and southeastern Australia (Eastman 1993). The remaining families Artedidraconidae, Bathydraconidae, Channichthyidae, Harpagiferidae and Nototheniidae are, with few exceptions, endemic to Antarctic waters and are usually referred to as the 'Antarctic clade' (e.g. Eastman 1993). Notothenioids dominate the Antarctic continental shelf and upper slope, accounting for approximately 46% of the species diversity and over 90% of the fish biomass (Eastman & Clarke 1998; Eastman 2005).

Antarctic waters are constrained by the Antarctic Circumpolar Current (ACC). The Antarctic Polar Front, the northern boundary of the ACC between 50°S and 60°S, acts as major oceanographic barrier, effectively isolating the Southern Ocean faunal assemblages from those of the Indian, Pacific and Atlantic oceans. Through the establishment of a thermally and oceanographically isolated area and the inhibition of faunal admixture, the Antarctic Polar Front is, hence, a likely driver of notothenioid evolution (Coppes Petricorena & Somero 2007). As a means to adapt to Southern Ocean environmental conditions, the Antarctic notothenioids evolved special anatomical and physiological features and, at the same time, lost traits no longer 'needed' in permanently cold waters: (i) The evolution of antifreeze glycoproteins is regarded as an evolutionary key innovation of notothenioids (Eastman 1993; Matschiner *et al.* 2011), facilitating permanent life in subzero temperate waters. (ii) All notothenioids lack a functional swim bladder. Several pelagic species, however, have evolved neutral buoyancy by a combination of skeletal mineralization and the accumulation of lipid deposits (Eastman 1993; Klingenberg & Ekau 1996). (iii) Some notothenioids have lost the classical heat-shock protein response (Place & Hofmann 2005; Clark *et al.* 2008). (iv) The Channichthyidae represent the only known vertebrate group that lacks erythrocytes in the adult state and that is unable to synthesize a functional version of the respiratory oxygen transporter haemoglobin (Ruud 1954; Near *et al.* 2006).

Here, we investigate niche evolution in notothenioids, using a set of 25 representative species (and 365 individuals) that belong to four of the five notothenioid families in the exceptionally species-rich Antarctic clade. Apparently, Antarctic notothenioids diversified along the benthic-pelagic axis in the absence of competition from other fish taxa (Eastman 1993, 2005). From a morphological perspective, this process termed 'pelagization' appears to have occurred independently in several clades (Klingenberg & Ekau 1996; Bargelloni *et al.* 2000).

We used isotopic signatures as indicators for ecological specialization to assess the diversity of lifestyles and feeding strategies/habits of the Antarctic clade, as has been done for adaptively radiating rockfishes (Ingram 2011), and to further test whether these strategies/habits evolved clade-specifically and unidirectionally or independently in several lineages. Stable isotope analysis (SIA) makes use of the fact that the C and N stable isotope signatures (δ^{13}C and δ^{15}N) of organisms are directly related to their diet. In general, the ratio of the heavier over the lighter stable isotope is greater in consumers than in food material and thus continuously increases with trophic level (TL; e.g. Hobson & Welch 1992; Hobson *et al.* 1994). This is particularly true for nitrogen, where N isotope fractionation leads to trophic shifts of 3–5‰ (DeNiro & Epstein 1978; Minagawa & Wada 1984; Post 2002). The C isotope fractionation is less pronounced during food chain processing, with a typical 1‰ increase per TL (Hobson & Welch 1992). Yet, carbon isotopic values can often be used to assess constraints on the primary carbon source, which can vary strongly between different feeding grounds (e.g. inshore vs. offshore and pelagic vs. benthic). Thus, while N isotope ratios can be used to predict the relative TL of an organism, its C isotopic composition yields valuable information with regard to its habitat (e.g. Hobson *et al.* 1994).

To reconstruct the evolution of ecological specialization in notothenioids, which has not been studied in detail, we established a new phylogeny of the studied species based on mitochondrial and nuclear markers [3148 base pairs (bp) in total]. This phylogeny extends previous work (e.g. Near & Cheng 2008) by the use of multiple nuclear markers and by the longest total sequence length used in notothenioid phylogenetics to date. Phylogeny and time estimation were fully integrated with SIA by the application of a disparity-through-time (DTT) analysis.

According to the results of earlier studies (Klingenberg & Ekau 1996; Eastman & McCune 2000), we expected to find evidence for independent colonization of ecological niches in different lineages. Furthermore, should previous descriptions of the notothenioid diversification as an adaptive radiation be appropriate, the pattern of average subclade disparity throughout the radiation could be expected to resemble those found in

other adaptive radiations like *Liolaemus* lizards (Harmon *et al.* 2003) or Tanganyikan cichlid fishes (Gonzalez-Voyer *et al.* 2009) and to be different from patterns observed in putative non-adaptive radiations, such as rats (Rowe *et al.* 2011).

Materials and methods

Sample collection

Sampling took place during three expeditions in the austral summer to the Scotia Sea: The ICEFISH 2004 cruise with RV Nathaniel B. Palmer (Jones *et al.* 2008), cruise ANT-XXIII/8 with RV Polarstern, and the 2008/09 US AMLR Survey with RV Yuzhmorgeologiya (Jones *et al.* 2009) (Fig. 1 and Table 1, Tables S1 and S2, Supporting information). White muscle tissue samples were preserved in 95% ethanol and stored at −20 °C for subsequent investigations. A total of 365 adult individuals of 25 Antarctic notothenioid species were processed for SIA. Molecular analyses were performed with 39 individuals of the same 25 species and three representatives of non-Antarctic notothenioid families serving as outgroups (Table 1).

DNA extraction, amplification, sequencing and alignment

Genomic DNA from approx. $10 mm^3$ white muscle tissues was extracted by proteinase K digestion, followed by sodium chloride extraction and ethanol precipitation. Marker selection was based on the genome-wide marker comparison of Li *et al.* (2007). We included a fast-evolving gene (*myh6*), a gene evolving at intermediate rates (*Ptr*) and a slowly evolving gene (*tbr1*). As a representative mitochondrial marker


Fig. 1 Sampling sites off the northern Antarctic Peninsula, the South Orkney Islands and the South Sandwich Islands. The solid line indicates the 1000 m depth contour.

(mtDNA), we used cytochrome *b* (cyt *b*), which had previously been proven suitable for phylogenetic analyses in notothenioids (Chen *et al.* 1998; Matschiner *et al.* 2011). Nuclear markers were amplified with the following primer pairs: myh6_F507/myh6_R1325, Ptr_F458/Ptr_R1248 and tbr1_F86/tbr1_R820 (Li *et al.* 2007); the amplification of cyt *b* was performed using the primers NotCytBf and H15915n (Matschiner *et al.* 2011). Sequences of the three outgroup species and *Pogonophryne scotti*, as well as *Ptr* sequences of *Notothenia coriiceps* and *Trematomus newnesi* were obtained from GenBank (see Data accessibility and Table S4, Supporting information).

The gene fragments were amplified using different polymerase chain reaction (PCR) protocols. Cyt *b*, *myh6* and *Ptr* PCR products were achieved using the Finnzymes' Phusion® High-Fidelity DNA Polymerase (Finnzymes). Individual reaction volumes contained 8.6 μL ddH$_2$0, 10.0 μL 2 × Phusion® Master Mix with HF Buffer [containing 0.04 U/μL Phusion® DNA Polymerase, 2 × Phusion® HF Buffer, 400 μM of each deoxynucleotides (dNTP)], 0.2 μL forward primer, 0.2 μL reverse primer and 1.0 μL DNA template. The PCR profiles included initial denaturation (30 s, 98 °C), followed by 30 (cyt *b*) or 40 cycles (*myh6*, *Ptr*) of denaturation (10 s, 98 °C), annealing (30 s, 56 °C) (53 °C for *Ptr*), extension (30 s, 72 °C) and a final extension phase (10 min, 72 °C). *Tbr1* amplification was achieved using REDTaq® DNA Polymerase (Sigma-Aldrich). The PCR mixes contained 5.5 μL ddH$_2$O, 1.25 μL 10× Taq buffer (Sigma-Aldrich), 1.0 μL MgCl$_2$, 1.25 μL dNTP mix, 1.0 μL forward primer, 1.0 μL reverse primer, 0.5 μL REDTaq® DNA Polymerase (Sigma-Aldrich) and 1.0 μL DNA template. Amplifications of *tbr1* were carried out using the following temperature profile: initial denaturation (2 min, 94 °C) followed by 32 thermocycles of denaturation (30 s, 94 °C), annealing (30 s, 57 °C), extension (1 min, 72 °C) and a final extension phase (7 min, 72 °C). All amplification products were purified using the ExoSAP-IT (USB) standard protocol, adding 0.5 μL ExoSAP-IT and 3.5 μL ddH$_2$O to 2.5 μL PCR templates, incubating (15 min, 37 °C; 15 min, 80 °C) and, in some cases, using the GenElute™ Gel Extraction Kit (Sigma-Aldrich). The purified PCR products were used as templates for cycle sequencing reactions with the BigDye® Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems), following the manufacturer's instructions. The reaction volumes included 0.5 μL primer, 1.0 μL BigDye® Terminator Reaction Mix (Applied Biosystems) and 3.0–6.5 μL purified DNA in a total volume of 8 μL. The nuclear markers were sequenced with one forward and reverse primer each. Sequencing of cyt *b* was additionally performed with two different forward primers: NotCytBf (Matschiner *et al.* 2011) and
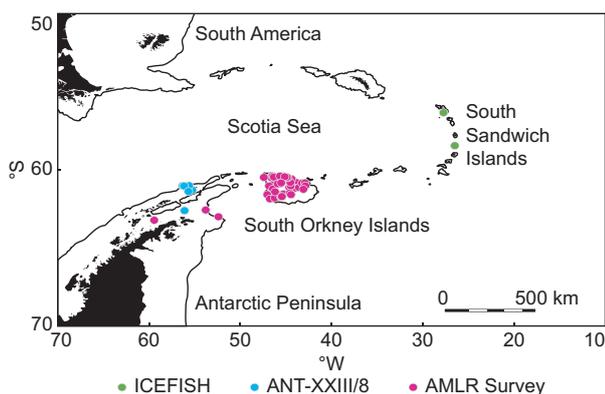
Sample	Location (*n*)	Lifestyle of adults
Bovichtidae		
Bovichtus diacanthus	Tristan da Cunha	
Pseudaphritidae		
Pseudaphritis urvillii	Victoria, Australia	
Eleginopidae		
Eleginops maclovinus	South America	
Nototheniidae		
Aethotaxis mitopteryx	AP (4), SO (7)	Pelagic[*,†,‡,§], benthopelagic[¶]
Dissostichus mawsoni	AP (2), SO (5)	Pelagic[†,§]
Gobionotothen gibberifrons	AP (10), SO (10)	Benthic[†,‡]
Lepidonotothen larseni	SO (10), SSI (10)	Semipelagic[†]
Lepidonotothen nudifrons	SO (10)	Benthic[†,§]
Lepidonotothen squamifrons	AP (10), SO (10)	Benthic[†]
Notothenia coriiceps	AP (10), SO (11)	Benthic[§]
Notothenia rossii	SO (11)	Semipelagic[†]
Pleuragramma antarcticum	AP (10), SO (10)	Pelagic[*,†,§]
Trematomus eulepidotus	AP (10), SO (10)	Epibenthic[*,†,‡]
Trematomus hansoni	SO (11)	Benthic[†,‡]
Trematomus newnesi	AP (10), SO (10)	Cryopelagic[†]
Trematomus nicolai	SO (6)	Benthic[*,†,‡,**,††], benthopelagic[‡‡]
Trematomus tokarevi	SO (11)	Benthic[††]
Artedidraconidae		
Pogonophryne barsukovi	SO (8)	Benthic[§§]
Pogonophryne scotti	SO (10)	Benthic[†,§§]
Bathydraconidae		
Gymnodraco acuticeps	AP (15)	Benthic[†]
Parachaenichthys charcoti	SO (11)	Benthic[†]
Channichthyidae		
Chaenocephalus aceratus	AP (10), SO (10)	Benthic[†,¶]
Chaenodraco wilsoni	AP (10)	Pelagic[***]
Champsocephalus gunnari	AP (11), SO (10)	Pelagic[†,¶]
Chionodraco rastrospinosus	AP (10), SO (10)	Benthic[†], benthopelagic[†††]
Cryodraco antarcticus	AP (10), SO (10)	Pelagic[†], benthic[¶]
Neopagetopsis ionah	AP (6), SO (6)	Pelagic[¶]
Pseudochaenichthys georgianus	SO (10)	Pelagic[†,¶], semipelagic[†]

Table 1 Sampled species with collection site, sample size for stable isotope analysis (*n*) and lifestyle of adult individuals. Lifestyle descriptions are often based on trawl depth and may not be definite.

[*]DeWitt *et al.* (1990); [†]Eastman (1993); [‡]Klingenberg & Ekau (1996); [§]Kock (1992); [¶]Kunzmann & Zimmermann (1992); [**]Kuhn *et al.* (2009); [††]La Mesa *et al.* (2004); [‡‡]Brenner *et al.* (2001); [§§]Lombarte *et al.* (2003); [¶¶]Kock (2005); [***]Kock *et al.* (2008); [†††]Hureau (1985b).
AP, Antarctic Peninsula, SO, South Orkney Islands, SSI, South Sandwich Islands.

cyt*b*centralF (5′- CYA CCC TNA CYC GYT TCT TTG C -3′), which was newly designed to bind at a central position of cyt *b* (bases 518–539 in cyt *b* of *Chionodraco rastrospinosus*). The reaction conditions were as follows: initial denaturation (1 min, 94 °C) followed by 25 cycles of denaturation (10 s, 94 °C), annealing (20 s, 52 °C) and elongation phase (4 min, 60 °C). Unincorporated BigDye® terminators were removed with the BigDye® XTerminator™ Purification Kit (Applied Biosystems). To this end, 14.5 μL ddH₂O, 22.5 μL SAM™ solution and 5.0 μL XTerminator™ beads were added to the sequencing products, then shaken (30 min, 2000 rpm), and finally centrifuged (2 min, 211 *g*). All sequences were read with an ABI3130*xl* Capillary Sequencer (Applied

Biosystems). Sequence reads were verified by eye, and forward and reverse fragments were assembled using CODONCODE ALIGNER v.3.5.6 (CodonCode Corporation).

All sequences were aligned per locus with the multiple sequence alignment program MAFFT v.6.717b (Katoh & Toh 2008). The alignments were trimmed in MESQUITE v.2.72 (Maddison & Maddison 2009) so that each alignment started and ended with codon triplets, and we also checked for stop codons. Alignments were concatenated and partitioned by molecule type and codon position to account for heterogeneity in evolutionary rates and substitution patterns. Thus, the first and second codon positions of mitochondrial cyt *b* ('mit12'), the third codon positions of mitochondrial cyt *b* ('mit3'), the

first and second codon positions of nuclear genes ('nuc12') and the third positions of nuclear genes ('nuc3') were used as separate partitions. In a second partitioning scheme, the data set was partitioned with respect to the four genes. The best-fitting models of molecular evolution for each of the eight partitions were estimated with the computer program jMODELTEST v.0.1.1 (Posada 2008), using the Bayesian information criterion (BIC; Schwarz 1978). Selected models were TPM2uf+G (*myh6*), K80+G (*Ptr*), HKY+I (*tbr1*), TrN+G+I (cyt *b*), HKY+I+G (mit12), K80+I (nuc12) and TrN+G (mit3, nuc3).

Phylogenetic analysis

Phylogenetic tree reconstructions were carried out using maximum-likelihood (ML) and Bayesian inference (BI) approaches. Maximum-likelihood phylogenetic inference was performed with both partitioning schemes, applying the respective models of molecular evolution for each partition, in a partition-enabled version of GARLI, GARLI-PART v.0.97 (Zwickl 2006). Heuristic searches were used to find the topology with the best likelihood score. The searches were conducted using automatic termination, after a maximum of 5 million generations, or, alternatively, after 10 000 generations without significant ($P < 0.01$) improvement in scoring topology. Bootstrap (BS) analysis was performed with 100 BS replicates, which were summarized using PAUP* v.4.0a110 (Swofford 2003). The non-Antarctic notothenioid species *Bovichtus diacanthus* was defined as outgroup on the basis of well-supported phylogenetic information (e.g. Near & Cheng 2008; Matschiner *et al.* 2011).

Bayesian phylogenetic analyses were performed with the software BEAST v.1.5.3 (Drummond & Rambaut 2007). For divergence date estimation, the separation of Bovichtidae, Pseudaphritidae and Eleginopidae from the Antarctic lineage (nodes A, B, and C in Fig. 3), as well as the initial diversification of the Antarctic clade (node D) were temporally constrained according to the results of Matschiner *et al.* (2011). Specifically, normal prior distributions were used for each of these splits to approximate highest posterior density (HPD) intervals found by Matschiner *et al.* (2011). Thus, the root of Notothenioidei (node A) was constrained with a mean divergence prior to 71.4 million years ago (Ma; 2.5% quantile: 89.1 Ma, 97.5% quantile: 53.8 Ma), and nodes B-D were constrained at 63.0 (79.5–46.6) Ma, 42.9 (56.5–29.4) Ma and 23.9 (31.3–16.4) Ma, respectively. While these time constraints generally agree with the interpretation of *Proeleginops grandeastmanorum* from the La Meseta Formation on Seymour Island (~40 Ma; Eastman & Grande 1991) as an early representative of the

eleginopid lineage (Balushkin 1994), we deliberately avoided using it as a time constraint owing to its debated taxonomical assignment (Near 2004). With the exception of outgroup relationships, which were used for time calibration, no topological constraints were applied. Divergence dates were estimated using the uncorrelated lognormal relaxed molecular clock and the reconstructed birth-death process as a tree prior (Gernhard 2008). Following Shapiro *et al.* (2006), we implemented the codon position-specific model of sequence evolution $HKY_{112} + CP_{112} + \Gamma_{112}$, but we furthermore tested $GTR_{112} + CP_{112} + \Gamma_{112}$ and the model combination selected by BIC for codon-specific partitions. For each of the three combinations, 10 independent analyses were performed with 20 million generations each. Replicates were combined in LOGCOMBINER v.1.5.3 (Drummond & Rambaut 2007) after removing the first 2 million generations of each run as burn-in. Convergence of run replicates was verified by effective sample sizes > 1200 for all parameters and by comparison of traces within and between replicates in TRACER v.1.5 (Rambaut & Drummond 2007). The three settings were compared with Bayes factors (BF), using the harmonic mean approach as implemented in TRACER. While we acknowledge that the harmonic mean estimator may be biased towards more parameter-rich models (Lartillot & Hervé 2006), we chose this approach owing to the lack of suitable alternatives. As the inclusion of multiple individuals per species may violate assumptions of constant diversification implicit in the birth–death tree prior, BI analyses were repeated with a reduced data set containing only one individual of each species.

Stable isotope analysis

In this study, approximately 10 mm^3 of white muscle tissue was used for the SIA. White muscle tissue is less variable with regard to the carbon and nitrogen isotope composition and has a longer retention time than other tissue types (Pinnegar & Polunin 1999; Quevedo *et al.* 2009). Samples were dried (24 h, 60 °C) and then ground in a Zirconia bead mill (30 min, 1800 bpm). Then, the sample powder was rinsed from the beads using 1 mL 99% ethanol, and the supernatant was evaporated (24 h, 60 °C). The ethanol treatment had no effect on subsequent carbon isotope analyses (e.g. Syväranta *et al.* 2008). For C and N isotope measurements, between 0.5 and 0.8 mg sample powder was filled into 5×9 mm tin capsules and introduced into an elemental analyser (Thermo Finnigan) coupled to a Finnigan Delta V Advantage Isotope Ratio Mass Spectrometer, with standard setup for N_2 and CO_2 analysis. Measurements were replicated for about 10% of the samples (42 samples). The isotopic composition is expressed in the

conventional delta notation as permil (‰) deviation vs. atmospheric N_2 (AIR) and carbonate standards (V-PDB): $\delta = [(R_{sample}/R_{standard}) - 1] \times 1000$, with R representing the ratio of the heavy to the light isotope (i.e. $^{13}C/^{12}C$ and $^{15}N/^{14}N$) in the sample and in the standard material, respectively. EDTA ($\delta^{13}C = -30.25‰$, $\delta^{15}N = -1.1‰$) and ammonium oxalate ($\delta^{13}C = -17.02‰$, $\delta^{15}N = 32.7‰$) were used as internal standards, calibrated against international nitrogen (IAEA-N1, IAEA-N2) and carbon (NBS22) standards. The analytical reproducibility based on replicate sample and standard measurements was better than $0.2‰$ for both $\delta^{13}C$ and $\delta^{15}N$. Isotope values are presented as mean ± standard deviation (SD). Variable lipid content can have a biasing effect on the interpretation of bulk C and N stable isotope data. In marine fish samples, this effect seems to be minor (Kiljunen *et al.* 2006; Logan *et al.* 2008), and hence, we did not perform a lipid removal step. Nevertheless, we performed a posteriori 'mathematical lipid correction' after the study of Logan *et al.* (2008). The correction, however, did not affect the species distribution pattern, and thus, only the uncorrected values are presented in this study. (The corrected data set is available upon request.)

Statistical analysis

The correlation of $\delta^{13}C$ and $\delta^{15}N$ was tested with a Pearson correlation, whereby we accounted for phylogenetic non-independence using phylogenetic independent contrast ('pic' function in the R package 'ape'; Paradis *et al.* 2004; R Development Core Team 2009). We tested for the effect of geographic sites on isotopic signatures by comparison of pooled $\delta^{13}C$ and $\delta^{15}N$ values between AP and SO (*t*-test). Here, only values from species with similar sample sizes at both locations were considered. Pairwise niche overlap between all families and additional comparisons of the nototheniid *Lepidonotothen–Trematomus* clade with the other families were tested with a multivariate analysis of variance (MANOVA). To assess the group overlap in isotopic signatures, we calculated Wilk's lambda (Wilk's λ) for each comparison.

We analysed the subdivision of ecological niche space throughout the radiation using the BI phylogeny (Fig. 3) and the averaged stable isotope data for each species. Average subclade disparity was calculated at each splitting event and plotted against time. A Brownian motion (BM) model of trait evolution was employed for comparison. Disparity-through-time analyses were conducted in R using the package 'geiger' (Harmon *et al.* 2008). Using 475 trees drawn from the posterior distribution of the BI analysis and 500 permutations of the stable isotope data, we assessed the robustness of

the observed pattern against phylogenetic uncertainty and intraspecific variation.

Results

Phylogenetic analysis

The alignments had lengths of 1099 bp (cyt *b*), 705 bp (*myh6*), 702 bp (*Ptr*) and 642 bp (*tbr1*), resulting in a total of 3148 bp with only 0.3% missing data. The *myh6* alignment contained a short insertion (6 bp) in the non-Antarctic outgroup *B. diacanthus*; these 6 bp were excluded from the following phylogenetic analyses. Sequences are available at GenBank under the accession numbers JF264479–JF264629. Bayes factors provided 'very strong' (Kass & Raftery 1995) evidence that the codon position-specific combination of substitution models selected by BIC yielded a better fit than both the $HKY_{112} + CP_{112} + \Gamma_{112}$ (log 10 BF 6.215) and $GTR_{112} + CP_{112} + \Gamma_{112}$ (log 10 BF 19.19) models.

Our ML and BI phylogenetic analyses produced identical topologies and confirmed the monophyly of the Antarctic clade with high support values (BS 100%; Fig. 2, Fig. S1, Supporting information). Yet, BS support and Bayesian posterior probability (BPP) were low at the base of the diversification of the Antarctic clade (but high at species-level relationships). In all cases, clustering of individuals from different populations of the same species was strongly supported (BS ≥ 93% and BPP = 1.00). The three families Artedidraconidae, Bathydraconidae and Channichthyidae were recovered as monophyletic, while the Nototheniidae appeared paraphyletic. An ancestral position was assigned to *Aethotaxis mitopteryx*. The monophyly of a clade containing *Lepidonotothen* and *Trematomus* was highly supported (BS 100% and BPP 1.00), and *Notothenia* appeared as the sister group to the more derived 'high-Antarctic clade', comprising the families Artedidraconidae, Bathydraconidae and Channichthyidae. Both the high-Antarctic clade and the channichthyid family were found monophyletic with BS 100% and BPP 1.00. The two artedidraconids, *P. barsukovi* and *P. scotti*, grouped together in all analyses (with high support values). Monophyly of the two bathydraconid representatives was weakly supported (BS 35% and BPP 0.67). Within the family of Channichthyidae, *Champsocephalus gunnari* was placed as sister species of all other representatives followed by a clade containing *Pseudochaenichthys georgianus* and *Neopagetopsis ionah* and a clade containing the four genera *Chionodraco*, *Chaenodraco*, *Chaenocephalus* and *Cryodraco*. The ML reconstruction with gene-specific partitions resulted in minor topological differences (Fig. S1, Supporting information). Reduction in the data set to one individual per species did not change the tree
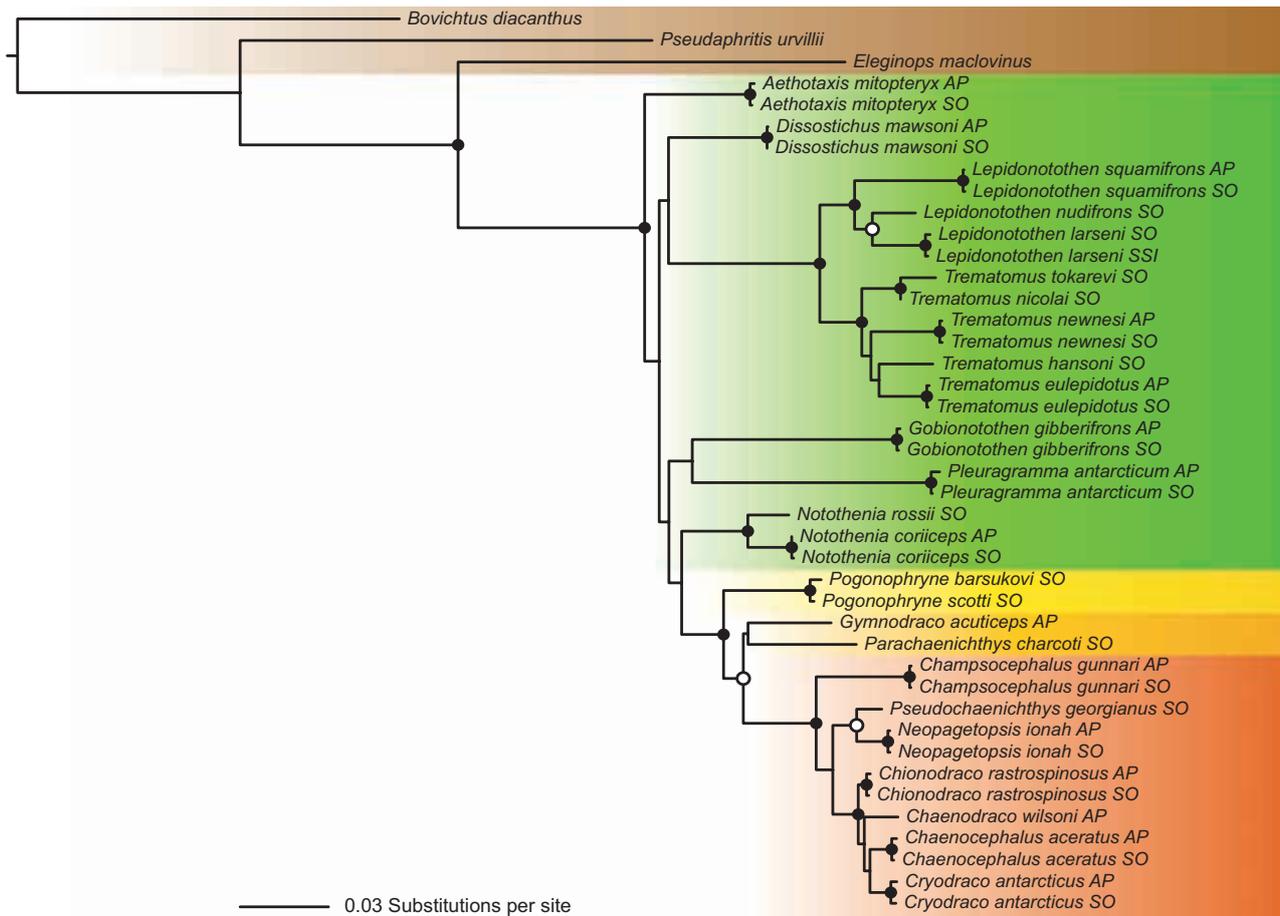
Fig. 2 Maximum-likelihood tree of the notothenioid phylogeny based on the codon position–specific partitioning scheme. Filled circles indicate strongly supported nodes, and moderately supported nodes are marked by open circles Bootstrap (BS ≥ 95 and BS ≥ 70). All species are coloured according to family: brown = non-Antarctic species, green = Nototheniidae, yellow = Artedidraconidae, orange = Bathydraconidae and red = Channichthyidae.

topology with the exception of *Dissostichus mawsoni*, which appeared basal to a group containing the high-Antarctic clade as well as *Nototheniia*, *Pleuragramma* and *Gobionotothen* and the relationships within the *Trematomus* genus (Fig. S1, Supporting information).

According to our time-calibrated phylogeny, diversification of the well-supported nototheniid clade combining *Lepidonotothen* and *Trematomus* began 12.0 Ma (95% HPD 16.4–7.9 Ma; node H) (Fig. 3). The high-Antarctic clade separated from the Nototheniidae around 18.6 Ma (95% HPD 24.0–13.4 Ma; node E). Within the high-Antarctic clade, artedidraconids separated from bathydraconids and channichthyids around 14.6 Ma (95% HPD 15.5–7.0 Ma; node F). The split between Bathydraconidae and Channichthyidae occurred around 2 million years later (12.5 Ma; 95% HPD 16.7–8.5 Ma; node G). The radiation of Channichthyidae, the most derived notothenioid family, began 7.7 Ma (95% HPD 10.6–5.0 Ma; node I).

Stable C and N isotope ratios

The stable carbon and nitrogen isotope composition for the 25 notothenioid species exhibited a comparatively large variability, with values between −27.8‰ and −19.7‰ for $\delta^{13}C$ and between 7.3‰ and 15.6‰ for $\delta^{15}N$ (Fig. 3). Mean values ranged between −25.4‰ and −21.9‰ for $\delta^{13}C$ (SD: 0.3‰ to 1.8‰) and 8.5‰ to 13.8‰ for $\delta^{15}N$ (SD: 0.2‰ to 1.7‰; Fig. 4). Intraspecific ranges of isotopic signatures span from 1.0‰ to 8.1‰ for $\delta^{13}C$ and from 0.4‰ to 5.7‰ for $\delta^{15}N$. Overall, mean intraspecific ranges ($\delta^{13}C$: 2.79‰, $\delta^{15}N$: 2.80‰) were small compared to interspecific ranges of isotopic signatures ($\delta^{13}C$: 8.12‰, $\delta^{15}N$: 8.29‰). The isotopic signatures of $\delta^{13}C$ and $\delta^{15}N$ correlated significantly (0.69; $P < 0.001$), and the correlation remained significant ($P < 0.01$) after correcting for phylogenetic non-independence. No significant difference between values from AP and SO locations was found ($P > 0.16$; t-test), even though the
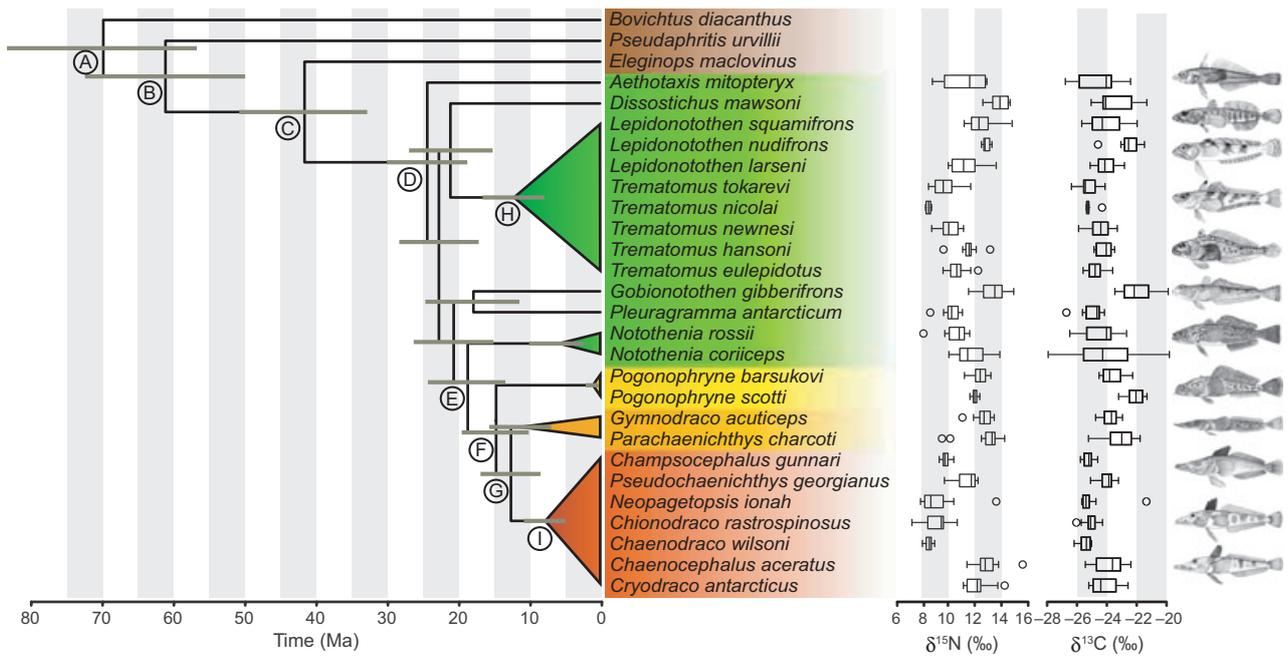
Fig. 3 Left: Time-calibrated phylogeny based on codon-specific partition, inferred with Bayesian inference. Time axis is given in million years ago and nodes labelled A-I are mentioned in the text. Grey node bars indicate upper and lower 95% HPD. All species are coloured according to family: brown = non-Antarctic species, green = Nototheniidae, yellow = Artedidraconidae, orange = Bathydraconidae and red = Channichthyidae. Right: Boxplot of stable isotope values of all included notothenioids. Representative habitus are illustrated at the right, from top to bottom: *Aethotaxis mitopteryx*[d], *Dissostichus mawsoni*[d], *Lepidonotothen nudifrons*[d], *Lepidonotothen larseni*[d], *Trematomus tokarevi*[d], *Gobionotothen gibberifrons*[d], *Notothenia rossii*[b], *Pogonophryne barsukovi*[c], *Gymnodraco acuticeps*[a], *Pseudochaenichthys georgianus*[e], *Chionodraco rastrospinosus*[e] and *Chaenocephalus aceratus*[e]. [a]Boulenger (1902); [b]DeWitt *et al.* (1990); [c]Eakin (1990); [d]Hureau (1985a); [e]Hureau (1985b).
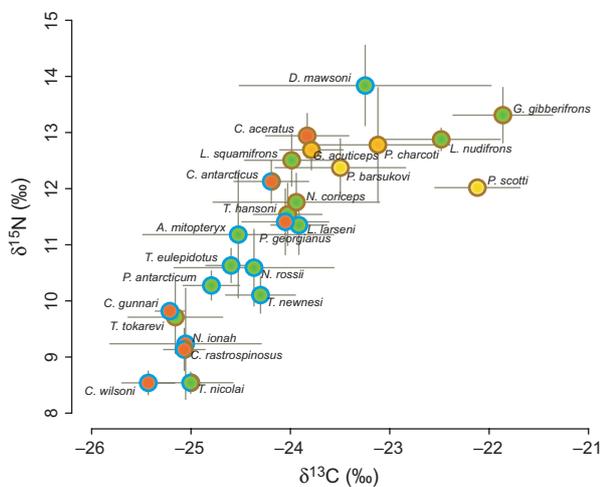
Fig. 4 Scatter plot of carbon and nitrogen isotopic values. Grey bars indicate 95% confidence intervals. All species are coloured according to family (brown: non-Antarctic species, green: Nototheniidae, yellow: Artedidraconidae, orange: Bathydraconidae, red: Channichthyidae), and strokes indicate corresponding lifestyle [blue = pelagic, benthopelagic, semipelagic and epibenthic; brown = benthic; and semicircles when references (Table 1) disagree].

mean values differed slightly (AP δ^{13}C: −24.37‰, SO δ^{13}C: −24.13‰; AP δ^{15}N: 11.30‰, SO δ^{15}N: 10.99‰).

With regard to inferred lifestyle patterns, our SIA data are consistent with previous studies (Hobson *et al.* 1994; Post 2002) in that species that are commonly classified as pelagic clustered around lower δ^{13}C values, while benthic species possessed relatively higher δ^{13}C signatures. However, there are notable exceptions to this: *D. mawsoni*, *C. rastrospinosus*, *Trematomus nicolai* and *T. tokarevi* (Fig. 4, Table 1 and Data S1, Supporting information). Most species had relatively high δ^{15}N signatures, indicating feeding at upper TL. The two well-represented families Nototheniidae and Channichthyidae covered a wide range of isotopic signatures, while bathydraconids and artedidraconids displayed a relatively low variability in both δ^{13}C and δ^{15}N (although the number of individuals was significantly lower). Overlap of the C and N isotope compositions as proxies for niche space was found in all pairwise comparisons (MANOVA) of the four Antarctic notothenioid families (Table 2). Wilk's λ was largest for comparisons of Nototheniidae with all other families (λ > 0.91; Table 2), and lower values were found for comparisons

Family 1	Family 2	Wilk's λ
Artedidraconidae	Nototheniidae	0.936
	Lepidonotothen–Trematomus clade	0.791
Bathydraconidae	Nototheniidae	0.913
	Lepidonotothen–Trematomus clade	0.818
Channichthyidae	Nototheniidae	0.930
	Lepidonotothen–Trematomus clade	0.932
Artedidraconidae	Bathydraconidae	0.681
Artedidraconidae	Channichthyidae	0.629
Bathydraconidae	Channichthyidae	0.781

Table 2 Pairwise niche overlap comparisons for the four Antarctic notothenioid families, performed with MANOVA (Wilk's λ)

including the lesser-represented families Artedidraconidae and Bathydraconidae (λ > 0.68). Notably, within-family variation resulted mostly from interspecific variation, instead of intraspecific variation, and closely related species with small intraspecific variation could be found at both ends of the ranges (e.g. *T. nicolai* and *Lepidonotothen nudifrons*; Fig. 3).

Using the DTT method, we assessed how the stable isotope space (as a proxy for ecological niche space) used by the whole clade was subdivided by smaller and smaller subclades as the radiation proceeded. We find positive deviations from the averaged neutral-evolution BM model, indicating larger overlap in niche space between subclades than would be expected if evolution proceeded neutrally (Fig. 5). This result was found to be robust against phylogenetic uncertainty and intraspecific variation by visual inspection of repeated DTT analyses.
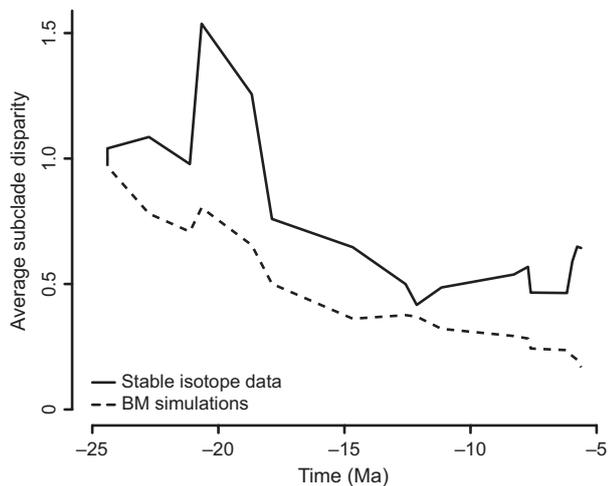

Fig. 5 Disparity-through-time plot for the stable isotopic signatures of Antarctic notothenioid fishes and Brownian motion simulations of character evolution. Time axis is given in million years ago.

Discussion

Phylogenetic relationships

Previous molecular phylogenetic analyses of notothenioids were based on mitochondrial DNA sequences (Bargelloni *et al.* 2000; Stankovic *et al.* 2002; Near 2004; Near *et al.* 2004), on a combination of mtDNA with a single nuclear gene (Near & Cheng 2008) or on morphological characters in addition to molecular data (Derome *et al.* 2002; Sanchez *et al.* 2007). The family-level phylogeny of notothenioids is thus relatively well established. Several questions remain, however, such as the position of the genus *Gobionotothen* (Near *et al.* 2004; Sanchez *et al.* 2007; Near & Cheng 2008) or whether Bathydraconidae are mono- or paraphyletic (e.g. Derome *et al.* 2002; Near & Cheng 2008).

In agreement with most previous studies (e.g. Near 2004; Near & Cheng 2008), our results support paraphyly of the family Nototheniidae. The low support values at the beginning of the Antarctic diversification are characteristic for rapid diversifications. Consequently, the basal position of *D. mawsoni* and the sister species relationships of *G. gibberifrons* and *Pleuragramma antarcticum* remain questionable. As in previous studies (Near 2004; Near & Cheng 2008), the three neutrally buoyant species *A. mitopteryx*, *D. mawsoni* and *P. antarcticum* diverged early within the Antarctic clade but did not cluster together. Phylogenetic relationships of the two genera *Notothenia* and *Lepidonotothen* are consistent with former studies (Bargelloni *et al.* 2000; Near & Cheng 2008). Also, the topology of the nototheniid subfamily Trematominae agrees with previous findings (Sanchez *et al.* 2007; Kuhn & Near 2009), except for *T. tokarevi* and *T. nicolai*, which appeared at basal positions in the phylogeny based on codon position–specific substitution models (Fig. 2, Fig. S1, Supporting information). The early split of the two included bathydraconid species relative to the divergence between Bathydraconidae and Channichthyidae

could indicate paraphyly of the former, as was concluded in previous studies (e.g. Derome *et al.* 2002; Near *et al.* 2004; Near & Cheng 2008). Resulting support values within the channichthyids were high, and the recovered topology was in complete agreement with the study of Derome *et al.* (2002). The three genera *Champsocephalus*, *Neopagetopsis* and *Pseudochaenichthys* seem to be well established as the most basal channichthyids (Chen *et al.* 1998; Near *et al.* 2003). In disagreement with former findings, *C. rastrospinosus* and *Chaenodraco wilsoni* did not cluster monophyletically (Chen *et al.* 1998). Near *et al.* (2003) also recovered these two species as paraphyletic but placed *Chaenocephalus aceratus* as the sister taxon to the genera *Cryodraco*, *Chaenodraco* and *Chionodraco,* which disagrees with our findings. Near & Cheng (2008) determined *C. aceratus* as the closest related species of *C. rastrospinosus.*

Inferred split dates (Fig. 3) roughly agree with those found by Near (2004) and Matschiner *et al.* (2011): Divergence estimates for the *Lepidonotothen–Trematomus* clade and the high-Antarctic clade were 12.0 (95% HPD 16.4–7.9) Ma and 18.6 (95% HPD 24.0–13.4) Ma, respectively, while Near (2004) reported them to be 14 ± 0.4 Ma and Matschiner *et al.* (2011) found these splits at 10.3 (95% HPD 15.2–6.1) Ma and 14.7 (95% HPD 20.0–9.9) Ma. According to our estimates, the radiation of the Channichthyidae began 7.7 (95% HPD 10.6–5.0) Ma ago, in good agreement with the estimates of Near (2004) (8.5 ± 0.3 Ma) and Matschiner *et al.* (2011) (6.2 Ma; 95% HPD 9.4–3.4 Ma).

Foraging ecology of notothenioids

So far, it has been shown that some particular feeding strategies are poorly represented or even absent in notothenioids, such as active skeleton-breaking predation (Clarke *et al.* 2004) or planktivory (Eastman & Grande 1989; Eastman 1993). The latter is probably due to restricted phytoplankton production during the austral winter (Clarke *et al.* 2004). The drawback of traditional dietary proxies (stomach content analyses and foraging observations) is that they only captures a snapshot of food uptake. Contrarily, SIA provides time-integrated information on the feeding 'ecology' for a period of weeks to years (McIntyre & Flecker 2006). Isotopic signatures could theoretically be influenced by geographic differences, sampling season and the age of sampled individuals, especially when ontogenic shifts occur in the investigated species. However, our sampling design accounted for these potential problems, as only adult specimens were collected, and all expeditions took place during austral summers. Also, most species were collected at the same two sampling locations, AP and SO, and populations from these two sites did not differ in isotopic signatures. Thus, the observed interspecific differences suggest ecological specialization rather than effects of geographical distribution or life history traits.

Our SIA data confirm that notothenioids occupy a wide variety of ecological niches (Figs 3 and 4). Comparatively high $\delta^{15}N$ values suggest that most investigated species reside at a high TL and may be considered tertiary consumers (see also Dunton 2001; Pakhomov *et al.* 2006). The wide range of the carbon stable isotope signatures reflects the notothenioids' variety in habitats along the benthic-pelagic axis (Fig. 4). However, our results are only partly congruent with the lifestyles and feeding reports based on stomach content analyses (Fig. 4, Table 1, Table S3 and Data S1, Supporting information).

At the family level, Nototheniidae are – in terms of habitat and feeding strategies – the most diverse clade among Antarctic notothenioids (La Mesa *et al.* 2004; this study) and include plankton, nekton and benthos feeders, as well as species that combine several feeding modes (Gröhsler 1994). The five included *Trematomus* species were differentiated in both isotopic signatures, thus indicating trophic niche separation (see also Brenner *et al.* 2001). Artedidraconids and bathydraconids represent the most benthic families among notothenioids (Fig. 4; Olaso *et al.* 2000; La Mesa *et al.* 2004). Their $\delta^{15}N$ values suggest feeding habits at higher TL (Olaso *et al.* 2000; Jones *et al.* 2009). The well-studied channichthyids clustered into three groups according to their diet (Fig. 4: *C. wilsoni, N. ionah, C. rastrospinosus* and *C. gunnari* at low TL; *P. georgianus* and *Cryodraco antarcticus* at intermediate TL; and *C. aceratus* at high TL; see also Kock 2005). Carbon signatures indicated a rather pelagic lifestyle for most channichthyid species, with the exception for *C. aceratus*, which we can classify as benthic top predator, in agreement with previous findings (Kock 2005; Reid *et al.* 2007).

The DTT plot (Fig. 5) indicates larger overlap of subclades in niche use than expected from a model of neutral evolution. This is characteristic for adaptive radiations (Harmon *et al.* 2003; Gonzalez-Voyer *et al.* 2009) and differs from patterns of putative nonadaptive radiations, which show a negative deviation from the averaged neutral-evolution BM model (e.g. Rowe *et al.* 2011). Taking into account the considerable variation in stable isotope signatures found in notothenioids as a whole (Fig. 4) – basically ruling out stasis in the evolution of niche use – as well as the robustness of this pattern against intraspecific variation, these results suggest convergent evolution in niche use between species of notothenioid subclades, especially between those clades separating around 20 Ma (Figs 3 and 5). This emphasizes the importance of ecological niche differentiation in the adaptive radiation of notothenioids.

Adaptive radiation and ecological diversification in notothenioids

Our integrative analyses, combining both the phylogenetic relationships and the isotopic signatures of 25 notothenioid species, reveal that ecological diversification into overlapping feeding niches has occurred multiple times in parallel in different notothenioid families (Figs 3 and 5). Using carbon and nitrogen stable isotope ratios as indicators of TL, feeding strategy and macrohabitat, we find great variation within, and substantial overlap between the more basal nototheniids and the derived channichthyids. The representatives of the benthic artedidraconids and bathydraconids also overlap and cluster at high TLs and δ^{13}C values. Our results further confirm partitioning of habitat and trophic resources within notothenioid fishes, indicating that diversification along the benthic–pelagic axis and to different TLs took place independently in at least two of five notothenioid families of the Antarctic clade (Nototheniidae and Channichthyidae; Fig. 3 and Table 2).

Convergent diversification in habitat and trophic ecology suggests interspecific competition and is a characteristic of adaptive radiations (e.g. Losos 1995; Schluter 2000). For example, *Anolis* lizards of the Caribbean have independently evolved four to six so-called ecomorphs on each of the four large islands of the Greater Antilles, including species specialized to live on grass, twigs, trunks and tree crowns. Variation in limb lengths of anole ecomorphs supports these different lifestyles, so that e.g. the trunk-ground ecomorph possesses relatively long legs adapted to running and jumping on broad surfaces, while the twig ecomorph has short legs and moves slowly on narrow surfaces (Losos 2009). In this context, diversification of notothenioids along the benthic-pelagic axis, as evidenced by their isotopic composition, and the respective adaptations in buoyancy (Eastman 1993) can be considered analogous to the *Anolis* diversification along the ground-tree axis. The notothenioid adaptive radiation shows further analogies to that of Caribbean anoles in terms of species richness (both around 120 species) and age (about 24 and 15–66 Ma, respectively) (Fig. 3; Eastman 2005; Nicholson *et al.* 2005; Losos 2009; Matschiner *et al.* 2011). Not all descendents of the *Anolis* radiation remained within the confined area of the radiation (Nicholson *et al.* 2005), and neither did the notothenioids: *Notothenia angustata*, *N. microlepidota* and the genus *Patagonotothen* secondarily escaped Antarctic waters and occur in New Zealand and South America (Eastman 2005). Moreover, both radiations were probably triggered by key innovations: subdigital toepads support the particular arboreality of *Anolis* lizards, whereas antifreeze glycoproteins in blood and tissues allow notothenioid survival in ice-laden

Antarctic waters (Chen *et al.* 1997; Losos 2009; Matschiner *et al.* 2011).

Compared to another well-studied adaptive radiation, that of cichlid fishes in East African lakes, the rate at which lineage formation seems to have occurred is much smaller in Antarctic notothenioids. In the Great Lakes of East Africa, cichlid fishes have diversified into at least 1500 species that differ greatly in naturally and sexually selected traits, including body shape, mouth morphology and colouration (Salzburger 2009). Comparison of cichlid species flocks between East African lakes, as well as mathematical models, have shown that larger habitats effectuate higher diversification rates, as they provide greater habitat heterogeneity and facilitate isolation by distance ('area effect'; Salzburger & Meyer 2004; Gavrilets & Vose 2005; Seehausen 2006). Different adaptive radiations may not be directly comparable as they depend on many ecological, genetic and developmental factors, with an important contribution of historical contingencies (Gavrilets & Losos 2009). Cichlids are known for their philopatry and low dispersal abilities (Danley & Kocher 2001; Salzburger & Meyer 2004), whereas most notothenioids have prolonged pelagic larval stages, enhancing long-range migration (Eastman 1993). Notothenioid populations are characterized by fragmented habitat, historical demographic fluctuations (Patarnello *et al.* 2011) and the absence of genetic structuring over large distances (Matschiner *et al.* 2009; and references therein), whereas many cichlid species posses significant population structuring even on extremely small scales (e.g. Arnegard *et al.* 1999; Rico & Turner 2002). Genetic differentiation over small scales has rarely been found in notothenioids (but see Clement *et al.* 1998). Eastman & McCune (2000) suggested that the smaller species number of notothenioids, compared with cichlid species flocks, could be explained by the absence of certain prime inshore habitats in the Southern Ocean. Alternatively, the notothenioid adaptive radiation may not yet have entered its final stage, namely the diversification with respect to communication. Streelman & Danley (2003) suggested a three-stage model of adaptive radiation (see also Danley & Kocher 2001), in which diversification first occurs with respect to macrohabitats, then with respect to microhabitats and finally with respect to communication (e.g. mating traits such as colouration; see also Gavrilets & Losos 2009). Full species richness would only be achieved through this final step. Streelman & Danley (2003) further suggested that divergence of habitat and trophic morphology is driven by natural selection, whereas diversification along the axis of communication is forced by sexual selection. It is as of yet unclear whether the radiation of notothenioids followed discrete stages. Here, we provide conclusive evidence that the

species are separated along the benthic-pelagic axis (i.e. according to macrohabitats; Figs 3 and 4) and probably also as a function of bottom topography and sediment types (Kock & Stransky 2000). Much less is known about microhabitat diversification, although our data suggest that closely related species do differ with respect to foraging strategies (e.g. genera *Lepidonotothen* and *Trematomus*; Figs 3 and 4). Recent evidence further indicates the possibility of divergence along Streelman and Danley's axis of communication, as egg guarding and parental care were observed in all major notothenioid lineages except within the Artedidraconidae (Kock *et al.* 2006; Barrera-Oro & Lagger 2010 and references therein).

On the other hand, because of the paucity of the Antarctic fossil record, it cannot be excluded that the notothenioid radiation has already surpassed its maximum species richness. It is an important characteristic that young adaptive radiations often 'overshoot' in terms of species number and that, generally, niche filling causes declining speciation rates (e.g. Seehausen 2006; Gavrilets & Losos 2009; Meyer *et al.* 2011). That notothenioids already underwent periods of 'overshooting' and niche filling could possibly explain the smaller diversity of Notothenioidei compared to the younger cichlid radiation in the East African Lakes. However, in this case, an early burst of diversification should have left its footprint in a 'bottom-heavy' phylogeny (Gavrilets & Vose 2005). A more extensive study, including many more representatives of the notothenoids, would be necessary to reconstruct the succession of their adaptive radiation.

Acknowledgements

References

Arnegard ME, Jeffrey AM, Danley PD *et al.* (1999) Population structure and colour variation of the cichlid fish *Labeotropheus fuelleborni* Ahl along a recently formed archipelago of rocky habitat patches in southern Lake Malawi. *Proceedings of the Royal Society B*, **266**, 119–130.

Balushkin AV (1994) Fossil notothenioid, and not gadiform, fish *Proeleginops grandeastmanorum* gen. nov. sp. nov. (Perciformes, Notothenioidei, Eleginopidae) from the late Eocene found in Seymour Island (Antarctica). *Journal of Ichthyology*, **34**, 298–307.

Bargelloni L, Marcato S, Zane L, Patarnello T (2000) Mitochondrial phylogeny of notothenioids: a molecular approach to Antarctic fish evolution and biogeography. *Systematic Biology*, **49**, 114–129.

Barrera-Oro ER, Lagger C (2010) Egg-guarding behaviour in the Antarctic bathydraconid dragonfish *Parachaenichthys charcoti*. *Polar Biology*, **33**, 1585–1587.

Boulenger GA (1902) Pisces. In: *Report on the Collections of Natural History Made in the Antarctic Regions During the Voyage of the 'Southern Cross'*. pp. 174–189, British Museum (Natural History), London, UK.

Brenner M, Buck BH, Cordes S *et al.* (2001) The role of iceberg scours in niche separation within the Antarctic fish genus *Trematomus*. *Polar Biology*, **24**, 502–507.

Chen L, DeVries AL, Cheng CHC (1997) Convergent evolution of antifreeze glycoproteins in Antarctic notothenioid fish and Arctic cod. *Proceedings of the National Academy of Sciences, USA*, **94**, 3817–3822.

Chen WJ, Bonillo C, Lecointre G (1998) Phylogeny of the Channichthyidae (Notothenioidei, Teleostei) based on two mitochondrial genes. In: *Fishes of Antarctica. A Biological Overvie* (eds di Prisco G, Pisano E, Clarke A), pp. 287–298. Springer-Verlag, Milan, Italy.

Clark MS, Fraser KPP, Burns G, Peck LS (2008) The HSP70 heat shock response in the Antarctic fish *Harpagifer antarcticus*. *Polar Biology*, **31**, 171–180.

Clarke A, Aronson RB, Crame JA, Gili JM, Blake DB (2004) Evolution and diversity of the benthic fauna of the Southern Ocean continental shelf. *Antarctic Science*, **16**, 559–568.

Clement O, Ozouf-Costaz C, Lecointre G, Berrebi P (1998) Allozymic polymorphism and phylogeny of the family Channichthyidae. In: *Fishes of Antarctica. A Biological Overvie* (eds di Prisco G, Pisano E, Clarke A), pp. 299–309. Springer Verlag Publishers, Milan, Italy.

Coppes Petricorena ZL, Somero GN (2007) Biochemical adaptations of notothenioid fishes: comparisons between cold temperate South American and New Zealand species and Antarctic species. *Comparative Biochemistry and Physiology*, **147A**, 799–807.

Danley PD, Kocher TD (2001) Speciation in rapidly diverging systems: lessons from Lake Malawi. *Molecular Ecology*, **10**, 1075–1086.

DeNiro MJ, Epstein S (1978) Influence of diet on the distribution of carbon isotopes in animals. *Geochimica et Cosmochimica Acta*, **42**, 495–506.

Derome N, Chen WJ, Dettaï A, Bonillo C, Lecointre G (2002) Phylogeny of Antarctic dragonfishes (Bathydraconidae, Notothenioidei, Teleostei) and related families based on their

anatomy and two mitochondrial genes. *Molecular Phylogenetics and Evolution*, **24**, 139–152.

DeWitt HH, Heemstra PC, Gon O (1990) Nototheniidae. In: *Fishes of the Southern Ocea* (eds Gon O, Heemstra PC), pp. 279–331. J.L.B. Smith Institute of Ichthyology, Grahamstown, South Africa.

Drummond AJ, Rambaut A (2007) BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evolutionary Biology*, **7**, 214.

Dunton KH (2001) δ^{15}N and δ^{13}C measurements of Antarctic Peninsula fauna: trophic relationships and assimilation of benthic seaweeds. *American Zoologist*, **41**, 99–112.

Eakin RR (1990) Artedidraconidae. In: *Fishes of the Southern Ocea* (eds Gon O, Heemstra PC), pp. 332–356. J.L.B. Smith Institute of Ichthyology, Grahamstown, South Africa.

Eakin RR, Eastman JT, Near TJ (2009) A new species and a molecular phylogenetic analysis of the Antarctic fish genus *Pogonophryne* (Notothenioidei: Artedidraconidae). *Copeia*, **4**, 705–713.

Eastman JT (1993) Antarctic fish biology: evolution in a unique environment. *Academic Press*, San Diego, California.

Eastman JT (2000) Antarctic notothenioid fishes as subjects for research in evolutionary biology. *Antarctic Science*, **12**, 276–287.

Eastman JT (2005) The nature of the diversity of Antarctic fishes. *Polar Biology*, **28**, 93–107.

Eastman JT, Clarke A (1998) A comparison of adaptive radiations of Antarctic fish with those of non Antarctic fish. In: *Fishes of Antarctica. A Biological Overvie* (eds di Prisco G, Pisano E, Clarke A), pp. 3–26. Springer-Verlag, Milan, Italy.

Eastman JT, Grande L (1989) Evolution of the Antarctic fish fauna with emphasis on the recent notothenioids. In: *Origins and Evolution of the Antarctic Biot* (ed Crame JA), pp. 241–252. Geological Society. London, UK.

Eastman JT, Grande L (1991) Late Eocene gadiform (Teleostei) skull from Seymour-Island, Antarctic Peninsula. *Antarctic Science*, **3**, 87–95.

Eastman JT, McCune AR (2000) Fishes on the Antarctic continental shelf: evolution of a marine species flock? *Journal of Fish Biology*, **57**, 84–102.

Gavrilets S, Losos JB (2009) Adaptive radiation: contrasting theory with data. *Science*, **323**, 732–737.

Gavrilets S, Vose A (2005) Dynamic patterns of adaptive radiation. *Proceedings of the National Academy of Sciences, USA*, **102**, 18040–18045.

Gernhard T (2008) The conditioned reconstructed process. *Journal of Theoretical Biology*, **253**, 769–778.

Gonzalez-Voyer A, Winberg S, Kolm N (2009) Distinct evolutionary patterns of brain and body size during adaptive radiation. *Evolution*, **63**, 2266–2274.

Gröhsler T (1994) Feeding habits as indicators of ecological niches: investigations of Antarctic fish conduced near Elephant Island in late autumn/winter 1986. *Archive of Fishery and Marine Research*, **42**, 17–34.

Harmon LJ, Schulte JA, Larson A, Losos JB (2003) Tempo and mode of evolutionary radiation in iguanian lizards. *Science*, **301**, 961–964.

Harmon LJ, Weir JT, Brock CD, Glor RE, Challenger W (2008) GEIGER: investigation evolutionary radiations. *Bioinformatics*, **24**, 129–131.

Hobson KA, Welch HE (1992) Determination of trophic relationships within a high Arctic marine food web using δ^{13}C and δ^{15}N analysis. *Marine Ecology Progress Series*, **84**, 9–18.

Hobson KA, Piatt JF, Pitocchelli J (1994) Using stable isotopes to determine seabird trophic relationships. *Journal of Animal Ecology*, **63**, 786–798.

Hureau JC (1985a) Family Channichthyidae – Icefishes. In: *FAO Species Identification Sheets for Fishery Purposes. Southern Ocean. CCAMLR Convention Area. Fishing Areas 48, 58 and 8* (eds Fischer W, Hureau JC), pp. 261–277. Food and Agriculture Organization of the United Nations, Rome, Italy.

Hureau JC (1985b) Family Nototheniidae – Antarctic rock cods. In: *FAO Species Identification Sheets for Fishery Purposes. Southern Ocean. CCAMLR Convention Area. Fishing Areas 48, 58 and 8* (eds Fischer W, Hureau JC), pp. 323–385. Food and Agriculture Organization of the United Nations, Rome, Italy.

Ingram T (2011) Speciation along a depth gradient in a marine adaptive radiation. *Proceedings of the Royal Society B*, **278**, 613–618.

Jones CD, Anderson ME, Balushkin AV *et al.* (2008) Diversity, relative abundance, new locality records and population structure of Antarctic demersal fishes from the northern Scotia Arc islands and Bouvetøya. *Polar Biology*, **31**, 1481–1497.

Jones CD, Damerau M, Deitrich K *et al.* (2009) Demersal finfish survey of the South Orkney Islands. In: *AMLR 2008/2009 Field Season Report: Objectives, Accomplishments and Tentative Conclusion* (ed Van Cise AM), pp. 49–66. U.S. Department of Commerce, La Jolla, CA. NOAA Technical Memorandum NMFS, NOAA-TM-NMFS-SWFSC-445.

Kass RE, Raftery AE (1995) Bayes factors. *Journal of the American Statistical Association*, **90**, 773–795.

Katoh K, Toh H (2008) Recent developments in the MAFFT multiple alignment program. *Briefings in Bioinformatics*, **9**, 286–298.

Kiljunen M, Grey J, Sinisalo T *et al.* (2006) A revised model for lipid-normalizing δ^{13}C values from aquatic organisms, with implications for isotope mixing models. *Journal of Applied Ecology*, **43**, 1213–1222.

Klingenberg CP, Ekau W (1996) A combined morphometric and phylogenetic analysis of an ecomorphological trend: pelagization in Antarctic fishes (Perciformes: Nototheniidae). *Biological Journal of the Linnean Society*, **59**, 143–177.

Kock KH (1992) Antarctic fish and fisheries. *Cambridge University Press*, Cambridge, UK.

Kock KH (2005) Antarctic icefishes (Channichthyidae): a unique family of fishes. A review, part I. *Polar Biology*, **28**, 862–895.

Kock KH, Stransky C (2000) The composition of the coastal fish fauna around Elephant Island (South Shetland Islands, Antarctica). *Polar Biology*, **23**, 825–832.

Kock KH, Pshenichnov LK, DeVries AL (2006) Evidence for egg brooding and parental care in icefish and other nototehnioids in the Southern Ocean. *Antarctic Science*, **18**, 223–227.

Kock KH, Pshenichnov L, Jones CD, Gröger J, Riehl R (2008) The biology of the spiny icefish *Chaenodraco wilsoni* Regan, 1914. *Polar Biology*, **31**, 381–393.

Kuhn KL, Near TJ (2009) Phylogeny of *Trematomus* (Notothenioidei: Nototheniidae) inferred from mitochondrial and nuclear gene sequences. *Antarctic Science*, **21**, 565–570.

Kuhn KL, Near TJ, Jones CD, Eastman JT (2009) Aspects of the biology and population genetics of the Antarctic nototheniid fish *Trematomus nicolai*. *Copeia*, **2**, 320–327.

Kunzmann A, Zimmermann C (1992) *Aethotaxis mitopteryx*, a high Antarctic fish with benthopelagic mode of life. *Marine Ecology Progress Series*, **88**, 33–40.

La Mesa M, Eastman JT, Vacchi M (2004) The role of notothenioid fish in the food web of the Ross Sea shelf waters: a review. *Polar Biology*, **27**, 321–338.

Lartillot N, Hervé P (2006) Computing Bayes Factors using thermodynamic integration. *Systematic Biology*, **55**, 195–207.

Li C, Ortí G, Zhang G, Lu G (2007) A practical approach to phylogenomics: the phylogeny of ray-finned fish (Actinopterygii) as a case study. *BMC Evolutionary Biology*, **7**, 44.

Logan JM, Jardine TD, Miller TJ *et al.* (2008) Lipid correction in carbon and nitrogen stable isotope analyses: comparison of chemical extraction and modeling methods. *Journal of Animal Ecology*, **77**, 838–846.

Lombarte A, Olaso I, Bozzano A (2003) Ecomorphological trends in the Artedidraconidae (Pisces: Perciformes: Notothenioidei) of the Weddell Sea. *Antarctic Science*, **15**, 211–218.

Losos JB (1995) Community evolution in Greater Antillean *Anolis* lizards: phylogenetic patterns and experimental tests. *Philosophical Transactions of the Royal Society of London B*, **349**, 69–75.

Losos JB (2009) *Lizards in an Evolutionary Tree: Ecology and Adaptive Radiation of Anoles*. University of California Press, Berkeley.

Maddison WP, Maddison DW (2009) MESQUITE: a modular system for evolutionary analysis. Version 2.72, Available from http://mesquiteproject.org (last accessed 7 September 2011).

Matschiner M, Hanel R, Salzburger W (2009) Gene flow by larval dispersal in the Antarctic notothenioid fish *Gobionotothen gibberifrons*. *Molecular Ecology*, **18**, 2574–2587.

Matschiner M, Hanel R, Salzburger W (2011) On the origin and trigger of the notothenioid adaptive radiation. *PLoS ONE*, **6**, e18911.

McIntyre PB, Flecker AS (2006) Rapid turnover of tissue nitrogen of primary consumers in tropical freshwaters. *Oecologia*, **148**, 12–21.

Meyer JR, Schoustra SE, Lachapelle J, Kassen R (2011) Overshooting dynamics in a model adaptive radiation. *Proceedings of the Royal Society B*, **278**, 392–398.

Minagawa M, Wada E (1984) Stepwise enrichment of ^{15}N along food chains: further evidence and the relation between δ^{15}N and animal age. *Geochimica et Cosmochimica Acta*, **48**, 1135–1140.

Near TJ (2004) Estimating divergence times of notothenioid fishes using a fossil-calibrated molecular clock. *Antarctic Science*, **16**, 37–44.

Near TJ, Cheng CHC (2008) Phylogenetics of notothenioid fishes (Teleostei: Acanthomorpha): inferences from mitochondrial and nuclear gene sequences. *Molecular Phylogenetics and Evolution*, **47**, 832–840.

Near TJ, Pesavento JJ, Cheng CHC (2003) Mitochondrial DNA, morphology and the phylogenetic relationships of Antarctic icefishes (Notothenioidei: Channichthyidae). *Molecular Phylogenetics and Evolution*, **28**, 87–98.

Near TJ, Pesavento JJ, Cheng CHC (2004) Phylogenetic investigations of Antarctic notothenioid fishes (Perciformes: Notothenioidei) using complete gene sequences of the mitochondrial encoded 16S rRNA. *Molecular Phylogenetics and Evolution*, **32**, 881–891.

Near TJ, Parker SK, Detrich HW (2006) A genomic fossil reveals key steps in hemoglobin loss by the Antarctic icefishes. *Molecular Phylogenetics and Evolution*, **23**, 2008–2016.

Nicholson KE, Glor RE, Kolbe JJ *et al.* (2005) Mainland colonization by island lizards. *Journal of Biogeography*, **32**, 929–938.

Olaso I, Rauschert M, De Broyer C (2000) Trophic ecology of the family Artedidraconidae (Pisces: Osteichthyes) and its impact on the eastern Weddell Sea benthic system. *Marine Ecology Progress Series*, **194**, 143–158.

Pakhomov EA, Bushula T, Kaehler S, Watkins BP, Leslie RW (2006) Structure and distribution of the slope fish community in the vicinity of the sub-Antarctic Prince Edward Archipelago. *Journal of Fish Biology*, **68**, 1834–1866.

Paradis E, Claude J, Strimmer K (2004) APE: analyses of phylogenetics and evolution in R language. *Bioinformatics*, **20**, 289–290.

Patarnello T, Verde C, di Prisco G, Bergelloni L, Zane L (2011) How will fish that evolved at constant sub-zero temperatures cope with global warming? Notothenioids as a case study. *Bioessays*, **33**, 260–268.

Pinnegar JK, Polunin NVC (1999) Differential fractionation of δ^{13}C and δ^{15}N among fish tissues: implications for the study of trophic interactions. *Functional Ecology*, **13**, 225–231.

Place SP, Hofmann GE (2005) Constitutive expression of a stress-inducible heat shock protein gene, *hsp70*, in phylogenetically distant Antarctic fish. *Polar Biology*, **28**, 261–267.

Posada D (2008) JMODELTEST: phylogenetic model averaging. *Molecular Biology and Evolution*, **25**, 1253–1256.

Post DM (2002) Using stable isotopes to estimate trophic position: models, methods and assumptions. *Ecology*, **83**, 703–718.

Quevedo M, Svanbäck R, Eklöv P (2009) Intrapopulation niche partitioning in a generalist predator limits food web connectivity. *Ecology*, **90**, 2263–2274.

R Development Core Team (2009) *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, Available from http://www.R-project.org (last accessed 7 September 2011).

Rambaut A, Drummond AJ (2007) TRACER V.1.5. Available from http://beast.bio.ed.ac.uk/Tracer (last accessed 7 September 2011).

Reid WDK, Clarke S, Collins MA, Belchier M (2007) Distribution and ecology of *Chaenocephalus aceratus* (Channichthyidae) around South Georgia and Shag Rocks (Southern Ocean). *Polar Biology*, **30**, 1523–1533.

Rico C, Turner GF (2002) Extreme microallopatric divergence in cichlid species from Lake Malawi. *Molecular Ecology*, **11**, 1585–1590.

Rowe KC, Aplin KP, Baverstock PR, Moritz C (2011) Recent and rapid speciation within limited morphological disparity in the genus *Rattus*. *Systematic Biology*, **60**, 188–203.

Ruud JT (1954) Vertebrates without erythrocytes and blood pigment. *Nature*, **173**, 848–850.

Salzburger W (2008) To be or not to be a hamlet pair in sympatry. *Molecular Ecology*, **17**, 1397–1400.

Salzburger W (2009) The interaction of sexually and naturally selected traits in the adaptive radiations of cichlid fishes. *Molecular Ecology*, **18**, 169–185.

Salzburger W, Meyer A (2004) The species flocks of East African cichlid fishes: recent advances in molecular phylogenetics and population genetics. *Naturwissenschaften*, **91**, 277–290.

Sanchez S, Dettaï A, Bonillo C *et al.* (2007) Molecular and morphological phylogenies of the Antarctic teleostean family Nototheniidae, with emphasis on the Trematominae. *Polar Biology*, **30**, 155–166.

Schluter D (2000) *The Ecology of Adaptive Radiation*. Oxford University Press, New York.

Schwarz G (1978) Estimating the dimension of a model. *The Annals of Statistics*, **6**, 461–464.

Seehausen O (2006) African cichlid fish: a model system in adaptive radiation research. *Proceedings of the Royal Society B: Biological Sciences*, **273**, 1987–1998.

Shapiro B, Rambaut A, Drummond AJ (2006) Choosing appropriate substitution models for the phylogenetic analysis of protein-coding sequences. *Molecular Biology and Evolution*, **23**, 7–9.

Simpson GG (1953) *The Major Features of Evolution*. Colombia University Press, New York.

Stankovic A, Spalik K, Kamler E, Borsuk P, Weglenski P (2002) Recent origin of sub-Antarctic notothenioids. *Polar Biology*, **25**, 203–205.

Streelman JT, Danley PD (2003) The stages of vertebrate evolutionary radiation. *Trends in Ecology and Evolution*, **18**, 126–131.

Swofford DL (2003) *PAUP*. Phylogenetic Analysis Using Parsimony (*and Other Methods)*, ed. 4.0a. Sinauer Associates, Sunderland, Massachusetts.

Syväranta J, Vesala S, Rask M, Ruuhijärvi J, Jones RI (2008) Evaluating the utility of stable isotope analyses of archived freshwater sample materials. *Hydrobiologia*, **600**, 121–130.

Zwickl DJ (2006) *Genetic algorithm approaches for the phylogenetic analysis of large biological sequence datasets under the maximum-likelihood criterion*. PhD. thesis, The University of Texas at Austin.

S.R. is interested in the diversity of animals, including their evolution and adaptations to the environment. M.Ma. works on the molecular processes underlying adaptive radiation. M.D. is interested in population ecology and evolutionary biology. M.F.L. is specialized on stable isotope biogeochemistry in aquatic environments. R.H. is interested in causes and pathways of adaptation and speciation in the sea. W.S. and M.Mu.'s research focuses on the understanding of the genetic basis of adaptation, evolutionary innovation and animal diversification, using the East Africa' cichlid radiations as main model system. The laboratory's homepage at http://www.evolution.unibas.ch/salzburger provides further details on the group's (research) activities.

Data accessibility

All DNA sequences from this study are available under Gen-Bank accessions: JF264479–JF264516 (cyt *b*); JF264517–JF264554 (*myh6*); JF264555–JF264590 (*Ptr*); and JF264591–JF264629 (*tbr1*). GenBank accession numbers for sequences of other studies are the following: *B. diacanthus* (HM049936; HM050034; HM050153; HM050214); *Eleginops maclovinus* (DQ526429; HM050045; HM050163; HM050225); *N. coriiceps* (HM050183); *P. urvillii* (HM049963; HM050074; HM050195; HM050258); *P. scotti* (HM049962; HM050072; HM050193); and *T. newnesi* (HM050204) (see Table S4, Supporting information). All stable isotope values are given in Table S5, (Supporting information).

Supporting information

Additional supporting information may be found in the online version of this article:

Fig. S1 Maximum-likelihood tree based on the codon position-specific partitioning with numbered nodes (1–19).

Table S1 Antarctic notothenioid samples with corresponding collection id (Table S2) and sample size (*n*) for stable isotope analysis.

Table S2 Collection id for all Antarctic notothenioid samples.

Table S3 Lifestyle and feeding for all included Antarctic notothenioid species.

Table S4 GenBank accession numbers for all used samples.

Table S5 Stable isotope values of all investigated species.

Data S1 Discussion of stable isotope analysis results of individual species.

Please note: Wiley-Blackwell are not responsible for the content or functionality of any supporting information supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.

Supporting Information

Parallel ecological diversification in Antarctic notothenioid fishes as evidence for adaptive radiation

Sereina Rutschmann[∗], **Michael Matschiner**[∗], **Malte Damerau**[†], **Moritz F. Lehmann**[‡], **Reinhold Hanel**[†], **and Walter Salzburger**[∗]

Addresses:

[∗]Zoological Institute, University of Basel, Vesalgasse 1, CH-4051 Basel, Switzerland

[†]Institute of Fisheries Ecology, Johann Heinrich von Thünen-Institute, Federal Research Institute for Rural Areas, Forestry and Fisheries, Palmaille 9, D-22767 Hamburg, Germany

[‡]Institute of Environmental Geosciences, University of Basel, Bernoullistrasse 30, CH-4056 Basel, Switzerland

Correspondence:

Walter Salzburger, Fax: +41 61 267 03 01, E-mail: walter.salzburger@unibas.ch, and Reinhold Hanel, Fax: +49 40 38 90 52 61, E-mail: reinhold.hanel@vti.bund.de

1 Figures

Fig. S1: ML tree based on the codon position-specific partitioning with numbered nodes (1-19). BS and BPP values for the corresponding nodes in ML and BI analyses are listed in the table below. All species are coloured according to family (see Fig. 2).
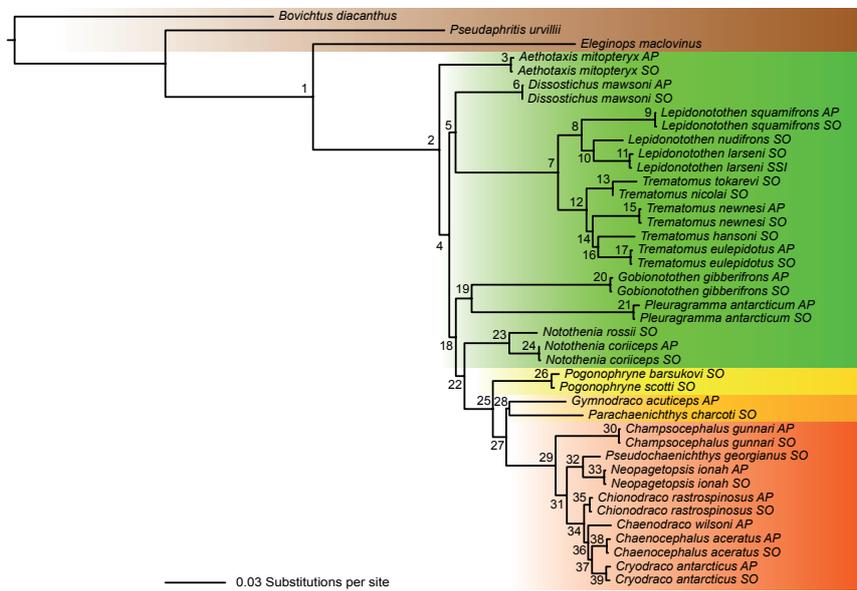
Node number	BS partition by codon position 1-2 ind. per species	BPP partition by codon position 1-2 ind. per species	BS partition by gene 1-2 ind. per species	BPP partition by codon position 1 ind. per species
1	100	*	100	*
2	100	*	100	*
3	100	1.00	100	-
4	26	0.38	**	0.39
5	11	0.30	**	**
6	100	1.00	100	-
7	100	1.00	100	1.00
8	96	1.00	95	1.00
9	100	1.00	100	-
10	75	0.99	82	0.98
11	100	1.00	100	-
12	100	1.00	98	1.00
13	100	1.00	100	1.00
14	51	0.70	**	**
15	100	1.00	100	-
16	53	0.80	46	0.61
17	100	1.00	100	-
18	30	0.73	**	0.68
19	40	0.90	**	0.80
20	100	1.00	100	-
21	100	1.00	100	-
22	65	0.99	72	0.99
23	100	1.00	100	1.00
24	100	1.00	100	-
25	100	1.00	100	1.00
26	100	1.00	100	1.00
27	86	1.00	77	1.00
28	35	0.67	34	0.76
29	100	1.00	100	1.00
30	100	1.00	100	-
31	58	1.00	57	1.00
32	78	1.00	79	1.00
33	100	1.00	100	-
34	94	1.00	87	1.00
35	93	1.00	92	-
36	68	0.94	66	0.94
37	66	1.00	78	0.99
38	100	1.00	100	-
39	100	1.00	100	-

* constrained as monophyletic; ** node not present due to topological differences; - node not present due to exclusion of taxa.

2   Tables

Tab. S1: Antarctic notothenioid samples with corresponding collection id (Table S2) and sample size (n) for stable isotope analysis

Samples	Collection id (n)
Nototheniidae	
Aethotaxis mitopteryx	56 (4), 11 (4), 47 (2), 49 (1)
Dissostichus mawsoni	56 (2), 11 (2), 17 (3)
Gobionotothen gibberifrons	56 (10), 10 (10)
Lepidonotothen larseni	9 (2), 10 (4), 36 (4), 1 (6), 2 (4)
Lepidonotothen nudifrons	17 (2), 18 (1), 40 (2), 44 (2), 46 (1), 53 (2)
Lepidonotothen squamifrons	54 (10), 16 (4), 17 (6)
Notothenia coriiceps	7 (10), 18 (3), 22 (1), 38 (1), 40(1), 41 (4), 50 (1)
Notothenia rossii	12 (1), 17 (5), 21 (2), 29 (1), 32 (1), 51 (1)
Pleuragramma antarcticum	56 (10), 42 (1), 49 (9)
Trematomus eulepidotus	54 (5), 55 (2), 56 (3), 20 (1), 22 (8), 28 (1)
Trematomus hansoni	15 (1), 16 (1), 23 (5), 24 (1), 26 (1), 27 (1), 30 (1)
Trematomus newnesi	8 (10), 13 (1), 39 (1), 41 (8)
Trematomus nicolai	11 (2), 32 (3), 37 (1)
Trematomus tokarevi	31 (1), 33 (2), 36 (2), 38 (1), 48 (3), 52 (1), n.a. (1)
Artedidraconidae	
Pogonophryne barsukovi	20 (2), 35 (1), 42 (1), 48 (2), 49 (2)
Pogonophryne scotti	25 (1), 27 (1), 34 (6), 42 (2)
Bathydraconidae	
Gymnodraco acuticeps	54 (1), 56 (14)
Parachaenichthys charcoti	13 (1), 17 (1), 40 (3), 43 (2), 45 (1), 53 (3)
Channichthyidae	
Chaenocephalus aceratus	3 (9), 5 (1), 10 (10)
Chaenodraco wilsoni	56 (10)
Champsocephalus gunnari	4 (5), 6 (6), 27 (8), 51 (2)
Chionodraco rastrospinosus	55 (1), 56 (9), 19 (8), 22 (2)
Cryodraco antarcticus	56 (10), 23 (3), 28 (7)
Neopagetopsis ionah	56 (6), 11 (3), 47 (1), 49 (1), 52 (1)
Pseudochaenichthys georgianus	14 (1), 16 (9)

Tab. S2: Collection id for all Antarctic notothenioid samples. AP, Antarctic Peninsula, SO, South Orkney Islands, SSI, South Sandwich Islands, with mean values for latitude, longitude and depth.

Id	Site	Latitude	Longitude	Depth	Id	Site	Latitude	Longitude	Depth
1	SSI	56°19'18"S	27°27'02"W	330 m	29	SO	61°33'52"S	45°15'32"W	259 m
2	SSI	58°27'11"S	26°12'51"W	270 m	30	SO	61°30'49"S	44°32'42"W	380 m
3	AP	61°20'44"S	55°15'23"W	350 m	31	SO	61°36'25"S	44°24'23"W	390 m
4	AP	61°15'23"S	54°50'10"W	152 m	32	SO	61°13'00"S	45°55'49"W	240 m
5	AP	60°58'59"S	55°11'08"W	299 m	33	SO	61°49'12"S	46°11'30"W	453 m
6	AP	60°59'19"S	55°53'18"W	203 m	34	SO	61°43'08"S	45°49'03"W	398 m
7	AP	61°00'20"S	55°43'40"W	96 m	35	SO	61°14'04"S	46°23'16"W	274 m
8	AP	62°33'48"S	55°41'52"W	162 m	36	SO	61°25'44"S	46°09'28"W	352 m
9	SO	60°26'15"S	46°17'46"W	142 m	37	SO	60°54'59"S	45°37'17"W	294 m
10	SO	60°25'46"S	46°25'07"W	142 m	38	SO	60°55'18"S	45°51'09"W	208 m
11	SO	60°30'53"S	46°35'08"W	457 m	39	SO	60°53'57"S	46°03'26"W	187 m
12	SO	60°24'06"S	46°30'57"W	220 m	40	SO	60°46'03"S	46°16'10"W	150 m
13	SO	60°28'58"S	46°21'53"W	106 m	41	SO	60°37'59"S	46°31'26"W	130 m
14	SO	60°26'37"S	45°38'53"W	237 m	42	SO	61°45'22"S	45°26'20"W	375 m
15	SO	60°26'32"S	45°16'51"W	497 m	43	SO	60°39'11"S	46°16'52"W	104 m
16	SO	60°29'22"S	45°08'06"W	350 m	44	SO	60°45'10"S	44°13'00"W	166 m
17	SO	60°31'53"S	44°45'24"W	310 m	45	SO	60°42'49"S	46°00'02"W	96 m
18	SO	60°49'16"S	44°29'27"W	172 m	46	SO	60°30'22"S	47°23'22"W	657 m
19	SO	60°36'31"S	44°20'33"W	211 m	47	SO	61°03'16"S	46°49'16"W	764 m
20	SO	61°03'06"S	42°49'45"W	425 m	48	SO	61°36'19"S	47°00'49"W	629 m
21	SO	60°51'29"S	42°52'18"W	359 m	49	SO	61°52'30"S	46°43'21"W	750 m
22	SO	60°52'13"S	43°11'46"W	336 m	50	SO	61°16'02"S	44°54'32"W	322 m
23	SO	61°17'30"S	43°05'25"W	469 m	51	SO	60°50'07"S	43°48'18"W	221 m
24	SO	61°08'57"S	43°31'56"W	455 m	52	SO	60°36'04"S	44°45'52"W	118 m
25	SO	61°02'38"S	44°42'50"W	254 m	53	SO	60°48'03"S	45°53'35"W	128 m
26	SO	61°07'55"S	44°35'22"W	314 m	54	AP	63°01'05"S	52°21'56"W	623 m
27	SO	61°08'01"S	44°13'59"W	337 m	55	AP	62°35'14"S	53°46'22"W	731 m
28	SO	61°11'05"S	43°56'44"W	426 m	56	AP	63°14'18"S	59°25'13"W	759 m

Tab. S3: Lifestyle and feeding for all included Antarctic notothenioid species. The listed feeding ecology was inferred from stomach content analyses (except for reference e, where it is unclear), and may not reflect the full diet.

Samples	Lifestyle	Feeding
Nototheniidae		
Aethotaxis mitopteryx	pelagic[b,d,g,h], benthopelagic[l]	gammarid, amphipod[l]
Dissostichus mawsoni	pelagic[d,h]	fish, misc. invert.[f]
Gobionotothen gibberifrons	benthic[d,g]	misc. invert., polychaete, salp, ophiuroid, krill, amphipod, isopod[f]
Lepidonotothen larseni	semipelagic[d]	misc. invert., krill, salp, mysid, amphipod[f]
Lepidonotothen nudifrons	benthic[d,h]	misc. invert., amphipod, polychaete, echinoderm, isopod, krill[f]
Lepidonotothen squamifrons	benthic[d]	salp, misc. invert., krill, fish, amphipod, polychaete, isopod[f]
Notothenia coriiceps	benthic[h]	krill, fish, misc. invert., salp[f]
Notothenia rossii	semipelagic[d]	fish, krill, salp, misc. invert., amphipod[f]
Pleuragramma antarcticum	pelagic[b,d,h]	krill, misc. invert.[f]
Trematomus eulepidotus	epibenthic[b,d,g]	krill, misc. invert., salp, fish, mysid, isopod[f]
Trematomus hansoni	benthic[d,g]	fish, misc. invert., krill, salp, octopus, isopod, mysid, amphipod[f]
Trematomus newnesi	cryopelagic[d]	krill, misc. invert., fish[f]
Trematomus nicolai	benthic[b,d,g,k,m], benthopelagic[a]	fish[f]
Trematomus tokarevi	benthic[m]	amphipod[f]
Artedidraconidae		
Pogonophryne barsukovi	benthic[n]	krill[f]
Pogonophryne scotti	benthic[d,n]	krill, fish, misc. invert., isopod[f]
Bathydraconidae		
Gymnodraco acuticeps	benthic[d]	krill[f]
Parachaenichthys charcoti	benthic[d]	fish, krill, misc. invert.[f]
Channichthyidae		
Chaenocephalus aceratus	benthic[d,i]	fish, krill, misc. invert., mysid[f]
Chaenodraco wilsoni	pelagic[j]	krill[e]
Champsocephalus gunnari	pelagic[d,i]	krill, fish[f]
Chionodraco rastrospinosus	benthic[d], benthopelagic[e]	fish, krill, misc. invert.[f]
Cryodraco antarcticus	pelagic[d], benthic[i]	fish, misc. invert., mysid, krill, amphipod[f]
Neopagetopsis ionah	pelagic[i]	fish, krill, misc. invert.[f]
Pseudochaenichthys georgianus	benthic[d,i], semipelagic[d]	fish, krill, misc. invert., mysid[f]

[a]Brenner *et al.* 2001; [b]DeWitt *et al.* 1990; [c]Eakin 1990; [d]Eastman 1993; [e]Hureau 1985b; [f]Jones *et al.* 2009; [g]Klingenberg & Ekau 1996; [h]Kock 1992; [i]Kock 2005; [j]Kock *et al.* 2008; [k]Kuhn *et al.* 2009; [l]Kunzmann & Zimmermann 1992; [m]La Mesa *et al.* 2004; [n]Lombarte *et al.* 2003.

Tab. S4: GenBank accession numbers for all used samples. AP, Antarctic Peninsula, SO, South Orkney Islands, SSI, South Sandwich Islands.

Species	Location	cyt *b*	myh6	Ptr	tbr1
Aethotaxis mitopteryx	AP	JF264479	JF264517	JF264555	JF264591
Aethotaxis mitopteryx	SO	JF264480	JF264518	JF264556	JF264592
Chaenocephalus aceratus	AP	JF264481	JF264519	JF264557	JF264593
Chaenocephalus aceratus	SO	JF264482	JF264520	JF264558	JF264594
Champsocephalus gunnari	AP	JF264483	JF264521	JF264559	JF264595
Champsocephalus gunnari	SO	JF264484	JF264522	JF264560	JF264596
Chaenodraco wilsoni	AP	JF264485	JF264525	JF264561	JF264597
Chionodraco rastrospinosus	AP	JF264486	JF264523	JF264562	JF264598
Chionodraco rastrospinosus	SO	JF264487	JF264524	JF264563	JF264599
Cryodraco antarcticus	AP	JF264488	JF264526	JF264564	JF264600
Cryodraco antarcticus	SO	JF264489	JF264527	JF264565	JF264601
Dissostichus mawsoni	AP	JF264490	JF264528	JF264566	JF264602
Dissostichus mawsoni	SO	JF264491	JF264529	JF264567	JF264603
Gobionotothen gibberifrons	AP	JF264492	JF264530	JF264568	JF264604
Gobionotothen gibberifrons	SO	JF264493	JF264531	JF264569	JF264605
Gymnodraco acuticeps	AP	JF264494	JF264532	JF264570	JF264606
Lepidonotothen larseni	SO	JF264495	JF264533	JF264571	JF264607
Lepidonotothen larseni	SS	JF264496	JF264534	JF264572	JF264608
Lepidonotothen nudifrons	SO	JF264497	JF264535	JF264573	JF264609
Lepidonotothen squamifrons	AP	JF264498	JF264536	JF264574	JF264610
Lepidonotothen squamifrons	SO	JF264499	JF264537	JF264575	JF264611
Neopagetopsis ionah	AP	JF264500	JF264538	JF264576	JF264612
Neopagetopsis ionah	SO	JF264501	JF264539	JF264577	JF264613
Notothenia coriiceps	AP	JF264503	JF264540	HM050183	JF264614
Notothenia coriiceps	SO	JF264502	JF264541	JF264578	JF264615
Notothenia rossii	SO	JF264504	JF264542	JF264579	JF264616
Parachaenichthys charcoti	SO	JF264505	JF264543	JF264580	JF264617
Pleuragramma antarcticum	AP	JF264506	JF264544	JF264581	JF264618
Pleuragramma antarcticum	SO	JF264507	JF264545	JF264582	JF264619
Pogonophryne barsukovi	SO	JF264508	JF264546	JF264583	JF264620
Pogonophryne scotti	SO	HM049962	HM050072	HM050193	JF264621
Pseudochaenichthys georgianus	SO	JF264509	JF264547	JF264584	JF264622
Trematomus eulepidotus	AP	JF264510	JF264548	JF264585	JF264623
Trematomus eulepidotus	SO	JF264511	JF264549	JF264586	JF264624
Trematomus hansoni	SO	JF264512	JF264550	JF264587	JF264625
Trematomus newnesi	AP	JF264513	JF264551	HM050204	JF264626
Trematomus newnesi	SO	JF264514	JF264552	JF264588	JF264627
Trematomus nicolai	SO	JF264515	JF264553	JF264589	JF264628
Trematomus tokarevi	SO	JF264516	JF264554	JF264590	JF264629

3   Text

Text S1: Discussion of SIA results of individual species.

Our results are only partly congruent with the lifestyles and feeding reports based on stomach content analyses (Fig. 4 and Tables 1, S3, Supporting Information). *Chionodraco rastrospinosus*, for example, has been described as a benthic (Eastman 1993) or benthopelagic (Hureau 1985) species but shows one of the lowest δ^{13}C values, suggesting a pelagic lifestyle. Our SIA results are, however, consistent with buoyancy assessments by Eastman & Sidell (2002), who reported low weight in seawater for *C. rastrospinosus*, which is indicative of a pelagic lifestyle. We also obtain conflicting results for *T. nicolai* and *T. tokarevi*, which are considered as benthic or benthopelagic species and as deep-water species, respectively (see Table 1 and references therein; Andriashev 1978). Our data suggest that both are pelagic species residing at low TLs (Fig. 4). Carbon isotopic signatures of *A. mitopteryx* indicate feeding on higher TL in disagreement with previous reports (Table S3, Supporting Information). Finally, *D. mawsoni* displays the greatest variation in δ^{13}C signatures and the highest mean δ^{15}N value, indicating a broad range of habitats along the benthic-pelagic axis and piscivorous feeding. This agrees with its characterization as one of the largest notothenioid species (up to 1.75 m in length) and a top predator (DeWitt *et al.* 1990). It has been suggested that individual specialization to different habitats is more common in predators due to higher intraspecific competition (Quevedo *et al.* 2009).

References

Andriashev AP (1978) *Trematomus tokarevi*, a new species of the family Nototheniidae (Pisces) from the abyssal waters near Antarctica. *Journal of Ichthyology*, **18**, 521–526.

DeWitt HH, Heemstra PC, Gon O (1990) Nototheniidae. In: *Fishes of the Southern Ocean.* (eds. Gon O, Heemstra PC). J.L.B. Smith Institute of Ichthyology, Grahamstown, South Africa.

Eastman JT (1993) *Antarctic Fish Biology: Evolution in a Unique Environment.* Academic Press, San Diego, USA.

Eastman JT, Sidell BD (2002) Measurements of buoyancy for some Antarctic notothenioid fishes from the South Shetland Islands. *Polar Biology*, **25**, 753–760.

Hureau JC (1985) Family Nototheniidae - Antarctic rock cods. In: *FAO Species Identification Sheets for Fishery Purposes. Southern Ocean. CCAMLR Convention Area. Fishing Areas 48, 58 and 88.* (eds. Fischer W, Hureau JC), 2, pp. 323-385. Food and Agriculture Organization of the United Nations, Rome, Italy.

Quevedo M, Svanbäck R, Eklöv P (2009) Intrapopulation niche partitioning in a generalist predator limits food web connectivity. *Ecology*, **90**, 2263–2274.

IV   Fieldwork

During the course of this doctoral work, samples were acquired in three fieldwork expeditions. In September 2009, potential sister groups of notothenioid fishes, including zoarcids and scorpaenoids, were collected as part of a two week expedition with RV Jan Mayen to the fjords of Svalbard and into sea ice north of Svalbard. In November 2009, notothenioids of the basal South American lineage *Eleginops maclovius* were sampled near Puerto Deseado, Argentina for DNA, RNA, and stomach content analysis. The largest number of notothenioid samples, however, was obtained during Antarctic expedition ANT-XXVII/3 with RV Polarstern in February-April 2011. Over 1500 tissue samples of 48 notothenioid and related species were taken at 56 stations between South Georgia, the South Orkney Islands, the South Shetland Islands, along the Antarctic Peninsula, in the Weddel Sea, and near Bouvet Island. These samples are currently used in ongoing efforts to generate a genus-level molecular phylogeny of notothenioid fishes, that is time-calibrated with the fossil record of non-notothenioid outgroups.

Sampling details of expedition ANT-XXVII/3 will be published as two seperate reports in *Berichte zur Polar- und Meeresforschung*:

4.1   Mintenbeck K, Damerau M, Hirse T, Knust R, Koschnick N, Matschiner M, Rath L: Biodiversity and zoogeography of demersal fish. In: The expedition of the research vessel "Polarstern" to the Antarctic in 2011 (ANT-XXVII/3) (Ed. Knust R). *Ber Polar Meeresforschg*. In press.

4.2   Damerau M, Hanel R, Matschiner M, Salzburger W: 3.2 Notothenioidei. In: The expedition of the research vessel "Polarstern" to the Antarctic in 2011 (ANT-XXVII/3) (Ed. Knust R). *Ber Polar Meeresforschg*. In press.

4.1 Biodiversity and zoogeography of demersal fish

Mintenbeck K, Damerau M, Hirse T, Knust R, Koschnick N, Matschiner M, Rath L

In: *The expedition of the research vessel "Polarstern" to the Antarctic in 2011 (ANT-XXVII/3).* In press.

4.1.1 Report: p. 360 - 362

Personal contribution:

In the report of Mintenbeck et al., I contributed to fieldwork and writing of the manuscript.

3.1 Biodiversity and zoogeography of demersal fish

Katja Mintenbeck (AWI), Malte Damerau (vTI), Timo Hirse (AWI), Rainer Knust (AWI), Nils Koschnick (AWI), Michael Matschiner (UNIBAS), Lena Rath (UHH-IHF)

Objectives

The fish fauna of the Southern Ocean is mainly composed of primarily bottom dwelling species belonging to the perciform suborder Notothenioidei. Composition of the demersal fish community, however, varies between different regions. There are three major abiotic factors determining the zoogeography and dispersal of species and thus composition and diversity of communities: geographical distance, oceanic currents, and ambient water temperature. The high Antarctic shelf is relatively isolated from other continental shelves, but islands such as those found along the Scotia Arc might serve as stepping stones for north- and/or southwards dispersal of fish species. Water temperature at the sea floor differs strongly between sub-Antarctic and high Antarctic shelf areas and ranges from about +4°C on the shelves of the northern Scotia Arc to -1.8°C on the high Antarctic shelf. The zoogeography of Southern Ocean species thus most likely also reflects their thermal tolerance window. Studies conducted during earlier expeditions already indicated that some species are widely distributed and are apparently able to cope with a wide temperature range while others are strongly limited in their latitudinal distribution. To investigate the zoogeography of species and the composition and diversity of communities, the demersal fish fauna was sampled from different parts of the Southern Ocean. These ecological studies are closely linked to studies on genetic population structure and experimental studies on physiological performance of different fish species depending on water temperature.

Work at sea

In total, 24 bottom trawls were carried out during the expedition. Sampling areas included the islands of the Scotia Arc, King George Island, the western (Larsen A, B, and C) and eastern Weddell Sea as well as Bouvet Island. Sampling was carried out on the shelf in water depth between 250 and 470m. Species were identified, and individuals were measured and weighted. Sex, maturity and liver weight were determined from subsamples. From common species otoliths were taken for age determinations, stomachs for diet analyses, and tissue samples from white muscle for stable isotope analyses and genetic studies. To account for differences in trawled area biomass, data of each haul were standardized to an area of 1000m^2. Mean biomass was calculated for each of the study areas. For the spatial comparison of fish communities, % biomass contribution of each species, Shannon diversity (H'), Pilous' evenness (J') and Bray Curtis similarity were calculated.

Preliminary results

Mean demersal fish biomass was high on the shelves of the Scotia Arc Islands, intermediate at King George Island, the eastern Weddell Sea and Bouvet Island and very low in the entire Larsen area (Table 3.1.1).

Table 3.1.1 Composition, mean biomass, diversity and evenness of demersal fish communities in the study areas

Area	No. of hauls	Species No.	Shannon Diversity H'	Evenness J' (Pilou)	Mean fish biomass [g *1000m^2]
Burdwood Bank	2	7	0.46	0.24	766.73
Shag Rocks	2	9	0.63	0.29	3059.03
South Georgia	2	10	1,75	0.76	3306.02
South Orkneys	2	16	1,88	0.68	997.03
King George Island	2	11	0,44	0.19	143.86
Larsen A	2	8	1.38	0.66	15.19
Larsen B	1	6	0,73	0.40	29.97

Larsen C	2	8	1.73	0.83	22.44
Eastern Weddell Sea	7	21	2.62	0.86	106.35
Bouvet Island	2	4	0.89	0.64	437.81

At Burdwood Bank the fish community was dominated by typical low latitude species such as *Micromesistius australis* (Gadidae) and the nototheniids *Patagonotothen guntheri* and *Dissostichus eleginoides*. At Shag Rocks 87% of fish biomass was contributed by *Lepidonotothen squamifrons* and 11% by *D. eleginoides*. Diversity and evenness were low in this northern part of the Scotia Arc. At South Georgia the catches were dominated by large individuals of the marbled notothen, *Notothenia rossii*. Other species contributing each more than 3% to overall biomass were the icefishes *Chaenocephalus aceratus*, *Champsocephalus gunnari* and *Pseudochaenichthys georgianus*, as well as the nototheniids *D. eleginoides* and *Gobionotothen gibberifrons*. At the South Orkneys *C. aceratus*, *C. gunnari* and *P. georgianus* were still represented with an overall biomass of 15%, but the dominating species in this area were *G. gibberifrons* and *Chionodraco rastrospinosus*. At King George Island diversity and evenness were low due to the high biomass of the Antarctic silverfish, *Pleuragramma antarcticum*, which accounted for 82% of fish biomass. Other typical components of the King George Island fish community were *C. rastrospinosus*, *Cryodraco antarcticum*, *Chaenodraco wilsoni*, *G. gibberifrons*, *L. squamifrons* and *Trematomus hansoni*.

In Larsen A and B the catches were dominated by *Gymnodraco acuticeps*, *P. antarcticum*, *Trematomus eulepidotus* and *T. scotti*. *Gobionotothen gibberfrons* accounted for 37% in Larsen A but was absent in Larsen B. The fish community in Larsen C was similar to Larsen B, with *C. wilsoni*, the cryopelagic fish *Pagothenia borchgrevinki,* and *T. hansoni* additionally contributing more than 10% each to the overall biomass. Highest species number, diversity and evenness were found on the eastern Weddell Sea shelf. In this area the fish community was dominated by the icefish *C. antarcticum* and several *Trematomus* species. Small *Artedidraco* spp. were common but didn't contribute much to overall biomass. At Bouvet Island only 4 fish species were found, *Lepidonotothen kempi*, *L. larseni*, *C. gunnari* and *C. aceratus*, with the latter 3 species dominating the community.

Cluster analysis of similarity between the different fish communities revealed two large clusters, one sub Antarctic and one high Antarctic cluster, with each two subgroups (Fig. 3.1.1). Within the high Antarctic cluster the communities of the Larsen area are most similar among each other. Similarity of this cluster to the King George Island and eastern Weddell Sea fish communities is comparatively low.
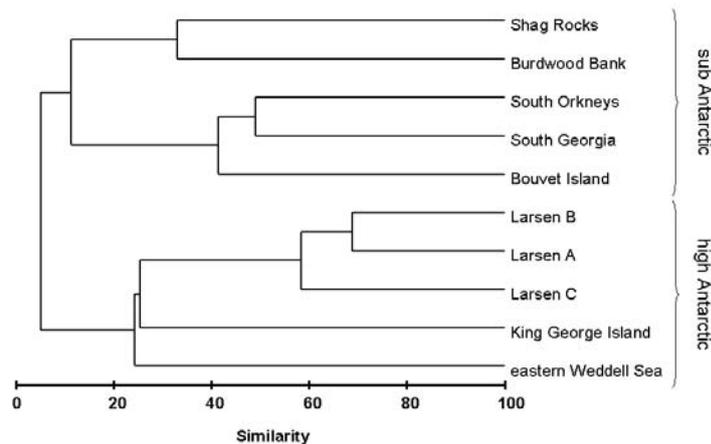
Fig. 3.1.1 Dendrogram of similarity (Bray Curtis similarity, 4th root transformation) of fish community composition between the different study areas

Within the sub-Antarctic cluster the fish community of Shag Rocks was most similar to that found at Burdwood Bank. A second group within the sub-Antarctic cluster includes the fish communities of the South Orkneys, South Georgia and Bouvet Island. All species found at Bouvet Island (see above) are also common in the Scotia Arc, and the similarity of the Bouvet fish fauna to the two more westerly two island shelves of South Georgia and the South Orkneys might be the result of a faunal drift via the Antarctic Circumpolar Current. However, whether ongoing gene flow occurs between fish populations along the Scotia Arc and between the South Orkneys/South Georgia and Bouvet Island or whether the populations are completely separated is still unknown for most species, and will be investigated as part of the project described in chapter 3.2.

Based on the results of the studies on species' zoogeography we could identify *Patagonotothen guntheri* as a typical sub-Antarctic species with its distribution limited to the northern (warmer) part of the Scotia Arc. In contrast, the distributional range of *Trematomus* spp. is limited to the cold waters of high Antarctic latitudes. *Lepidonotothen squamifrons* is a species with a wide latitudinal distribution from Shag Rocks down to King George Island, and thus occupies water masses with strongly differing temperatures. These three nototheniid fish species were chosen for studies on temperature dependent physiological performance.

4.2 Notothenioidei

Damerau M, Hanel R, Matschiner M, Salzburger W

In: *The expedition of the research vessel "Polarstern" to the Antarctic in 2011 (ANT-XXVII/3).* In press.

4.2.1 Report: p. 364

Personal contribution:

In the report of Damerau et al., I contributed to fieldwork and writing of the manuscript.

3.2 Notothenioidei

Malte Damerau (vTI), Reinhold Hanel* (vTI), Michael Matschiner (UNIBAS), Walter Salzburger* (UNIBAS) *not on board

Objectives

Since the cooling of the Southern Ocean approximately 20 million years ago, a unique ichthyofauna evolved on the shelves of the Antarctic continent and adjacent islands showing low species diversity and high levels of endemism. The majority of fish are bottom dwelling and belong to the suborder Notothenioidei (Perciformes). Their larvae usually develop pelagically over an extended period of several months. During this time, larvae may be dispersed over large distances by strong prevailing current systems, including the Antarctic Circumpolar Current that encircles the Antarctic continent. Indeed, high genetic homogeneity and low differentiation among populations is often found even for species with circum-Antarctic distributions, highlighting the role of protracted larval phases for gene flow. On the other hand, larvae are often found to be retained in neritic waters by local gyres. Also, oceanic fronts and strong currents may act as barriers hindering gene flow by larval dispersal or migration.

In our study, we compare the genetic population structures along the Scotia Arc region of selected notothenioid species with differing life-history strategies and larval durations to elucidate the role of protracted larval phases and prevailing current systems in population structuring and, moreover, the influence of ecology and gene flow on the ongoing adaptive radiation of notothenioids in the Southern Ocean. We further prepare an extensive phylogenetic dataset to assess notothenioid relationships at the genus-level.

Work at sea

Demersal notothenioids were collected at Burdwood Bank, Shag Rocks, South Georgia, South Orkneys, South Shetlands, Weddell Sea (Larsen and Bendex areas) and Bouvet Island employing bottom trawls and Agassiz trawls. In addition, pelagic species and larvae were caught using Rectangular Midwater trawls and bentho-pelagic nets. After each haul, fish were identified, sorted by species and individual biological data (total length, standard length, weight and if possible gutted weight, liver weight, sex, maturity stage, gonad weight) taken. For population genetic and phylogenetic analyses, muscle tissue of up to 80 specimen per species and area was sampled and stored in 96% ethanol. For selected species, additional tissue samples of muscle, liver and brain were taken for RNA analyses and stored at -20°C in RNA later solution. Otoliths of the blackfin icefish *Chaenocephalus aceratus* were collected at South Georgia for later age determination.

Preliminary results

During ANT-XXVII/3 we collected tissue at 56 stations for genetic analyses from 1500 individuals and 48 different species of the four notothenioid families Artedidraconidae (10 species), Bathydraconidae (6), Channichthyidae (13) and Nototheniidae (19). In addition, tissue was sampled from gadids (1), liparids (1), macrourids (1), muraenolepids (4) and myctophids (1) that can be used as outgroups in phylogenetic analyses. Overall, we sampled 1519 individuals that represent important contributions to our existing population genetic and phylogenetic sample sets. The obtained specimens will allow comparative population genetic analyses of *Gobionotothen gibberifrons*, *Champsocephalus gunnari*, *Chaenocephalus aceratus*, *Pseudochaenichthys georgianus*, and *Chionodraco rastrospinosus* between the Antarctic Peninsula, the South Orkney Islands, and South Georgia. In addition, we were able to add eight notothenioid genera to our sample set, (*Dacodraco*, *Akarotaxis*, *Prionodraco*, *Histiodraco*, *Dolloidraco*, *Cygnodraco*, *Patagonotothen*, and *Cottoperca*) which will allow more robust phylogenetic inference and molecular dating of the notothenioid radiation.

V Discussion

This doctoral work has led to a number of important insights into the adaptive radiation of notothenioid fishes. It has shown that extant notothenioid species are characterized by high levels of gene flow by larval dispersal even between distant populations, if these are connected by oceanic currents (Matschiner et al. 2009). A comparative analysis including the most abundant species of the Scotia Arc has suggested that high levels of gene flow are not limited to few notothenioid species, but that it represents a common phenomenon among Notothenioidei (Damerau et al., submitted). Thus, the question remains how notothenioids could have obtained their present species diversity despite strong genetic homogenization. There are three possible solutions to this conundrum.

First, the notothenioid species richness could have evolved before the onset of strong oceanic currents, or before a pelagic larval stage became a common characteristic of notothenioid life histories. Both options are unlikely. The divergence date estimates for a large number of notothenioid species are far younger than those for the onset of the ACC (Matschiner et al. 2011, Rutschmann et al. 2011), and pelagic larval stages are very common among notothenioid families including the ancestral bovichtids (Eastman 1993, Balushkin 2000), and thus seem to have evolved long before the Antarctic diversification of notothenioids.

Second, gene flow by larval dispersal may be a temporary phenomenon, interrupted by periods that are sufficiently long for speciation events. Such interruptions could possibly be provided by glacial maxima, during which the notothenioid habitat is reduced to a limited number of ice-free pockets of the continental shelves, and to sub-Antarctic islands (Barnes & Conlan 2007). During these periods, notothenioid population sizes may be smaller, separated by longer distances, and the strength of the ACC may be reduced due to glaciation of the Scotia Ridge. However, whether glacial maxima are long enough (~100 000 yr) for the accumulation of sufficient genetic diferences and for speciation events remains speculative. If glacial maxima contribute to notothenioid species richness, then the speciation rates of notothenioids should have increased since the onset of glacial maxima in the Early Pleistocene. This hypothesis can be tested with a time-calibrated species level phylogeny of notothenioids.

Third, it is possible that gene flow does not necessarily - as it is usually assumed - impede, but instead stimulate speciation by allowing adaptive alleles to spread throughout a species range to those populations that profit most from it (Nosil 2008, Niemiller et al. 2008). In fact, mathematical models have shown that intermediate levels of gene flow allow the greatest adaptive divergence (Gavrilets & Vose 2005, Garant et al. 2007). How the special properties of unidirectional gene flow affect adaptation remains unclear, but could be tested by extensions of the applied mathematical models.

Further insights into the adaptive radiation of Antarctic notothenioids are provided by the phylogenetic analyses that are part of this doctoral work. Using non-notothenioid fossil calibrations and a relaxed molecular clock, it was shown that antifreeze glycoproteins emerged at a time of cooling of Antarctic water, which supports their role as a key innovation that endowed ecological opportunity to notothenioids, and thus triggered their radiation. According to the ecological theory of adaptive radiation, speciation rates are initially high, spurred by

divergent natural selection, and slow down at later stages of the radiation, due to filling of vacant niches and decreasing ecological opportunity (Simpson 1953, Schluter 2000, Kassen et al. 2004, Rabosky 2009). Support for this hypothesis has recently accumulated and comes from mathematical models of adaptive radiation (Gavrilets & Vose 2005, Gavrilets & Losos 2009), experimental diversification in laboratory conditions (Kassen et al. 2004, Meyer et al. 2011), from comparison of parallel adaptive radiations of different ages (Gillespie 2004, Seehausen 2006), and from time-calibrated molecular phylogenies of the extant radiations of North American wood warblers (Rabosky & Lovette 2008a, 2008b) and island lizards (Mahler et al. 2010, Rabosky & Glor 2010), demonstrating how speciation rates decrease with ecological opportunity. Further support comes from the fossil record, showing that speciation rates of well-investigated faunas, including North American mammals (Alroy 1999) and marine invertebrates (Alroy 2008), are highly density-dependent.

Besides decreasing speciation rates, there are at least two reasons why increasing extinction rates could also contribute to the observed slowing of species accumulation in adaptive radiation. First, explosive diversification in the early stages of adaptive radiations should lead to smaller mean population sizes, which are therefore increasingly prone to extinction (Levontin 1979, Quental & Marshall 2010). Second, strong natural selection that is characteristic for adaptive radiations may lead to competitive replacement of variants within the same niche class and to loss of intermediate phenotypes (Meyer et al. 2011).

Due to the coupled effects of decreasing speciation rates and increasing extinction, species richness may peak shortly after the initial burst of diversification (apparently as early as 0.5 Myr in the parallel radiations of cichlid fishes (Seehausen 2006)), and subsequently fall back to lower levels of equilibrium diversity. This phenomenon termed 'overshooting ' seems counterintuitive, as adaptive radiation is often associated with the continual increase of diversity that is observed in phylogenetic inference of extant taxa. Nevertheless, patterns of overshooting are confirmed with mathematical models of adaptive radiation (Gavrilets & Vose 2005, Gavrilets & Losos 2009), and with experimental adaptive radiation in laboratory conditions, and are interpreted as persistent divergent selection, leading to increased extinction (Meyer et al. 2011).

The fossil record strongly supports the theory of overshooting. By comparison of taxonomic and morphological diversity in the paleocoic radiations of trilobites and blastoids, it could be shown that morphology continues to diversify for long periods even despite severe losses in species richness (Foote 1993). If these diversifications were extant, they would likely be interpreted as ongoing adaptive radiations, even millions of years after peaks in species richness had been reached.

In summary, three key questions concerning the adaptive radiation of Antarctic notothenioid fishes remain: Have glacial maxima of the Pleistocene contributed to increased speciation rates in Antarctic notothenioids? How does unidirectional gene flow by larval dispersal affect speciation rates? And finally, did the notothenioid radiation overshoot early in its history, after an initial burst of speciation? Further insights into these questions can be gained from extensive analyses of diversification rates that will become available in the future, through ongoing sequencing efforts using notothenioid samples that were recently obtained during Antarctic expedition ANT-XXVII/3 with RV Polarstern.

References

Alroy J (1999) The fossil record of North American mammals: evidence for a Paleocene evolutionary radiation. *Syst Biol* 48: 107–118.

Alroy J (2008) Dynamics of origination and extinction in the marine fossil record. *Proc Natl Acad Sci USA* 105: 11536–11542.

Balushkin AV (2000) Morphology, classification, and evolution of notothenioid fishes of the Southern Ocean (Notothenioidei, Perciformes). *J Ichthyol* 40: S74–S109.

Barnes DKA, Conlan KE (2007) Disturbance, colonization and development of Antarctic benthic communities. *Philos Trans R Soc London B Biol Sci* 362: 11–38.

Damerau M, Matschiner M, Salzburger W, Hanel R: Comparative population genetics of seven notothenioid fish species reveals high levels of gene flow along ocean currents in the southern Scotia Arc, Antarctica. Submitted.

Eastman JT (1993) Antarctic Fish Biology: Evolution in a Unique Environment. Academic Press, Inc., San Diego, CA.

Foote M (1993) Discordance and concordance between morphological and taxonomix diversity. *Paleobiology* 19: 185–204.

Garant D, Forde SE, Hendry AP (2007) The multifarious effects of dispersal and gene flow on contemporary adaptation. *Funct Ecology* 21: 434–443.

Gavrilets S, Losos JB (2009) Adaptive radiation: contrasting theory with data. *Science* 323: 732–737.

Gavrilets S, Vose A (2005) Dynamic patterns of adaptive radiation. *Proc Natl Acad Sci USA* 102: 18040–18045.

Gillespie RG (2004) Community assembly through adaptive radiation in Hawaiian spiders. *Science* 303: 356–359.

Kassen R, Llewellyn M, Rainey PB (2004) Ecological constraints on diversification in a model adaptive radiation. *Nature* 431: 984–988.

Levontin JS (1979) A theory of diversity equilibrium and morphological evolution. *Science* 204: 335–336.

Mahler DL, Revell LJ, Glor RE, Losos JB (2010) Ecological opportunity and the rate of morphological evolution in the diversification of Greater Antillean anoles. *Evolution* 64: 2731–2745.

Matschiner M, Hanel R, Salzburger W (2009) Gene flow by larval dispersal in the Antarctic notothenioid fish *Gobionotothen gibberifrons*. *Mol Ecol* 18: 2574–2587.

Matschiner M, Hanel R, Salzburger W (2011) On the origin and trigger of the notothenioid adaptive radiation. *PLoS ONE* 6: e18911.

Meyer JR, Schoustra SE, Lachapelle J, Kassen R (2011) Overshooting dynamics in a model adaptive radiation. *Proc R Soc Lond B Biol Sci* 278: 392–398.

Niemiller ML, Fitzpatrick BM, Miller BT (2008) Recent divergence with gene flow in Tennessee cave salamanders (Plethodontidae: *Gyrinophilus*) inferred from gene genealogies. *Mol Ecol* 17: 2258–2275.

Nosil P (2008) Speciation with gene flow could be common. *Mol Ecol* 17: 2103–2106.

Quental TB, Marshall C (2010) Diversity dynamics: molecular phylogenies need the fossil record. *Trends Ecol Evol* 25: 434–441.

Rabosky DL (2009) Ecological limits on clade diversification in higher taxa. *Am Nat* 173: 662–674.

Rabosky DL, Glor RE (2010) Equilibrium speciation dynamics in a model adaptive radiation of island lizards. *Proc Natl Acad Sci USA* 107: 22178–22183.

Rabosky DL, Lovette IJ (2008a) Explosive evolutionary radiations: decreasing speciation or increasing extinction through time? *Evolution* 62: 1866–1875.

Rabosky DL, Lovette IJ (2008b) Density-dependent diversification in North American wood warblers. *Proc R Soc Lond B Biol Sci* 275: 2363–2371.

Rutschmann S, Matschiner M, Damerau M, et al. (2011) Parallel ecological diversification in Antarctic notothenioid fishes as evidence for adaptive radiation. *Mol Ecol* 20: 4707-4721.

Schluter D (2000) The Ecology of Adaptive Radiation. Oxford University Press, New York.

Seehausen O (2006) African cichlid fish: a model system in adaptive radiation research. *Proc R Soc Lond B Biol Sci* 273: 1987–1998.

Simpson GG (1953) The Major Features of Evolution. Columbia University Press, New York.

Acknowledgements

First and foremost I am most grateful to

Ass. Prof. Dr. Walter Salzburger

for inviting me to do a PhD in his new lab in Basel, and for the many years of invaluable support that I have since received from him.

Further, I am particularly thankful to

Dr. Reinhold Hanel

for continuous support with advice, sample provision, and fieldwork arrangements, without which this doctoral work would not have been possible.

I am grateful to

Prof. Dr. Thorsten Reusch

for discussions and his agreement to review this thesis.

I would like to thank

all members of the WalterLab

for being great colleagues and friends, for sharing thoughts, beers, and foosball matches with me.

In particular, I am thankful to **Moritz Muschick** for many insightful discussions, to **Sereina Rutschmann** for the excellent collaboration and the highly successful outcome of her master thesis, to **Brigitte Aeschbach** for being an incredibly friendly and helpful contact person for everybody in the lab, to **Nico Boileau** and **Eveline Diepeveen** for being great office mates, refreshing my highschool algebra, and for dutch sweets, to **Lukas Zimmermann** for generous software provision, to **Fabio Cortesi** for many performed PCRs and reminding me to "write the paper", to **Hugo Gante**, **Britta Meyer**, **Laura Baldo**, **Daniel Berner**, and **Zuzana Musilová** for relaxed clock, supertree, phylogenomics, and simulations discussions, and to **Julia De Maddalena** for forcing me to learn *BEAST.

Furthermore, I am thankful to

all participants of Polarstern expedition ANT-XXVII/3

for making this expedition an unforgettable experience.

Especially, I'ld like to thank **Malte Damerau**, **Katja Mintenbeck**, **Lena Rath**, **Nils Koschnick**, and **Timo Hirse** for sharing all those hours in the fish lab, 1500 samples, and well-deserved beer with me, **Rainer Knust** for being a most caring and responsible cruise leader, and **Shobhit Agrawal** and **Chester Sands** for discussions and input to my postdoctoral grant application.

I owe special thanks to

Julia M.I. Barth

not only for being the most important person in my life, but also for valuable comments on individual studies and this thesis.

It's my pleasure to thank

Jago Wallenschus

for support and company during fieldwork on Svalbard.

Discussions with

Sebastian Höhna

contributed greatly to phylogenetic analyses and the development of the R package 'ageprior'.

Further important contributions to this thesis came from

Florian Odroniz

who made me aware of VolkswagenStiftung fellowships, and encouraged me to apply for one. Also, he introduced me to Macintosh computers and to the Ruby programming language, which both greatly facilitated this work.

Curriculum Vitae: **Michael Matschiner**

Michael Matschiner
Vogesenstrasse 73
CH-4056 Basel
Switzerland
T +41 61 2670302
M +41 78 6799871
michael.matschiner@unibas.ch
www.evolution.unibas.ch/salzburger/team/mmatschiner

Personal data

Born: August 19, 1979
Birthplace: Neumarkt in der Oberpfalz, Germany
Nationality: German

Education

University of Basel, Switzerland | 2008 ongoing
Graduate studies in the group of Prof. Walter Salzburger. Subject of the PhD thesis is the adaptive radiation of Antarctic notothenioid fishes. Funded by a Volkswagen Foundation PhD Scholarship.

University of Konstanz, Germany | 2002-2008
Diploma studies of Biological Sciences (German equivalent to Master studies).
Main subjects: Bioinorganic Chemistry, Molecular Evolutionary Biology, Microbial Ecology, and Bioinformatics.
2007-2008: Diploma thesis externally in the group of Prof. Walter Salzburger, University of Basel. Title: 'Adaptive radiation in an extreme environment: Population genetics of Antarctic notothenioids'. Mark: 1.0 (range 1 to 5, 1 = highest, 5 = lowest).
2007: Diploma examination. Subjects examined: Bioinformatics, Plant Physiology. Mark: 1.0 (range 1 to 5, 1 = highest, 5 = lowest).
2005: Intermediate diploma, mark: 2.7 (range 1 to 5, 1 = highest, 5 = lowest).
2002-2007: Active member of the student council.

Courses

University Centre in Svalbard, Longyearbyen, Norway | 2009
Six-week course on the Marine Benthos of Svalbard, including a two-week field work cruise with RV Jan Mayen to the north of Svalbard and into sea ice. Application of a wide range of sampling methods including pelagic and bottom trawls, Agassiz trawls, Sneli and RP-sledges, MPS, WP-3 nets, grab samplers, deployment of moorings and CTDs. Onboard taxonomic work, predation experiments, analysis of benthos imagery and CTD data.
Faculty: Jørgen Berge, Bjørn Gulliksen, Geir Johnson, Paul Renaud, Fredrik Pleijel and others.
Written report and final exams, overall mark: A (range A to F, A = highest, F = lowest).

Guarda workshop, Guarda, Switzerland | 2009
Course on conceptual scientific thinking and writing of grant proposals.
Faculty: Peter and Rosemary Grant, Richard Lenski, and others.
Written report, passed.

Workshop on Molecular Evolution, Český Krumlov, Czech Republic | 2009

Three-week course on theory and practice of phylogenetic analysis using PAUP*, MrBayes, BEAST, Modeltest, GARLI, MAFFT, Tracer, PAML, BLAST, LAMARC, Migrate-n, gsi, AWTY.

Faculty: the authors of many of the above software applications, including Michael Cummings, Dave Swofford, Andrew Rambaut, Fredrik Ronquist, Peter Beerli, and others.

University of Lausanne, Switzerland | 2008

Course: Foundations of the theory of speciation. Covered speciation and fitness landscapes, models of sympatric speciation, ecological speciation, and adaptive radiation.

Faculty: Sergey Gavrilets.

Swiss Institute for Bioinformatics, Lausanne, Switzerland | 2007

Course: Phylogeny and Evolution using Bioinformatics. Covered reading and understanding of phylogenetic trees, measures of confidence, patterns of evolution.

Faculty: Marc Robinson-Rechavi, Nicolas Salamin, Nicolas Gaultier, and others.

Swiss Institute for Bioinformatics, Basel, Switzerland | 2007

Course: Advanced Programming for Biologists. Covered database programming (SQL), PHP interfaces, web services.

Faculty: Laurent Falquet, Lorenza Bordoli, and others.

Final exam, passed.

University Centre in Svalbard, Longyearbyen, Norway | 2006

Six-week course on Marine and Terrestrial Arctic Microbiology, including a one-week cruise with RV Lance to Ny Ålesund and Fram Strait. Quantitative and qualitative analysis of seawater (≤ 1600 m) and soil samples. Applied methods included measurements of methane and carbon flux, nitrogen fixation, organic matter content, ammonium and nitrate content, DNA fragment amplification, amplified rDNA restriction analysis (ARDRA), Fluorescent in situ hybridization, microbial counts.

Faculty: Rolf Olsen, Bjørn Solheim, Steve Coulson, and others.

Written report and final exam, mark: C (range A to F, A = highest, F = lowest).

Fieldwork experience

Antarctica | 2011

Three months on board RV Polarstern (ANTXXVII-3). Collection of notothenioid fish samples for DNA, RNA, stable isotope, and morphological analysis at South Georgia, the South Orkney Islands, the Antarctic Peninsula, Dronning Maud Land, and Bouvetøya.

Patagonia, Argentina | 2009

One week collection of notothenioid fish samples for DNA, RNA, and stomach content analysis.

Svalbard, Norway | 2009

Two weeks onboard RV Jan Mayen, to fjords of Svalbard and into sea ice on 81°N. Collection of quantitative benthos samples and ice fauna using bottom trawls, Agassiz trawls, Sneli and RP-sledges, MPS, WP-3 nets, and grab samplers. Onboard species identification, dissection, stomach content analysis, and predation experiments. Extraction of tissue samples for DNA and stable isotope analysis.

Lake Tanganyika, Zambia | 2008

Three weeks. Collection of cichlid fishes using gill nets and dives. On site dissection and extraction of tissue samples for DNA, RNA, stable isotope, and stomach content analysis.

Svalbard, Norway | 2006

One week onboard RV Lance. Collection of soil samples from various locations at the westcoast of Spitsbergen, and water samples from the Fram Strait. On site measurements of carbon flux.

Teaching

Workshop on Molecular Evolution, Český Krumlov, Czech Republic | 2010

Based on my participation in 2009 (see above), I was invited to work as a Teaching Assistant in the Workshop on Molecular Evolution. Topics covered: UNIX, Emacs, FASTA, BLAST, MAFFT, PAUP*, jModelTest, GARLI, LAMARC, Migrate-n, MrBayes, PAML, gsi, BEAST, Tracer.

Supervision of students

Master thesis

Sereina Rutschmann (2011) Parallel ecological diversification in Antarctic notothenioid fishes as evidence for adaptive radiation.

Publications

Papers

Palandačič A, Matschiner M, Sloj A: Genetic evidence for long-range subterranean dispersal of fish (Cyprinidae: *Delminichthys adspersus*). submitted.

Damerau M, Matschiner M, Salzburger W, Hanel R: Comparative population genetics of seven notothenioid fish species in the southern Scotia Arc, Antarctica. submitted.

Rutschmann S, Matschiner M, Damerau M, Muschick M, Lehmann MF, Hanel R, Salzburger W (2011) Parallel ecological diversification in Antarctic notothenioid fishes as evidence for adaptive radiation. *Molecular Ecology* 20:4707-4721.

Matschiner M, Hanel R, Salzburger W (2011) On the origin of the notothenioid adaptive radiation. *PLoS ONE* 6:e18911.

Matschiner M, Hanel R, Salzburger W (2009) Gene flow by larval dispersal in the Antarctic notothenioid fish *Gobionotothen gibberifrons*. *Molecular Ecology* 18:2574-2587.

Matschiner M, Salzburger W (2009) TANDEM: integrating automated allele binning into genetics and genomics workflows. *Bioinformatics*, 25:1982-1983.

Book chapters

Matschiner M, Hanel R, Salzburger W (2010) Phylogeography and speciation processes in marine fishes and fishes from large freshwater lakes. In: *Phylogeography: concepts, intraspecific patterns and speciation processes* (ed Rutgers DS), pp. 1-29. Nova Science Publishers, New York.

Presentations

Invited talks

The origin of the notothenioid adaptive radiation. 10/29/2010, IFM-Geomar, University of Kiel.

Scientific visualization. 10/13/2010, Zurich University of Arts.

Population genetics meets oceanography: larval dispersal across the Scotia Sea, Antarctica. 07/27/2009, Institute of Zoology, University of Graz.

Conference talks

Matschiner M (2011) Fossil-constrained dating of adaptive radiations. 08/25/2011, 13th Congress of the European Society for Evolutionary Biology, Tübingen.

Matschiner M, Hanel R, Salzburger W (2010) A new dating of the notothenioid radiation. 05/10/2010, Second Status Symposium in Evolutionary Biology, Volkswagen Foundation, Fraueninsel, Lake Chiemsee.

Prizes and awards

Prix de Quervain, 2011. Swiss Committee on Polar and High Altitude Research. 5000 CHF.

Research and travel grants

Diversity through time in adaptive radiations. Swiss National Science Foundation Fellowship for prospective researchers. 65 050 CHF.

13th Congress of the European Society for Evolutionary Biology, Tübingen, 2011. University of Basel travel grant. 360 CHF.

Entstehung der Artenvielfalt Antarktischer Eisfische. Freiwillige Akademische Gesellschaft research grant. 7 500 CHF.

Guarda workshop, Guarda, Switzerland, 2009. University of Basel travel grant. 276 CHF.

Annual Meeting of the Society for Molecular Biology and Evolution, Barcelona, Spain, 2008. University of Basel travel grant. 430 CHF.

Adaptive radiation in an extreme environment - genetic population structure in Antarctic notothenioids. Volkswagen Foundation PhD Scholarship. 138 800 €.

Service as reviewer

PLoS ONE
Molecular Ecology
BMC Evolutionary Biology
Journal of Heredity
Marine Biology