



**Cite this article:** Friedman M, Keck BP, Dornburg A, Eytan RI, Martin CH, Hulsey CD, Wainwright PC, Near TJ. 2013 Molecular and fossil evidence place the origin of cichlid fishes long after Gondwanan rifting. *Proc R Soc B* 280: 20131733.  
<http://dx.doi.org/10.1098/rspb.2013.1733>

Received: 4 July 2013

Accepted: 29 August 2013

**Subject Areas:**

evolution, genetics, palaeontology

**Keywords:**

biogeography, dispersal, fossil record, molecular clock, vicariance

**Authors for correspondence:**

Matt Friedman

e-mail: [matt.friedman@earth.ox.ac.uk](mailto:matt.friedman@earth.ox.ac.uk)

Thomas J. Near

e-mail: [thomas.near@yale.edu](mailto:thomas.near@yale.edu)

<sup>†</sup>Present address: Department of Environmental Science, Policy, and Management, University of California, Berkeley, CA 94720, USA.

Electronic supplementary material is available at <http://dx.doi.org/10.1098/rspb.2013.1733> or via <http://rspb.royalsocietypublishing.org>.

# Molecular and fossil evidence place the origin of cichlid fishes long after Gondwanan rifting

Matt Friedman<sup>1</sup>, Benjamin P. Keck<sup>2</sup>, Alex Dornburg<sup>3</sup>, Ron I. Eytan<sup>3</sup>, Christopher H. Martin<sup>4,†</sup>, C. Darrin Hulsey<sup>2</sup>, Peter C. Wainwright<sup>4</sup> and Thomas J. Near<sup>3</sup>

<sup>1</sup>Department of Earth Sciences, University of Oxford, South Parks Road, Oxford OX1 3AN, UK

<sup>2</sup>Department of Ecology and Evolutionary Biology, University of Tennessee, Knoxville, TN 37996, USA

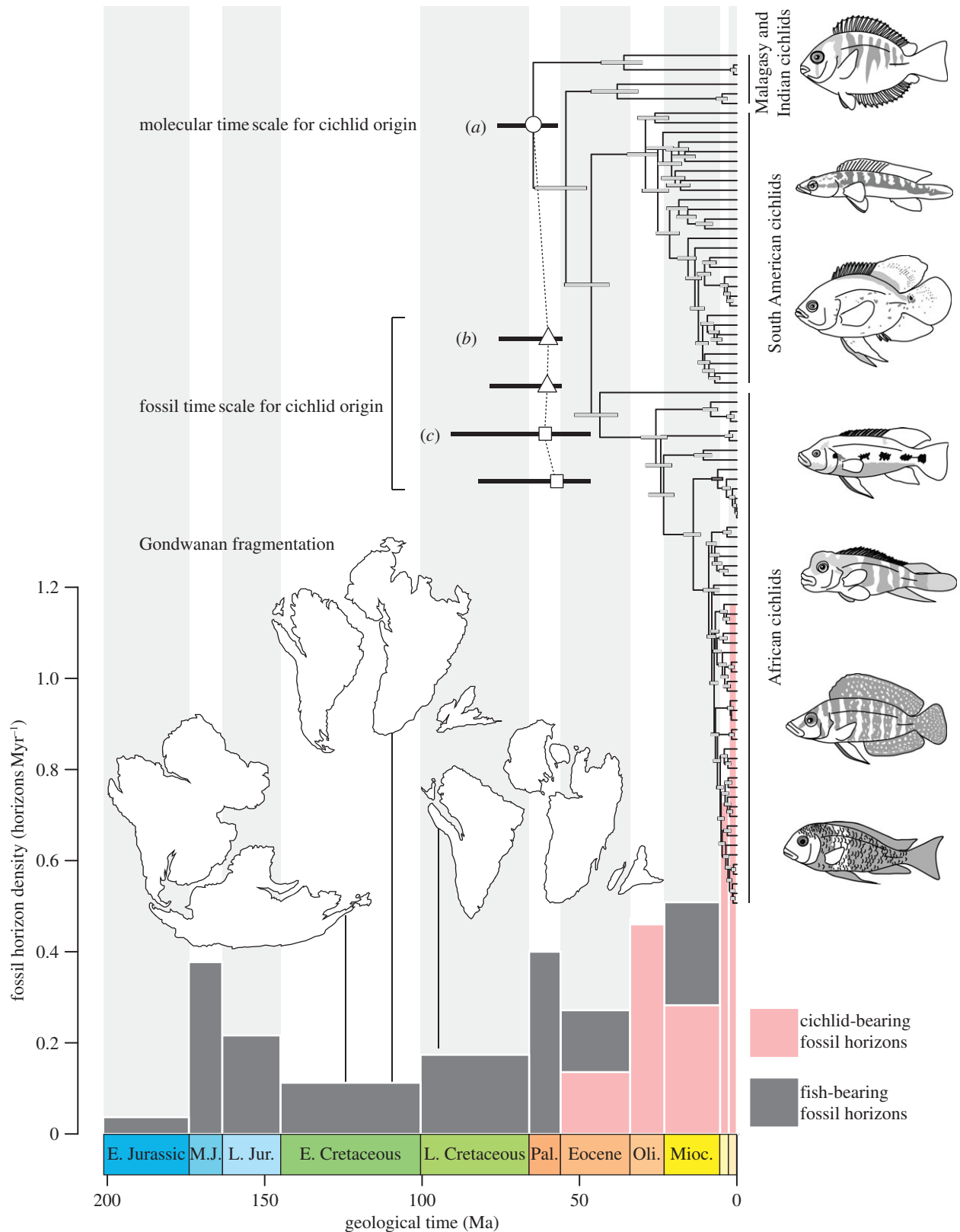
<sup>3</sup>Department of Ecology and Evolutionary Biology and Peabody Museum of Natural History, Yale University, New Haven, CT 06520, USA

<sup>4</sup>Department of Evolution and Ecology, University of California, Davis, CA 95616, USA

Cichlid fishes are a key model system in the study of adaptive radiation, speciation and evolutionary developmental biology. More than 1600 cichlid species inhabit freshwater and marginal marine environments across several southern landmasses. This distributional pattern, combined with parallels between cichlid phylogeny and sequences of Mesozoic continental rifting, has led to the widely accepted hypothesis that cichlids are an ancient group whose major biogeographic patterns arose from Gondwanan vicariance. Although the Early Cretaceous (*ca* 135 Ma) divergence of living cichlids demanded by the vicariance model now represents a key calibration for teleost molecular clocks, this putative split pre-dates the oldest cichlid fossils by nearly 90 Myr. Here, we provide independent palaeontological and relaxed-molecular-clock estimates for the time of cichlid origin that collectively reject the antiquity of the group required by the Gondwanan vicariance scenario. The distribution of cichlid fossil horizons, the age of stratigraphically consistent outgroup lineages to cichlids and relaxed-clock analysis of a DNA sequence dataset consisting of 10 nuclear genes all deliver overlapping estimates for crown cichlid origin centred on the Palaeocene (*ca* 65–57 Ma), substantially post-dating the tectonic fragmentation of Gondwana. Our results provide a revised macroevolutionary time scale for cichlids, imply a role for dispersal in generating the observed geographical distribution of this important model clade and add to a growing debate that questions the dominance of the vicariance paradigm of historical biogeography.

## 1. Introduction

Cichlid fishes, along with Darwin's finches and Caribbean *Anolis* lizards, represent a key vertebrate model system for understanding the evolutionary assembly of biodiversity [1,2]. Despite the group's prominence in biological research, a consistent macroevolutionary time scale and biogeographic history for cichlids has remained elusive [3–6]. For nearly four decades, the study of deep cichlid evolutionary history has been dominated by vicariance models of biogeography that link the present-day distribution of the group to the tectonic fragmentation of the supercontinent of Gondwana during the mid to late Mesozoic (*ca* 135–90 Ma; figure 1) [8,9]. Continued investigation of cichlid intrarelationships, including phylogenetic analysis of molecular sequence data, has shown congruence between the order of divergences among geographically restricted cichlid clades and proposed sequences of continental break-up [10]. The vicariance hypothesis of cichlid historical biogeography has become so entrenched that the rifting history of Gondwana is routinely used to calibrate teleost molecular clocks [3,6], with the consequence that this hypothetical scenario



**Figure 1.** Congruent molecular and palaeontological time scales place the origin of cichlid fishes in the Late Cretaceous–Eocene interval, substantially post-dating Gondwanan rifting. (a) Molecular phylogeny for Cichlidae calibrated using fossils belonging to non-cichlid groups (full phylogeny provided in electronic supplementary material, figures S1 and S2). (b) Bayesian point estimates and 95% CIs for the timing of cichlid origin based on the distribution of cichlid fossils and the availability of freshwater sedimentary deposits of Triassic–Recent age on Gondwanan landmasses that bear articulated fish remains. The top estimate is derived from the record of landmasses inhabited by extant cichlids, and the bottom estimate is derived from the record of all Gondwanan landmasses. The density of all Gondwanan horizons bearing articulated freshwater fish fossils is indicated by the histogram at the bottom of the figure (densities including disarticulated material are given in electronic supplementary material). Grey bars indicate total horizon density. Pink bars indicate the density of the subset of fossil fish horizons that bear cichlids. (c) Bayesian point estimates and 95% CIs for the timing of cichlid origin based on successive fossil outgroups to the clade. The two estimates reflect competing hypotheses for the earliest fossil examples of some outgroups. The top estimate is based on the oldest proposed outgroup ages and the bottom estimate is based on the youngest proposed outgroup ages. Cichlid illustrations, from top to bottom: *Etroplus*, *Crenicichla*, *Astronotus*, *Hemichromis*, *Steatocranus*, *Altolamprologus* and *Tropheus*. Continental arrangements based on palaeogeographic reconstructions by R. Blakey, originally presented in [7].

now directly influences estimates of the evolutionary time scale for more than half of all modern vertebrate diversity.

If the distribution of modern cichlids is attributable purely to Gondwanan break-up, then it necessarily follows that the common ancestor of all living cichlids originated no later than the final separation between Madagascar–India and South America–Africa–Arabia. Current geological evidence places this continental fragmentation event in the Early Cretaceous (*ca* 135 Ma) [11]. However, the stratigraphically oldest fossil cichlids are Eocene in age (approx. 46 Ma) [12,13], implying a gap of approximately 90 Myr in the early history of the group. This, along with the absence of Early Cretaceous fossils belonging to more inclusive and taxonomically diverse clades that contain cichlids, has led some to abandon the orthodoxy of Mesozoic vicariance in favour of Cenozoic dispersal to explain the present-day distribution of cichlids [12,14].

The complete absence of fossil cichlids from many former Gondwanan landmasses would seem equally problematic for the vicariance hypothesis, but has received surprisingly little attention. For example, the Australian fossil record contains several fish-bearing freshwater deposits of Mesozoic and Cenozoic age, but no fossil cichlid is known from the continent. While it is clear that assembly of the compositionally distinctive Australian freshwater fish fauna has a complex history stemming from isolation, aridification and marine invasion coupled with the persistence of ancient lineages [15], this complexity does not undermine the prediction of the vicariance model that cichlids should have been widely distributed across Gondwanan landmasses during the Mesozoic [9].

These palaeontological arguments have been dismissed as ‘non-evidence’ by advocates of cichlid vicariance [16]. Some authors have even suggested that the fossil record supports the notion of cichlids deep within the Mesozoic, citing the high probability of non-preservation for freshwater taxa of Cretaceous age [9] or inferring that the advanced morphology of the earliest fossil cichlids implies a long, and as yet unsampled, palaeontological history of the group [9,10,13]. The seemingly ambiguous signal of palaeontological data with respect to the question of cichlid origin is symptomatic of a qualitative approach to an inherently quantitative problem. Invoked stratigraphic gaps are neither *ad hoc* contrivances nor trivial inconveniences to be dismissed as non-evidence; they are hypotheses amenable to statistical interrogation.

In order to provide a robust time scale for cichlid diversification and select between competing biogeographic hypotheses, we applied three semi-independent approaches in estimating the age of crown-group Cichlidae. Our first two methods are palaeontological, and draw on (i) the distribution of fossil horizons yielding cichlids and those that might plausibly yield cichlids (i.e. fish-bearing freshwater deposits on former Gondwanan landmasses) [17], and (ii) the stratigraphic distribution of more inclusive teleost lineages (meaning clades of higher taxonomic rank) that contain cichlids [18]. These techniques relate directly to two contrasting arguments that emerge repeatedly in palaeontological debates concerning the chronology of cichlid evolution: either that the record of freshwater fishes generally, and cichlids specifically, is sufficiently poor that the absence of Mesozoic cichlid fossils is unsurprising, or that the minimum age of origin for a series of more inclusive lineages of teleost fishes precludes the origin of cichlids deep within the Mesozoic. Significantly, these methods share only one common feature in their calculations: both are necessarily constrained by the minimum age for cichlids as imposed by

the oldest fossil example(s) of the group. As an independent assessment of the divergence times estimated from palaeontological data, we conducted a relaxed-molecular-clock analysis for cichlids and Ovalentaria [19], a percomorph lineage that includes cichlids. Our dataset includes 10 protein-coding nuclear genes for 89 cichlids and 69 non-cichlid species of Percomorpha.

## 2. Material and methods

### (a) Estimating time of evolutionary origin using the distribution of cichlid-bearing fossil horizons

One method of estimating credible intervals (CIs) on stratigraphic durations draws on the number of fossil horizons within the sampled range of the group of interest. The simplest approach assumes that fossil horizons are distributed at random [20,21], but the potential for fossil recovery undoubtedly varies over time as a consequence of a heterogeneous rock record. Marshall [17] developed a more general method that permits non-uniform preservation by using an empirically informed function that quantifies potential for fossil recovery. We have applied this logic in conjunction with a Bayesian approach that provides a statistically appropriate framework for discussing the probability of clade origin within certain stratigraphic intervals [20]. Our results are conditioned on the prior assumption that cichlids are post-Palaeozoic in age (i.e. they originated in the Triassic or later), which is consistent with the fossil record and does not exclude the possibility of Gondwanan vicariance.

We assembled a database of known fossil occurrences of cichlids on Gondwanan landmasses based on the literature (see electronic supplementary material). Different geological formations (or localities where there is no formalized lithostratigraphic framework) were assumed to represent distinct sampling horizons. The function for the potential recovery of fossil cichlids was estimated by tabulating the number of sedimentary horizons (formations or localities) that meet three key criteria. First, candidate deposits must be present on former Gondwanan landmasses. Second, candidate deposits must represent freshwater environments. Third, candidate deposits must have the potential to yield fossils of cichlids, were this group present. Sites yielding fish fossils (including but not restricted to cichlids) meet this final criterion. The nature of fossils (articulated or fragmentary) from sites satisfying these conditions was also recorded.

Because of uncertainty surrounding age assessments, uniform recovery potential was assumed within each epoch-level stratigraphic bin, with relative recovery potential given by the number of candidate horizons present in a given interval divided by its duration. Ambiguity surrounds the age of many freshwater deposits. In this study, imprecisely dated deposits are given their oldest plausible age. This approach systematically biases analysis towards older age estimates for the time of clade origin, thereby providing a more generous test of the Gondwanan vicariance hypothesis.

These data were used to generate point estimates and 95% CIs for cichlid origin based on (i) the fossil records of Gondwanan landmasses currently inhabited by cichlids (South and Central America including the Caribbean, Africa, Madagascar, India, Arabia; cichlid fossils are known from all of these regions except Madagascar and India), and (ii) these records combined with those of Australia and Antarctica, former Gondwanan landmasses that lack cichlids but would be predicted to have once been inhabited by the group under the vicariance hypothesis. For both, we calculated CIs based on the record of all cichlid fossils and estimated range extensions based on articulated cichlid remains alone combined with appropriate recovery potential functions generated from the subset of deposits that yield complete fish specimens. This modified procedure is more conservative and reflects the very real possibility that the earliest cichlids might be



recognized only on the basis of articulated remains, as their isolated fragments might be too character-poor, too generalized or both to permit reliable taxonomic attributions.

## (b) Estimating time of evolutionary origin using the distribution of ages of outgroups to cichlids

Hedman [18] devised a Bayesian approach for constraining the time of origin of a clade based on the distribution of stratigraphic ages of successive outgroups. This method requires that outgroups appear in the fossil record in an order matching phylogeny and that they pre-date or are contemporary with the first appearance of the focal clade. Such perfect congruence is rare in empirical examples, and we adopt a proposed solution that conservatively excludes inconsistent ages.

An account of the outgroups used in analysis is provided in the electronic supplementary material. In some cases, there is disagreement surrounding the identity of the earliest representatives of these lineages. To accommodate uncertainty, two sets of calculations were completed: one using the oldest proposed minimum age for a clade and the other applying the youngest. Collectively, these paired analyses provide upper and lower estimates of CIs for divergence times given present understanding of both the fossil record and teleost interrelationships. These age estimates are conditioned on a prior assumption that divergence occurred after a user-specified hard upper bound. This bound applies to the divergences of all groups considered, not only the focal clade. We have therefore selected the Carboniferous (Serpukhovian) *Discoserra*, a putative stem neopterygian [22], as defining an upper bound of 322.8 Ma (see the electronic supplementary material).

## (c) Collection of sequence data, phylogenetic analyses and relaxed molecular clocks

Standard phenol–chloroform extraction protocol or Qiagen DNeasy Blood and Tissue kits were used to isolate DNA from tissue biopsies sampled from 158 species of percomorph teleosts that included 89 species of Cichlidae (electronic supplementary material, table S6). Previously published PCR primers (see the electronic supplementary material) were used to amplify and sequence exons from 10 nuclear genes (*ENC1*, *Glyt*, *myh6*, *plagl2*, *Ptr*, *rag1*, *SH3PX3*, *sreb2*, *tbr1* and *zic1*). Amplified gene copies were cleaned and used as templates for DNA cycle sequencing. Alignments of the DNA sequences from the individual genes were constructed from the inferred amino acid sequences. Thirty data partitions were designated that corresponded to the three separate codon positions for each of the 10 protein-coding genes. A phylogeny of the aligned DNA matrix was inferred using maximum-likelihood and relaxed-clock analyses using a random local molecular-clock model in the computer program BEAST v. 1.6 (figure 1; electronic supplementary material, figure S1) [23,24]. DNA sequences are deposited on GenBank KF556709–KF557487. Aligned gene sequences used in phylogenetic analyses, phylogenetic trees resulting from RAxML and BEAST analyses, files formatted for BEAST analyses and files used to estimate the age of cichlids using palaeontological data are available from the dryad digital repository (<http://dx.doi.org/10.5061/dryad.48f62>). Fossil-based age constraints were applied to 10 nodes in the percomorph phylogeny (electronic supplementary material, figure S2).

## (d) Fossil calibration age priors

For each fossil calibration prior, we identify the calibrated node in the percomorph phylogeny, list the taxa that represent the first occurrence of the lineage in the fossil record, describe the character states that justify the phylogenetic placement of the fossil taxon, provide information on the stratigraphy of the rock formation(s)

bearing the fossil, give the absolute age estimate for the fossil, outline the prior age setting in the BEAST relaxed-clock analysis and provide any additional notes on the calibration [25]. Each calibration is numbered and the phylogenetic placement of the calibration is highlighted in the electronic supplementary material, figure S3. Full justification of our calibrations is given in the electronic supplementary material.

Because we look to provide a critical test of competing models of cichlid biogeography, we have not assumed Gondwanan vicariance *a priori* and did not use the timing of the fragmentation history of this supercontinent to inform calibrations in the relaxed-molecular-clock analyses. Furthermore, we have not included any internal calibrations within Cichlidae, so that our relaxed-molecular-clock estimate of the evolutionary time scale for the group is truly independent of its fossil record, which contributes to our palaeontological estimates of divergence times (see §2).

## 3. Results

Our three approaches to estimating a time scale of cichlid origin and diversification yield overlapping CIs that diverge significantly from the predictions made by the Gondwanan vicariance biogeographic hypothesis, and are discussed in turn in §3*a,b* (figure 1).

### (a) Palaeontological time scales for cichlid evolution

The distribution of cichlid-bearing fossil horizons, combined with an empirically informed function describing fossil recovery potential, indicates an age of origin for cichlids in the Late Cretaceous or Palaeocene. If only the records of landmasses that are currently inhabited by cichlids are considered, the time of origin of the clade is estimated as 59.2 Ma (95% CI: 56.1–67.6 Ma). By contrast, a slightly younger age estimate of 57.8 Ma (95% CI: 56.1–62.4 Ma) is obtained if the record of all Gondwanan landmasses is considered. Restricting the scope of analysis to consider articulated remains alone provides a more conservative means of estimating the time of origin for cichlids, because early members of this group might not be recognized on the basis of less diagnostic skeletal debris. Point estimates for the timing of cichlid origin under this approach do not change drastically from those obtained using the entirety of the cichlid fossil record, but the upper bounds of the CIs do increase by more than 10 Ma. Depending on the scope of geographical analysis, we estimate the time of cichlid origin based only on articulated remains as ranging from 59.8 Ma (95% CI: 56.1–75.1 Ma; landmasses inhabited by modern cichlids) to 60.2 Ma (95% CI: 56.1–77.8 Ma; all Gondwanan landmasses). The Gondwanan vicariance hypothesis requires a pre-Eocene record of cichlids that is roughly 10–30 times worse than their recorded fossil history, with rescaled recovery potentials conditioned on point estimates for the origin of the group at 135 Ma ranging from 2.8–3.3% (all fossils) to 6.6–6.9% (articulated fossils only) of their original values. Classical confidence intervals deliver similar results to the Bayesian estimates (see electronic supplementary material, table S2).

Analysis of outgroup ages provides broadly similar estimates for the timing of cichlid origin to those derived from the distribution of cichlid fossil horizons, in terms of both the magnitude of point estimates and the degree of uncertainty surrounding them. We find a mean age of 60.7 Ma (95% CI: 46.8, 90.1 Ma) using the oldest possible fossil ages for outgroups. The time scale for cichlid origin is predictably more

**Table 1.** Posterior molecular age estimates for major lineages of Cichlidae. Ages refer to crown groups.

clade	mean age (Ma)	95% highest posterior density interval (Ma)
Cichlidae	64.9	57.3–76.0
Etioplineae (India and Madagascar)	36.0	30.3–42.2
Ptychochrominae (Madagascar)	38.2	31.7–46.4
unnamed Afro-American clade	46.4	40.9–54.9
Cichlinae (neotropics)	29.2	25.5–34.8
Pseudocrenilabrinae (Africa)	43.7	38.2–51.6
unnamed east African clade	8.0	6.9–9.5
most recent common ancestor of Lake Malawi and Lake Victoria radiations	2.3	1.7–3.1
Crater Lake Barombi Mbo (Cameroon)	1.4	0.8–2.3

recent using the youngest possible fossil ages for outgroups, but only slightly so, with a mean age of 57.0 Ma (95% CI: 46.8–81.2 Ma). Using this same approach, it is also possible to determine probable times of origin for a series of more inclusive clades that contain Cichlidae: Ovalentaria, Percomorpha, Acanthopterygii, Acanthomorpha, Eurypterygii, Euteleostei and Teleostei. This exercise implies that no crown acanthomorph lineages are likely to be sufficiently ancient to have vicariant Gondwanan distributions, as we estimate the age of the group as between 106.4 Ma (95% CI: 98.5–132.2 Ma) and 109.2 Ma (95% CI: 98.5–136.0 Ma). The most restrictive group containing cichlids that we can date with this method and which is of sufficient apparent antiquity to have been affected by the initial rifting of Gondwana is Eurypterygii, the radiation containing Acanthomorpha, Myctophiformes and Aulopiformes [25]. Our estimates for the time of origin for this major teleost clade range between 131.1 Ma (95% CI: 104.9–163.2 Ma) and 142.1 Ma (95% CI: 126.2–166.2 Ma).

### (b) A molecular time scale for cichlid evolution

The phylogeny of Ovalentaria and the major cichlid lineages inferred from the 10 nuclear genes is similar to previous molecular and morphological analyses [8,10,19], with Etioplineae (India, Madagascar) resolved as the earliest-diverging clade and Ptychochrominae (Madagascar) as the sister lineage to the unnamed clade that contains the African (Pseudocrenilabrinae) and Neotropical (Cichlinae) cichlid lineages (figure 1; electronic supplementary material, figure S1). The 10 nuclear gene phylogeny preserves the parallels between patterns of cichlid interrelationships and the fragmentation history of Gondwana that has led to the prominence of vicariance biogeographic scenarios for this lineage [9]. However, the Bayesian random local molecular-clock analyses yield age estimates for the origin of cichlids consistent with those derived from analysis of fossils alone (figure 1 and table 1; electronic supplementary material, figure S1).

Based on the timing of Gondwanan fragmentation events, crown cichlids should occur in the Early Cretaceous or Late Jurassic [3,6,9,10]; however, the Bayesian random local molecular-clock analyses place the origin of the modern cichlid radiation near the Cretaceous–Palaeogene boundary (figure 1 and table 1; electronic supplementary material, figure S1), with a mean age estimate of 64.9 Ma (95% CI: 57.3–76.0 Ma). The estimated age of the most recent common ancestor

(MRCA) of cichlids and their sister lineage, *Pholidichthys*, is also younger (mean: 103.7 Ma; 95% CI: 92.0–118.4 Ma) than the initial rifting of Gondwana at approximately 135 Ma [11]. The mean estimated age of the MRCA of the African and Neotropical cichlids was 46.4 Ma (95% CI: 40.9–54.9 Ma), post-dating the final separation of Africa and South America by more than 40 Myr. The cichlid time tree confirms ages estimated in previous studies for the east African [26] (mean: 8.0 Ma; 95% CI: 6.9–9.5 Ma) and Cameroon crater lake Barombi Mbo [27] radiations (mean: 1.4 Ma; 95% CI: 0.8–2.3 Ma), verifying relatively young ages for these remarkable examples of adaptive radiation (figure 1 and table 1; electronic supplementary material, figure S2). The age estimate in the 10 nuclear gene inferred time tree closest to the timing of Gondwanan fragmentation is that of the inclusive (mean: 123.5 Ma; 95% CI: 111.4–136.2 Ma), but unnamed, percomorph clade that contains more than one-quarter of all living vertebrate species (approx. 16 570 species), including cichlids (see electronic supplementary material, figure S2).

## 4. Discussion

### (a) Congruence between palaeontological and molecular time scales for cichlid evolution

The application of two contrasting palaeontological approaches in calculating temporal range extensions yields strikingly congruent time scales for cichlid evolution. Both methods provide point estimates for the origin of the group that range between 57 and 60 Ma (Palaeocene), and strongly reject the possibility that crown cichlids are sufficiently old to have been affected by the initial rifting of Gondwana. Instead, upper limits for the origin of cichlids lie consistently within the late Late Cretaceous. This congruence is particularly compelling because the methods that yielded these comparable results share only one similarity in their calculations: both are constrained by the minimum age for cichlids as imposed by the oldest fossil example of the group.

Our molecular time tree provides a mean estimate for the timing of cichlid origin in the Palaeocene, but cannot reject the possibility that the group arose as early as the Late Cretaceous. This result is consistent with other recent molecular-clock estimates for the origin of cichlids that do not assume Gondwanan vicariance for the group *a priori*, and which range

in age from Late Cretaceous to Eocene [28–33]. In terms of point estimates and surrounding uncertainty, our revised molecular time scale is entirely consistent with the ages derived from analyses of the fossil record alone (figure 1). It is important to note that our relaxed-molecular-clock analysis shares no palaeontological data in common with either our analysis of the distribution of cichlid-bearing fossil horizons or our database of outgroup-based age constraints. We interpret the convergence of these three semi-independent approaches, which all deliver age estimates for cichlids that are within error of one another, as a consequence of genuine evolutionary signal that strongly contradicts the time scales for cichlids demanded by hypotheses of Gondwanan vicariance.

## (b) The timing of cichlid diversification: congruence and incongruence

Our estimates for the time of cichlid origin are congruent not only with one another, but also with previous molecular time scales for the evolution of this group that do not assume a Gondwanan vicariance scenario *a priori* [4,5,28–33]. The oldest such estimates from previous work are early Late Cretaceous [3], pre-dating our proposed time of origin by roughly 35–45 Myr. However, these more ancient dates derived from analysis of mitochondrial sequences, which are characterized by high rates of nucleotide substitution that might bias clock analyses towards older estimated times of divergence [25,34–36].

Generally, the only molecular-clock analyses to deliver time scales consistent with the predictions of the vicariance hypothesis were themselves calibrated using a combination of age constraints from the fossil record and Gondwanan fragmentation events [3,5,6,37]. There is no published relaxed-molecular-clock analysis that results in an Early Cretaceous or Jurassic origin of cichlids that is independent of the ages implied by the timing of the fragmentation of western Gondwana.

Our palaeontological time scales for Cichlidae constrain only the origin of the group, but our time-calibrated phylogeny permits investigation of the timing of deep divergences within the clade (table 1). We estimate the divergence of South American and African cichlids as Eocene, with the origin of the African cichlid crown within the same interval. This is consistent with the placement of the middle Eocene (approx. 46 Ma) †*Mahengechromis* as an early crown pseudocrenilabrine [12]. By contrast, our estimated Eocene–Oligocene age for the South American cichlid crown contradicts published interpretations of the fossil cichlids from the ‘Faja Verde’ level of the Luján Formation of Argentina. These fossils are often cited as early–middle Eocene in age [13,38,39], leading to calibration minima of 49 Ma in recent molecular clock studies [6]. However, the hard minimum for the age of these fossils is 33.9 Ma, which derives from radiometric dating of overlying tuff layers [40]. This more appropriate minimum age estimate only partially reconciles our time scale with previous phylogenetic interpretations of the Luján cichlids, each of which has been placed within the South American crown in association with specific cichline tribes (†*Protocara* as either a geophagine or a stem member of an unnamed clade comprising Chaetobranchini, Geophagini, Cichlasomatini and Heroini [39,41]; †*Gymnogeophagus eocenicus* as phylogenetically nested within a living genus [38]; and †*Plesioheros* as a crown heroine [13]). It is difficult to evaluate confidence in the evolutionary relationships proposed for these fossils because published analyses using morphological data do not provide support for

nodes in accompanying phylogenies (e.g. bootstrap resampling scores or Bayesian posterior probabilities). We also note that some phylogenetic hypotheses derive from successive reweighting exercises [39], while others assume restricted placement of fossil species prior to analysis [13]. There is no doubt that Luján cichlids are significant on account of their antiquity and geographical provenance. However, in the absence of demonstrably robust phylogenetic placements of these fossil lineages within a group well known for convergent morphological evolution [42], their exact implications for the timing of major events in cichlid evolution are likely to remain ambiguous.

## (c) Comparison with other putative examples of Gondwanan vicariance

Among vertebrates assumed to have limited dispersal ability across marine barriers, cichlids are not unique in showing a broad distribution across southern landmasses combined with a fossil record that commences long after the tectonic break-up of Gondwana. Several groups of freshwater fishes, reptiles, mammals and plants show disjunct distributions, with members present in South America and Africa, but only a few instances seem definitively attributable to drift-based vicariance [14,43,44]. Instead, molecular clock analyses for a range of groups with apparent vicariant distributions across southern continents [45–48] paint a picture of widespread ‘pseudo-congruence’, where similar biogeographic patterns originate at different times that may be disjunct with the age of specific palaeogeographic events [49].

Our consistent time scales for cichlid evolution reject Gondwanan vicariance as a viable mechanism for the modern distribution of the group, but they demand what can only be considered a series of highly unlikely trans-oceanic dispersal events. Like the fossil record, the salinity tolerance of cichlids has been subjected to contrasting interpretations; it has been cited as both consistent [12] and inconsistent [10] with marine dispersal. Experimental evidence points to high salinity tolerance in some cichlids [50,51], but the fact that no cichlid inhabits the open ocean indicates that long-distance marine migration is improbable. Dispersal across the south Atlantic would appear to be especially unlikely, given that it measured roughly 1000 km [52] in width by the time of the inferred divergence between South American and African cichlids in the Eocene (figure 1). Despite the presence of a substantial marine barrier, it is clear that at least two groups of terrestrial mammals—primates and hystricognath rodents—dispersed from Africa to South America at approximately this time [53]. More generally, there is strong evidence from other animal groups and plants for surprisingly high levels of biotic interchange between South America and Africa throughout the Late Cretaceous and Palaeogene [54,55]. Geological evidence indicates the presence of a chain of now-submerged islands across the south Atlantic during the Palaeogene [52]. These islands coincided with strong east-to-west palaeocurrents across the south Atlantic and both have been invoked as key elements of a selective dispersal route from Africa to South America during the Eocene [12,52]. It is also possible that freshwater plumes, such as that produced by the modern Congo River [56], provided corridors of brackish surface water that could have permitted migration by freshwater taxa across a narrower marine barrier during the Palaeogene.



Coeval examples of migration provide circumstantial evidence for the possibility of trans-Atlantic dispersal, and geographical factors during the Eocene would appear to have facilitated the crossing, but cichlid migration across the south Atlantic and other marine barriers nevertheless remains an extraordinary claim. However, the evolutionary time scale inferred for cichlids on the basis of both fossils and molecules demands that this hypothesis is given serious consideration rather than being dismissed *a priori*.

Our estimation of consistent palaeontological and molecular ages for the origin of cichlids adds to a growing number of studies reporting close congruence between divergence time estimates from 'rocks' and 'clocks', in cases where these approaches had previously delivered wildly different evolutionary time scales [57]. This convergence would seem to signal the end of an era dominated by debates on the relative merits of molecular and fossil data, permitting molecular

biologists and palaeontologists to move forward on addressing questions related to the timing of major events underpinning the origin of modern biodiversity.

**Acknowledgements.** We thank J. Friel of the Cornell University Museum of Vertebrates, USA, and A. C. Bentley and E. O. Wiley of the Biodiversity Institute of the University of Kansas, USA for generous gifts of tissue specimens. G. Watkins-Colwell assisted with museum collections. G. Lloyd and D. Bellwood commented on previous versions of this work, and we thank A. Murray and M. Vences for their insightful reviews and J. Hutchinson for editorial assistance.

**Funding statement.** This research was supported by the Peabody Museum of Natural History and grants awarded to P.C.W. and T.J.N. from the National Science Foundation, USA (DEB-0444842, DEB-0716155, DEB-0717009, DEB-0732642, ANT-0839007, DEB-1060869, DEB-1061806 and DEB-1061981). M.F. is supported by the Natural Environment Research Council (NERC NE/I005536/1). C.H.M. was supported by the National Geographic Society (Young Explorers Grant) and the American Philosophical Society (Lewis and Clark Field Grant).

## References

- Wagner CE, Harmon LJ, Seehausen O. 2012 Ecological opportunity and sexual selection together predict adaptive radiation. *Nature* **487**, 366–369. (doi:10.1038/nature11144)
- Kocher TD. 2004 Adaptive evolution and explosive speciation: the cichlid fish model. *Nat. Rev. Genet.* **5**, 288–298. (doi:10.1038/nrg1316)
- 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. (doi:10.1186/1471-2148-8-215)
- Vences M, Freyhof J, Sonnenberg R, Kosuch J, Veith M. 2001 Reconciling fossils and molecules: Cenozoic divergence of cichlid fishes and the biogeography of Madagascar. *J. Biogeogr.* **28**, 1091–1099. (doi:10.1046/j.1365-2699.2001.00624.x)
- Genner MJ, Seehausen O, Lunt DH, Joyce DA, Shaw PW, Carvalho GR, Turner GF. 2007 Age of cichlids: new dates for ancient lake fish radiations. *Mol. Biol. Evol.* **24**, 1269–1282. (doi:10.1093/molbev/msm050)
- López-Fernández H, Arbour JH, Winemiller KO, Honeycutt RL. 2013 Testing for ancient adaptive radiations in neotropical cichlid fishes. *Evolution* **67**, 1321–1337.
- Blakey RC. 2008 Gondwana paleogeography from assembly to breakup—a 500m.y. odyssey. *Geol. Soc. Am. Spec. Pap.* **441**, 1–28. (doi:10.1130/2008.2441(01))
- Stiassny MJ. 1991 Phylogenetic intrarelationships of the family Cichlidae: an overview. In *Cichlid fishes: behaviour, ecology, and evolution* (ed. MHA Keenleyside), pp. 1–35. London, UK: Chapman & Hall.
- Chakrabarty P. 2004 Cichlid biogeography: comment and review. *Fish Fish.* **5**, 97–119. (doi:10.1111/j.1467-2979.2004.00148.x)
- Sparks JS, Smith WL. 2004 Phylogeny and biogeography of cichlid fishes (Teleostei: Perciformes: Cichlidae). *Cladistics* **20**, 501–517. (doi:10.1111/j.1096-0031.2004.00038.x)
- Jokat W, Boebel T, König M, Meyer U. 2003 Timing and geometry of early Gondwana breakup. *J. Geophys. Res.* **108**, 2428. (doi:10.1029/2002JB001802)
- Murray AM. 2001 The fossil record and biogeography of the Cichlidae (Actinopterygii: Labroidae). *Biol. J. Linn. Soc.* **74**, 517–532.
- Perez PA, 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. (doi:10.1590/S1679-62252010000300008)
- Lundberg JG. 1993 African-South American freshwater fish clades and continental drift: problems with a paradigm. In *Biological relationships between Africa and South America* (ed. P Goldblatt), pp. 156–199. New Haven, CT: Yale University Press.
- Unmack PJ. 2001 Biogeography of Australian freshwater fishes. *J. Biogeogr.* **28**, 1053–1089. (doi:10.1046/j.1365-2699.2001.00615.x)
- Sparks JS, Smith WL. 2005 Freshwater fishes, dispersal ability, and nonevidence: 'Gondwana Life Rafts' to the rescue. *Syst. Biol.* **54**, 158–165. (doi:10.1080/10635150590906019)
- Marshall CR. 1997 Confidence intervals on stratigraphic ranges with nonrandom distributions of fossil horizons. *Paleobiology* **23**, 165–173.
- Hedman MM. 2010 Constraints on clade ages from fossil outgroups. *Paleobiology* **36**, 16–31. (doi:10.1666/0094-8373-36.1.16)
- Wainwright PC, Smith WL, Price SA, Tang KL, Sparks JS, Ferry LA, Kuhn KL, Eytan RI, Near TJ. 2012 The evolution of pharyngognath: a phylogenetic and functional appraisal of the pharyngeal jaw key innovation in Labroidae and beyond. *Syst. Biol.* **61**, 1001–1027. (doi:10.1093/sysbio/sys060)
- Strauss D, Sadler PM. 1989 Classical confidence intervals and Bayesian probability estimates for ends of local taxon ranges. *Math. Geol.* **21**, 411–427. (doi:10.1007/BF00897326)
- Marshall CR. 1990 Confidence intervals on stratigraphic ranges. *Paleobiology* **16**, 1–10.
- Hurley IA *et al.* 2007 A new time-scale for ray-finned fish evolution. *Proc. R. Soc. B* **274**, 489–498. (doi:10.1098/rspb.2006.3749)
- Drummond AJ, Suchard MA. 2010 Bayesian random local clocks, or one rate to rule them all. *BMC Biol.* **8**, 114. (doi:10.1186/1741-7007-8-114)
- Drummond AJ, Rambaut A. 2007 BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evol. Biol.* **7**, 214. (doi:10.1186/1471-2148-7-214)
- Near TJ, Eytan RI, Dornburg A, Kuhn KL, Moore JA, Davis MP, Wainwright PC, Friedman M, Smith WL. 2012 Resolution of ray-finned fish phylogeny and timing of diversification. *Proc. Natl Acad. Sci. USA* **109**, 13 698–13 703. (doi:10.1073/pnas.1206625109)
- Koblmüller S, Egger B, Sturmbauer C, Sefc KM. 2010 Rapid radiation, ancient incomplete lineage sorting and ancient hybridization in the endemic Lake Tanganyika cichlid tribe Tropheini. *Mol. Phylogenet. Evol.* **55**, 318–334. (doi:10.1016/j.ympev.2009.09.032)
- Schliwen UK, Klee B. 2004 Reticulate speciation in Cameroonian crater lake cichlids. *Front. Zool.* **1**, 1–12. (doi:10.1186/1742-9994-1-5)
- Alfaro ME, Santini F, Brock C, Alamillo H, Dornburg A, Rabosky DL, Carnevale G, Harmon LJ. 2009 Nine exceptional radiations plus high turnover explain species diversity in jawed vertebrates. *Proc. Natl Acad. Sci. USA* **106**, 13 410–13 414. (doi:10.1073/pnas.0811087106)
- 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**, 164. (doi:10.1186/1471-2148-9-194)
- Near TJ *et al.* 2013 Phylogeny and tempo of diversification in the superradiation of spiny-rayed fishes. *Proc. Natl Acad. Sci. USA* **110**, 12 738–12 743. (doi:10.1073/pnas.1304661110)

31. Crottini A, Madsen O, Poux C, Strauss A, Vieites DR, Vences M. 2012 Vertebrate time-tree elucidates the biogeographic pattern of a major biotic change around the K-T boundary in Madagascar. *Proc. Natl Acad. Sci. USA* **109**, 5358–5363. (doi:10.1073/pnas.1112487109)
32. Betancur RR *et al.* 2013 The tree of life and a new classification of bony fishes. *PLoS Curr.* (doi:10.1371/currents.tol.53ba26640df0cace75bb165c8c26288)
33. McMahan CD, Chakrabarty P, Sparks JS, Smith WL, Davis MP. 2013 Temporal patterns of diversification across global cichlid biodiversity (Acanthomorpha: Cichlidae). *PLoS ONE* **8**, e71162. (doi:10.1371/journal.pone.0071162)
34. Brandley MC, Wang Y, Guo X, Nieto Montes de Oca A, Feria-Ortiz M, Hikida T, Ota H. 2011 Accommodating heterogeneous rates of evolution in molecular divergence dating methods: an example using intercontinental dispersal of *Plestiodon* (*Eumeces*) lizards. *Syst. Biol.* **60**, 3–15. (doi:10.1093/sysbio/syq045)
35. Phillips MJ. 2009 Branch-length estimation bias misleads molecular dating for a vertebrate mitochondrial phylogeny. *Gene* **441**, 132–140. (doi:10.1016/j.gene.2008.08.017)
36. Zheng Y, Peng R, Kuro-O M, Zeng X. 2011 Exploring patterns and extent of bias in estimating divergence time from mitochondrial DNA sequence data in a particular lineage: a case study of salamanders (Order Caudata). *Mol. Biol. Evol.* **28**, 2521–2535. (doi:10.1093/molbev/msr072)
37. Miya M *et al.* 2010 Evolutionary history of anglerfishes (Teleostei: Lophiiformes): a mitogenomic perspective. *BMC Evol. Biol.* **10**, 58. (doi:10.1186/1471-2148-10-58)
38. Malabarba MC, Malabarba LR, Del Papa C. 2010 *Gymnogeophagus eocenicus*, n. sp. (Perciformes: Cichlidae), an Eocene cichlid from the Luján Formation in Argentina. *J. Vert. Paleontol.* **30**, 341–350. (doi:10.1080/02724631003618348)
39. Malabarba MC, Zuleta O, del Papa C. 2006 *Proterocara argentina*, a new fossil cichlid from the Luján Formation, Eocene of Argentina. *J. Vert. Paleontol.* **26**, 267–275. (doi:10.1671/0272-4634(2006)26[267:PAANFC]2.0.CO;2)
40. del Papa C, Kirschbaum A, Powell J, Brod A, Hongn F, Pimentel M. 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. (doi:10.1016/j.jsames.2009.06.004)
41. Smith WL, Chakrabarty P, Sparks JS. 2008 Phylogeny, taxonomy, and evolution of Neotropical cichlids (Teleostei: Cichlidae: Cichlinae). *Cladistics* **24**, 625–641. (doi:10.1111/j.1096-0031.2008.00210.x)
42. Rüber L, Adams DC. 2001 Evolutionary convergence of body shape and trophic morphology in cichlids from Lake Tanganyika. *J. Evol. Biol.* **14**, 325–332. (doi:10.1046/j.1420-9101.2001.00269.x)
43. de Queiroz A. 2005 The resurrection of oceanic dispersal in historical biogeography. *Trends Ecol. Evol.* **20**, 68–73. (doi:10.1016/j.tree.2004.11.006)
44. Beaulieu JM, Tank DC, Donoghue MJ. 2013 A Southern Hemisphere origin for campanulid angiosperms, with traces of the break-up of Gondwana. *BMC Evol. Biol.* **13**, 80. (doi:10.1186/1471-2148-13-80)
45. Gamble T, Bauer AM, Colli GR, Greenbaum E, Jackman TR, Vitt LJ, Simons AM. 2011 Coming to America: multiple origins of New World geckos. *J. Evol. Biol.* **24**, 231–244. (doi:10.1111/j.1420-9101.2010.02184.x)
46. Rowe DL, Dunn KA, Adkins RM, Honeycutt RL. 2010 Molecular clocks keep dispersal hypotheses afloat: evidence for trans-Atlantic rafting by rodents. *J. Biogeogr.* **37**, 305–324. (doi:10.1111/j.1365-2699.2009.02190.x)
47. Cook LG, Crisp MD. 2005 Not so ancient: the extant crown group of *Nothofagus* represents a post-Gondwanan radiation. *Proc. R. Soc. B* **272**, 2535–2544. (doi:10.1098/rspb.2005.3219)
48. Knapp M, Stockler K, Havell D, Delsuc F, Sebastiani F, Lockhart PJ. 2005 Relaxed molecular clock provides evidence for long-distance dispersal of *Nothofagus* (Southern Beech). *PLoS Biol.* **3**, 38–43. (doi:10.1371/journal.pbio.0030014)
49. Donoghue MJ, Moore BR. 2003 Toward an integrative historical biogeography. *Integr. Comp. Biol.* **43**, 261–270. (doi:10.1093/icb/43.2.261)
50. Suresh AV, Lin CK. 1992 Tilapia culture in saline waters: a review. *Aquaculture* **106**, 201–226. (doi:10.1016/0044-8486(92)90253-H)
51. Uchida K, Kaneko T, Miyazaki H, Hasegawa S, Hirano T. 2000 Excellent salinity tolerance of Mozambique tilapia (*Oreochromis mossambicus*): elevated chloride cell activity in the branchial and opercular epithelia of the fish adapted to concentrated seawater. *Zool. Sci.* **17**, 149–160. (doi:10.2108/Zsj.17.149)
52. de Oliveira FB, Molina EC, Marroig G. 2009 Paleogeography of the South Atlantic: a route for primates and rodents into the New World? In *South American primates* (eds PA Garber, A Estrada, JC Bicca-Marques, EW Heymann, KB Strier), pp. 55–68. New York, NY: Springer.
53. Loss-Oliveira L, Aguiar BO, Schrago CG. 2012 Testing synchrony in historical biogeography: the case of New World primates and Hystriognathi rodents. *Evol. Bioinform.* **8**, 127–137. (doi:10.4137/Ebo.S9008)
54. Ezcurra MD, Agnolin FL. 2012 A new global palaeobiogeographical model for the Late Mesozoic and Early Tertiary. *Syst. Biol.* **61**, 553–566. (doi:10.1093/sysbio/syr115)
55. Morley RJ. 2000 *Origin and evolution of tropical rain forests*, p. 362. Chichester, NY: Wiley.
56. Measey GJ, Vences M, Drewes RC, Chiari Y, Melo M, Bourles B. 2007 Freshwater paths across the ocean: molecular phylogeny of the frog *Ptychocheilichthys newtoni* gives insights into amphibian colonization of oceanic islands. *J. Biogeogr.* **34**, 7–20. (doi:10.1111/j.1365-2699.2006.01589.x)
57. Erwin DH, Laflamme M, Tweedt SM, Sperling EA, Pisani D, Peterson KJ. 2011 The Cambrian conundrum: early divergence and later ecological success in the early history of animals. *Science* **334**, 1091–1097. (doi:10.1126/science.1206375)



Electronic Supplementary Material for:

## Molecular and Fossil Evidence Place the Origin of Cichlids Long After Gondwanan Rifting

Matt Friedman<sup>1\*</sup>, Benjamin P. Keck<sup>2</sup>, Alex Dornburg<sup>3</sup>, Ron I Eytan<sup>3</sup>, Christopher H. Martin<sup>4</sup>, C. Darrin Hulsey<sup>5</sup>, Peter C. Wainwright<sup>4</sup>, & Thomas J. Near<sup>3\*</sup>

<sup>1</sup>. *Department of Earth Sciences, University of Oxford, South Parks Road, Oxford OX1 3AN, UK.*

<sup>2</sup>. *Department of Ecology & Evolutionary Biology, University of Tennessee, TN 37996 U.S.A.*

<sup>3</sup>. *Department of Ecology and Evolutionary Biology and Peabody Museum of Natural History, Yale University, New Haven, CT 06520 U.S.A.*

<sup>4</sup>. *Department of Evolution and Ecology, University of California, Davis, CA 95616 U.S.A.*

<sup>5</sup>. *Department of Ecology and Evolution, University of Tennessee, Knoxville, TN 37996, U.S.A.*

*\*Authors for correspondence: matt.friedman@earth.ox.ac.uk, thomas.near@yale.edu*

## 1. PALAEOONTOLOGICAL ESTIMATES OF TIME OF EVOLUTIONARY ORIGIN

### (a) *Estimating times of evolutionary origin using fossil occurrences of the focal lineage*

Strauss & Sadler (1989) provide a Bayesian framework for inferring credible intervals on stratigraphic ranges, but this has not been as widely adopted as their classical approach to calculating confidence intervals (reviewed in Marshall, 1990). However, it is their Bayesian solution that permits legitimate statements about the probability of a stratigraphic range (or evolutionary lineage) commencing within a particular interval, and it is this approach that we adopt here. For completeness, we have also calculated classical confidence intervals for the same problem (Strauss & Sadler, 1989; Marshall, 1990), and show that these are broadly comparable to Bayesian credible intervals. As in Strauss & Sadler's (1989: figure 3) worked examples, we find that Bayesian methods deliver older point estimates and upper bounds for times of origin than the classical approach.

The Bayesian estimates of Strauss & Sadler (1989) are conditioned only on a prior claim that the endpoint of the range of a lineage lies within a certain interval. Here, we impose the prior that the origin of cichlids occurred after the Palaeozoic (252.2 Ma), which is consistent with both the fossil record and the Gondwanan hypothesis of cichlid biogeography. We have found this method is relatively insensitive to variation on the hard prior provided that it substantially predates the probable time of origin of the group.

Strauss & Sadler's (1989) method assumes uniform preservation with respect to time (see their equation 23), a condition that is almost certainly violated in the case of cichlids where the number of palaeoenvironmentally and taphonomically suitable sampling horizons decreases substantially with increasing age of geological intervals

(figure 1). Marshall (1997) provided a logical framework for dealing with this problem in the context of classical confidence intervals. Where preservation potential is uniform over time, confidence intervals can be expressed simply as the product of an observed taxon range and a scaling factor derived from the number of fossil-bearing horizons and desired level of confidence. Marshall's solution to cases in which preservation potential is non-uniform involves recasting the problem in terms of recovery potential rather than time, and requires an estimate of a function describing recovery potential over time (figure 1).

The area under this curve between the first and last occurrence of a lineage is the analogue of lineage duration in the special case of uniform preservation, with the confidence interval measured as the product of this summed recovery potential and the scaling factor reflecting the number of fossil-bearing horizons and desired level of confidence. This confidence interval is measured in terms of recovery potential rather than time, and must be rescaled in order to be interpretable within an evolutionary and stratigraphic framework. This is achieved by integrating the function for recovery potential from the oldest fossil in question to some earlier point at which the area under the curve is equal to the confidence interval, as measured in terms of recovery potential. This point, which corresponds to a date predating the oldest fossil for the focal lineage, is the confidence limit for the first appearance of the group measured in terms of time. It follows that instances with uniform preservation potential represent a special case where accrued recovery potential increments linearly with elapsed time, meaning that rescaling of confidence intervals in terms of recovery potential is unnecessary.

We have applied the same logic to Strauss & Sadler's (1989) Bayesian formulation in order to accommodate heterogeneity in the fossil record of freshwater



fishes on Gondwanan landmasses from the Triassic to the Recent. For their equations 22-27, we have measured the observed range of cichlids and the prior not in terms of time, but rather in terms of summed recovery potential. In our example, these are equal to the areas under our empirical estimate of the recovery potential function between: (i) the Recent and the oldest fossil (the analogue of the observed lineage duration in the uniform case); and (ii) the oldest fossil and the lower limit of the stratigraphic window of consideration (the analogue of the prior belief for the age of the base of the interval in which the first appearance occurs). Values for four variants of the recovery potential function are given in table S1. Posterior distributions, as well as point estimates and credible intervals, for lineage endpoints are expressed in terms of accrued recovery potential. Because all environmentally and taphonomically appropriate horizons within the Eocene time bin appear to predate the oldest cichlid fossils, we consider recovery potential in this bin prior to the first appearance of cichlids to be zero, rather than the average fossil horizon density in the Eocene. This is a generous test of the hypothesis of Gondwanan vicariance, as it yields older age estimates for cichlids than those calculated using a recovery potential function that does not acknowledge the presence of the early Eocene gap in the fossil record.

We calculated medians and two-tailed 95% credible intervals for the time of origin of cichlids directly from posterior distributions, and point estimates using equation 27 in Strauss & Sadler (1989). These values are initially calculated in terms of recovery potential, and subsequently converted to a temporal scale in light of the empirical function for recovery potential.

Observed cichlid duration and maximum age priors (measured in terms of recovery potential and time) are given in table S2, along with point estimates (means) and 95% credible intervals (measured in terms of time). Equivalent values using classical confidence intervals (Strauss & Sadler, 1989; Marshall, 1990, 1997) are provided in table S3.

Interval	Gondwanan	Gondwanan <sup>A</sup>	Modern range	Modern range <sup>A</sup>
Holocene	598.2906	0	598.2906	0
Pleistocene	20.96029	1.164461	20.96029	1.164461
Pliocene	7.285974	0.728597	7.285974	0.728597
Miocene	3.107871	0.508561	2.938351	0.452054
Oligocene	1.287948	0.459982	0.827967	0.459982
Eocene	0.633484	0.271493	0.40724	0.226244
Paleocene	1.4	0.4	0.7	0.4
Late Cretaceous	0.985507	0.173913	0.811594	0.173913
Early Cretaceous	0.314607	0.11236	0.247191	0.089888
Late Jurassic	0.648649	0.216216	0.216216	0.216216
Middle Jurassic	0.283019	0.377358	0.188679	0.188679
Early Jurassic	0.220588	0.036765	0.073529	0.036765
Late Triassic	0.207715	0.118694	0.207715	0.118694
Middle Triassic	0.245902	0.163934	0.245902	0.163934
Early Triassic	1.4	1.2	1	0.6

Table S1. Empirical recovery functions, based on the distribution of sedimentary horizons with taphonomic and geographic profiles consistent with yielding cichlids. Column names refer to the geographic scope of analysis. ‘Modern’ indicates that only the fossil records of regions currently inhabited by cichlids are considered. ‘Gondwanan’ indicates that the fossil records all former Gondwanan landmasses are considered. Superscript ‘A’ indicates that only deposits yielding articulated remains are considered. Units are in horizons·Ma<sup>-1</sup>. Source data are provided at the Dryad Digital Repository (<http://dx.doi.org/10.5061/dryad.48f62>).



<b>Geographic /taphonomic partion</b>	<b><math>N_{\text{horizons}}</math></b>	<b>Duration (cumulative recovery potential)</b>	<b>Maximum age prior (cumulative recovery potential)</b>	<b>Point estimate (Ma)</b>	<b>Lower 0.95 credible interval (Ma)</b>	<b>Upper 0.95 credible interval (Ma)</b>
Gondwanan	70	164	264	57.75	56.05	62.44
Gondwanan <sup>A</sup>	18	25	61	60.17	56.10	77.84
Modern range	70	151	220	59.22	56.09	67.60
Modern range <sup>A</sup>	18	23	53	59.83	56.09	75.05

Table S2. Bayesian point estimates and credible intervals for the timing of cichlid origin.

Values for duration and the maximum age prior are given in terms of recovery potential (measured as the number of appropriate fossil-bearing horizons). Point estimates are given by equation 26 in Strauss & Sadler (1989). Credible intervals are calculated from the posterior distribution specified by equation 26 in Strauss & Sadler (1989).

<b>Geographic /taphonomic partion</b>	<b><math>N_{\text{horizons}}</math></b>	<b>Duration (cumulative recovery potential)</b>	<b>Point estimate (Ma)</b>	<b>Lower 0.95 confidence interval (Ma)</b>	<b>Upper 0.95 confidence interval (Ma)</b>
Gondwanan	70	164	57.70	56.04	58.75
Gondwanan <sup>A</sup>	18	25	59.68	56.09	77.34
Modern range	70	151	59.13	56.08	61.49
Modern range <sup>A</sup>	18	23	59.38	56.09	75.05

Table S3. Maximum likelihood point estimates and classical confidence intervals for the timing of cichlid origin. Values for duration are given in terms of recovery potential (measured as the number of appropriate fossil-bearing horizons). Point estimates are given by equation 8 in Strauss & Sadler (1989). Confidence intervals are calculated using equation 3 in Marshall (1997).

<b>Geographic/taphonomic partition</b>	<b>Bayesian</b>	<b>Classical</b>
Gondwanan	0.02706	0.01201
Gondwanan <sup>A</sup>	0.06622	0.06900
Modern range	0.0330	0.01624
Modern range <sup>A</sup>	0.06927	0.07306

Table S4. Factors by which the pre-Eocene preservation potential of cichlids must be less than their observed fossil record to allow for a time of origin that is clearly consistent with the Gondwanan vicariance hypotheses (i.e., a 50% probability that the true age of the group exceeds 135 Ma). In all cases, the pre-Eocene fossil record of cichlids must be at least an order of magnitude worse than the observed Eocene and younger record, with some cases demanding that the record be nearly two orders of magnitude worse.



***(b) Estimating times of evolutionary origin using the ages of outgroups to the focal lineage***

Hedman (2010) proposed a simple Bayesian approach to estimating the time of origin for a group of interest based on the ages of outgroups to the focal clade. Two variants of this approach were proposed. The first makes use of all outgroups, and assigns each stratigraphically inconsistent outgroup a minimum time of origin based on the oldest taxon that shares a more recent common ancestor with the focal lineage. The second disregards these stratigraphically inconsistent outgroups, and it is this more conservative method that we use here. Both variants use a user-specified maximum bound on clade ages ( $t_0$ ), which is set to some value believed to exceed any of the divergences concerned. Hedman (2010) has shown that this outgroup-based method is relatively insensitive to variation in  $t_0$ .

The logic underlying this approach is simple. The time of origin for the deepest branching outgroup is constrained between  $t_0$  and the first appearance of that lineage in the fossil record. The time of divergence is considered uniformly distributed over this interval. The time of origin for the second deepest branching outgroup is constrained between its first appearance in the fossil record and the time of divergence of the deepest branching outgroup. The time of origin for the second deepest outgroup is considered uniformly distributed over this span, but the upper limit on the age varies depending on the time of origin for the deepest outgroup. The time of origin for the second deepest outgroup is evaluated based on all possible ages for the deepest outgroup, cumulatively generating a non-uniform distribution of ages. This approach is repeated up the

phylogenetic backbone, until a probability distribution for the age of the focal group can be produced.

We have used the script provided by Hedman (2010) for the statistical computing language R (R Core Team, 2013) for calculating posterior age distributions for cichlids. Hedman's algorithm evaluates the posterior age distribution at discrete time steps; 1,000 such steps were used in our analyses. A series of phylogenetic hypotheses focused on different parts of the ray-finned fish tree were consulted in order to generate an outgroup sequence. For the relationships of crown-group teleosts, we relied principally on the molecular phylogeny presented by Near et al. (2012). The branching sequence of outgroups to the teleost crown is based on trees presented by Patterson (1977), Patterson & Rosen (1977), Arratia (1997), and Arratia & Tischlinger (2010). There is uncertainty surrounding the earliest representatives of particular lineages. In such cases, we have considered both the youngest and oldest candidates. We analyzed minimum and maximum outgroup ages separately, providing two mean estimates and associated credible intervals for the time of cichlid origin. For both analyses, we have specified  $t_0$  as 322.8 Ma, based on the stem neopterygian †*Discoserra*. Justifications for this hard upper bound and all outgroup ages are provided below:

**Cichlidae.** *Taxon.* †*Mahengechromis plethos* and other members of the †*Mahengechromis* species flock (Murray 2000, 2001a, b). *Geological horizon and locality.* Mahenge, Singida Plateau, Tanzania. *Stratigraphy and age estimate.* U-Pb dating of a zircon crystal provides a well-constrained age estimate of  $45.83 \pm 0.17$  Ma for the kimberlite at Mahenge (Harrison et al., 2001), but the lacustrine

sediments yielding †*Mahengechromis* overlie these intrusive units and must postdate them. Harrison et al. (2001) provide a convincing argument, based on other crater lakes, that kimberlite emplacement likely predates lacustrine deposition and complete basin infill by only 0.2 to 0.1 Ma. Here we adopt an age of 46.0 Ma for the deposits at Mahenge (cf. Harrison et al., 2001). *Phylogenetic placement and additional notes.* Soft-tissue anatomy and other details not preserved in fossils represent the most definitive cichlid synapomorphies (Stiassny, 1981). Identification of †*Mahengechromis* as a cichlid is supported by structure of the lower pharyngeal jaw, details of squamation, and meristic counts (Murray, 2000, 2001a, b). A single supraneural and ctenoid scales place †*Mahengechromis* in crown-group Cichlidae as a pseudocrenilabrine (Murray 2001a: fig. 3). Eocene cichlids from Argentina are sometimes cited as the earliest representatives of this family, but the minimum age constraint for these specimens is younger than that for †*Mahengechromis*. The Argentine cichlids are discussed in more detail in the body of the text.

**Outgroup 1: Atherinomorpha.** *Taxon.* †*Ramphexocetus volans* (Bannikov et al., 1985). *Geological horizon and locality.* ‘Calcarei nummulitici’ of Bolca, Italy. *Stratigraphy and age estimate.* Medizza (1975) assigned the ‘Calcarei nummulitici’ deposits to NP 14 on the basis of calcareous nannoplankton, while the larger foraminifera place the unit within SBZ 11 (Papazzoni & Trevisani, 2006). NP 14 and SBZ 11 overlap narrowly in the late Ypresian, and the age of the fish beds at Bolca can be approximately correlated with the base of NP 14, which is dated as

49.11 Ma (Gradstein et al., 2012). *Phylogenetic placement and additional notes.*

†*Ramphexocetus* can be placed within the atherinomorph subgroup Beloniformes on the basis of its caudal fin (Rosen & Parenti, 1981), which has an expanded ventral lobe that contains more principal fin-rays than the dorsal lobe (Bannikov et al., 1985). The greatly elongated dentary of †*Ramphexocetus* places the genus within the Exocetoidei, a suborder of Beloniformes (Rosen & Parenti, 1981).

**Outgroup 2: Unnamed percomorph clade containing Syngnathiformes,**

**Tetraodontiformes (oldest minimum).** *Taxon.* †*Plectocretacicus clarae* (tetraodontiform). *Geological horizon and locality.* Sannine Limestone, Hakel, Lebanon. *Stratigraphy and age estimate.* The ammonite †*Mantelliceras mantelli* is known from the local subdivision of the Lebanese Cenomanian that contains the Hakel horizon (Zummofen, 1926; Dalla Vecchia et al., 2002). †*M. mantelli* defines the earliest complete ammonite zone of the Cenomanian, the top of which can be dated as approximately 98.0 Ma (Gradstein et al., 2012). *Phylogenetic placement and additional notes.* †*Plectocretacicus* has been placed as a stem tetraodontiform in a manual cladistic solution presented by Tyler & Sorbini (1996), and a formal analysis of a morphological character matrix presented by Tyler & Santini (2002). This genus is reconstructed placed in a clade on the tetraodontiform stem along with two other taxa: †*Cretatriacanthus* and †*Protriacanthus*. These three stem tetraodontiforms differ strikingly in their morphology, and some are linked to tetraodontiforms by features present in other groups of fishes. For this reason, we also adopt a minimum age for

Tetraodontiformes based on †*Cretatriacanthus*, the youngest of the †plectocretacicoids (see below).

**Outgroup 2: Unnamed percomorph clade containing Syngnathiformes,**

**Tetraodontiformes (youngest minimum).** *Taxon.* †*Cretatriacanthus guidottii* (tetraodontiform; Tyler & Sorbini, 1996) and †*Gasterorhamphosus zuppichinii* (syngnathiform; Sorbini, 1981). *Geological horizon and locality.* ‘Calcarei Melissano’ (of historical usage) at Canale near Nardò, Italy. *Stratigraphy and age estimate.* The fish beds at Nardò yield the calcareous nannofossil †*Uniplanarus trifidus* (reported as reported as †*Quadrum trifidum* by Medizza & Sorbini, 1980). The first appearance of this species marks the beginning of Calcareous Nannoplankton Zone CC23, and it makes its last appearance in the middle of CC24. The top of CC24 is roughly equivalent to the top of the †*Baculites clinolobatus* Zone of the Western Interior Seaway, which contains a bentonite horizon dated as  $70.08 \pm 0.37$  Ma (Gradstein et al., 2012). From this we estimate an age of 69.71 Ma. *Phylogenetic placement and additional notes.*

†*Cretatriacanthus* has been placed as a stem tetraodontiform in a manual cladistic solution presented by Tyler & Sorbini (1996), and a formal analysis of a morphological character matrix presented by Tyler & Santini (2002). It is the geologically youngest representative of †Plectocretacicoidei, a putative clade of stem tetraodontiforms. †*Gasterorhamphosus* has not been included in a formal cladistic analysis, but it presents numerous derived characters of Syngnathiformes including a tubular snout, absence of an anal-fin spine, an enlarged and serrated



dorsal-fin spine, absence of ribs, an enlarged posterodorsal process of the cleithrum, a rod-like anteroventral process of the coracoids, and simple pectoral rays (Pietsch, 1978; Orr, 1995).

**Outgroup 3: Beryciformes.** *Taxon.* †*Stichocentrus liratus*, †*Stichocentrus elegans*, †*Stichocentrus spinulosus* (Holocentroidei); †*Stichopteryx lewisi*, †*Lissoberyx dayi*, †*Lissoberyx arambourgi*, †*Microcapros libanicus*, †*Libanoberyx spinulosus* (Trachichthyoidei) (Forey et al., 2003). *Geological horizon and locality.* Sannine Limestone, Hajula, Lebanon. *Stratigraphy and age estimate.* The Hajula horizon lies below that of Hakel within the Sannine limestone, but its age is constrained by the same biostratigraphic marker: the ammonite †*Mantelliceras mantelli*. The top of the †*M. mantelli* Zone can be dated as approximately 98.0 Ma (Gradstein et al., 2012). *Phylogenetic placement and additional notes.* Many of the beryciforms from Hajula have not been included in any formal cladistic analyses. However, they all bear pelvic-fin spines, which are a synapomorphy of Acanthopterygii (Beryciformes plus Percomorpha; Johnson & Patterson, 1993).

**Outgroup 4: non-acanthopterygian Acanthomorpha.** *Taxon.* †*Aipichthys minor* (Forey et al., 2003). *Geological horizon and locality.* Sannine Limestone, Hajula, Lebanon. *Stratigraphy and age estimate.* The Hajula horizon lies below that of Hakel within the Sannine limestone, but its age is constrained by the same biostratigraphic marker: the ammonite †*Mantelliceras mantelli*. The top of the †*M.*

*mantelli* Zone can be dated as approximately 98.0 Ma (Gradstein et al., 2012).

*Phylogenetic placement and additional notes.* †*Aipichthys* is considered an acanthomorph based on the derived presence of dorsal- and anal-fin spines (Patterson, 1993).

**Outgroup 5: †Ctenothrissiformes.** *Taxon.* †*Ctenothrissa signifer* (Forey et al., 2003). *Geological horizon and locality.* Sannine Limestone, Hajula, Lebanon.

*Stratigraphy and age estimate.* The Hajula horizon lies below that of Hakel within the Sannine limestone, but its age is constrained by the same biostratigraphic marker: the ammonite †*Mantelliceras mantelli*. The top of the †*M. mantelli* Zone can be dated as approximately 98.0 Ma (Gradstein et al., 2012). *Phylogenetic placement and additional notes.* †Ctenothrissiformes is placed on the acanthomorph stem in the manual cladistic solution presented by Rosen (1973).

**Outgroup 6: Aulopiformes (oldest minimum).** *Taxon.* †*Atolvorator*

*longipectoralis* (Gallo & Cohelo, 2008). *Geological horizon and locality.*

Coquero Seco Formation, Sergipe-Alagoas Basin, Brazil (Gallo & Cohelo, 2008).

*Stratigraphy and age estimate.* Palynological data, based on 15 species of spores and 20 species of pollen grains, indicate that the Coquero Seco Formation can be assigned to the lower Jiquía, a Brazilian regional stage that is correlated with the Barremian (Gallo & Cohelo, 2008). The end of the Barremian is dated approximately as 125 Ma (Gradstein et al., 2012). *Phylogenetic placement and additional notes.* Gallo & Cohelo (2008) provide a verbal argument for placement

of †*Atolvorator* within Aulopiformes, but do not include the genus within a formal cladistic analysis (but see Silva & Gallo, 2011). As such, we regard †*Atolvorator* as an upper estimate for the timing of the origin of Aulopiformes, subject to future investigation. †*Apateodus* (see below) is the earliest undoubted aulopiform, and represents the youngest possible minimum date for the origin of the group.

**Outgroup 6: Aulopiformes (youngest minimum).** *Taxon.* †*Apateodus glyphodus*. *Geological horizon and locality.* Gault Clay Formation, Folkstone, Kent, UK. *Stratigraphy and age estimate.* The Gault Clay Formation has been divided into a series of beds, but the collection horizon(s) for †*Apateodus* from this deposit are generally not recorded for historical material in museum collections. This unit spans the †*Hoplites dentatus* to †*Mortoniceras inflatum* ammonite zones, making it middle to late Albian in age (Gale & Owen, 2010). The top of the †*M. inflatum* Zone has been assigned an interpolated numerical age of 103.13 Ma (Gradstein et al., 2012). *Phylogenetic placement and additional notes.* †*Apateodus* is recognized as a member of the extinct aulopiform suborder †Enchodontoidei. Its position within this radiation has been corroborated by formal cladistic analysis of morphological character data (Silva & Gallo, 2011).

**Outgroup 7: non-eurypterygian Euteleostei.** *Taxon.* †*Leptolepides haerteisi*. *Geological horizon and locality.* Solnhofen Formation, Zandt Member, Zandt, Bavaria, Germany (Arratia, 1997). *Stratigraphy and age estimate.* Based on

ammonite biostratigraphy, the Zandt Member of the Solnhofen Formation is constrained to the †*Lithacoceras eigeltiense* Horizon of the †*L. riedense* Subzone of the †*Hybonoticeras hybonotum* subzone (Schweigert, 2007), indicating an earliest Tithonian age. The top of the †*H. hybonotum* subzone corresponds approximately to Chron M22.n25, which has been dated as  $150.94 \pm 0.05$  Ma (Gradstein et al., 2012). This yields an age estimate of 150.89 Ma.

*Phylogenetic placement and additional notes.* †*Leptolepides haerteisi* belongs to an extinct radiation of late Mesozoic teleosts (†Orthogonikleithridae) that is resolved as the sister group of Salmonoidei plus Esocoidei in cladistic analyses of morphological data (Arratia, 1997; Arratia, 1999, 2000; Arratia & Tischlinger, 2010).

**Outgroup 8: Elopomorpha.** *Taxon.* †*Anaethalion zapporum*. *Geological horizon and locality.* Rögling Formation, village of Schamhaupten, near Eichstätt, Bavaria, Germany (Arratia, 2000). *Stratigraphy and age estimate.* The Rögling Formation underlies the more famous Solnhofen Formation, and is assigned to the Malm Epsilon division of the Franconian Jura (Vihol & Zapp, 2007). The beds at Schamhaupten can be constrained to the †*Neochetoceras rebouletianum* Horizon within the †*Lithacoceras ulmense* subzone of the †*Hybonoticeras beckeri* Ammonite Zone, indicating a latest Kimmeridgian age (Schweigert, 2007; Gradstein et al., 2012). The top of the Kimmeridgian is dated to  $152.1 \pm 0.9$  Ma (Gradstein et al., 2012), from which we derive an age of 151.2 Ma. *Phylogenetic placement and additional notes.* †*Anaethalion* is placed within Elopomorpha

based on formal cladistic analyses of morphological character data (Arratia, 1997, 2000).

**Outgroup 9: †Ichthyodectiformes.** *Taxon.* †*Occithrissops willsoni*. *Geological horizon and locality.* Sundance Formation, Stockade Beaver Shale Member, near Hulett, Wyoming (Schaeffer & Patterson, 1984). *Stratigraphy and age estimate.* Ammonite biostratigraphy indicates a late Bathonian age for the Stockdale Beaver Shale Member of the Sundance Formation (Imlay 1980; Schaeffer & Patterson, 1984). The top of the Bathonian is dated as  $166.1 \pm 1.2$  Ma, from which we derive an age of 164.9 Ma (Gradstein et al., 2012). *Phylogenetic placement and additional notes:* Cavin et al. (2012) place †*Occithrissops* within †Ichthyodectiformes on the basis of a formal cladistic analysis of a morphological character matrix. The position of †Ichthyodectiformes as a crownward radiation on the teleost stem is supported by manual and formal, matrix-based cladistic analyses (e.g., Patterson & Rosen, 1977; Arratia & Vihol, 2010).

**Outgroup 10: unnamed plesion.** *Taxon.* †*Leptolepis coryphaenoides*. *Geological horizon and locality.* Posidonia Shale, ‘Tafelfliens’ unit, Germany. *Stratigraphy and age estimate.* †*Leptolepis* occurs throughout the Posidonia Shale (Hauff & Hauff, 1981), but its lowest occurrence is reported in an informal unit that spans the Pliensbachian-Toarcian boundary (Röhl et al., 2001). The upper limit of this unit is contained within the †*Protogrammoceras paltum* Subzone of the †*Dactylioceras teunicostatum* Ammonite Zone. Gradstein et al. (2012) indicate an

interpolated age estimate for the top of the zone of 181.7 Ma. *Phylogenetic placement and additional notes.* The placement of †*Leptolepis* on the teleost stem crownward of *Proleptolepis* and *Pholidophorus* is supported by the formal cladistic analysis of a morphological character matrix presented by Arratia & Thies (2001).

**Outgroup 11: unnamed plesion.** *Taxon.* †*Proleptolepis furcata*. *Geological horizon and locality.* Charmouth Mudstone Formation, Black Ven Mudstone Member, Dorset, UK (Forey et al., 2010). *Stratigraphy and age estimate.* Material of †*Proleptolepis furcata* derives from the †*Asteroceras obtusum* Zone of the Charmouth Mudstone Formation. The top of the †*A. obtusum* Zone is dated approximately via interpolation as 193.81 Ma (Gradstein et al., 2012). *Phylogenetic placement and additional notes.* The formal cladistic solution presented by Arratia & Thies (2001) places †*Proleptolepis* on the teleost stem, more removed from extant teleosts than †*Leptolepis*, but crownward of †*Pholidophorus*.

**Outgroup 12: unnamed plesion (oldest minimum).** *Taxon.* †*Pholidophorus latiusculus*. *Geological horizon and locality.* Seefeld Formation, Austria (Nybelin, 1966). *Stratigraphy and age estimate.* The localities yielding †*Pholidophorus latiusculus* are considered to lie within the Norian, the top of which is dated as approximately 208.4 Ma (Gradstein et al., 2012). *Phylogenetic placement and additional notes.* †*Pholidophorus latiusculus* is included in a cladistic assessment



of early teleost interrelationships by Arratia and Thies (2001), where it is reconstructed as the sister taxon of †*P. bechei*. However, this species has not been included in subsequent, more extensive analyses (e.g., Arratia & Tischlinger, 2010). In including this species as an upper estimate, we allow for the possibility that †*Pholidophorus* is not monophyletic (cf. Patterson, 1977).

**Outgroup 12: unnamed plesion (youngest minimum).** *Taxon.* †*Pholidophorus bechei* (Arratia 1997; Arratia & Tischlinger 2010). *Geological horizon and locality.* Charmouth Mudstone Formation, Black Ven Mudstone Member, Dorset, UK (Forey et al., 2010). *Stratigraphy and age estimate.* Exact collecting horizons are not known for historical material of †*Pholidophorus bechei*. On the basis of lithology, Forey et al. (2010) proposed that material available to them derived from the †*Arnioceras semicostatum* Zone to †*Asteroceras obtusum* Zone of the Charmouth Mudstone Formation. The top of the †*A. obtusum* Zone is dated approximately via interpolation as 193.81 Ma (Gradstein et al., 2012). *Phylogenetic placement and additional notes.* The position of †*Pholidophorus bechei* as a deep branch along the teleost stem has been corroborated by numerous formal analyses of morphological character data (Arratia, 1997; Arratia & Thies, 2001; Arratia & Tischlinger, 2010).

**Outgroup 13: Holostei.** *Taxon.* †*Watsonulus eugnathoides*. *Geological horizon and locality.* Middle Sakamena Formation, Sakamena Group, Ambilombe Bay, Madagascar (Olsen, 1984). *Stratigraphy and age estimate.* Precise age estimates

for the Sakamena Formation, and the rich vertebrate assemblage it yields, have long been elusive. This unit has been correlated with Induan-Olenekian ('Scythian') elsewhere in Africa (Catuneanu et al., 2005). Gradstein et al. (2012) date the top of the Olenekian as 247.1 Ma. *Phylogenetic placement and additional notes.* The placement of †*Watsonulus* as a holostean is supported by multiple cladistic analyses of morphological character data (Grande & Bemis, 1998; Gardiner et al., 1996; Grande, 2010). We have re-examined type material of †*Brachydegma caelatum*, and are unconvinced that this taxon is a crown-group holostean or even a crown-group neopterygian (*contra* Hurley et al., 2007; see also Near et al., 2012).

**Upper bound on calculations.** We have set the upper bound for divergence estimates at the Mississippian-Pennsylvanian boundary ( $323.2 \pm 0.4$  Ma, yielding a point estimate of 322.8 Ma), based on the Serpukhovian stem-group neopterygian †*Discoserra pectinodon* (Hurley et al., 2007). In practice, Hedman's (2010) method is relatively insensitive to the choice of this upper bound.

**Sequence of outgroup ages (youngest minima):**

247.1, 193.81, 193.81, 181.7, 164.9, 151.2, 150.89, 103.13, 98.0, 98.0, 98.0, 69.71, 49.11, 46.0

**Sequence of outgroup ages (oldest minima):**

247.1, 208.4, 193.81, 181.7, 164.9, 151.2, 150.89, 125, 98.0, 98.0, 98.0, 98.0, 49.11, 46.0

Clade	Age estimate and 95% CI (Ma) based on youngest minima	Age estimate and 95% CI (Ma) based on oldest minima
Cichlidae	56.98 (46.83, 81.19)	60.68 (46.83, 90.06)
Ovalentaria	67.74 (50.48, 94.04)	75.14 (50.75, 100.89)
Percomorpha	86.08 (70.98, 105.94)	100.97 (98.45, 113.75)
Acanthopterygii	102.33 (98.45, 119.83)	103.71 (98.45, 123.20)
Acanthomorpha	106.43 (98.45, 132.20)	109.20 (98.45, 136.03)
Eurypterygii	131.11 (104.89, 163.16)	142.13 (126.19, 166.18)
Euteleostei	158.87 (151.23, 180.83)	159.10 (151.23, 182.04)
Crown Teleostei	166.73 (152.06, 196.03)	167.19 (152.06, 197.92)

Table S5. Age estimates (mean of posterior age distributions) and 95% credible intervals for cichlids and a series of more inclusive teleost clades based on the two outgroup sequences described above. Note that age estimates for Cichlidae, Ovalentaria, Percomorpha, and Acanthopterygii are inconsistent with Gondwanan vicariance in any of these groups.

## 2. MOLECULAR SYSTEMATICS

**(a) Collection of DNA sequence data and phylogenetic analyses.** Standard phenol-chloroform extraction protocol or Qiagen DNeasy Blood and Tissue kits were used to isolate DNA from tissue biopsies sampled from 158 species of Percomorpha that included 89 species of Cichlidae (table S6). Previously published PCR primers were used to amplify and sequence an exon from each of nine nuclear genes (*ENC*, *Glyt*, *myh6*, *plagl2*, *Ptr*, *rag1*, *SH3PX3*, *sreb2*, *tbr1*, and *zic1*) (Li et al., 2007; Lopez et al., 2004). The genes were aligned by eye using the inferred amino acid sequences. No frame mutations or DNA substitutions that resulted in stop codons were observed in the aligned DNA sequences. The combined ten gene dataset contained 7,887 base pairs.

Thirty data partitions were designated that corresponded to the three separate codon positions for each of the ten protein coding genes. A GTR+G substitution model was used in a partitioned maximum likelihood analysis using the computer program RAxML 7.2.6 (Stamatakis, 2006), run with the *-D* option. Support for nodes in the RAxML tree was assessed with a thorough bootstrap analysis (option *-f i*) with 1,000 replicates.

**(b) Specimens used for DNA sequencing.**

Species	Family	Catalogue Information
<i>Seriola dumerili</i>	Carangidae	YFTC 17741
<i>Trachinotus carolinus</i>	Carangidae	YFTC 11458
<i>Echeneis naucrates</i>	Echeneidae	YFTC 11539
<i>Rachycentron canadum</i>	Rachycentridae	YFTC 17742
<i>Centropomus armatus</i>	Centropomidae	KU 40318
<i>Sphyrna barracuda</i>	Sphyrnidae	YFTC 11488
<i>Labidesthes sicculus</i>	Atherinopsidae	YPM 20505, YFTC 13167
<i>Atherinomorus lacunosus</i>	Atherinidae	YFTC 12611
<i>Melanotaenia</i> sp.	Melanotaeniidae	YPM 20539, YFTC 13566
<i>Rheocles wrightae</i>	Bedotiidae	WLS 517
<i>Fundulus heteroclitus</i>	Fundulidae	Genbank
<i>Gambusia affinis</i>	Poeciliidae	INHS 38579, YFTC 250
<i>Oryzias latipes</i>	Adrianichthyidae	Genbank
<i>Arrhamphus sclerolepis</i>	Hemiramphidae	WLS 516
<i>Dermogenys collettei</i>	Hemiramphidae	WLS 515
<i>Xenentodon cancila</i>	Belonidae	YFTC 11548
<i>Platybelone argala</i>	Belonidae	YPM 21768, YFTC 16172
<i>Strongylura marina</i>	Belonidae	YFTC 11543
<i>Monocirrhus polyacanthus</i>	Polycentridae	YFTC 11553
<i>Polycentrus schomburgki</i>	Polycentridae	YFTC 11554
<i>Pholidichthys leucotaenia</i>	Pholidichthidae	YFTC 11546
<i>Ambassis urotaenia</i>	Chandidae	YPM 20542, YFTC 13569
<i>Mugil cephalus</i>	Mugilidae	YFTC 738
<i>Cymatogaster aggregata</i>	Embiotocidae	INHS 45419, YFTC 1055
<i>Rhacochilus vacca</i>	Embiotocidae	YFTC 11438
<i>Embiotoca jacksoni</i>	Embiotocidae	YFTC 11437
<i>Embiotoca lateralis</i>	Embiotocidae	YFTC 11440
<i>Hyperprosopon argenteum</i>	Embiotocidae	YFTC 11511
<i>Congrogadus subducens</i>	Congrogadidae	WLS 505
<i>Plesiops coeruleolineatus</i>	Plesiopidae	WLS 1163
<i>Chromis cyanea</i>	Pomacentridae	YFTC 11489
<i>Microspathodon bairdii</i>	Pomacentridae	YFTC 11481
<i>Stegastes leucostictus</i>	Pomacentridae	YPM 21762, YFTC 16177
<i>Pseudochromis fridmani</i>	Pseudochromidae	YFTC 12089
<i>Pholidochromis cerasiae</i>	Pseudochromidae	WLS 1162
<i>Ogilbynia novaehollandae</i>	Pseudochromidae	WLS 1161
<i>Labracinus cyclophthalmus</i>	Pseudochromidae	YFTC 12090
<i>Grama loreto</i>	Grammatidae	WLS 506
<i>Diademichthys lineatus</i>	Gobiesocidae	YFTC 11521
<i>Gobiesox maeandricus</i>	Gobiesocidae	YFTC 17736
<i>Ophioblennius atlanticus</i>	Blenniidae	USNM 349074
<i>Meiacanthus grammistes</i>	Blenniidae	YFTC 12593
<i>Gasterosteus aculeatus</i>	Gasterosteidae	YFTC 12993
<i>Gasterosteus wheatlandi</i>	Gasterosteidae	YFTC 12991
<i>Lycodes diapterus</i>	Zoarcidae	YFTC 13952
<i>Siganus spinus</i>	Siganidae	KU T4159
<i>Selenotoca multifasciata</i>	Scatophagidae	ROM 68452
<i>Diodon holocanthus</i>	Diodontidae	YFTC 11512
<i>Tetraodon muris</i>	Tetraodontidae	YFTC 11549
<i>Aracana aurita</i>	Aracnidae	YFTC 11513

---

<i>Ostracion cubicus</i>	Ostraciidae	YFTC 11509
<i>Triacanthodes anomalus</i>	Triacanthodidae	MM 201
<i>Masturus lanceolatus</i>	Molidae	YFTC 18814
<i>Mola mola</i>	Molidae	MM 910
<i>Ranzania laevis</i>	Molidae	YFTC 22151
<i>Pervagor janthinosoma</i>	Monacanthidae	YFTC 11514
<i>Cantherhines pullus</i>	Monacanthidae	YPM 21753, YFTC 16217
<i>Abalistes stellatus</i>	Balistidae	YFTC 11516
<i>Luvarus imperialis</i>	Luvaridae	MCZ 159566
<i>Naso lituratus</i>	Acanthuridae	KU T5440
<i>Acanthurus nigricans</i>	Acanthuridae	YFTC 12500
<i>Gazza minuta</i>	Leiognathidae	YFTC 18187
<i>Leiognathus equulus</i>	Leiognathidae	YFTC 18188
<i>Chaetodon striatus</i>	Chaetodontidae	YPM 21761, YFTC 16186
<i>Chelmon rostratus</i>	Chaetodontidae	YFTC 11467
<i>Cichla temensis</i>	Cichlidae	WLS 511
<i>Etroplus maculatus</i>	Cichlidae	YFTC 12081
<i>Herichthys cyanoguttatus</i>	Cichlidae	Genbank
<i>Heros efasciatus</i>	Cichlidae	YFTC 3249
<i>Heterochromis multidentis</i>	Cichlidae	WLS 520
<i>Oreochromis niloticus</i>	Cichlidae	Genbank
<i>Paratilapia polleni</i>	Cichlidae	YFTC 12082
<i>Paretroplus maculatus</i>	Cichlidae	WLS 501
<i>Paretroplus dambabe</i>	Cichlidae	YFTC 17924, YPM 23180
<i>Ptychochromis grandidieri</i>	Cichlidae	WLS 521
<i>Ptychochromis oligacanthus</i>	Cichlidae	YFTC 18167, YPM 23177
<i>Retroculus xinguensis</i>	Cichlidae	HLF 1202
<i>Astronotus sp.</i>	Cichlidae	YFTC 21011
<i>Mikrogeophagus ramirezi</i>	Cichlidae	YFTC 18358, YPM 23155
<i>Geophagus iporangensis</i>	Cichlidae	YFTC 18364
<i>Dicrossus filamentosus</i>	Cichlidae	YFTC 19886
<i>Biotoecus opercularis</i>	Cichlidae	YFTC 18402
<i>Satanoperca jurupari</i>	Cichlidae	YFTC 21017
<i>Crenicichla sveni</i>	Cichlidae	YFTC 18177, YPM 23163
<i>Cleithracara maronii</i>	Cichlidae	YFTC 18176, YPM 23151
<i>Laetacara curviceps</i>	Cichlidae	YFTC 18357, YPM 23157
<i>Cichlasoma boliviense</i>	Cichlidae	YFTC 18357, YPM 23161
<i>Adinocara rivulatus</i>	Cichlidae	YFTC 18173, YPM 23160
<i>Pterophyllum leopoldi</i>	Cichlidae	YFTC 18181, YPM 23150
<i>Hypselecara temporalis</i>	Cichlidae	YFTC 18178, YPM 23159
<i>Uaru amphiacanthoides</i>	Cichlidae	YFTC 18159, YPM 23154
<i>Mesonauta insignis</i>	Cichlidae	YFTC 18179, YPM 23152
<i>Mesonauta festivus</i>	Cichlidae	YFTC 18161, YPM 23153
<i>Mesonauta egregius</i>	Cichlidae	YFTC 19867
<i>Mesonauta mirificus</i>	Cichlidae	YFTC 19870, YPM 23890
<i>Nandopsis tetracanthus</i>	Cichlidae	No voucher
<i>Parachromis managuensis</i>	Cichlidae	YFTC 21015
<i>Cryptoheros sajica</i>	Cichlidae	YFTC 18369
<i>Tomocichla sieboldii</i>	Cichlidae	YFTC 18172, YPM 23164
<i>Caquetaia kraussii</i>	Cichlidae	CUMV 90808
<i>Rocio octofasciatum</i>	Cichlidae	YFTC 21012
<i>Thorichthys meeki</i>	Cichlidae	YFTC 21016
<i>Cichlasoma salvini</i>	Cichlidae	YFTC 21013
<i>Hemichromis elongatus</i>	Cichlidae	CUMV 93163
<i>Hemichromis letourneuxi</i>	Cichlidae	CUMV 94559
<i>Hemichromis cf. bimaculatus</i>	Cichlidae	YFTC 21014

---

<i>Tylochromis mylodon</i>	Cichlidae	CUMV 91216
<i>Tylochromis lateralis</i>	Cichlidae	CUMV 91565
<i>Benitochromis batesii</i>	Cichlidae	CUMV 93617
<i>Nanochromis parilus</i>	Cichlidae	YFTC 18170, YPM 23218
<i>Sarotherodon galilaeus</i>	Cichlidae	CUMV 91585
<i>Oreochromis tanganicae</i>	Cichlidae	YFTC 19267
<i>Myaka myaka</i>	Cichlidae	YFTC 19238
<i>Stomatepia pindu</i>	Cichlidae	CUMV 93584
<i>Konia eisentrauti</i>	Cichlidae	CUMV 93584
<i>Pungu maclareni</i>	Cichlidae	YFTC 18558
<i>Steatocranus ubanguiensis</i>	Cichlidae	CUMV 91448
<i>Steatocranus tinanti</i>	Cichlidae	YFTC 15695, YPM 23891
<i>Gobiocichla ethelwynnae</i>	Cichlidae	YFTC 15700
<i>Tilapia mariae</i>	Cichlidae	CUMV 90069
<i>Tilapia ruweti</i>	Cichlidae	YFTC 18368
<i>Tilapia sparrmanii</i>	Cichlidae	CUMV 91213
<i>Boulengerochromis microlepis</i>	Cichlidae	YFTC 18180, YPM 23186
<i>Bathybates fasciatus</i>	Cichlidae	CUMV 88627
<i>Neolamprologus toae</i>	Cichlidae	CUMV 89315
<i>Neolamprologus brevis</i>	Cichlidae	YFTC 17921, YPM 23185
<i>Altolamprologus calvus</i>	Cichlidae	YFTC 18141, YPM 23198
<i>Cyphotilapia frontosa</i>	Cichlidae	YFTC 17918, YPM 23183
<i>Reganochromis calliurus</i>	Cichlidae	YFTC 17922, YPM 23182
<i>Benthochromis tricoti</i>	Cichlidae	YFTC 18144, YPM 23189
<i>Plecodus straeleni</i>	Cichlidae	CUMV 88634
<i>Haplotaxodon microlepis</i>	Cichlidae	CUMV 88641
<i>Paracyprichromis brienii</i>	Cichlidae	YFTC 18142, YPM 23211
<i>Cyprichromis leptosoma</i>	Cichlidae	YFTC 17917, YPM 23203
<i>Cardiopharynx schoutedeni</i>	Cichlidae	YFTC 18145, YPM 23213
<i>Aulonocranus dewindti</i>	Cichlidae	YFTC 18143, YPM 23209
<i>Ophthalmotilapia ventralis</i>	Cichlidae	YFTC 18147, YPM 23210
<i>Eretmodus cyanostictus</i>	Cichlidae	YFTC 17923, YPM 23181
<i>Spathodus marlieri</i>	Cichlidae	CUMV 93679
<i>Orthochromis luongoensis</i>	Cichlidae	CUMV 91742
<i>Chetia mola</i>	Cichlidae	CUMV 91753
<i>Pseudocrenilabrus sp.</i>	Cichlidae	CUMV 91757
<i>Tropheus duboisi</i>	Cichlidae	YFTC 17920, YPM 23184
<i>Petrochromis polyodon</i>	Cichlidae	CUMV 82912
<i>Limnotilapia dardennii</i>	Cichlidae	CUMV 82854
<i>Gnathochromis pfefferi</i>	Cichlidae	CUMV 88652
<i>Pseudosimochromis curvifrons</i>	Cichlidae	CUMV 89316
<i>Serranochromis angusticeps</i>	Cichlidae	CUMV 91238
<i>Haplochromis brauschi</i>	Cichlidae	No voucher
<i>Astatotilapia burtoni</i>	Cichlidae	YFTC 17919, YPM 23215
<i>Haplochromis bloyeti</i>	Cichlidae	CUMV 93967
<i>Mbipia lutea</i>	Cichlidae	YFTC 18165, YPM 23188
<i>Haplochromis latifasciatus</i>	Cichlidae	YFTC 18362, YPM 23216
<i>Tyrannochromis nigriventer</i>	Cichlidae	YFTC 18164, YPM 23199
<i>Chilotilapia rhoadesii</i>	Cichlidae	YFTC 18155, YPM 23197
<i>Cyrtocara moorii</i>	Cichlidae	YFTC 18157, YPM 23207

Table S6. Specimens for DNA sequencing. Institutional abbreviations are as follows:

CUMV, Cornell University Museum of Vertebrates, Cornell University, Ithaca, New



York, USA; HLF, Hernán López-Fernández tissue collection; INHS, Illinois Natural History Survey, University of Illinois, Champaign, Illinois, USA; KU, University of Kansas, Lawrence, Kansas, USA; MCZ, Museum of Comparative Zoology, Harvard University, Cambridge, Massachusetts, USA; MM, Masaki Miya tissue collection; ROM, Royal Ontario Museum, Toronto, Ontario, Canada; WLS, W. Leo Smith tissue collection; YFTC, Yale Fish Tissue Collection, Yale University, New Haven, Connecticut, USA; YPM, Yale Peabody Museum, Yale University, New Haven, Connecticut, USA.

### 3. MOLECULAR ESTIMATES OF TIME OF EVOLUTIONARY ORIGIN

(a) *Relaxed-molecular clock analyses.* Divergence times of the sampled percomorph lineages were estimated using the random local clock (RLC) model of molecular evolutionary rate heterogeneity implemented in the computer program BEAST v. 1.6.1 (Drummond et al., 2006; Drummond & Rambaut, 2007; Drummond & Suchard, 2010). The nucleotide substitution models for the ten gene dataset were partitioned by gene and codon as in the MrBayes analysis above, but the random local clock models were partitioned by gene. Ten lognormal calibration priors from the fossil record of ray-finned fishes were used in the RLC analyses and are detailed below. A birth-death speciation prior was used for branching rates in the phylogeny. The BEAST analyses were run ten times with each run consisting of  $2.0 \times 10^8$  generations, sampling at every 10,000 generations. The resulting trees and log files from each of the ten runs were combined using the computer program LogCombiner v. 1.6.1 (<http://beast.bio.ed.ac.uk/LogCombiner>). Convergence of model parameter values and estimated node-heights to their optimal posterior distributions was assessed by plotting the marginal posterior probabilities versus the generation state in the computer program Tracer v. 1.5 (<http://beast.bio.ed.ac.uk/Tracer>). Effective sample size (ESS) values were calculated for each parameter to ensure adequate mixing of the MCMC (ESS>200). The posterior probability density of the combined tree and log files was summarized as a maximum clade credibility tree using TreeAnnotator v. 1.6.1 (<http://beast.bio.ed.ac.uk/TreeAnnotator>). The mean and 95% highest posterior density estimates of divergence times and the posterior probabilities of inferred clades were

visualized on the using the computer program FigTree v. 1.3.1

(<http://beast.bio.ed.ac.uk/FigTree>).

**(b) Fossil-based age priors for relaxed molecular clock analysis.** Here we provide, for each fossil calibration prior, the identity of the calibrated node in the teleost phylogeny, the taxa that represent the first occurrence of the lineage in the fossil record, a description of the character states that justify the phylogenetic placement of the fossil taxon, information on the stratigraphy of the rock formations bearing the fossil, the absolute age estimate for the fossil, outline the prior age setting used in the BEAST relaxed clock analysis, and provide any additional notes on the calibration. Less detailed information is provided for five of the calibrations used in a previous study of actinopterygian divergence times, as all of the information and prior settings for these calibrations is found in Near et al. (2012). Each calibration is numbered and the phylogenetic placement of the calibration is highlighted in figure S3.

**Calibration 1.** *Node:* Stem lineage Sphyraenidae, dating the MRCA of *Sphyraena* and *Centropomus*. *First occurrence:* †*Sphyraena bolcensis*, Pesciara beds of ‘Calcarei nummulitici’, Bolca, Italy. *Resolution in phylogenetic analyses:* none. *Character states:* three ‘T’-shaped, sutured predorsals or spineless pterygiophores (Johnson, 1986); elongate gape; upper jaw non-protrusible; enlarged fangs on dentary, premaxilla, and palatine. *Stratigraphy:* upper Ypresian [NP14] (Papazzoni & Trevisani, 2006). *Absolute age estimate:* 50.0 Ma (Papazzoni & Tervisani, 2006). *Prior setting:* a lognormal prior with the mean = 0.672 and the

standard deviation = 0.8 to set 50.0 Ma as the minimal age offset and 57.3 Ma as the 95% soft upper bound. The upper bound is based on the calculation of FA<sub>95</sub> following Marshall (2008).

**Calibration 2.** *Node:* Stem lineage Acanthurinae, dating the MRCA of Acanthurinae and Nasinae. *First occurrence:* †*Proacanthurus tenius*, ‘Pesciara beds of ‘Calcari nummulitici’, Bolca, Italy. *Resolution in phylogenetic analyses:* none. *Character states:* caudal peduncle bears folding spine. *Stratigraphy:* upper Ypresian [NP14] (Papazzoni & Trevisani, 2006). *Absolute age estimate:* 50.0 Ma (Papazzoni & Trevisani, 2006). *Prior setting:* a lognormal prior with the mean = 0.672 and the standard deviation = 0.8 to set 50.0 Ma as the minimal age offset and 57.3 Ma as the 95% soft upper bound. The upper bound is based on the calculation of FA<sub>95</sub> following Marshall (2008).

**Calibration 3.** Chaetodontidae *cf. Chaetodon* (tholichthys-stage larva), calibration number 30 in Near et al. (2012). *Prior setting:* a lognormal prior with the mean = 0.165 and standard deviation = 0.8 to set 30.1 Ma as the minimal age offset and 34.5 Ma as the 95% soft upper bound.

**Calibration 4.** †*Gornylistes prodigiosus*, calibration number 35 in Near et al. (2012). *Prior setting:* a lognormal prior with the mean = 0.37 and standard deviation = 0.8 to set 37.2 Ma as the minimal age offset and 42.6 Ma as the 95% soft upper bound.

**Calibration 5.** †*Archaeus oblongus*, calibration number 23 in Near et al. (2012).

*Prior setting:* a lognormal prior with the mean = 0.776 and standard deviation = 0.8 to set 55.8 Ma as the minimal age offset and 63.9 Ma as the 95% soft upper bound.

**Calibration 6.** †*Opisthomyzon glaronensis* and an unnamed echeneid cf.

*Echeneis*, calibration number 24 in Near et al. (2012). *Prior setting:* a lognormal prior with the mean = 0.165 and standard deviation = 0.8 to set 30.1 Ma as the minimal age offset and 34.5 Ma as the 95% soft upper bound.

**Calibration 7.** *Node:* Crown lineage *Gasterosteus*, dating the MRCA of

*Gasterosteus aculeatus* and *G. wheatlandi*. *First occurrence:* *Gasterosteus* cf. *aculeatus*, Monterey Formation, California, USA (Bell et al., 2009). *Resolution in phylogenetic analyses:* none. *Character states:* two elongate, free dorsal fin spines, complete series of lateral plates (Bell et al., 2009). *Stratigraphy:* Seravallian, lowermost Mohnian regional stage (Bell et al., 2009). *Absolute age estimate:* 13 Ma (Bell et al., 2009). *Prior setting:* an exponential prior with the mean = 1.17 to set 13.0 Ma as the minimal age offset and 16.5 Ma as the 95% soft upper bound. The upper bound is based on the calculation of FA<sub>95</sub> following Marshall (2008).

**Calibration 8.** †*Avitoluvarus diana*, †*Avitoluvarus mariannae*, †*Kushlukia*

*permira*, and †*Luvarus necopinatus*, calibration number 25 in Near et al. (2012).

*Prior setting:* a lognormal prior with the mean = 0.776 and standard deviation =

0.8 to set 55.8 Ma as the minimal age offset and 63.9 Ma as the 95% soft upper bound.

**Calibration 9.** *Node:* Crown lineage Molidae, dating the MRCA of *Ranzania* and a clade containing *Mola* and *Masturus*. *First occurrence:* †*Austromola angerhoferi*, Ebelsberg Formation, Pucking, Austria (Gregorova et al., 2009). *Resolution in phylogenetic analyses:* parsimony analysis of 57 morphological characters resolves †*Austromola* as the sister lineage of a clade containing *Mola* and *Masturus* (Gregorova et al., 2009). *Character states:* basal pterygiophores of dorsal and anal fins expanded proximally; first basal pterygiophore of anal fin concave and encloses distal end of haemal spines of first two caudal vertebrae; posteriormost dorsal and anal fin pterygiophores inserted approximately normal to body axis; pectoral fin with rounded outline; deep body; bone spongy and weakly ossified (Santini & Tyler, 2002; Gregorova et al., 2009). *Stratigraphy:* Aquitanian, upper Egerian regional stage (Gregorova et al., 2009). *Absolute age estimate:* 22 Ma (Gregorova et al., 2009). *Prior setting:* a lognormal prior with the mean = 1.4055 and the standard deviation = 0.8 to set 22.0 Ma as the minimal age offset and 37.2 Ma as the 95% soft upper bound.

**Calibration 10.** *Node:* Crown lineage Leiognathidae, dating the MRCA of *Leiognathus* and *Gazza*. †*Euleiognathus tottori*, calibration number 32 in Near et al. (2012). *Prior setting:* a lognormal prior with the mean = 1.602 and standard

deviation = 0.8 to set 11.6 Ma as the minimal age offset and 23.1 Ma as the 95% soft upper bound.



### 3. SUPPLEMENTARY REFERENCES

- Arratia, G. 1997. Basal teleosts and teleostean phylogeny. *Palaeo Ichthyologica*, **7**, 1-168.
- Arratia, G. 1999. The monophyly of Teleostei and stem-group teleosts, p. 265-334. In *Mesozoic fishes 2 – systematics and fossil record* (eds G. Arratia & H.-P. Schultze) pp. 265-334. Munich, Germany: Verlag Dr. Friedrich Pfeil.
- Arratia, G. 2000. Remarkable teleostean fishes from the Late Jurassic of southern Germany and their phylogenetic relationships. *Mitteilungen aus dem Museum für Naturkunde in Berlin, Geowissenschaftliche, Reihe*, **3**, 137-179.
- Arratia, G. & Thies, D. 2001. A new teleost (Osteichthyes, Actinopterygii) from the Early Jurassic *Posidonia* shale of northern Germany. *Mitteilungen aus dem Museum für Naturkunde in Berlin, Geowissenschaftliche, Reihe*, **4**, 167-187.
- Arratia, G. & Tischlinger, H. 2010. The first record of Late Jurassic crossognathiform fishes from Europe and their phylogenetic importance for teleostean phylogeny. *Fossil Record*, **13**, 317-341.
- Bannikov, A., Parin, N. V. & Pinna, G. 1985. *Rhamphexocetus volans*, gen. et sp. nov.: a new beloniform fish (Beloniformes, Exocetoidei) from the Lower Eocene of Italy. *Journal of Ichthyology*, **25**, 150-155.
- Bell, M. A., Stewart, J. D. & Park, P. J. 2009. The world's oldest fossil threespine stickleback fish. *Copeia*, **2009**, 256-65.
- Catuneanu, O., Wopfer, H., Eriksson, P.G., Carincross, B., Rubidge, B.S., Smith, R.M.H. & Hancox, P.J. (2005) The Karoo basins of south-central Africa. *Journal of African Earth Sciences*, **43**, 211-253.

- Cavin, L., Forey, P. L. & Giersch, S. 2012. Osteology of *Eubiodectes libanicus* (Pictet & Humbert, 1866) and some other ichthyodectiformes (Teleostei): phylogenetic implications. *Journal of Systematic Palaeontology*, **11**, 115-177.
- Dalla Vecchia, F. M., Venturini, S. & Tentor, M. 2002. The Cenomanian (Late Cretaceous) *Konservat-Lagerstätte* of en Nammoûra, (Kesrouâne Province), northern Lebanon. *Bolletino della Società Paleontologica Italiana*, **41**, 51-68.
- Drummond, A. J. & Rambaut, A. 2007. BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evol Biology*, **7**, 214.
- Drummond, A. J. & Suchard, M. A. 2010. Bayesian random local clocks, or one rate to rule them all. *BMC Biology*, **8**, 114.
- Drummond, A. J., Ho, S. Y. W., Phillips, M. J. & Rambaut, A. 2006. Relaxed phylogenetics and dating with confidence. *PLoS Biology*, **4**, 699-710.
- Forey, P.L., Longbottom, A. & Mulley, J. 2010. Fishes—bony fishes. In *Fossils from the lower Lias of the Dorset coast* (eds A.R. Lord & P.G. Davis), pp. 341-371. London: The Palaeontological Association.
- Forey, P. L., Lu, Y., Patterson, C. & Davies, C. E. 2003. Fossil fishes from the Cenomanian (Upper Cretaceous) of Namoura, Lebanon. *Journal of Systematic Palaeontology*, **4**, 227-330.
- Gale, A. S. & Owen, H. G. 2010. Introduction to the Gault. In *Fossils of the Gault Clay* (eds J. R. Young, A. S. Gale, R. I. Knight & A. B. Smith), pp. 1-15. London: The Palaeontological Association.
- Gallo, V. & Coelho, P. M. 2008. First occurrence of an aulopiform fish in the Barremian of the Sergipe-Algoas Basin, northeastern Brazil. In *Mesozoic fishes 4—homology*

- and phylogeny* (eds G. Arratia, H.-P. Schultze & M.V.H. Wilson), pp. 351-371. Munich, Germany: Verlag Dr. Friedrich Pfeil.
- Gardiner, B. G., Maisey, J. G. & Littlewood, D. T. J. 1996. Interrelationships of basal neopterygians. In *Interrelationships of fishes* (eds M. L. J. Stiassny, L. R. Parenti, & G. D. Johnson), pp. 117-146. San Diego: Academic Press.
- Gradstein, F. M., Ogg, J. G., Schmitz, M., & Ogg, G. 2012. *The geologic time scale 2012*. Amsterdam: Elsevier.
- Grande, L. 2010. An empirical synthetic pattern study of gars (Lepisosteiformes) and closely related species, based mostly on skeletal anatomy. The resurrection of the Holostei. *American Society of Ichthyologists and Herpetologists Special Publication*, **6**, 1-871.
- Grande, L. & Bemis, W.E. (1998) A comprehensive phylogenetic study of amiid fishes (Amiidae) based on comparative skeletal anatomy. An empirical search for interconnected patterns of natural history. *Society of Vertebrate Paleontology, Memoir*, **4**, 1-690.
- Gregorova, R., Schultz, O., Harzhauser, M., Kroh, A. & Coric, S. 2009. A giant early Miocene sunfish from the north alpine Foreland Basin (Austria) and its implication for molid phylogeny. *Journal of Vertebrate Paleontology*, **29**, 359-71.
- Harrison, T., Msuya, C. P., Murray, A. M., Jacobs, B. F., Báez, A. M., Mundil, R. & Ludwig, K. R. 2001. Paleontological investigations at the Eocene locality of Mahenge in north-central Tanzania, East Africa. In *Eocene biodiversity: Unusual occurrences and rarely sampled habitats* (ed. G. F. Gunell) pp. 39-74. New York: Kluwer Academic/Plenum Publishers.

- Hauff, B. & Hauff, R. B. 1981 Das Holzmadenbuch. Fellbach, Germany: REPRO-DRUCK.
- Hedman, M. M. 2010. Constraints on clade ages from fossil outgroups. *Paleobiology*, **36**, 16-31.
- Hurley, I.A., Mueller, R.L., Dunn, K.A., Schmidt, E.J., Friedman, M., Ho, R.K., Prince, V.E., Yang, Z., Thomas, M.G. & Coates, M.I. (2007) A new time-scale for ray-finned fish evolution. *Proceedings of the Royal Society B*, **274**, 489-498.
- Imlay, R. W. 1980. Jurassic paleobiogeography of the conterminous United States and its continental setting. *U. S. Geological Survey Professional Paper*, **1062**, 1-134.
- Johnson, G. D. 1986. Scombroid phylogeny: an alternative hypothesis. *Bulletin of Marine Science*, **39**, 1-41.
- Johnson, G. D. & Patterson, C. 1993. Percomorph phylogeny: a survey of acanthomorphs and a new proposal. *Bulletin of Marine Science*, **52**, 554-626.
- Li, C. H., Ortí, G., Zhang, G. & Lu, G. Q. 2007. A practical approach to phylogenomics: the phylogeny of ray-finned fish (Actinopterygii) as a case study. *BMC Evol Biology*, **7**, 44.
- Lopez, J. A., Chen, W. J. & Ortí, G. 2004. Esociform phylogeny. *Copeia*, **2004**, 449-464.
- Marshall, C. R. 1990. Confidence intervals on stratigraphic ranges. *Paleobiology*, **16**, 1-10.
- Marshall, C. R. 1997. Confidence intervals on stratigraphic ranges with nonrandom distributions of fossil horizons. *Paleobiology*, **23**, 165-173.

- Marshall, C. R. 2008. A simple method for bracketing absolute divergence times on molecular phylogenies using multiple fossil calibration points. *American Naturalist*, **171**, 726-42.
- Medizza, F. 1975. Il nannoplancton calcareo della Pesciara di Bolca (Monti Lessini). *Studi e Ricerche sui Giacimenti Terziari di Bolca*, **2**, 433-444.
- Medizza, F. & Sorbini, L. 1980. Il giacimento del Salento (Lecce). In *I vertebrati fossili italiani – Catalogo della Mostra*, pp. 131-134. Verona, Italy: Museo Civico di Storia Naturale, Verona.
- Murray, A. M. 2000. Eocene cichlid fishes from Tanzania, East Africa. *Journal of Vertebrate Paleontology*, 20:651-664.
- Murray, A. M. 2001a. The oldest fossil cichlids (Teleostei: Perciformes): indication of a 45 million-year-old species flock. *Proceedings of the Royal Society, Series B*, **268**, 679-684.
- Murray, A. M. 2001b. The fossil record and biogeography of the Cichlidae (Actinopterygii: Labroidei). *Biological Journal of the Linnean Society*, **74**, 517-532.
- Near, T. J., Eytan, R. I., Dornburg, A., Kuhn, K. L., Moore, J. A., Davis, M. P., Wainwright, P. C., Friedman, M. & Smith, W. L. 2012. Resolution of ray-finned fish phylogeny and timing of diversification. *Proceedings of the National Academy of Sciences of the USA*, **109**, 13698-13703.
- Nybelin, O. 1966. On certain Triassic and Liassic representatives of the family Pholidophoridae s. str. *Bulletin of the British Museum (Natural History): Geology*, **11**, 351-432.

- Olsen, P. E. 1984. The skull and pectoral girdle of the parasemionotid fish *Watsonulus eugnathoides* from the Early Triassic Sakamena Group of Madagascar, with comments on the relationships of the holostean fishes. *Journal of Vertebrate Paleontology*, **4**, 481-499.
- Orr, J. W. 1995. Phylogenetic relationships of the gasterosteiform fishes (Teleostei: Acanthomorpha). PhD thesis, University of Washington, Seattle, USA.
- Papazzoni, C. A. & Trevisani, E. 2006. Facies analysis, palaeoenvironmental reconstruction, and biostratigraphy of the “Pesciara di Bolca” (Verona, northern Italy): An early Eocene Fossil-Lagerstätte. *Palaeogeography, Palaeoclimatology, Palaeoecology*, **242**, 21-35.
- Patterson, C. 1977. The contribution of paleontology to teleostean phylogeny. In *Major patterns in vertebrate evolution* (eds M. K. Hecht, P. C. Goody & B. M. Hecht), pp. 579-643. New York: Plenum Press.
- Patterson, C. 1993. An overview of the early fossil record of acanthomorphs. *Bulletin of Marine Science*, **52**, 29-59.
- Patterson, C. & Rosen, D.E. 1977. Review of ichthyodectiform and other Mesozoic teleost fishes and the theory and practice of classifying fossils. *Bulletin of the American Museum of Natural History*, **158**, 81-172.
- Pietsch, T. W. 1978. Evolutionary relationships of the sea moths (Teleostei: Pegasidae) with a classification of gasterosteiform families. *Copeia*, **1978**, 517-529.
- R Core Team. 2013. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. See: <http://www.R-project.org>

- Röhl, H.-J., Schmid-Röhl, A., Oschmann, W., Frimmel, A. & Schwark, L. 2001 The Posidonia Shale (Lower Toarcian) of SW-Germany: an oxygen-depleted ecosystem controlled by sea level and palaeoclimate. *Palaeogeography, Palaeoclimatology, Palaeoecology*, **165**, 27–52.
- Rosen, D. E. 1973. Interrelationships of higher euteleostean fishes. In *Interrelationships of fishes* (eds P. H. Greenwood, R. S. Miles & C. Patterson), pp. 397-513. London: Academic Press.
- Rosen, D. E. & Parenti, L. R. 1981. Relationships of *Oryzias* and the groups of atherinomorph fishes. *American Museum Novitates*, **2719**, 1-25.
- Santini, F. & Tyler, J. C. 2004. Phylogeny of the ocean sunfishes (Molidae, Tetraodontiformes), a highly derived group of teleost fishes. *Italian Journal of Zoology*, **69**, 37-43.
- Schaeffer, B. & Patterson, C. 1984. Jurassic fishes from the western United States, with comments on Jurassic fish distribution. *American Museum Novitates*, **2796**, 1-86.
- Schweigert, G. 2007. Ammonite biostratigraphy as a tool for dating Upper Jurassic lithographic limestones from South Germany – first results and open questions. *Neues Jahrbuch für Geologie und Paläontologie, Abhandlungen*, **245**, 117-125.
- Silva, H. M. A. & Gallo, V. 2011. Taxonomic review and phylogenetic analysis of Enchodontoidei (Teleostei: Aulopiformes). *Anais da Academia Brasileira de Ciências*, **83**, 483-511.
- Sorbini, L. 1981. The Cretaceous fishes of Nardò. 1°. Order Gasterosteiformes (Pisces). *Bollettino del Museo Civico di Storia Naturale di Verona*, **8**, 1-27.

- Stamatakis, A. 2006. RAxML-VI-HPC: Maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics*, **22**, 2688-2690.
- Stiassny, M. L. J. 1981. The phyletic status of the family Cichlidae (Pisces: Perciformes): a comparative anatomical investigation. *Netherlands Journal of Zoology*, **31**, 275-314.
- Strauss, D. & Sadler, P. M. 1989. Classical confidence intervals and Bayesian probability estimates for ends of local taxon ranges. *Mathematical Geology*, **21**, 411-427.
- Tyler, J. C. & Santini, F. 2002. Review and reconstructions of the tetraodontiform fishes from the Eocene of Monte Bolca, Italy, with comments on related Tertiary taxa. *Studi e Rieche sui Giacimenti Terziari di Bolca, Museo Civico di Storia Naturale di Verona*, **9**, 47-119.
- Tyler, J. C. & Sorbini, L. 1996. New superfamily and three new families of tetraodontiform fishes from the Upper Cretaceous: The earliest and most morphologically primitive plectognaths. *Smithsonian Contributions to Paleobiology*, **82**, 1-59.
- Zumoffen, G. 1926. *Géologie du Liban*. Paris: Barrère



#### 4. SUPPLEMENTARY FIGURES

Figure S1. Phylogeny of 89 species of Cichlidae inferred from a partitioned maximum-likelihood analysis of DNA sequences from ten nuclear genes. Filled black circles identify clades supported with a bootstrap score of 100%, filled grey circles identify clades with a bootstrap score between 99 and 90%, and unfilled circles identify clades supported with a bootstrap score between 89% and 70%.

Figure S2. Time-calibrated maximum clade credibility Bayesian inferred phylogeny of 158 species of Percomorpha, showing only the portion containing 89 sampled species of Cichlidae, based on ten nuclear genes and 10 fossil age constraints. Bars represent the posterior distribution of divergence time estimates. Grey bars identify nodes supported with Bayesian posterior probabilities (BPP)  $\geq 0.95$ , while white bars mark nodes with BPP  $< 0.95$ . The time-calibrated tree is shown with ages given in millions of years.

Figure S3. Time-calibrated maximum clade credibility Bayesian inferred phylogeny of 158 species of Percomorpha based on ten nuclear genes and 10 fossil age constraints. Bars represent the posterior distribution of divergence time estimates. Grey bars identify nodes supported with Bayesian posterior probabilities (BPP)  $\geq 0.95$ , while white bars mark nodes with BPP  $< 0.95$ . Nodes with age priors taken from the fossil record are numbered and specific information on calibrations are provided above in section 3. Calibration labels are placed on the branch leading to the node if it would completely obscure the bar depicting the posterior distribution. The time-calibrated tree is shown with ages given in millions of years.

- = 100% bootstrap replicates
- ◐ = 99-90% bootstrap replicates
- = 89-70% bootstrap replicates

