Preprint uploaded to bioRχiv on 08 October 2019
Version 3, © 2019 Laurin, Lapauze, Marjanović CC-BY 4.0

Peer Community In Paleontology

# What do ossification sequences tell us about the origin of extant amphibians?

**Michel Laurin[1], Océane Lapauze[1], David Marjanović[2]**

[1]CR2P (Centre de Recherche sur la Paléodiversité et les Paléoenvironments; UMR 7207), CNRS/MNHN/ Sorbonne Université, Muséum national d'Histoire naturelle, Département Histoire de la Terre, 57 rue Cuvier, F-75231 Paris cedex 05, France; *michel.laurin@mnhn.fr*
[2]Museum für Naturkunde (Leibniz Institute for Evolutionary and Biodiversity Research), Invalidenstraße 43, D-10115 Berlin, Germany; *david.marjanovic@gmx.at*

**ORCID**
0000-0003-2974-9835 (Michel Laurin)
0000-0001-9720-7726 (David Marjanović)

**ABSTRACT**

The origin of extant amphibians has been studied using several sources of data and methods, including phylogenetic analyses of morphological data, molecular dating, stratigraphic data, and integration of ossification sequence data, but a consensus about their affinities with Paleozoic tetrapods has failed to emerge. We have compiled five datasets to assess the relative support for six competing hypotheses about the origin of extant amphibians: a monophyletic origin among temnospondyls, a monophyletic origin among lepospondyls, a diphyletic origin among both temnospondyls and lepospondyls, a diphyletic origin among temnospondyls alone, and two variants of a triphyletic origin, in which anurans and urodeles come from different temnospondyl taxa while caecilians come from lepospondyls and are either closer to anurans and urodeles or to amniotes. Our datasets comprise ossification sequences of up to 107 terminal taxa and up to eight cranial bones, and up to 65 terminal taxa and up to seven appendicular bones, respectively. Among extinct taxa, only two or three temnospondyl can be analyzed simultaneously for cranial data, but this is not an insuperable problem because each of the six tested hypotheses implies a different position of temnospondyls and caecilians relative to other sampled taxa. For appendicular data, more extinct taxa can be analyzed, including some lepospondyls and the finned tetrapodomorph *Eusthenopteron*, in addition to temnospondyls. The data are analyzed through maximum likelihood, and the AICc (corrected Akaike Information Criterion) weights of the six hypotheses allow us to assess their relative support. By an unexpectedly large margin, our analyses of the cranial data support a monophyletic origin among lepospondyls; a monophyletic origin among temnospondyls, the current near-consensus, is a distant second. All other hypotheses are exceedingly unlikely according to our data. Surprisingly, analysis of the appendicular data supports triphyly of extant amphibians within a clade that unites lepospondyls and temnospondyls, contrary to all phylogenies based on molecular data and recent trees based on paleontological data, but this conclusion is not very robust.

# INTRODUCTION

Paleontologists have been studying the origin of the extant amphibian clades for more than a century. Early studies generally proposed an origin of at least some extant amphibians from temnospondyls. Cope (1888) initially suggested that batrachians (anurans and urodeles) derived from temnospondyls (a large clade of limbed vertebrates known from the Early Carboniferous to the Early Cretaceous) because he believed that the batrachian vertebral centrum was an intercentrum, the dominant central element of temnospondyls. Later, Watson (1940) argued that anurans were derived from temnospondyls because of similarities (mostly in the palate) between the temnospondyl "*Miobatrachus*" (now considered a junior synonym of *Amphibamus*) and anurans. Monophyly of extant amphibians (Lissamphibia) was proposed by Parsons and Williams (1962, 1963), an idea that was accepted more quickly by herpetologists than by paleontologists. Lissamphibian monophyly was supported by (among a few other character states) the widespread occurrence of pedicellate, bicuspid teeth. The subsequent discovery of such teeth in the amphibamid temnospondyl *Doleserpeton* (Bolt 1969) reinforced the widespread acceptance of an origin of Lissamphibia from within temnospondyls (e.g., Schoch and Milner 2004). Recently, this hypothesis, referred to as the temnospondyl hypothesis or TH for short (Fig. 1c), has been supported by several phylogenetic analyses based on phenotypic data matrices (e.g. Ruta and Coates 2007; Sigurdsen and Green 2011; Maddin et al. 2012; Pardo et al. 2017a, b: fig. S6; Mann et al. 2019).

Other hypotheses about the origin of extant amphibians have been available in the literature for nearly as long a time (see Schoch and Milner 2004 for a historical review). These were initially formulated especially for the urodeles and caecilians, which are less similar to temnospondyls and lack a tympanic middle ear (which is present in most anurans and often inferred for at least some temnospondyls but absent in lepospondyls). Thus, Steen (1938) highlighted similarities in the palate (broad cultriform process of the parasphenoid) and cheek (loss of several bones) between lysorophian lepospondyls and urodeles. Carroll and Currie (1975) and Carroll and Holmes (1980) argued that the exant amphibians had three distinct origins among early stegocephalians; while they accepted an origin of anurans among temnospondyls, they suggested that urodeles and caecilians originated from two distinct groups of lepospondyls (*Rhynchonkos* for caecilians, Hapsidopareiidae for urodeles). Later, based mostly on developmental similarities between the temnospondyl *Apateon* and urodeles, Carroll (2001, 2007) and Fröbisch et al. (2007) proposed another hypothesis involving a triphyletic origin of lissamphibians, with an origin of anurans and urodeles from two distinct temnospondyl groups, while the caecilians would remain in the lepospondyl clade. This is what we call the polyphyly hypothesis (PH). We have tested two versions. One (here called PH1; Fig. 1e) was cautiously suggested by Fröbisch et al. (2007); it agrees with the paleontological consensus in placing all or most lepospondyls closer to Amniota than to Temnospondyli (Fig. 1b; Sigurdsen and Green 2011; Pardo et al. 2017a, b: fig. S6; Marjanović and Laurin 2019; Clack et al. 2019; Mann et al. 2019). The other (PH2; Fig. 1f) is modified to make Lissamphibia monophyletic with respect to Amniota, a fact we consider demonstrated beyond reasonable doubt by multiple phylogenetic analyses of molecular data (Fig. 1a; Irisarri et al. 2017; Feng et al. 2017; and references cited therein); this comes at the expense of contradicting the paleontological consensus, which was not yet established when Milner (1993: 16–18, fig. 5B) argued for something like the PH2 as one of two more or less equal possibilities. Anderson (2007) and Anderson et al. (2008) found lissamphibian diphyly, specifically a monophyletic, exclusive Batrachia among the temnospondyls while keeping the caecilians among the lepospondyls (DH1; Fig. 1g). Pardo et al. (2017b: fig. 2, S7) presented a similar hypothesis, with batrachians and caecilians having separate origins within the temnospondyls (DH2; Fig. 1h); we should point out, however, that their dataset contained only temnospondyls and lissamphibians, and while they found the DH2 using Bayesian inference, it was only one of four equally parsimonious results (see Marjanović and Laurin 2019 for this fact and a discussion of Bayesian analysis of paleontological datasets). Further, a monophyletic origin of all extant amphibians among lepospondyls has also been proposed (Laurin 1998; Pawley 2006: appendix 16; Marjanović and Laurin 2009, 2013a, 2019). This will be referred to below as the lepospondyl hypothesis (LH; Fig. 1d).

Phylogenetic analyses of molecular data cannot distinguish the TH, the PH2, the DH2 or the LH from each other by topology (Fig. 1) because all of these imply lissamphibian monophyly with respect to amniotes, and molecular data are not available from any other tetrapodomorphs. Several other types of data and methods have, however, been used to try to discriminate between the various hypotheses on the origin of extant amphibians. In addition to classical phylogenetic analyses of morphological data matrices, these include the use of molecular dating (Zhang et al. 2005; Marjanović and Laurin 2007; Pardo et al. 2017b) and stratigraphic data (Marjanović and Laurin 2008) to compare the inferred divergence dates between the three main extant amphibian clades on the basis of molecular data with predictions based on the fossil record under the TH and the LH on one side

and the PH and the DH on the other. However, developmental data, in the form of ossification sequences, have been the second-most frequently used (after classic morphological data) to argue for particular phylogenetic hypotheses. These data include mainly cranial (e.g. Schoch 2002, 2006; Schoch and Carroll 2003; Schoch and Milner 2004; Anderson 2007; Carroll 2007; Germain and Laurin 2009) and autopodial ossification sequences (e.g. Fröbisch et al. 2007, 2015). Ossification sequences of other parts of the skeleton, like the vertebrae, shoulder girdle and scales, are also documented in a few Paleozoic stegocephalians (e.g. Carroll et al. 1999; Witzmann and Schoch 2006; Anderson 2007; Carroll 2007; Olori 2013), not to mention finned tetrapodomorphs (Cloutier 2009), but these have played a minor role in the controversy about the origin of extant amphibians. Recently, Danto et al. (2019) concluded that vertebral ossification sequences varied too quickly and could not be used to assess the origin of lissamphibians. This study relies on both cranial and appendicular ossification sequences and compares their implications for tetrapod phylogeny.

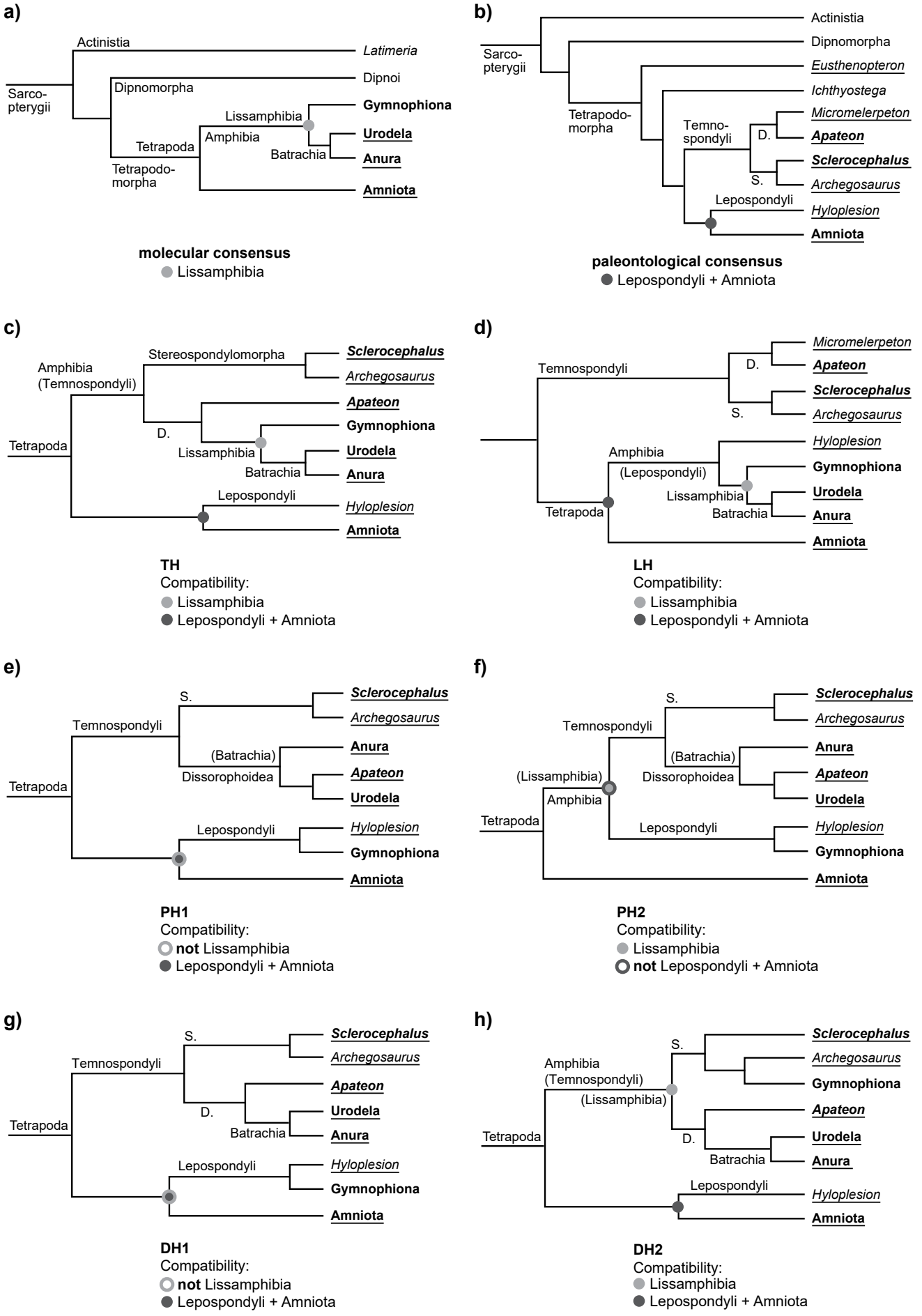
# MATERIAL AND METHODS

## Ossification sequence data

From all the literature we could access, we compiled the most extensive database on ossification sequences for osteichthyans that exists to date. The most useful sources for extant taxa included compilations: Harrington et al. (2013) for amphibians, Weisbecker and Mitgutsch (2010) for anurans, Hugi et al. (2012) for squamates, Maxwell et al. (2010) for birds, and Koyabu et al. (2014) and Weisbecker (2011) for mammals. The cranial and appendicular sequences of Permian temnospondyls (the stereospondylomorphs *Sclerocephalus* and *Archegosaurus*, the non-branchiosaurid "branchiosaur" *Micromelerpeton* and the branchiosaurids "*Melanerpeton*" *humbergense*, *Apateon caducus* and *A. pedestris*) were assembled from several references cited in the Appendix; note that the two *Apateon* species are each represented by two different sequences scored after populations from two separate paleo-lakes (Erdesbach and Obermoschel) in which both species occur. Appendicular ossification sequences of the lepospondyls *Microbrachis* and *Hyloplesion* are incorporated from Olori (2013), that for the finned tetrapodomorph *Eusthenopteron* was combined from Cote et al. (2002) and Leblanc and Cloutier (2005).

All sources of our sequence data can be found in the Appendix. The sequences themselves and the phylogenetic trees corresponding to the tested hypotheses are included in the supplements, which are posted on the bioRχiv page from which this paper is available. The sequences were not used to generate the tree topology or the branch lengths (which represent evolutionary time); the tree is compiled from published sources (provided below) which did not use any ossification sequences in their phylogenetic analyses.

**Figure 1 (next page).** Hypotheses on the relationships of the extant amphibian clades since the late 20th century. The names of terminal taxa sampled here for cranial characters are in boldface, those sampled for appendicular characters are underlined; the names of larger clades are placed toward the right end of a branch if they have minimum-clade (node-based) definitions, to the left if they have maximum-clade (branch-based) definitions. Names in parentheses would, given that phylogenetic hypothesis, not be used, but replaced by synonyms. Among terminal taxa, "*Melanerpeton*" *humbergense*, sampled for appendicular characters, is not shown, but is always the sister-group of *Apateon*; *Microbrachis*, likewise sampled for appendicular characters, is not shown either, but is always the sister-group of *Hyloplesion*; *Eusthenopteron* is not shown in c)–h), where it forms the outgroup (b)). See text for *Micromelerpeton* and for references. The first two trees (a, b) show the current consensus; the other trees (c–h) show the various tested paleontological hypotheses. Abbreviations: D., Dissorophoidea; S., Stereospondylomorpha. a) Consensus of the latest phylogenetic analyses of molecular ; all named clades are therefore extant. Note the monophyly of the extant amphibians (Lissamphibia, marked with a light gray dot) with respect to Amniota. b) Consensus of all analyses of Paleozoic limbed vertebrates, omitting the extant amphibian clades. Note the monophyly of "lepospondyls" + amniotes (marked with a dark gray dot). c) TH: "temnospondyl hypothesis". Lissamphibia nested among dissorophoid temnospondyls. Compatible with both a) and b) (gray dots). d) LH: "lepospondyl hypothesis". Lissamphibia nested among "lepospondyls"; consequently, temnospondyls are not crown-group tetrapods. Compatible with both a) and b) (gray dots). e) PH1: "polyphyly hypothesis", first variant. Urodela as dissorophoid temnospondyls close to *Apateon*, Anura as a separate clade of dissorophoid temnospondyls, Gymnophiona as "lepospondyls". Compatible with b) (dark gray dot) but not with a) (light gray circle). f) PH2: "polyphyly hypothesis", second variant. Like PH1, but with restored monophyly of extant amphibians with respect to amniotes (light gray dot; see a)) at the expense of compatibility with the paleontological consensus concerning the position of temnospondyls, lepospondyls, and amniotes (dark gray circle; see b)). g) DH1: "diphyly hypothesis", first variant. Batrachia as dissorophoid temnospondyls, Gymnophiona as "lepospondyls". Compatible with b) (dark gray dot) but not with a) (light gray circle). h) DH2: "diphyly hypothesis", second variant. Batrachia as dissorophoid temnospondyls, Gymnophiona as stereospondylomorph temnospondyls. Compatible with both a) and b).


a)
Actinistia
Latimeria
Dipnoi
Sarco-
pterygii
Dipnomorpha
Gymnophiona
Lissamphibia
Urodela
Amphibia
Tetrapoda
Batrachia
Anura
Tetrapodo-
morpha
Amniota
molecular consensus
Lissamphibia
b)
Actinistia
Dipnomorpha
Sarco-
pterygii
Eusthenopteron
Ichthyostega
Tetrapodo-
morpha
Micromelerpeton
Temno-
spondyli
D.
Apateon
Sclerocephalus
S.
Archegosaurus
Lepospondyli
Hyloplesion
Amniota
paleontological consensus
Lepospondyli + Amniota
c)
Stereospondylomorpha
Sclerocephalus
Amphibia
(Temnospondyli)
Archegosaurus
Apateon
Gymnophiona
D.
Tetrapoda
Lissamphibia
Urodela
Batrachia
Anura
Lepospondyli
Hyloplesion
Amniota
TH
Compatibility:
Lissamphibia
Lepospondyli + Amniota
d)
Micromelerpeton
Temnospondyli
D.
Apateon
Sclerocephalus
S.
Archegosaurus
Hyloplesion
Amphibia
Gymnophiona
(Lepospondyli)
Lissamphibia
Urodela
Tetrapoda
Batrachia
Anura
Amniota
LH
Compatibility:
Lissamphibia
Lepospondyli + Amniota
e)
S.
Sclerocephalus
Temnospondyli
Archegosaurus
Anura
(Batrachia)
Tetrapoda
Dissorophoidea
Apateon
Urodela
Lepospondyli
Hyloplesion
Gymnophiona
Amniota
PH1
Compatibility:
not Lissamphibia
Lepospondyli + Amniota
f)
S.
Sclerocephalus
Temnospondyli
Archegosaurus
Anura
(Batrachia)
(Lissamphibia)
Dissorophoidea
Apateon
Amphibia
Urodela
Tetrapoda
Lepospondyli
Hyloplesion
Gymnophiona
Amniota
PH2
Compatibility:
Lissamphibia
not Lepospondyli + Amniota
g)
S.
Sclerocephalus
Temnospondyli
Archegosaurus
Apateon
D.
Urodela
Tetrapoda
Batrachia
Anura
Lepospondyli
Hyloplesion
Gymnophiona
Amniota
DH1
Compatibility:
not Lissamphibia
Lepospondyli + Amniota
h)
S.
Sclerocephalus
Amphibia
(Temnospondyli)
Archegosaurus
Gymnophiona
(Lissamphibia)
Apateon
Tetrapoda
D.
Urodela
Batrachia
Anura
Lepospondyli
Hyloplesion
Amniota
DH2
Compatibility:
Lissamphibia
Lepospondyli + Amniota

The software we used to compute AICc weights, the CoMET module (Lee et al. 2006) for Mesquite 3.6 (Maddison and Maddison 2018), cannot handle missing data. This unfortunately meant we had to discard much information. In order to keep as many taxa as possible in the analysis, we first compiled a matrix (not shown) of 244 taxa and 213 characters. All of these characters are positions of skeletal elements (cranial, appendicular, axial and others) in ossification sequences, standardized between 0 and 1 following Germain and Laurin (2009), as explained below. Of these, we kept characters that were scored in the Paleozoic taxa in our initial database, and extant taxa that were scored for the same sets of characters. This resulted in two initial datasets, one of cranial and one of appendicular sequences (it was not possible to include both sets of sequences together because this would have left too few taxa in the matrix).

In the end, however, we were left with three overlapping cranial datasets. The largest cranial dataset we could make, dataset 2 of Table 1, has 105 taxa (103 extant, plus the two species of *Apateon* scored from Erdesbach) and seven characters: the appearance times of the premaxilla, maxilla, nasal, parietal, pterygoid, exoccipital and squamosal bones. It lacks *Sclerocephalus*, which cannot be scored for the appearance time of the squamosal. This is unfortunate because *Sclerocephalus* is one of only three extinct taxa for which a usable cranial ossification sequence is known at all, and further because it occupies a special place in the DH2, according to which it lies on the caecilian stem. We attempted to compensate for this deficiency by assembling two more cranial datasets: dataset 1, which contains 107 taxa (104 extant, *Apateon* spp. from Erdesbach, and *Sclerocephalus*) but only six characters by lacking the squamosal, and dataset 5, which includes 84 taxa (81 extant, *Apateon* spp. from Erdesbach, and *Sclerocephalus*) and eight cranial characters (the vomer and the frontal bone are added to the six of dataset 1).

For the appendicular characters, in addition to dataset 3 which contains seven characters (humerus, radius, ulna, ilium, femur, tibia and fibula) and 62 taxa (54 extant, *Apateon* spp. from Obermoschel, *Sclerocephalus*, *Archegosaurus*, *Micromelerpeton*, *Hylopleslon*, *Microbrachis* and *Eusthenopteron*), another (dataset 4) includes only four characters (radius, ulna, ilium, and femur), but it features 65 sequences, the additional data being *Apateon* spp. from Erdesbach and "*Melanerpeton*" *humbergense*. See Table 1 for a list of these datasets and the supplements for the datasets themselves.

The data loss in these various datasets is not as severe as it may first seem, because most of the characters that have been excluded from these analyses had less than 10% scored cells (sometimes less than 1%), and most of them could not be scored for any temnospondyl or lepospondyl, so they could not have helped resolve the main question examined in this study.

The order in which the sampled cranial bones ossify varies substantially in our sample of taxa, but based on simple (not phylogenetically-weighted) average position, the frontal appears first, followed closely by the premaxilla, parietal, and maxilla (in close succession), and then by the squamosal, exoccipital, pterygoid, and last by the nasal. However, all of these bones ossify first (among these bones; not necessarily in the whole skeleton) in at least one of the included taxa. Among the appendicular bones, there is more variability; all ossify first in at least one of the 62 sampled taxa, and three (radius, ulna and ilium) ossify last in at least one taxon.

Due to the homology problems between the skull bones of tetrapods and actinopterygians and missing data, we had to omit all actinopterygians from our analyses. As cranial ossification sequences remain poorly documented for extant finned sarcopterygians, except perhaps lungfish, whose skull bones seem mostly impossible to homologize (Criswell 2015), our analyses of those data are restricted to limbed vertebrates. However, for appendicular data, we were able to include the Devonian tristichopterid *Eusthenopteron foordi*.

Unfortunately, the only cranial ossification sequence available for any supposed lepospondyl, that of the aïstopod *Phlegethontia longissima*, is documented from only three ossification stages (Anderson et al. 2003; Anderson 2007). This poses a problem for our analysis method, which assumes that character evolution can be modeled as Brownian motion; this assumption is decreasingly realistic as the number of character states (sequence positions) decreases, because the resulting distribution deviates increasingly from that of a continuous character. Furthermore, some recent anatomical restudies and phylogenetic analyses suggest that aïstopods are not lepospondyls, but early-branching stem-stegocephalians (Pardo et al. 2017a, 2018; Mann et al. 2019; Clack et al. 2019).

The low taxon sample is more limiting for this analysis than the low character sample. However, as explained below, the absence of lepospondyl sequences in our cranial dataset does not preclude testing the six hypotheses (TH, PH1, PH2, DH1, DH2, LH; see above or Figure 1 for the explanation of these abbreviations) because each of these six hypotheses makes different predictions about where temnospondyls and caecilians fit relative to other taxa. Thus, in

the absence of lepospondyls in our dataset, the tests of these hypotheses are somewhat indirect and inference-based, but they remain possible. Our tests based on appendicular data include two lepospondyls (*Hyloplesion longicostatum* and *Microbrachis pelikani*), but the absence of caecilians in that dataset proves more limiting than the absence of lepospondyls in the cranial dataset because the TH, DH1 and DH2 become indistinguishable (Fig. 1c, g, h). However, the presence of the temnospondyl *Micromelerpeton* allows us to test two variants of the TH/DH distinguished by the monophyly (e.g. Ruta and Coates 2007) or polyphyly (e.g. Schoch 2018) of "branchiosaurs" (the temnospondyls *Apateon*, "*Melanerpeton*" *humbergense* and *Micromelerpeton*).

**Table 1.** List of datasets used in this paper. All are subsets of our global compilation that were selected to meet the requirement of the method used (missing data cannot be handled). The temnospondyl species *Apateon caducus* and *A. pedestris* are included in all datasets, but scored after populations from two different paleo-lakes in which both species occur.

| Dataset number | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| Type of characters | cranial | cranial | appendicular | appendicular | cranial |
| Number of characters | 6 | 7 | 7 | 4 | 8 |
| Number of taxa | 107 | 105 | 62 | 65 | 84 |
| *Sclerocephalus* | yes | no | yes | yes | yes |
| Source of data for *Apateon* | Erdesbach | Erdesbach | Obermoschel | Erdesbach and Obermoschel | Erdesbach |
| Additional Paleozoic taxa | None | None | *Archegosaurus*, *Micromelerpeton*, *Hyloplesion*, *Microbrachis*, *Eusthenopteron* | *Archegosaurus*, *Micromelerpeton*, "*Melanerpeton*" *humbergense*, *Hyloplesion*, *Microbrachis*, *Eusthenopteron* | None |
| Table in which it is used | 2, 5 | 3, 6 | 4, 8 | 4, 9 | 7 |

## Sensitivity analysis for sequence polymorphism

Given the potential impact of intraspecific variability in ossification sequence on inferred nodal sequences and heterochrony (Olori 2013; Sheil et al. 2014), we compiled two consensus sequences for *Apateon caducus* and *A. pedestris* each, representing two localities where both species occur, the paleo-lakes of Erdesbach (Schoch 2004) and Obermoschel (Werneburg 2018). Based on dataset 4 (see Table 1), we incorporated these into a global and two separate analyses (one analysis per locality) to determine the impact of the observed variability. As detailed above, incorporating the sequences from Erdesbach reduced the number of characters from seven to only four because the software used cannot handle missing data (see above and below), but this information loss is compensated by the great increase in number of sequences from extinct taxa (eleven instead of two, when counting the sequences of *Apateon* from both localities separately) and the fact that this includes some lepospondyls (see below). It would have been even better to perform a sensitivity analysis incorporating variability for all taxa for which such information was available, but given the scope and nature of our study, this would have been exceedingly time-consuming and is best left for the future.

## Standardization of the data

Given that various taxa differ in their numbers of bones and that the resolution of the sequences is also variable between taxa, these data needed to be standardized to make comparisons and computations meaningful, as suggested by Germain and Laurin (2009). Note that we performed this standardization on the complete dataset of characters, before filtering for data completeness. This complete dataset (not shown) includes 213 cranial, appendicular and other characters, but no taxon is scored for all characters, because that matrix has much missing data. For instance, the most completely scored taxon, *Amia calva*, still has 57.4% missing data (more than half),

which indicates that 92 characters were scored for this taxon, including several ties (the resolution was 41 positions, so they varied by increments of 0.025 or 2.5% of the recorded ontogeny). We did not re-standardize after filtering characters out because we believe that the initial standardization better reflects the relative position of events in development than a standardization based on only seven events in ontogeny. Because of this, some characters in the reduced datasets lack states 0 or 1 for some taxa. This is simply because the first or last events in the ontogenetic sequence were filtered out. Thus, we used the position in the sequence (from first to last, in the complete dataset) and standardized this relative sequence position between 0 and 1 using the formula given by Germain and Laurin (2009). The standardized sequence position ($X_s$) is:

$$X_s = (X_i - X_{min})/(X_{max} - X_{min}),$$

where:

$X_i$ is the position of a given bone in the sequence

$X_{min}$ is the lowest position in the sequence (generally denoted 0 or 1)

$X_{max}$ is the highest position in the sequence (for instance, if there are 20 bones, $X_{min}$ is 1 and the sequence is completely resolved, $X_{max} = 20$).

This yields a standardized scale that varies between 0 and 1 for each taxon, in which 0 and 1 are the positions of the first and last events in the sequence, respectively. For instance, for *Ambystoma maculatum* (an extant urodele), in the original dataset, the first events (tied) were the ossification of premaxilla, vomer, dentary and coronoid (standardized position: 0); the last event was the articular (standardized position: 1), and there is a resolution of 12 positions (hence, increments of 0.0909 or 1/11). However, in the final dataset of 7 charcters, the articular is absent; hence, the first bone in the sequence is the premaxilla, at a standardized position of 0, and the last is the nasal, as a standardized position of 0.8181 because all events in position 1 (articular) and 0.9091 (stapes) have been filtered out.

We also experimented with using size (skull length) or developmental stage as standards, but this led to lower sequence resolution because body size is not available for all sequence positions and for all taxa (results not shown), so we worked only with sequences standardized by position. Given that our data filtering procedure retains few data (only six, seven or eight characters for the cranial dataset, and four or seven characters for the postcranial dataset), it is important to use the method that discards the least amount of data, and this was achieved by using sequence position. We do not imply that standardizing by size is not recommended in general. On the contrary, if good body size data were available for all taxa and all developmental stages, this should be a better strategy, and only having access to absolute time should be even better. However, practical limitations of data availability prevent us from using these methods now.

Our ossification sequence data (reduced dataset of four to eight characters) of extant and extinct taxa, and the phylogenetic trees we used, are available in the supplements.

## Analysis methods

To discriminate between the six hypotheses about the origin of extant amphibians, two methods are available: direct phylogenetic analysis of the sequence data, and comparisons of the tree length (number of steps in regular parsimony, squared length in squared-change parsimony, likelihood, or similar measures) of various trees selected a priori to represent these hypotheses (in these trees, only the position of caecilians and extinct taxa, here temnospondyls and lepospondyls, varies). We used both approaches but expected the second to perform much better because relatively few data are available, and thus, phylogenetic analysis of such data is unlikely to provide a well-resolved tree.

For the first approach, we first transformed the standardized sequence positions back into discrete characters using formulae in a spreadsheet and scaled the characters so that the highest state in all would be 9. This ensures that each character has equal weight in the analysis, regardless of its variability in the ossification sequence. The characters were ordered to reflect the assumed evolutionary model (ontogenetic timing is a quantitative character that was discretized) and because for such characters, ordering yields better results (Rineau et al. 2015, 2017; see discussion in Marjanović & Laurin 2019). The resulting data matrices (one for cranial and another for appendicular characters, both with seven characters each) were analysed using parsimony in PAUP* 4.0a165 (Swofford 2019). We used the TBR (tree bisection-reconnection) branch swapping algorithm and performed a search with 50 random addition replicates (or several such searches, for the cranial data) while holding two trees at each step and with a maximum number of trees set at one million. For cranial data, the main search lasted about 100 hours on a MacBook Pro Retina with a 2.5 GHz iCore 7 quadri-core processor and 16 GB RAM. The exact search time cannot be reported because PAUP* crashed after saving the trees to a file for one of the longest runs (several analyses were made, over

several days), but before the log could be saved. The analysis of the seven appendicular characters was much faster (27 minutes and a half), presumably because that matrix has fewer taxa (62 instead of 105).

For the second approach (comparison of fit of various trees selected a priori to reflect previously published hypotheses), we used the CoMET module (Lee et al. 2006) for Mesquite 3.6 (Maddison and Maddison 2018) to test the relative fit of the data on trees representing the six hypotheses. CoMET calculates the likelihood and the AIC (Akaike Information Criterion) of nine evolutionary models given continuous data and a tree. Note that our data only represent an approximation of continuous data; if standardization had been performed on developmental time or body size, the data would actually have been continuous. Standardization was carried out using sequence position because of data limitation problems, so the data actually follow a decimalized meristic scale. However, the difference between these situations decreases as the number of sequence positions increases, and our global scale includes up to 41 positions (and an average of 10.9 positions), so our data should approximate a continuous distribution sufficiently well for our analyses to be valid. This consideration prevents us from adding the highly apomorphic aïstopod *Phlegethontia*, for which only three cranial ossification stages are known (Anderson et al. 2003; Anderson 2007); moreover, five of the seven bones included in our analyses appear in the last two of these stages, and two of the relevant bones (parietal and exoccipital) are not present as separate ossifications, which would create additional missing data. In that case, the very low number of stages would create strong departures from the assumption of continuous data. This would probably create statistical artifacts, and the uncertainty about the position of *Phlegethontia* (Pardo et al. 2017a, 2018; Marjanović and Laurin 2019; Clack et al. 2019) would complicate interpretation of the results.

The nine models evaluated by CoMET are obtained by modifying the branch lengths of the reference tree. Thus, branches can be set to 0 (for internal branches only, to yield a non-phylogenetic model), to 1 (equal or speciational model), left unchanged from their original length (gradual evolution in our case, where the original lengths represent geologic time), or set free and evaluated from the data (free model). This can be applied to internal and/or external branches, and various combinations of these yield nine models (Lee et al. 2006: fig. 1). Among these nine models two have been frequently discussed in the literature and are especially relevant. The first is gradual evolution, in which branch lengths (here representing evolutionary time) have not been changed. The second is the speciational model, in which all branches are set to the same length because changes are thought to occur at speciation events, which are typically equated with cladogeneses in evolutionary models (Bokma et al. 2016).This model has some similarities with Eldredge and Gould's (1972) punctuated equilibria (though a model with one internal branch stemming from each node set to 0 and the other set to 1 would be even closer to the original formulation of that model). In this study, we assessed the fit of six of the nine models covered by CoMET; the other three (the punctuated versions of distance [original branch length], equal and free) in which the one of each pair of daughter-lineages has a branch length of zero, could not be assessed due to problems in the current version of CoMET and possibly the size of our dataset.

Provided that the same evolutionary model is optimal for all compared phylogenetic hypotheses (this condition is met, as shown below), the AIC weights of the various trees under that model can be used to assess the support for each tree. In such comparisons, the topology is part of the evolutionary model, and the data are the sequences. These comparisons can show not only which tree is best supported, but how many times more probable the best tree is compared to the alternatives. This quantification is another reason to prefer this approach over a phylogenetic analysis (performed below, but with the poor results that we anticipated), which can at best yield a set of trees showing where the extinct taxa most parsimoniously fit (if we had dozens of characters, this might be feasible). Comparisons with other hypotheses through direct phylogenetic analysis are not possible. Given the small sample size (which here is the number of characters), we computed the corrected AIC (AICc) and the AICc weights using the formulae given by Anderson and Burnham (2002) and Wagenmakers and Farrell (2004).

Our tests make sense only in the presence of a phylogenetic signal in the data. In addition to the test of evolutionary model in CoMET mentioned above (which tests non-phylogenetic as well as phylogenetic models), we performed a test based on squared-change parsimony (Maddison 1991) and random taxon reshuffling (Laurin 2004). For this test, we compared the length of the LH (lepospondyl hypothesis; Fig. 1d) reference tree (with and without *Sclerocephalus*) to a population of 10,000 random trees produced by taxon reshuffling.

It could be argued that using other methods (in addition to the method outlined above) would have facilitated comparisons with previous studies. However, the two main alternative methods, event-pair cracking with Parsimov (Jeffery et al. 2005) and

Parsimov-based genetic inference (PGI; Harrison and Larsson 2008), have drawbacks that led us to not using them. Our objections against event-pair cracking with Parsimov were detailed by Germain and Laurin (2009). In short, that method requires an unnecessary decomposition of sequences into event pairs, and it cannot incorporate absolute timing information (in the form of time, developmental stage or body size, for instance) or branch length information. More importantly, the simulations performed by Germain and Laurin (2009) showed that event-pair cracking with Parsimov yields more artefactual change and has lower power to detect real sequence shifts. That method is also problematic when trying to infer ancestral sequences and can lead to impossible ancestral reconstructions (e.g. A occurs before B, B occurs before C, and C occurs before A), as had been documented previously (Schulmeister and Wheeler 2004: 55). This would create problems when trying to compare the fit of the data on various phylogenetic hypotheses. The performance of Parsimov-based genetic inference (PGI; Harrison and Larsson 2008) has not been assessed by simulations, but it rests on an edit cost function that is contrary to our working hypothesis (that the timing of developmental events can be modeled with a bounded Brownian motion model, which is assumed by continuous analysis). More specifically, Harrison and Larsson (2008: 380) stated that their function attempts to minimize the number of sequence changes, regardless of the magnitude of these changes. We believe that disregarding the size of changes is unrealistic, as shown by the fact that Poe's (2006) analyses of thirteen empirical datasets rejected that model (which he called UC, for unconstrained change) in favor of the model we accept (AJ for adjacent states, which favors small changes over large ones). Furthermore, analyses of ossification sequence data using techniques for continuous data as done here (see above) have been performed by an increasingly large number of studies (e.g., Skawiński and Borczyk 2017; Spiekman and Werneburg 2017; Werneburg and Geiger 2017, just to mention papers published in 2017), so the issue of ease of comparisons of our results with other studies is not as serious as it would have been only a few years ago, and it should be decreasingly so in the future.

## Reference phylogenies

We built a reference timetree that attempts to capture the established consensus (Fig. 2; see the next paragraphs for the sources). The tree was compiled in Mesquite versions up to 3.6 (Maddison and Maddison 2018) and time-calibrated using the Stratigraphic Tools module for Mesquite (Josse et al. 2006). For consistency and to avoid the effects of gaps in the fossil record, we used molecular divergence dates whenever possible. The tree had to be time-scaled because many of the evolutionary models that we fit on the tree in the first series of tests (to determine which evolutionary model can be used to compare the fit of the hypotheses) use branch lengths to assess model fit. Note that our procedure requires estimating divergence times between all taxa (geological ages of all nodes). When taxa are pruned, branch lengths are adjusted automatically. The main sources we used for topology and divergence times (and hence branch lengths) are as follows:

The phylogeny of lissamphibians follows the work of Jetz and Pyron (2018). However, several other sources have been used for the temporal calibration of the tree: Germain and Laurin (2009) was used for the urodeles, whereas Feng et al. (2017), supplemented by Bossuyt and Roelants (2009) and Pyron (2014), was used for the anurans as well as more rootward nodes (Batrachia, Lissamphibia, Tetrapoda; also Amniota). Marjanović and Laurin (2013b) was used for the Ranidae, Ceratophryidae and Hylidae.

The sediments that have preserved the temnospondyls *Apateon* and *Sclerocephalus* are not easy to correlate with each other or with the global chronostratigraphic scale. Combining stratigraphic information from Schoch (2014a), Schneider et al. (2015) and Werneburg (2018), we have placed all three sampled species (*A. pedestris*, *A. caducus*, *S. haeuseri*) at the Sakmarian/Artinskian stage boundary (Permian; 290.1 Ma ago); combining stratigraphic information from Schneider et al. (2015) with the phylogeny in Schoch (2014a), we have tentatively placed the divergence between the two *Apateon* species (which are not sister-groups: Schoch 2014a) at the Kasimovian/Gzhelian stage boundary (Carboniferous; 303.7 Ma ago). The age of the last common ancestor of *Apateon* and *Sclerocephalus* depends strongly on temnospondyl phylogeny, which remains unresolved (Pardo et al. 2017b; Marjanović and Laurin 2019; and numerous references in both); as a compromise between the various options, we have provisionally placed it at the boundary between the Early and the Late Carboniferous (Serpukhovian/ Bashkirian, 323.2 Ma ago) where applicable.

We sampled many extant amniotes to achieve broad coverage of Tetrapoda. For the birds, Pons et al. (2005) was used for the Laridae, Wang et al. (2013) for the Phasianidae and Gonzales et al. (2009) for the Anatidae. The temporal calibration was taken from Prum et al. (2015) as recommended by Berv and Field (2017); gaps were filled in using the database www.birdtree.org.

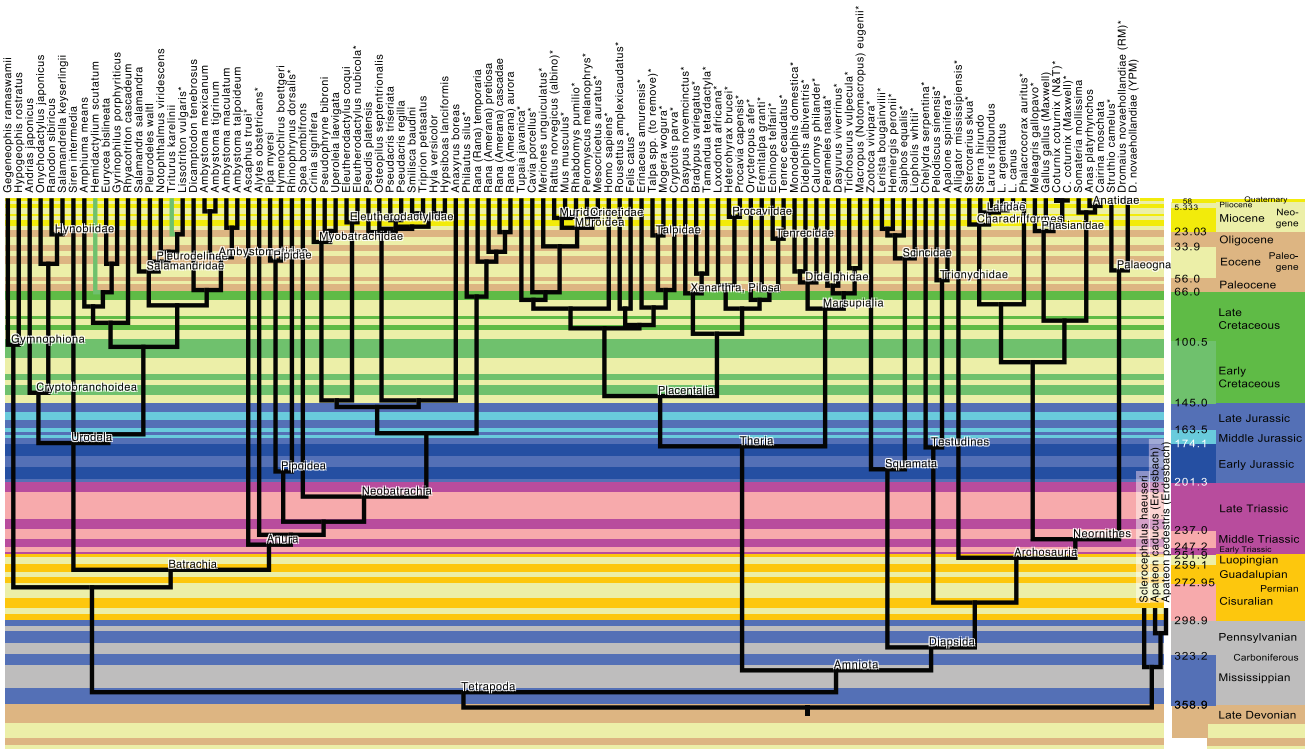



**Figure 2**. Reference phylogeny used for some of the analyses, illustrating the LH (lepospondyl hypothesis) of lissamphibian origins. The tree was time-calibrated, but analyses showed that branch lengths are irrelevant, given that the best model is speciational (Tables 2–4). Main sources for topology and divergence times: Reeder (2003); Brandley et al. (2005); Pons et al. (2005); Lecompte et al. (2008); Bossuyt and Roelants (2009); Germain and Laurin (2009); Hugall et al. (2007); Gonzales et al. (2009); Meredith et al. (2011); Sterli et al. (2013); Wang et al. (2013); Marjanović and Laurin (2013b, 2019); Pyron (2014); Rabosky et al. (2014); Schoch (2014a); Prum et al. (2015); Zhuang et al. (2015); Tarver et al. (2016); Feng et al. (2017); Irisarri et al. (2017); Lu et al. (2017); Pardo et al. (2017b); Jetz and Pyron (2018). The colored bands represent geological stages from the international geological timescale (Ogg et al. 2016).

Several papers, mainly Tarver et al. (2016), were used for the phylogeny and divergence times of mammals. For the Muridae, three references were used: Lecompte et al. (2008), Zhuang et al. (2015), and Lu et al. (2017) for the position of two taxa: *Mesocricetus auratus* and *Peromyscus melanophrys.* Other species were placed following the work of Meredith et al. (2011), which also gives divergence times. We caution, however, that all available molecular dates for Paleogene and earlier mammal nodes are controversial and may be overestimates (Berv and Field 2017; Phillips and Fruciano 2018).

Three references were also used to integrate squamates in the phylogenetic tree and for the calibration of divergence times: Brandley et al. (2005), Rabosky et al. (2014), Reeder (2003). Sterli et al. (2013) was used for turtles.

For turtles, there is now a near-consensus that they are diapsids, a hypothesis that is not necessarily incompatible with an origin among “parareptiles” (Laurin and Piñeiro 2017). Thus, following most recent phylogenetic analyses of molecular data (e.g., Hugall et al. 2007; Irisarri et al. 2017), we have inserted them as the sister-group of Archosauria.

We disagree with several of the calibration dates in Irisarri et al. (2017), which often appear unreasonably old. For instance, they place the divergence between caecilians and batrachians and the divergence between anurans and urodeles in the Early Carboniferous, around 330 and 320 Ma, respectively, but our thorough analyses of the fossil record, with due consideration of its incompleteness, suggest significantly more recent dates, in the Permian (Marjanović and Laurin 2007, 2008, 2013b). This is not surprising because some of the dating constraints used by Irisarri et al. (2017: table S8) are wrong. For instance, they enforced a minimal divergence age between cryptodiran and pleurodiran turtles of 210 Ma (Late Triassic), but all analyses of the last fifteen years (e.g. Sterli et al. 2013, 2018) strongly suggest that the oldest known turtles that fit within this dichotomy date from the Late Jurassic, less than 165 Ma. The divergence between humans and armadillos (boreotherian and xenarthran placentals) was constrained to the middle of the Cretaceous (95.3–113 Ma), based on outdated literature that assigned a

wide variety of stem-eutherians to highly nested positions in the placental crown; there are currently no clear placentals known from any Cretaceous sediments even as young as 66 Ma (see e.g. Halliday et al. 2015, 2016; Davies et al. 2017; Phillips and Fruciano 2018), barely half the age of the older end of the constraint range. Conversely, the divergence between diapsids (hence sauropsids) and synapsids had a minimal age constraint of 288 Ma (Early Permian), which is much too young given the presence of sauropsids (and presumed synapsids) in Joggins, in sediments that have recently been dated (Carpenter 2015) around 317–319 Ma (early Late Carboniferous). Thus, we have not used divergence dates from that source.

To discriminate among the hypotheses on lissamphibian origins, we inserted the temnospondyl *Apateon* in the tree where each predicts that it should be (Fig. 1c–h). Thus, according to the TH (temnospondyl hypothesis; Fig. 1c), *Apateon* lies on the lissamphibian stem. Under the LH (lepospondyl hypothesis; Fig. 1d), *Apateon* lies on the tetrapod stem. Under both versions of the DH (diphyly hypothesis; Fig. 1g, h), *Apateon* lies on the batrachian stem. Under both versions of the PH (polyphyly hypothesis; Fig. 1e, f), *Apateon* lies on the caudate stem. Within the DH and the PH, both versions of each differ in the position of Gymnophiona. Thus, despite the absence of any lepospondyl in our cranial ossification sequence datasets, our taxonomic sample allows us to test all these competing hypotheses. The appendicular datasets allow more direct tests of some of these hypotheses because they include two lepospondyl taxa, which were likewise placed in trees representing the tested hypotheses (Fig. 1).

*Sclerocephalus* is the sister-group of *Apateon* under the LH (Fig. 1d), immediately rootward of it (on the lissamphibian stem) under the TH (Fig. 1c) and likewise (but on the batrachian stem) under the DH1 (Fig. 1g), on the caecilian stem under the DH2 (Fig. 1h) and the sister-group of Batrachia (including *Apateon*) under both versions of the PH (Fig. 1e, f).

"*Melanerpeton*" *humbergense* (appendicular data only) is the sister-group of *Apateon* in all trees, except under the hypothesis of branchiosaur paraphyly; *Eusthenopteron* (appendicular data only) forms the outgroup in all trees.

The lepospondyls *Microbrachis* and *Hyloplesion*, from both of which only appendicular data are available, form an exclusive clade (Marjanović and Laurin 2019; Clack 2019). This clade is the sister-group of Lissamphibia (represented only by Batrachia) under the LH (because caecilians are lacking from the appendicular datasets), of Amniota under the TH and both versions of the DH (these three cannot be distinguished due to the absence of caecilians) as well as under the PH1, and of Temnospondyli (including Batrachia) under the PH2 (see the legend of Figure 1 for an explanation of these abbreviations).

The temnospondyl *Micromelerpeton*, from which likewise only appendicular data are available, forms the sister-group of *Apateon* under the LH. The uncertainty over its phylogenetic position within Dissorophoidea (as the sister-group to the rest, including anurans and urodeles: e.g. Schoch 2018; as the sister-group of *Apateon* + "*Melanerpeton*" *humbergense*: e.g. Ruta & Coates 2007; Marjanović and Laurin 2019) generates two versions of the TH/DH1/DH2 tree for the appendicular dataset. We tested both of these versions against that dataset, for a total of five trees.

To ensure that our analyses were not biased in favor of a given hypothesis, and in case that a continuous evolutionary model were favored, we initially adjusted the branch lengths such that the sum of branch lengths was equal between the compared topologies and that the root was approximately at the same age (in this case in the Tournaisian, the first stage of the Carboniferous). This was done for the trees used to compare the hypotheses using the cranial dataset because if a model incorporating (variable) branch length information had been selected, and if the trees representing the various hypotheses had not all had the same total length (the sum of all branch lengths), the resulting distortions in branch lengths created around the extinct taxa (whose height compared to extant taxa is specified by their geological age) would have introduced another variable influencing the AICc. But given that the selected model ignores branch lengths, this precaution turned out to be superfluous. We have therefore not made these time-consuming adjustments to the additional trees we generated later to analyze the appendicular data.

## RESULTS

In the phylogenetic analysis of cranial data, a single tree island of 22,077 trees of 438 steps was found, only once, so there might be more trees of that length and perhaps even shorter trees. Initially, an island of 22,075 trees was found; we swapped on each of these in a subsequent run, which only recovered two additional trees. Given that slightly longer trees did not differ much from those that we obtained, the low quality of the results (poor congruence with the established consensus about the monophyly of major clades such as squamates, birds, mammals and turtles) and the fact that about four full days of

computer time had been spent on analysis of the cranial data, we did not pursue that search further. As expected, the strict consensus tree is poorly resolved (Fig. 3). The majority-rule consensus (not shown, but available in the supplements available on the bioRχiv web page) is more resolved but not necessarily better because much of the additional resolution contradicts the established consensus. For the appendicular matrix, 22,757 trees of 164 steps were found. Their strict consensus (Fig. 4) deviates even more from the established consensus than the tree obtained from cranial data.

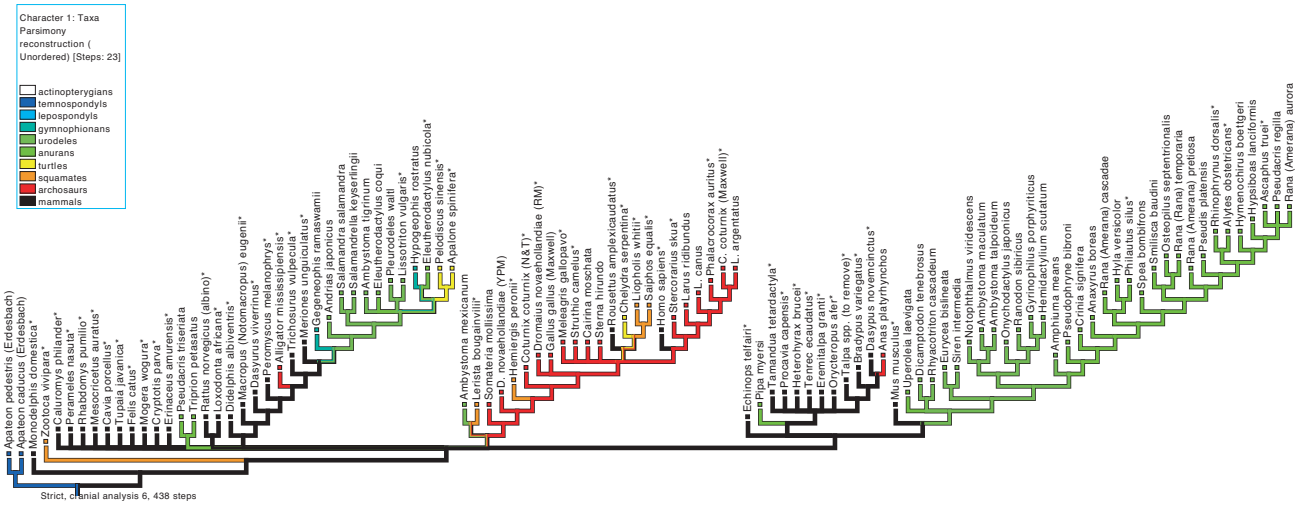



**Figure 3.** Strict consensus of the most parsimonious trees obtained by analyzing cranial dataset 2, which is comprised of 105 taxa and seven characters (see Table 1). Note that several higher taxa whose monophyly is well-established are para- or polyphyletic here.

This visual assessment of phylogenetic signal through an examination of the consensus trees (Figs. 3, 4) is congruent with the test based on squared-change parsimony and random taxon reshuffling (Laurin 2004). Indeed, the latter indicates that the phylogenetic signal in the cranial data is fairly strong, with a probability of less than 0.0001 that the observed covariation between the data and the tree reflects a random distribution (none of the 10,000 random trees generated were as short as the reference tree). However, it is weaker, with a probability of 0.0017, for the appendicular data.

The speciational model of evolution, in which all branch lengths are equal, has overwhelming support among cranial data, whether or not the Permian temnospondyl *Sclerocephalus* (Table 2) or the squamosal (Table 3) are included (including *Sclerocephalus* adds a second temnospondyl genus, but given that the timing of ossification of the squamosal is unknown in *Sclerocephalus*, including it requires excluding the squamosal from the analysis as described in the Methods section); the five other examined models have AICc weights < $10^{-11}$. For the appendicular data, the speciational model also has the most support, but that support is not as strong and varies depending on which dataset is analyzed (seven characters or four) and under which phylogenetic hypothesis. In three of the four tests performed, support for the second-best model, the non-phylogenetic/equal model, varied between 5% and 19% (Table 4).

Two main conclusions can be drawn from these tests (Tables 2–4). First, given that both of the best-supported models imply equal branch lengths, actual time represented by branches can be ignored, so we compare support of the six competing topologies using only the best-supported model (speciational). This simplifies the discussion, because it means that the original branch lengths are irrelevant (under that model, all branch lengths are equal); unfortunately, the branch length (evolutionary time) data were needed to reach this conclusion. Thus, the only remaining variable is the topology. Second, model fitting, along with the test based on squared-change parsimony and random taxon reshuffling, indicates that the phylogenetic signal in the cranial data is strong, but that it is noticeably weaker in the appendicular data (this is shown mostly by the non-negligible support for the non-phylogenetic/equal model). Thus, comparisons of the fit of the various phylogenetic hypotheses for the cranial data should be more reliable than for the appendicular data. However, given that for several Paleozoic taxa (most importantly both of the sampled lepospondyls), comparisons can be performed only for the appendicular data, these were performed as well.

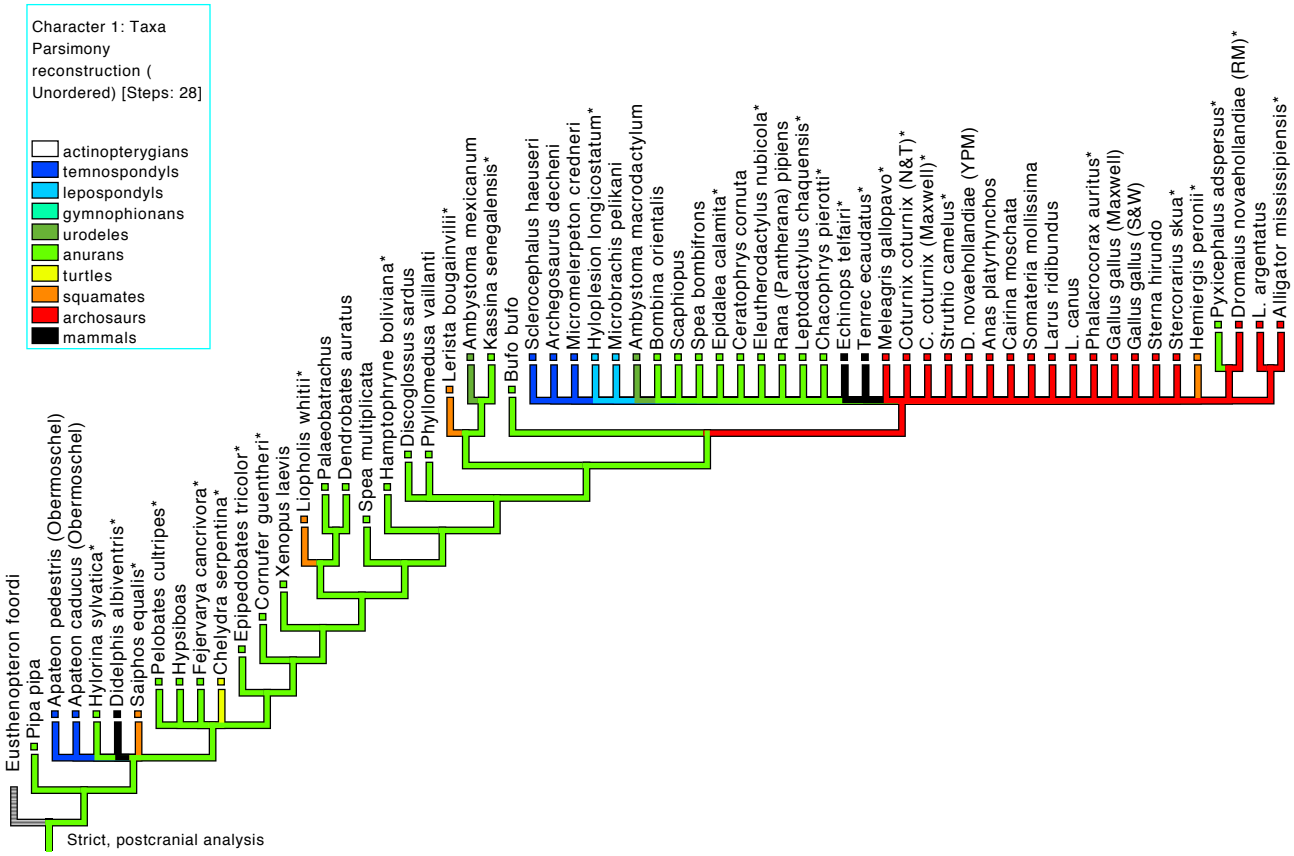



**Figure 4.** Strict consensus of the most parsimonious trees obtained by analyzing appendicular dataset 3, which is comprised of 62 taxa and seven characters (see Table 1). The phylogenetic signal in these data seems to be lower than in the cranial data.

Using the speciational model, the AICc weights of the six compared topologies indicate that there is strong support in the cranial data for the LH (lepospondyl hypothesis), with an AICc weight of 0.9885 when *Sclerocephalus* is included (Table 5) and 0.8848 when the squamosal is included instead (Table 6). Of the other topologies, the TH (temnospondyl hypothesis) was by far the best supported, with an AICc weight of 0.01144 (with *Sclerocephalus*) or 0.1056 (with the squamosal), which is 86.44 or 8.38 times less than for the LH. Both versions of the DH (diphyly hypothesis) and of the PH (polyphyly hypothesis) have negligible support (AICc weights < 0.01 when the squamosal is included, < 0.0001 when *Sclerocephalus* is included). The least support is found for the PH2 when *Sclerocephalus* is included, and for the DH1 when the squamosal is included. In both cases, the recently proposed DH2 (Pardo et al. 2017b) fares second-worst by a small margin. Notably, the DH1 contradicts the modern consensus on lissamphibian monophyly (Fig. 1g), while the PH2 and the DH2 fulfill this constraint from the molecular but not the paleontological point of view, having lissamphibian monophyly with respect to amniotes but not with respect to temnospondyls (Fig. 1f, h).

A slightly different dataset is used (only 84 taxa, but eight cranial characters – excluding the squamosal but including the frontal and the vomer – and *Apateon* sequences for both species from Erdesbach rather than Obermoschel) provides even stronger support for the LH, with an AICc weight of 0.9935 (Table 7). The next best-supported topology, which simultaneously represents the TH, DH1 and DH2 (due to the absence of caecilians from this dataset), has an AICc weight of only 0.0065.

**Table 2.** Support (AICc and AICc weights) for six evolutionary models given our reference tree (LH) and dataset 1 (see Table 1), which comprises six cranial characters (nasal, parietal, squamosal, maxilla, pterygoid, and exoccipital) scored in 107 taxa, including the temnospondyl *Sclerocephalus*. This was performed on the tree representing the LH (lepospondyl hypothesis), but doing this on other trees leads to similar results. Numbers presented with four significant digits; best values in boldface. "Distance" refers to keeping the original branch lengths (which represent evolutionary time), "equal" sets all branch lengths (internal and terminal) to 1, "free" infers them from the data. Abbreviations: k, number of estimable parameters; l, likelihood; wi, weight; $\Delta_i$, difference of AICc from that of the Pure-Phylogenetic / Equal model.

| Evolutionary model | AIC | l | k | AICc | $\Delta_i$ AICc | wi(AICc) |
|---|---|---|---|---|---|---|
| Pure-Phylogenetic / Distance | −584.4 | 293.2 | **1** | −583.4 | 641.2 | 5.85 E−140 |
| Pure-Phylogenetic / Equal (speciational) | **−1225.6** | **613.8** | **1** | **−1224.6** | **0** | **1.000** |
| Pure-Phylogenetic / Free | 2.000 E10 | −1.000 E10 | 486 | 2.000 E10 | 2.000 E10 | < E−165 |
| Non-Phylogenetic / Distance | −473.6 | 237.8 | **1** | −472.6 | 752.0 | 4.97 E−164 |
| Non-Phylogenetic / Equal | −959.9 | 481.0 | **1** | −958.9 | 265.7 | 2.02 E−58 |
| Non-Phylogenetic / Free | 2.000 E10 | −1.000 E10 | 244 | 2.000 E10 | 2.000 E10 | < E−165 |

**Table 3.** Support (AICc and AICc weights) for six evolutionary models given our reference tree (LH) and dataset 2 (see Table 1), which comprises seven cranial characters (nasal, parietal, squamosal, premaxilla, maxilla, pterygoid, and exoccipital) and 105 taxa, excluding *Sclerocephalus*. Abbreviations and boldface as in Table 2.

| Evolutionary model | AIC | L | k | AICc | Δi AICc | wi(AICc) |
|---|---|---|---|---|---|---|
| Pure-Phylogenetic / Distance | −715.9 | 359.0 | 1 | −714.9 | 683.5 | < E−26 |
| Pure-Phylogenetic / Equal | −1399.5 | 700.7 | 1 | −1398.5 | 0 | 1.000 |
| Pure-Phylogenetic / Free | 2.000 E10 | −1.000 E10 | 306 | 2.000 E10 | 2.000 E10 | 0 |
| Non-Phylogenetic / Distance | −580.6 | 291.3 | 1 | −579.6 | 818.8 | < E−26 |
| Non-Phylogenetic / Equal | −1106.0 | 554.0 | 1 | −1105.0 | 293.5 | 2.278 E−98 |
| Non-Phylogenetic / Free | 2.000 E10 | −1.000 E10 | 244 | 2.000 E10 | 2.000 E10 | < E−26 |

**Table 4**. AICc weights showing relative support for six evolutionary models given various appendicular datasets (3 and 4; see Table 1) and various hypotheses. Because of the number of analyses presented below, only the AICc weights are presented (best values in boldface). Abbreviations: DH, diphyly hypothesis (both versions); LH, lepospondyl hypothesis; TH, temnospondyl hypothesis.

| Evolutionary model | **7 characters, LH** | **7 characters, LH** | **4 characters, LH** | **4 characters, TH/DH** |
|---|---|---|---|---|
| Pure-Phylogenetic / Distance | 5.1857 E−149 | 2.340 E−70 | 1.227 E−52 | 2.646 E−52 |
| Pure-Phylogenetic / Equal | **1** | **0.9335** | **0.94459** | **0.8139** |
| Pure-Phylogenetic / Free | < E−179 | 1.598 E−277 | 4.012 E−158 | 3.002 E−155 |
| Non-Phylogenetic / Distance | 7.515 E−179 | 4.843 E−52 | 2.162 E-42 | 7.262 E−42 |
| Non-Phylogenetic / Equal | 2.14914 E−64 | 6.648 E−02 | 5.541 E−02 | 0.1861 |
| Non-Phylogenetic / Free | < E−179 | < E−179 | < E−179 | < E−179 |

The appendicular data are available in far more Paleozoic taxa than the cranial data; these include *Sclerocephalus haeuseri*, *Archegosaurus decheni*, and the non-branchiosaurid "branchiosaur" *Micromelerpeton credneri* among temnospondyls, the lepospondyls *Hyloplesion longicaudatum* and *Microbrachis pelikani*, and the tristichopterid finned stem-tetrapodomorph *Eusthenopteron foordi,* in addition to the same two species of *Apateon* as for the cranial datasets, *A. caducus* and *A. pedestris*. Analysis of these data (seven characters: humerus, radius, ulna, ilium, femur, tibia and fibula) yields surprising results, with the PH2 having the most support, with an AICc weight of 0.7978 when using the dataset of seven bones (Table 8). The TH, DH1 and DH2 with "branchiosaur" monophyly are collectively (they cannot be distinguished with that taxonomic sample) the second-best hypotheses with that dataset, with an AICc weight of only 0.1874. The least-supported hypothesis with these data is the TH with "branchiosaur" polyphyly.

Using the other postcranial dataset with only four bones (radius, ulna, ilium, and femur) but with more taxa (notably the branchiosaurid temnospondyl "*Melanerpeton*" *humbergense*) shows that intraspecific variation in the postcranial ossification sequences

of *Apateon* do not significantly impact our assessment of the support for various hypotheses. Whether both sequences of *Apateon* (from the Erdesbach and Obermoschel localities, which represent separate paleo-lakes) are included (treated as if they were distinct taxa, such as subspecies), or whether either one of these is used in isolation, the PH2 retains the highest support, with AICc weights of 0.62 to 0.65. The LH is a distant second, at 0.20–0.23, but still well ahead of the TH/DH and the PH1, which all receive AICc weights between 0.03 and 0.06 (Table 9).

**Table 5.** Support (AIC and AICc weights) for the six topologies, reflecting the six hypotheses about the origin of extant amphibians, under the speciational model (called Pure-Phylogenetic / Equal in Tables 2–4), with dataset 1 (see Table 1), which includes six cranial characters (nasal, parietal, squamosal, maxilla, pterygoid, and exoccipital) and 107 taxa (including, among Paleozoic taxa, *Apateon* and *Sclerocephalus*). Abbreviations and boldface as in Table 2, except $\Delta_i$: difference of AICc from that of the LH. Hypotheses from top to bottom: LH: monophyletic origin from lepospondyls; TH: monophyletic origin among temnospondyls; DH1: diphyletic origin, caecilians from lepospondyls and batrachians from temnospondyls, as in Anderson et al. (2008); DH2: diphyletic origin (batrachians and caecilians from different temnospondyls: Pardo et al. 2017b); PH1: triphyletic (polyphyletic) origin with anurans and urodeles from different temnospondyls, caecilians from lepospondyls, and lepospondyls closer to Amniota than to Batrachia (Fröbisch et al. 2007); PH2: triphyletic (polyphyletic) origin as above, but with lepospondyls and caecilians closer to temnospondyls than to amniotes (Milner 1993), reflecting the well-established lissamphibian monophyly among extant taxa (e.g. Irisarri et al. 2017; Feng et al. 2017).

| Hypothesis | AIC | L | AICc | $\Delta_i$ AICc | wi(AICc) |
|---|---|---|---|---|---|
| TH | −1217 | 609.4 | −1215 | 8.919 | 0.01144 |
| LH | **−1226** | **613.8** | **−1224** | **0** | **0.9885** |
| DH1 | −1204 | 602.9 | −1202 | 21.90 | 1.738 E−05 |
| DH2 | −1195 | 598.3 | −1193 | 31.01 | 1.827 E−07 |
| PH1 | −1194 | 597.9 | −1192 | 31.86 | 1.196 E−07 |
| PH2 | −1193 | 597.4 | −1191 | 32.89 | 7.143 E−08 |

**Table 6.** Support (AIC and AICc weights) for the six topologies, reflecting the six hypotheses about the origin of extant amphibians, for dataset 2 (see Table 1), which includes seven cranial characters (nasal, parietal, squamosal, premaxilla, maxilla, pterygoid, and exoccipital) and 105 taxa, excluding *Sclerocephalus* (among Paleozoic taxa, only *Apateon* is present). Abbreviations, boldface and hypotheses as in Tables 2 and 5.

| Hypothesis | AIC | L | AICc | $\Delta_i$ AICc | wi(AICc) |
|---|---|---|---|---|---|
| TH | −1395 | 698.6 | −1394 | 4.251 | 0.1056 |
| LH | **−1399** | **700.7** | **−1398** | **0** | **0.8848** |
| DH1 | −1384 | 693.1 | −1383 | 15.203 | 4.42 E−4 |
| DH2 | −1385 | 693.6 | −1384 | 14.315 | 6.89 E−4 |
| PH1 | −1387 | 694.5 | −1386 | 12.404 | 1.792 E−3 |
| PH2 | −1390 | 695.8 | −1388 | 9.792 | 6.615 E−3 |

**Table 7.** Support for the various hypotheses about amphibian origins for dataset 5 (see Table 1), which includes eight cranial characters (frontal added) and 84 taxa, with *Apateon* sequences from Erdesbach (in addition to *Sclerocephalus* among Paleozoic taxa). Abbreviations, boldface and hypotheses as in Tables 2 and 5. Because of the taxon sample, only three topologies can be tested.

| Hypothesis | AIC | L | AICc | $\Delta_i$ AICc | wi(AICc) |
|---|---|---|---|---|---|
| LH | **−1296** | **649.0** | **−1294** | **0** | **0.9935** |
| TH, DH1, DH2 | −1286 | 644.0 | −1284 | 10.061 | 6.493 E−3 |
| PH | −1274 | 638.0 | −1272 | 22.038 | 1.628 E−5 |

# DISCUSSION

## Phylogenetic signal

In his discussion of previous phylogenetic conclusions from ossification sequences (e.g. Schoch and Carroll 2003), Anderson (2007) noted that ossification sequences seemed to abound in symplesiomorphies and in autapomorphies of terminal taxa, while potential synapomorphies were scarce. This pessimism was seemingly confirmed by Schoch (2006) in a paper that was published after Anderson's (2007) book chapter had gone to press: not only were many similarities in the cranial ossification sequences

across Osteichthyes found to be symplesiomorphies, but a phylogenetic analysis of cranial ossification sequences did not recover Mammalia, Sauropsida, Amniota or Lissamphibia as monophyletic. Along with these results, Schoch (2006) dismissed another: the position of the temnospondyl *Apateon caducus* (the only included extinct taxon) outside the tetrapod crown-group, i.e. the lepospondyl hypothesis on lissamphibian origins (LH).

While ossification sequences alone may not provide enough data for a phylogenetic analysis, as shown by our results (Fig. 3, 4), there is clearly a phylogenetic signal because the taxa are not randomly scattered over the tree. Specifically, our datasets (with much larger taxon samples than in Schoch 2006) fit some tree topologies much better than others. Both the tests using CoMET and squared-change parsimony with random taxon reshuffling overwhelmingly support the presence of a strong phylogenetic signal in the cranial data; the null hypothesis of the absence of a phylogenetic signal can be rejected in both cases, given that it has a probability of $< 10^{-97}$ for the cranial and $< 10^{-4}$ for the appendicular dataset. We conclude that the cranial dataset contains a strong phylogenetic signal, and are therefore cautiously optimistic about future contributions of ossification sequences to phylogenetics. We are less optimistic about the appendicular sequence data, which both tests suggest contains less phylogenetic signal.

The sizable effect on nodal estimates and inferred heterochronies of intraspecific variation found by Sheil et al. (2014) in lissamphibians could raise doubts about the robustness of our findings. We have been able to incorporate infraspecific variability in only two terminal taxa (*Apateon caducus* and *A. pedestris*), but *Apateon* has played a prominent role in discussions about the significance of cranial ossification sequences on the origins of extant amphibians (Schoch and Carroll 2003; Schoch 2006; Germain and Laurin 2009). Thus, incorporation of intraspecific variability in *Apateon* is presumably much more important than in extant taxa, even though variability in the latter would obviously add to the analysis and should be tackled in the future. The variability in *Apateon* should be exempt from two sources of artefactual variability in ossification sequences discussed by Sheil et al. (2014), namely the way in which the specimens were collected (there can be no lab-raised specimens in long-extinct taxa) and the fixing method used (in this case, fossilization under quite consistent taphonomic conditions). The finding that the results are very similar no matter whether we used the *Apateon* sequences from Erdesbach, Obermoschel, or both, we find very similar results (Table 9), is reassuring. In this case, intraspecific variation has negligible impact. However, future studies should attempt to assess the effect of more generalized incorporation of infraspecific variability (in a greater proportion of the OTUs).

Of course, these results do not preclude functional or developmental constraints from applying to the same data. This phenomenon has been documented, among other taxa, in urodeles, whose development has often been compared with that of temnospondyls (e.g. Schoch 2006; Schoch and Carroll 2003; Fröbisch et al. 2007, 2015; Germain and Laurin 2009). For instance, Vorobyeva and Hinchliffe (1996) documented the larval functional constraints linked to early forelimb use that may cause an early development of manual digits 1 and 2, compared with other tetrapods, as briefly discussed below. However, in the case of our seven cranial characters, there is no evidence of functional constraints. This is a little-investigated topic, but all these bones apparently form a single developmental module of the urodele skull (Laurin 2014). For the appendicular data, functional constraints might explain the more subdued phylogenetic signal, but this will have to be determined by additional research.

The finding that the postcranial characters that we analyzed contain relatively little phylogenetic signal may raise doubts about the claims that have been made about the phylogenetic implications of other such data. Specifically, Carroll et al. (1999) stated that the neural arches ossify before the centra in frogs and temnospondyls, but not in salamanders, caecilians or lepospondyls. When it was found that the centra do ossify first in a few cryptobranchoid salamanders, Carroll (2007: 30) took this as “strong evidence that the most primitive crown-group salamanders had a sequence of vertebral development that is common to frogs and labyrinthodonts [including temnospondyls] (but distinct from that of lepospondyls)”. In fact, apart from tail regeneration in *Hyloplesion* and *Microbrachis* (where the centra ossify before the neural arches: Olori 2015; Fröbisch et al. 2015; van der Vos et al. 2017), only one incompletely ossified vertebral column (referred to *Utaherpeton*) is known of any putative lepospondyl. “In this specimen, […] five neural arches […] have ossified behind the most posterior centrum.” (Carroll and Chorn 1995: 40–41) Carroll’s (2007: 85) claim that “the centra always ossified prior to the arches” in lepospondyls is therefore rather puzzling.

Fröbisch et al. (2007, 2015) pointed out that the first two digital rays (digits, metapodials and distal carpals/tarsals) ossify before the others (“preaxial polarity”) in salamanders and the temnospondyls *Apateon*, *Micromelerpeton* and *Sclerocephalus*, while

the fourth ossifies first ("postaxial polarity") in amniotes, frogs and "probably" (Fröbisch et al. 2015: 233, 234) the lepospondyls *Microbrachis* and *Hyloplesion*. This latter inference, however, is based only on a delay in the ossification of the fifth ray that is shared specifically with sauropsid amniotes (Olori 2015). Ossification sequences (however partial) of the other four rays in any lepospondyl are currently limited to the tarsus of *Batropetes*, which clearly shows preaxial polarity (Glienke 2015: fig. 6O–S; Marjanović and Laurin 2019), and that of the putative (but see Clack et al. 2019) lepospondyl *Sauropleura*, in which likewise the second distal tarsal ossified before all others (Marjanović and Laurin 2019). Outside of temno- and lepospondyls, Marjanović and Laurin (2013, 2019) presented evidence that preaxial polarity is plesiomorphic, widespread and dependent on the use of the still developing limbs for locomotion, which would explain why it was independently lost in amniotes and frogs and reduced (the second ray still forms first, but the delays between the rays are much reduced so that all form nearly at the same time) in direct-developing salamanders as well as in th elimb regeneration of terrestrial postmetamorphic salamanders (Kumar et al. 2015). It may be relevant here that the PH2 (Fig. 1f), favored by our appendicular data, groups exactly those sampled taxa in a clade that are known to have preaxial polarity in limb development. To sum up, neither our own analyses nor the previous works that we cited above demonstrated conclusively that ossification sequences of postcranial elements provide reliable clues about the origin of extant amphibians.

**Table 8.** Support (AICc weights) for the various hypotheses about amphibian origins according to dataset 3 (see Table 1), which features seven appendicular characters (humerus, radius, ulna, ilium, femur, tibia and fibula) and 62 taxa, including several Paleozoic taxa (the temnospondyls *Archegosaurus decheni* and *Micromelerpeton credneri*, the lepospondyls *Hyloplesion longicaudatum* and *Microbrachis pelikani*, and the tristichopterid *Eusthenopteron foordi*) in addition to *Apateon* (two species, *A. caducus* and *A. pedestris*) and *Sclerocephalus haeuseri*. The *Apateon* sequences come from Obermoschel. Abbreviations, boldface and hypotheses as in Table 5, except that the TH and both variants of the DH become indistinguishable, but the phylogenetic position of the "branchiosaur" *Micromelerpeton* can be tested.

| Hypothesis | AIC | l | AICc | $\Delta_i$ AICc | wi(AICc) |
|---|---|---|---|---|---|
| LH | −885.0 | 443.5 | −884.2 | 11.808 | 2.177 E−3 |
| TH, DH (branchiosaur monophyly) | −881.1 | 441.6 | −880.3 | 2.897 | 0.1874 |
| TH, DH (branchiosaur polyphyly) | −886.4 | 444.2 | −885.6 | 15.754 | 3.027 E−4 |
| PH1 | −888.5 | 445.3 | −887.7 | 8.341 | 0.01232 |
| PH2 | **−896.9** | **449.4** | **−896.1** | **0.000** | **0.7978** |

**Table 9.** Effect of the intraspecific variability in ossification sequences of *Apateon* on the support (AICc weight; best values in boldface) for the various hypotheses about amphibian origins. The dataset (number 4; Table 1) includes only four appendicular bones (radius, ulna, ilium, and femur) and 63 to 65 taxa but it allows testing the impact of intraspecific variability in ossification sequences in *Apateon*, which are documented in two localities (Erdesbach and Obermoschel). Because of the number of tests presented (15: five topologies x three sets of sequences), only the AICc weights are given. In all tests, the following Paleozoic taxa are present: *Sclerocephalus haeuseri*, *Archegosaurus decheni*, "*Melanerpeton*" *humbergense*, *Micromelerpeton credneri*, *Apateon* (two species, *A. caducus* and *A. pedestris*) among temnospondyls, *Hyloplesion longicaudatum* and *Microbrachis pelikani* among lepospondyls, and the tristichopterid *Eusthenopteron foordi*. For abbreviations of the hypotheses, see Table 5.

| Hypothesis | Erdesbach and Obermoschel | Erdesbach | Obermoschel |
|---|---|---|---|
| LH | 0.21407 | 0.20169 | 0.22657 |
| TH, DH (branchiosaur monophyly) | 0.05492 | 0.05265 | 0.05532 |
| TH, DH (branchiosaur polyphyly) | 0.03713 | 0.04285 | 0.03342 |
| PH1 | 0.05653 | 0.05491 | 0.05638 |
| PH2 | **0.63735** | **0.64790** | **0.62832** |

In contrast, we are reasonably confident about our results on the cranial ossification sequences. Given the phylogenetic signal we have found in our cranial datasets, we think that ossification sequence data should eventually be added to phenotypic datasets for analyses of tetrapod phylogeny. Indeed, an analysis of amniote phylogeny using data from organogenesis sequences (coded using event-pairing in Parsimov) already exists (Werneburg and Sánchez-Villa-

gra 2009). The usefulness of such data for phylogenetic inference was further tested, with encouraging results, by Laurin and Germain (2011), and the present analysis adds additional support for it.

### Indirect support for the lepospondyl hypothesis from temnospondyls

The strong support for the lepospondyl hypothesis that we have found in cranial data is surprising because cranial ossification sequence data, especially those of the Permo-Carboniferous temnospondyl *Apateon*, have often been claimed to contradict the LH (lepospondyl hypothesis, Fig. 1d). Similarities between *Apateon* and extant urodeles, in particular the supposedly "primitive" hynobiid *Ranodon*, have often been emphasized (Schoch and Carroll 2003; Schoch and Milner 2004; Carroll 2007; Schoch 2014b). However, other studies have already raised doubts about some of these claims (e.g. Schoch 2006; Anderson 2007; Germain and Laurin 2009). Schoch (2006) and Anderson (2007) concluded that most characters shared between *Apateon* and urodeles were plesiomorphies. Germain and Laurin (2009) also demonstrated that, far from being very similar to the ancestral urodele morphotype (contra Schoch and Carroll 2003 or Carroll 2007), the cranial ossification sequence of *Apateon* was statistically significantly different from that of the hypothetical last common ancestor of all urodeles (as suspected by Anderson 2007). However, these earlier studies did not clearly show which of the various hypotheses on lissamphibian origins the ossification sequences of *Apateon* spp. – or the newly available partial sequence (Werneburg 2018) of the phylogenetically distant temnospondyl *Sclerocephalus* – supported most. This is what we have attempted to do here.

Unfortunately, the development of lepospondyls is too poorly documented to be incorporated into the cranial analyses, but we included two lepospondyls in analyses of appendicular data. These analyses weakly favor a polyphyletic origin of extant amphibians, with both temno- and lepospondyls in the amphibian clade, a hypothesis that has not been advocated seriously for decades (Milner 1993: fig. 5B) as far as we know. However, given the moderate phylogenetic signal in these data, we view these results with skepticism. Olori (2011), using event-pairing with Parsimov (Jeffery et al. 2005) and PGi (Harrison and Larsson 2008), analyzed lepospondyl postcranial ossification sequences and concluded that support for the three hypotheses that she tested (TH/DH with two different positions for *Micromelerpeton*, and LH) did not differ significantly. By contrast, our analyses of the postcranial data indicate a stronger support for polyphyly (PH2) than for the TH/DH, which is only a distant second (Table 8) or third (behind PH2 and LH; Table 9) depending on the analyses. Olori (2011) performed no statistical test of phylogenetic signal of her data, though a related test (performing phylogenetic analyses on the data) yielded trees (Olori, 2011: fig. 5.5–5.7) that are largely incongruent with the established consensus, in which most large taxa (Mammalia, Testudines, Lissamphibia, etc.) are para- or polyphyletic. Olori's (2011) results, like ours, support the conclusion that the phylogenetic signal in postcranial ossification sequence data is low.

Given the current limitations in the availability of developmental data in Paleozoic stegocephalians, we hope to have demonstrated that cranial ossification sequences of amniotes, lissamphibians and temnospondyls provide support for the LH that is independent of the phylogenetic analyses of Laurin (1998), Pawley (2006: appendix 16) or Marjanović and Laurin (2009, 2019). This independence is important because the cranial ossification sequence data cannot rival the morphological data in terms of data availability, simply because growth sequences of extinct taxa are rare (Sánchez-Villagra 2012), but having a fairly independent line of evidence to investigate a major evolutionary problem is clearly advantageous. We hope that the modest methodological progress made in this study will stimulate the search for fossilized ontogenies (Cloutier 2009; Sánchez-Villagra 2010).

## ACKNOWLEDGEMENTS

Jennifer Olori, two anonymous reviewers and the editor Robert Asher made helpful comments that improved this paper. D. M. would further like to thank Ralf Werneburg for an electronic reprint of his 2018 paper, Nadia Fröbisch for discussion of limb development in salamanders, and Daniel Field for discussion of molecular divergence times and the fossil record.

## ADDITIONAL INFORMATION

### Funding

This work was supported by the Centre National de la Recherche Scientifique and the French Ministry of Research (unnumbered recurring grants to the CR2P, for ML).

### Competing interests

None that we are aware of.

### Author contributions

ML designed the study, supervised the data collection, analyzed the data and wrote much of the draft; OL collected most of the ossification sequence

data; DM added data to our database (mostly of Paleozoic taxa), updated the timetrees, participated in the writing and drafted Figure 1.

### Data availability

All data used in this study can be downloaded from the same bioRχiv page as this preprint.

### Supplementary information

Data matrices and trees used in this analysis can be downloaded in NEXUS format for Mesquite from: https://www.biorxiv.org/content/10.1101/352609v3.supplementary-material

## REFERENCES


Anderson J.S. 2007. Incorporating ontogeny into the matrix: a phylogenetic evaluation of developmental evidence for the origin of modern amphibians. Pages 182–227 in Anderson J.S., Sues H.-D., editors. Major transitions in vertebrate evolution. Bloomington: Indiana University Press.

Anderson D.R., Burnham K.P. 2002. Avoiding pitfalls when using information-theoretic methods. J. Wildl. Manag. 66:912–918.

Anderson J.S., Carroll R.L., Rowe T.B. 2003. New information on *Lethiscus stocki* (Tetrapoda: Lepospondyli: Aistopoda) from high-resolution computed tomography and a phylogenetic analysis of Aistopoda. Can. J. Earth Sci. 40:1071–1083.

Anderson J.S., Reisz R.R., Scott D., Fröbisch N.B., Sumida S.S. 2008. A stem batrachian from the Early Permian of Texas and the origin of frogs and salamanders. Nature 453:515–518.

Berv J.S., Field D.J. 2017 (printed 2018). Genomic signature of an avian Lilliput Effect across the K-Pg Extinction. Syst. Biol. 67:1–13.

Bokma F., Godinot M., Maridet O., Ladevèze S., Costeur L.c., Solé F., Gheerbrant E., Peigné S., Jacques F., Laurin M. 2016. Testing for Depéret's Rule (body size increase) in mammals using combined extinct and extant data. Syst. Biol. 65:98–108.

Bolt J.R. 1969. Lissamphibian origins: possible protolissamphibian from the Lower Permian of Oklahoma. Science 166:888–891.

Bossuyt F., Roelants K. 2009. Frogs and toads (Anura). Pages 357–364 in Hedges S.B., Kumar S., editors. The Timetree of Life. New York: Oxford University Press.

Brandley M.C., Schmitz A., Reeder T.W. 2005. Partitioned Bayesian analyses, partition choice, and the phylogenetic relationships of scincid lizards. Syst. Biol. 54:373–390.

Carpenter D.K., Falcon-Lang H.J., Benton M.J., Grey M. 2015. Early Pennsylvanian (Langsettian) fish assemblages from the Joggins Formation, Canada, and their implications for palaeoecology and palaeogeography. Palaeontology 58:661–690.

Carroll R.L. 2001. The origin and early radiation of terrestrial vertebrates. J. Paleontol. 75:1202–1213.

Carroll R.L. 2007. The Palaeozoic ancestry of salamanders, frogs and caecilians. Zool. J. Linn. Soc. 150 (suppl. 1):1–140.

Carroll R.L., Chorn J. 1995. Vertebral development in the oldest microsaur and the problem of "lepospondyl" relationships. J. Vert. Paleont. 15:37–56.

Carroll R.L., Currie P.J. 1975. Microsaurs as possible apodan ancestors. Zool. J. Linn. Soc. 57: 229–247.

Carroll R.L., Holmes R. 1980. The skull and jaw musculature as guides to the ancestry of salamanders. Zool. J. Linn. Soc. 68:1–40.

Carroll R.L., Kuntz A., Albright K. 1999. Vertebral development and amphibian evolution. Evol. Dev. 1:36–48.

Cope E.D. 1888. On the intercentrum of the terrestrial Vertebrata. Trans. Am. Phil. Soc. 16:243–253.

Cloutier R. 2009 (printed 2010). The fossil record of fish ontogenies: Insights into developmental patterns and processes. Semin. Cell Dev. Biol. 21:400–413.

Criswell K.E. 2015. The comparative osteology and phylogenetic relationships of African and South American lungfishes (Sarcopterygii: Dipnoi). Zool. J. Linn. Soc. 174:801-858.

Danto M., Witzmann F., Kamenz S., Fröbisch N. 2019. How informative is vertebral development for the origin of lissamphibians? J. Zool. (Lond.) 307: 292–305.

Davies T.W., Bell M.A., Goswami A., Halliday T.J.D. 2017. Completeness of the eutherian mammal fossil record and implications for reconstructing mammal evolution through the Cretaceous/Paleogene mass extinction. Paleobiology 43:521–536.

Eldredge N., Gould S.J. 1972. Punctuated equilibria: an alternative to phyletic gradualism. Pages 82–115 in Schopf T.J.M., editor. Models in Paleobiology. San Francisco: Freeman, Cooper & Company.

Feng Y.-J., Blackburn D.C., Liang D., Hillis D.M., Wake D.B., Cannatella D.C., Zhang P. 2017. Phylogenomics reveals rapid, simultaneous diversification of three major clades of Gondwanan frogs at the Cretaceous–Paleogene boundary. Proc. Natl. Acad. Sci. USA E5864–E5870.

Fröbisch N.B., Carroll R.L., Schoch R.R. 2007. Limb ossification in the Paleozoic branchiosaurid *Apateon* (Temnospondyli) and the early evolution of preaxial dominance in tetrapod limb development. Evol. Dev. 9:69–75.

Fröbisch N.B., Bickelmann C., Olori J.C., Witzmann F. 2015. Deep-time evolution of regeneration and preaxial polarity in tetrapod limb development. Nature 527:231–234.

Germain D., Laurin M. 2009. Evolution of ossification sequences in salamanders and urodele origins assessed through event-pairing and new methods. Evol. Dev. 11:170–190.

Glienke S. 2015. Two new species of the genus *Batropetes* (Tetrapoda, Lepospondyli) from the Central European Rotliegend (basal Permian) in Germany. J. Vert. Paleont. 35:e918041.

Gonzalez J., Düttmann H., Wink M. 2009. Phylogenetic relationships based on two mitochondrial genes and hybridization patterns in Anatidae. J. Zool. 279:310–318.

Halliday T.J.D., Upchurch P., Goswami A. 2015 (printed 2017). Resolving the relationships of Paleocene placental mammals. Biol. Rev. 92:551–550.

Halliday T.J.D., Upchurch P., Goswami A. 2016. Eutherians experienced elevated evolutionary rates in the immediate aftermath of the Cretaceous–Palaeogene mass extinction. Proc. R. Soc. B 283: 20153026.

Harrington S.M., Harrison L.B., Sheil C.A. 2013. Ossification sequence heterochrony among amphibians. Evol. Dev. 15:344–364.

Harrison L.B., Larsson H.C.E. 2008. Estimating evolution of temporal sequence changes: a practical approach to inferring ancestral developmental sequences and sequence heterochrony. Syst. Biol. 57:378–387.

Hugall A.F., Foster R., Lee M.S.Y. 2007. Calibration choice, rate smoothing, and the pattern of tetrapod diversification according to the long nuclear gene RAG-1. Syst. Biol. 56:543–563.

Hugi J., Hutchinson M.N., Koyabu D., Sánchez-Villagra M.R. 2012. Heterochronic shifts in the ossification sequences of surface- and subsurface-dwelling skinks are correlated with the degree of limb reduction. Zoology 115:188–198.

Irisarri I., Baurain D., Brinkmann H., Delsuc F., Sire J.-Y., Kupfer A., Petersen J., Jarek M., Meyer A., Vences M., Philippe H. 2017. Phylotranscriptomic consolidation of the jawed vertebrate timetree. Nature Ecol. Evol. 1:1370–1378.

Jeffery J.E., Bininda-Emonds O.R.P., Coates M.I., Richardson M.K. 2005. A new technique for identifying sequence heterochrony. Syst. Biol. 54:230–240.

Jetz W., Pyron R.A. 2018. The interplay of past diversification and evolutionary isolation with present imperilment across the amphibian tree of life. Nat. Ecol. Evol. 2:850–858.

Josse S., Moreau T., Laurin M. 2006. Stratigraphic tools for Mesquite, version 1.0. http://mesquiteproject.org/packages/stratigraphicTools/

Koyabu D., Werneburg I., Morimoto N., Zollikofer C.P.E., Forasiepi A.M., Endo H., Kimura J., Ohdachi S.D., Son N.T., Sánchez-Villagra M.R. 2014. Mammalian skull heterochrony reveals modular evolution and a link between cranial development and brain size. Nat. Commun. 5:3625.

Kumar A., Gates P.B., Czarkwiani A., Brockes JP. 2015. An orphan gene is necessary for preaxial digit formation during salamander limb development. Nat. Commun. 6:8684.

Laurin M. 1998. The importance of global parsimony and historical bias in understanding tetrapod evolution. Part I. Systematics, middle ear evolution, and jaw suspension. Ann. Sci. Nat., Zool., 13e Sér. 19:1–42.

Laurin M. 2004. The evolution of body size, Cope's rule and the origin of amniotes. Syst. Biol. 53:594–622.

Laurin M. 2014. Assessment of modularity in the urodele skull: an exploratory analysis using ossification sequence data. J. Exp. Zool. B (Mol. Dev. Evol.) 322:567–585.

Laurin M., Germain D. 2011. Developmental characters in phylogenetic inference and their absolute timing information. Syst. Biol. 60:630–644.

Laurin M., Piñeiro G. 2017. A reassessment of the taxonomic position of mesosaurs, and a surprising phylogeny of early amniotes. Front. Earth Sci. 5:88.

Lecompte E., Aplin K., Denys C., Catzeflis F., Chades M., Chevret P. 2008. Phylogeny and biogeography of African Murinae based on mitochondrial and nuclear gene sequences, with a new tribal classification of the subfamily. BMC Evol. Biol. 8:199.

Lee C., Blay S., Mooers A.Ø., Singh A., Oakley T.H. 2006. CoMET: A Mesquite package for comparing models of continuous character evolution on phylogenies. Evol. Bioinformatics Online 2:193–196. https://labs.eemb.ucsb.edu/oakley/todd/software/comet

Lu T., Zhu M., Yi C., Si C., Yang C., Chen H. 2017. Complete mitochondrial genome of the gray red-backed vole (*Myodes rufocanus*) and a complete estimate of the phylogenetic relationships in Cricetidae. Mitochondrial DNA Part A 28:62-64.

Maddin H.C., Jenkins F.A. Jr., Anderson J.S. 2012. The braincase of *Eocaecilia micropodia* (Lissamphibia, Gymnophiona) and the origin of caecilians. PLOS ONE 7:e50743.

Maddison W.P. 1991. Squared-change parsimony reconstructions of ancestral states for continuous-valued characters on a phylogenetic tree. Syst. Zool. 40:304–314.

Maddison W.P., Maddison D.R. 2018. Mesquite: a modular system for evolutionary analysis. Version 3.6. http://mesquite.wikispaces.com

Mann A., Pardo J.D., Maddin H.C. 2019. *Infernovenator steenae*, a new serpentine recumbirostran from the 'Mazon Creek' *Lagertätte* [sic] further clarifies lysorophian origins. Zool. J. Linn. Soc. online early (12 pp.).

Marjanović D., Laurin M. 2007. Fossils, molecules, divergence times, and the origin of lissamphibians. Syst. Biol. 56:369–388.

Marjanović D., Laurin M. 2008. Assessing confidence intervals for stratigraphic ranges of higher taxa: The case of Lissamphibia. Acta Palaeont. Pol. 53:413–432.

Marjanović D., Laurin M. 2009. The origin(s) of modern amphibians: a commentary. Evol. Biol. 36: 336–338.

Marjanović D., Laurin M. 2013a. The origin(s) of extant amphibians: a review with emphasis on the "lepospondyl hypothesis". Geodiversitas 35:207–272.

Marjanović D., Laurin M. 2013b (printed 2014). An updated paleontological timetree of lissamphibians, with comments on the anatomy of Jurassic crown-group salamanders (Urodela). Hist. Biol. 26:535–550.

Marjanović D., Laurin M. 2019. Phylogeny of Paleozoic limbed vertebrates reassessed through revision and expansion of the largest published relevant data matrix. PeerJ 6:e5565.

Maxwell E.E., Harrison L.B., Larsson H.C.E. 2010. Assessing the phylogenetic utility of sequence heterochrony: evolution of avian ossification sequences as a case study. Zoology 113:57–66.

Meredith R.W., Janečka J.E., Gatesy J., Ryder O.A., Fisher C.A., Teeling E.C., Goodbla A., Eizirik E., Simão T.L.L., Stadler T., Rabosky D.L., Honeycutt R.L., Flynn J.J., Ingram C.M., Steiner C., Williams T.L., Robinson T.J., Burk-Herrick A., Westerman M., Ayoub N.A., Springer M.S., Murphy W.J. 2011. Impacts of the Cretaceous terrestrial revolution and KPg extinction on mammal diversification. Science 334:521–524.

Milner A.R. 1993. The Paleozoic relatives of lissamphibians. Herpetol. Monogr. 7:8–27.

Ogg J.G., Ogg G., Gradstein F.M. 2016. A concise geologic time scale: 2016. Elsevier, Amsterdam.

Olori J.C. 2011. The evolution of skeletal development in early tetrapods: anatomy and ontogeny of microsaurs (Lepospondyli) [doctoral thesis]. Austin: University of Texas at Austin. http://hdl.handle.net/2152/ETD-UT-2011-05-3535

Olori J.C. 2013. Ontogenetic sequence reconstruction and sequence polymorphism in extinct taxa: an example using early tetrapods (Tetrapoda: Lepospondyli). Paleobiology 39:400–428.

Olori J.C. 2015. Skeletal morphogenesis of *Microbrachis* and *Hyloplesion* (Tetrapoda: Lepospondyli), and implications for the developmental patterns of extinct, early tetrapods. PLOS ONE 10:e0128333.

Pardo J.D., Szostakiwskyj M., Ahlberg P.E., Anderson J.S. 2017a. Hidden morphological diversity among early tetrapods. Nature 546:642–645.

Pardo J.D., Small B.J., Huttenlocker A.K. 2017b. Stem caecilian from the Triassic of Colorado sheds light on the origins of Lissamphibia. Proc. Natl. Acad. Sci. U.S.A. 114:E5389–E5395.

Parsons T.S., Williams E.E. 1962. The teeth of Amphibia and their relation to amphibian phylogeny. J. Morph. 110:375–389.

Parsons T.S., Williams E.E. 1963. The relationships of the modern Amphibia: A re-examination. Q. Rev. Biol. 38:26–53.

Pawley K. 2006. The postcranial skeleton of temnospondyls (Tetrapoda: Temnospondyli) [doctoral thesis]. Melbourne: La Trobe University. http://hdl.handle.net/1959.9/405644

Poe S. 2006. Test of von Baer's law of the conservation of early development. Evolution 60: 2239–2245.

Pons J.-M., Hassanin A., Crochet P.-A. 2005. Phylogenetic relationships within the Laridae (Charadriiformes: Aves) inferred from mitochondrial markers. Mol. Phyl. Evol. 37:686–699.

Prum R.O., Berv J.S., Dornburg A., Field D.J., Townsend J.P., Lemmon A.M., Lemmon A.R. 2015. A comprehensive phylogeny of birds (Aves) using targeted next-generation DNA sequencing. Nature 526:569–573.

Pyron R.A. 2014. Biogeographic analysis reveals ancient continental vicariance and recent oceanic dispersal in amphibians. Syst. Biol. 63:779–797.

Rabosky D.L., Donnellan S.C., Grundler M., Lovette I.J. 2014. Analysis and visualization of complex macroevolutionary dynamics: an example from Australian scincid lizards. Syst. Biol. 63:610–627.

Reeder T.W. 2003. A phylogeny of the Australian *Sphenomorphus* group (Scincidae: Squamata) and the phylogenetic placement of the crocodile skinks (*Tribolonotus*): Bayesian approaches to assessing congruence and obtaining confidence in maximum likelihood inferred relationships. Mol. Phyl. Evol. 27: 384–397.

Rineau V., Grand A., Zaragüeta R., Laurin M. 2015. Experimental systematics: sensitivity of cladistic methods to polarization and character ordering schemes. Contr. Zool. 84:129–148.

Rineau V., Zaragüeta i Bagils R., Laurin M. 2018. Impact of errors on cladistic inference: simulation-based comparison between parsimony and three-taxon analysis. Contr. Zool. 87:25–40.

Ruta M., Coates M.I. 2007. Dates, nodes and character conflict: addressing the lissamphibian origin problem. J. Syst. Palaeontol. 5:69–122.

Sánchez M. 2012. Embryos in Deep Time: The Rock Record of Biological Development. U. of California Press, Berkeley.

Sánchez-Villagra M.R. 2010. Contributions on fossilised ontogenies: The rock record of vertebrate development. Semin. Cell Dev. Biol. 21:399.

Schneider J.W., Werneburg R., Rößler R., Voigt S., Scholze F. 2015. Example for the description of basins in the CPT Nonmarine-Marine Correlation Chart – Thuringian Forest Basin, East Germany. Permophiles 61:29–35.

Schoch R.R. 2002. The formation of the skull in Paleozoic and extant amphibians. Paleobiology 28: 378–396.

Schoch R.R. 2004. Skeleton formation in the Branchiosauridae: a case study in comparing ontogenetic trajectories. J. Vert. Paleont. 24:309–319.

Schoch R.R. 2006. Skull ontogeny: developmental patterns of fishes conserved across major tetrapod clades. Evol. Dev. 8:524–536.

Schoch R.R. 2014a. First evidence of the branchiosaurid temnospondyl *Leptorophus* in the Early Permian of the Saar-Nahe Basin (SW Germany). N. Jb. Geol. Paläont. Abh. 272:225–236.

Schoch R.R. 2014b. Amphibian skull evolution: the developmental and functional context of simplification, bone loss and heterotopy. J. Exp. Zool. (Mol. Dev. Evol.) 322B:619–630.

Schoch R.R. 2018 (printed 2019). The putative lissamphibian stem-group: phylogeny and evolution of the dissorophoid temnospondyls. J. Paleont. 93: 37–156.

Schoch R.R., Carroll R.L. 2003. Ontogenetic evidence for the Paleozoic ancestry of salamanders. Evol. Dev. 5:314–324.

Schoch R.R., Milner A.R. 2004. Structure and implications of theories on the origin of lissamphibians. Pages 345–377 in Arratia G., Wilson M.V.H., Cloutier R., editors. Recent Advances in the Origin and Early Radiation of Vertebrates. Munich: Dr. Friedrich Pfeil.

Schulmeister S., Wheeler W.C. 2004. Comparative and phylogenetic analysis of developmental sequences. Evol. Dev. 6:50–57.

Sheil C.A., Jorgensen M., Tulenko F., Harrington S. 2014. Variation in timing of ossification affects inferred heterochrony of cranial bones in Lissamphibia. Evol. Dev. 16:292–305.

Sigurdsen T., Green D.M. 2011. The origin of modern amphibians: a re-evaluation. Zool. J. Linn. Soc. 162:457–469.

Skawiński T., Borczyk B. 2017. Evolution of developmental sequences in lepidosaurs. PeerJ 5:e3262.

Spiekman S.N., Werneburg I. 2017. Patterns in the bony skull development of marsupials: high variation in onset of ossification and conserved regions of bone contact. Sci. Rep. 7:43197.

Steen M.C. 1938. On the fossil Amphibia from the Gas Coal of Nýřany and other deposits in Czechoslovakia. Proc. Zool. Soc. Lond. 108:205–283.

Sterli J., Pol D., Laurin M. 2013. Incorporating phylogenetic uncertainty on phylogeny-based paleontological dating and the timing of turtle diversification. Cladistics 29:233–246.

Sterli J., de la Fuente M.S., Rougier G.W. 2018. New remains of *Condorchelys antiqua* (Testudinata) from the Early-Middle Jurassic of Patagonia: anatomy, phylogeny, and paedomorphosis in the early evolution of turtles. J. Vert. Paleont. 38:e1480112.

Swofford D.L. 2019. PAUP*: Phylogenetic Analysis Using Parsimony (*and other methods). Version 4.0a165. [Expired; the current version 4.0a166, which will expire on 1 February 2020, can be downloaded from http://phylosolutions.com/paup-test/.]

Tarver J.E., dos Reis M., Mirarab S., Moran R.J., Parker S., O'Reilly J.E., King B.L., O'Connell M.J., Asher R.J., Warnow T., Peterson K.J., Donoghue P.C.J., Pisani D. 2016. The interrelationships of placental mammals and the limits of phylogenetic inference. Genome Biol. Evol. 8:330–344.

van der Vos W., Witzmann F., Fröbisch N.B. 2017. Tail regeneration in the Paleozoic tetrapod *Microbrachis pelikani* and comparison with extant salamanders and squamates. J. Zool. 304:34–44.

Vorobyeva E.I., Hinchliffe J.R. 1996. Developmental pattern and morphology of *Salamendrella keyserlingii* limbs (Amphibia, Hynobiidae) including some evolutionary aspects. Russ. J. Herpetol. 1:68–81.

Wagenmakers E.-J., Farrell S. 2004. AIC model selection using Akaike weights. Psychon. Bull. Rev. 11:192–196.

Wang N., Kimball R.T., Braun E.L., Liang B., Zhang Z. 2013. Assessing phylogenetic relationships among Galliformes: a multigene phylogeny with expanded taxon sampling in Phasianidae. PLOS ONE 8:1–12.

Watson D.M.S. 1940. The origin of frogs. Trans. R. Soc. Edinburgh 60:195–231.

Weisbecker V. 2011. Monotreme ossification sequences and the riddle of mammalian skeletal development. Evolution 65:1323–1335.

Weisbecker V., Mitgutsch C. 2010. A large-scale survey of heterochrony in anuran cranial ossification patterns. J. Zool. Syst. Evol. Research 48:332–347.

Werneburg R. 2018 (for 2017). Earliest 'nursery ground' of temnospondyl amphibians in the Permian. Semana 32:3–42.

Werneburg I., Geiger M. 2017. Ontogeny of domestic dogs and the developmental foundations of carnivoran domestication. J. Mammal. Evol. 24:323–343.

Werneburg I., Sánchez-Villagra M.R. 2009. Timing of organogenesis support basal position of turtles in the amniote tree of life. BMC Evol. Biol. 9:82.

Witzmann F., Schoch R.R. 2006. Skeletal development of the temnospondyl *Acanthostomatops vorax* from the Lower Permian Döhlen Basin of Saxony. Trans. R. Soc. Edinburgh: Earth Sci. 96: 365–385.

Zhang P., Zhou H., Chen Y.-Q., Liu Y.-F., Qu L.-H. 2005. Mitogenomic perspectives on the origin and phylogeny of living amphibians. Syst. Biol. 54:391–400.

Zhuang L., Bluteau G., Trueb B. 2015. Phylogenetic analysis of receptor FgfrL1 shows divergence of the C-terminal end in rodents. Comp. Biochem. Physiol. B 186:43–50.

**Appendix 1: Sources of data for ossification sequences.**

Empty cells indicate that these data are unavailable. Three methods were examined, and we used the one for which most data were available (position in the ossification sequence, last column).

| | Standardization method (data type used) | | |
|---|---|---|---|
| **Taxa** | **Ontogenetic stages** | **Snout-vent length (mm)** | **Ossification sequence position** |
| **Actinopterygii** | | | |
| *Amia calva* | | Grande and Bemis 1998 | Grande and Bemis 1998 |
| *Clarias gariepinus* | | Adriaens and Verraes 1998 | Adriaens and Verraes 1998 |
| *Danio rerio* | | Cubbage and Mabee 1996 | Cubbage and Mabee 1996 |
| *Oryzias latipes* | Langille and Hall 1987 | | |
| **Tristichopteridae** | | | |
| *Eusthenopteron foordi* | | Cote et al. 2002; Leblanc and Cloutier 2005 | Cote et al. 2002; Leblanc and Cloutier 2005 |
| **Temnospondyli** | | | |
| *Archegosaurus decheni* | | Witzmann 2006 | Witzmann 2006 |
| *Apateon caducus* (Erdesbach) | Schoch 2004 | Schoch 2004 | Schoch 2004 |
| *Apateon caducus* (Obermoschel) | | Werneburg 2018 | Werneburg 2018 |
| *Apateon pedestris* (Erdesbach) | Schoch 2004 | | Schoch 2004 |
| *Apateon pedestris* (Obermoschel) | | Werneburg 2018 | Werneburg 2018 |
| "*Melanerpeton*" *humbergense* | Schoch 2004 | | Schoch 2004 |
| *Micromelerpeton credneri* | | Boy 1995; Lillich and Schoch 2007; Witzmann and Pfretzschner 2009; Schoch 2009 | Boy 1995; Lillich and Schoch 2007; Witzmann and Pfretzschner 2009; Schoch 2009 |
| *Sclerocephalus haeuseri* | Lohmann and Sachs 2001; Schoch 2003; Schoch and Witzmann 2009; Werneburg 2018 | Lohmann and Sachs 2001; Schoch 2003; Schoch and Witzmann 2009; Werneburg 2018 | Lohmann and Sachs 2001; Schoch 2003; Schoch and Witzmann 2009; Werneburg 2018 |
| **Lepospondyli** | | | |
| *Hyloplesion longicaudatum* | | Olori 2013 | Olori 2013 |
| *Microbrachis pelikani* | | Olori 2013 | Olori 2013 |
| **Gymnophiona** | | | |
| *Gegeneophis ramaswamii* | Müller et al. 2005 | | Harrington et al. 2013 |
| *Hypogeophis rostratus* | Müller 2006 | | Harrington et al. 2013 |
| **Urodela** | | | |
| *Aneides lugubris* | | Wake et al. 1983 | Wake et al. 1983 |
| *Ambystoma macrodactylum* | | | Harrington et al. 2013 |
| *Ambystoma maculatum* | Moore 1989 | | Harrington et al. 2013 |
| *Ambystoma mexicanum* | | Laurin and Germain 2011 | Harrington et al. 2013 |
| *Ambystoma talpoideum* | Reilly 1987 | Reilly 1987 | Reilly 1987 |
| *Ambystoma texanum* | | Laurin and Germain 2011 | Harrington et al. 2013 |

| Taxa | Standardization method (data type used) | | |
|---|---|---|---|
| | **Ontogenetic stages** | **Snout-vent length (mm)** | **Ossification sequence position** |
| *Ambystoma tigrinum* | | | Harrington et al. 2013 |
| *Amphiuma means* | | | Harrington et al. 2013 |
| *Andrias japonicus* | | | Harrington et al. 2013 |
| *Bolitoglossa subpalmata* | | | Ehmcke and Clemen 2000 |
| *Dicamptodon tenebrosus* | | | Harrington et al. 2013 |
| *Eurycea bislineata* | | | Harrington et al. 2013 |
| *Gyrinophilus porphyriticus* | | | Harrington et al. 2013 |
| *Hemidactylium scutatum* | | | Harrington et al. 2013 |
| *Lissotriton vulgaris* | | Laurin and Germain 2011 | Harrington et al. 2013 |
| *Necturus maculosus* | | | Harrington et al. 2013 |
| *Notophthalmus viridescens* | Reilly 1986 | Reilly 1986 | Harrington et al. 2013 |
| *Onychodactylus japonicus* | | | Harrington et al. 2013 |
| *Pleurodeles waltl* | | | Harrington et al. 2013 |
| *Ranodon sibiricus* | | | Harrington et al. 2013 |
| *Salamandra salamandra* | | | Harrington et al. 2013 |
| *Salamandrella keyserlingii* | | | Harrington et al. 2013 |
| *Siren intermedia* | Reilly and Altig 1996 | Reilly and Altig 1996 | Reilly and Altig 1996 |
| *Triturus karelinii* | | | Harrington et al. 2013 |
| **Anura** | | | |
| *Alytes obstetricans* | | | Yeh 2002 |
| *Ascaphus truei* | | | Harrington et al. 2013 |
| *Anaxyrus boreas* | | | Gaudin 1978 |
| *Bombina orientalis* | | | Harrington et al. 2013 |
| *Bufo bufo* | | | Harrington et al. 2013 |
| *Cornufer guentheri* | | | Harrington et al. 2013 |
| *Ceratophrys cornuta* | | | Harrington et al. 2013 |
| *Chacophrys pierotti* | | | Harrington et al. 2013 |
| *Crinia signifera* | | | Harrington et al. 2013 |
| *Dendrobates auratus* | de Sá and Hill 1998 | de Sá and Hill 1998 | Harrington et al. 2013 |
| *Discoglossus sardus* | | | Pugener and Maglia 1997 |
| *Eleutherodactylus coqui* | | | Harrington et al. 2013 |
| *Eleutherodactylus nubicola* | | | Harrington et al. 2013 |
| *Epidalea calamita* | | | Harrington et al. 2013 |
| *Epipedobates tricolor* | de Sá and Hill 1998 | de Sá and Hill 1998 | Harrington et al. 2013 |
| *Fejervarya cancrivora* | | | Harrington et al. 2013 |
| *Hamptophryne boliviana* | | | Harrington et al. 2013 |
| *Hyla versicolor* | | | Harrington et al. 2013 |
| *Hylorina sylvatica* | | | Harrington et al. 2013 |
| *Hymenochirus boettgeri* | | | de Sá and Swart 1999 |
| *Hypsiboas lanciformis* | de Sá 1988 | de Sá 1988 | de Sá 1988 |
| *Kassina senegalensis* | | | Harrington et al. 2013 |
| *Leptodactylus chaquensis* | | | Harrington et al. 2013 |

| Taxa | Standardization method (data type used) | | |
|---|---|---|---|
| | Ontogenetic stages | Snout-vent length (mm) | Ossification sequence position |
| *Osteopilus septentrionalis* | | | Trueb 1966 |
| *Palaeobatrachus sp.* | | | Harrington et al. 2013 |
| *Pelobates cultripes* | | | Harrington et al. 2013 |
| *Philautus silus* | | | Harrington et al. 2013 |
| *Phyllomedusa vaillanti* | | | Harrington et al. 2013 |
| *Pipa myersi* | | | Yeh 2002 |
| *Pipa pipa* | | Trueb et al. 2000 | Harrington et al. 2013 |
| *Pseudacris regilla* | | | Harrington et al. 2013 |
| *Pseudacris triseriata* | | | Harrington et al. 2013 |
| *Pseudis platensis* | | | Harrington et al. 2013 |
| *Pseudophryne bibronii* | | | Harrington et al. 2013 |
| *Pyxicephalus adspersus* | | | Harrington et al. 2013 |
| *Rana (Amerana) aurora* | | | Harrington et al. 2013 |
| *Rana (Amerana) cascadae* | | | Harrington et al. 2013 |
| *Rana (Amerana) pretiosa* | | | Harrington et al. 2013 |
| *Rana (Rana) temporaria* | | | Harrington et al. 2013 |
| *Rana (Pantherana) pipiens* | | | Kemp and Hoyt 1969 |
| *Rhinophrynus dorsalis* | | | Harrington et al. 2013 |
| *Shomronella jordanica* | | | Harrington et al. 2013 |
| *Smilisca baudini* | | | Harrington et al. 2013 |
| *Spea bombifrons* | Wiens 1989 | Wiens 1989 | Wiens 1989 |
| *Spea multiplicata* | | | Harrington et al. 2013 |
| *Triprion petasatus* | | | Harrington et al. 2013 |
| *Uperoleia laevigata* | | | Harrington et al. 2013 |
| *Xenopus laevis* | | | Harrington et al. 2013 |
| **Mammalia** | | | |
| *Bradypus variegatus* | | | Hautier et al. 2011 |
| *Cavia porcellus* | | | Hautier et al. 2013 |
| *Choloepus didactylus* | | | Hautier et al. 2011 |
| *Cryptotis parva* | | | Koyabu et al. 2011 |
| *Cyclopes didactylus* | | | Hautier et al. 2011 |
| *Dasypus novemcinctus* | | | Hautier et al. 2011 |
| *Dasyurus viverrinus* | | | Hautier et al. 2013 |
| *Didelphis albiventris* | | de Oliveira et al. 1998 | de Oliveira et al. 1998 |
| *Echinops telfairi* | | | Werneburg et al. 2013 |
| *Elephantulus rozeti* | | | Hautier et al. 2013 |
| *Eremitalpa granti* | | | Hautier et al. 2013 |
| *Erinaceus amurensis* | | | Koyabu et al. 2011 |
| *Felis silvestris* | | | Sánchez-Villagra et al. 2008 |
| *Homo sapiens* | | | Hautier et al. 2013 |
| *Heterohyrax brucei* | | | Hautier et al. 2013 |
| *Loxodonta africana* | | | Hautier et al. 2012 |

| Taxa | Standardization method (data type used) | | |
|---|---|---|---|
| | Ontogenetic stages | Snout-vent length (mm) | Ossification sequence position |
| *Macropus eugenii* | | | Hautier et al. 2013 |
| *Macroscelides proboscideus* | | | Hautier et al. 2013 |
| *Manis javanica* | | | Hautier et al. 2013 |
| *Meriones unguiculatus* | | Yukawa et al. 1999 | Yukawa et al. 1999 |
| *Mesocricetus auratus* | | | Hautier et al. 2013 |
| *Mogera wogura* | | | Koyabu et al. 2011 |
| *Monodelphis domestica* | | | Hautier et al. 2013 |
| *Mus musculus* | | | Hautier et al. 2013 |
| *Ornithorhynchus anatinus* | | | Weisbecker 2011 |
| *Orycteropus afer* | | | Hautier et al. 2013 |
| *Perameles nasuta* | | | Hautier et al. 2013 |
| *Peromyscus melanophrys* | | | Hautier et al. 2013 |
| *Procavia capensis* | | | Hautier et al. 2013 |
| *Rattus norvegicus* | | | Hautier et al. 2013 |
| *Rhabdomys pumilio* | | | Hautier et al. 2013 |
| *Rousettus amplexicaudatus* | | | Hautier et al. 2013 |
| *Sus scrofa* | | | Hautier et al. 2013 |
| *Tachyglossus aculeatus* | | | Weisbecker 2011 |
| *Talpa* spp. | | | Sánchez-Villagra et al. 2008 |
| *Tenrec ecaudatus* | | | Werneburg et al. 2013 |
| *Tamandua tetradactyla* | | | Hautier et al. 2011 |
| *Tarsius spectrum* | | | Hautier et al. 2013 |
| *Trichosurus vulpecula* | Weisbecker et al. 2008 | | Hautier et al. 2013 |
| *Tupaia javanica* | | | Hautier et al. 2013 |
| **Squamata** | | | |
| *Lacerta vivipara* | | | Hautier et al. 2013 |
| *Lerista bougainvillii* | | Hugi et al. 2012 | Hugi et al. 2012 |
| *Liopholis whitii* | | Hugi et al. 2012 | Hugi et al. 2012 |
| *Hemiergis peronii* | | Hugi et al. 2012 | Hugi et al. 2012 |
| *Saiphos equalis* | | Hugi et al. 2012 | Hugi et al. 2012 |
| **Crocodylia** | | | |
| *Alligator mississipiensis* | Rieppel 1993a | | Rieppel 1993a |
| **Aves** | | | |
| *Anas platyrhynchos* | | | Maxwell et al. 2010 |
| *Cairina moschata* | | | Maxwell et al. 2010 |
| *Coturnix coturnix* | | | Maxwell et al. 2010 |
| *Coturnix coturnix* (N&T) | | | Maxwell et al. 2010 |
| *Dromaius novaehollandiae* | | | Maxwell et al. 2010 |
| *Dromaius novaehollandiae* (YPM) | | | Maxwell et al. 2010 |
| *Gallus gallus* | | | Maxwell et al. 2010 |

| Taxa | Standardization method (data type used) | | |
|---|---|---|---|
| | Ontogenetic stages | Snout-vent length (mm) | Ossification sequence position |
| *Gallus gallus* (S&W) | | | Maxwell et al. 2010 |
| *Larus argentatus* | | | Maxwell et al. 2010 |
| *Larus canus* | | | Maxwell et al. 2010 |
| *Larus ridibundus* | | | Maxwell et al. 2010 |
| *Meleagris gallopavo* | | | Maxwell et al. 2010 |
| *Phalacrocorax auritus* | | | Maxwell et al. 2010 |
| *Somateria mollissima* | | | Maxwell et al. 2010 |
| *Stercorarius skua* | | | Maxwell et al. 2010 |
| *Sterna hirundo* | | | Maxwell et al. 2010 |
| *Struthio camelus* | | | Maxwell et al. 2010 |
| **Testudines** | | | |
| *Apalone spinifera* | | | Sánchez-Villagra et al. 2008 |
| *Chelydra serpentina* | Rieppel 1993b | Rieppel 1990, 1993b | Rieppel 1993b |
| *Macrochelys temminckii* | | | Sánchez-Villagra et al. 2008 |
| *Pelodiscus sinensis* | | | Sánchez-Villagra et al. 2008 |

## Appendix references


Adriaens D., Verraes W. 1998. Ontogeny of the osteocranium in the African Catfish, *Clarias gariepinus* Burchell (1822) (Siluriformes: Clariidae): ossification sequence as a response to functional demands. J. Morph. 235:183–237.

Boy J. 1995. Über die Micromelerpetontidae (Amphibia: Temnospondyli). 1. Morphologie und Paläoökologie des *Micromelerpeton credneri* (Unter-Perm; SW-Deutschland). Paläont. Z. 69:429–457.

Cote S., Carroll R., Cloutier R., Bar-Sagi L. 2002. Vertebral development in the Devonian sarcopterygian fish *Eusthenopteron foordi* and the polarity of vertebral evolution in non-amniote tetrapods. J. Vertebr. Paleontol. 22:487–502.

Cubbage C.C., Mabee P.M. 1996. Development of the cranium of paired fins in the Zebrafish *Danio rerio* (Ostariophysi, Cyprinidae). J. Morph. 229:121–160.

Ehmcke J., Clemen G. 2000. The structure and development of the skull of Costa Rican plethodontid salamanders (Amphibia: Urodela). Ann. Anat. 182: 537–547.

Gaudin A.J. 1978. The sequence of cranial ossification in the California toad, *Bufo boreas* (Amphibia, Anura, Bufonidae). J. Herpetol. 12:309–318.

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. J. Vertebr. Paleontol. 18(1, suppl.; Soc. of Vert. Paleont. Memoir 4):688 pages.

Hautier L., Weisbecker V., Goswami A., Knight F., Kardjilov N., Asher R. J. 2011. Skeletal ossification and sequence heterochrony in xenarthran evolution. Evol. Dev. 13:460–476.

Hautier L., Stansfield F. J., Allen W. R., Asher R. J. 2012. Skeletal development in the African elephant and ossification timing in placental mammals. Proc. R. Soc. B 279:2188–2195.

Hautier L., Bennett N.C., Viljoen H., Howard L., Milinkovitch M.C., Tzika A.C., Goswami A., Asher R.J. 2013. Patterns of ossification in southern versus northern placental mammals: mammal skeletogenesis. Evolution 67:1994–2010.

Harrington S.M., Harrison L.B., Sheil C.A. 2013. Ossification sequence heterochrony among amphibians. Evol. Dev. 15:344–364.

Hugi J., Hutchinson M.N., Koyabu D., Sánchez-Villagra M.R. 2012. Heterochronic shifts in the ossification sequences of surface- and subsurface-dwelling skinks are correlated with the degree of limb reduction. Zoology 115:188–198.

Kemp N. E. and Hoyt J. 1969. Sequence of ossification in the skeleton of growing and metamorphosing tadpoles of *Rana pipiens*. J. Morphol. 129:415–444.

Koyabu D., Endo H., Mitgutsch C., Suwa G., Catania K. C., Zollikofer C. P. E., Oda S., Koyasu K., Ando M., Sánchez-Villagra M. R. 2011. Heterochrony and developmental modularity of cranial osteogenesis in lipotyphlan mammals. EvoDevo 2:1–18.

Laurin M., Germain D. 2011. Developmental characters in phylogenetic inference and their absolute timing information. Syst. Biol. 60:630–644.

Langille R.M., Hall B.K. 1987. Development of the head skeleton of the Japanese medaka, *Oryzias latipes* (Teleostei). J. Morphol. 193:135–158.

Leblanc J., Cloutier R. 2005. Developmental modularity and saltatory ontogeny in the Late Devonian osteolepiform *Eusthenopteron foordi*. Pp. 32–84 in Leblanc J: Précisions sur l'anatomie de l'ostéolépiforme *Eusthenopteron foordi* du Dévonien supérieur de Miguasha, Québec. Mémoire de maîtrise ( ~ M.Sc. thesis), Université du Québec à Rimouski. Available at http://semaphore.uqar.ca/283/1/Joel_Leblanc_aout2005.pdf

Lillich R., Schoch R. 2007. Finally grown up—the significance of adult *Micromelerpeton* [abstract]. J. Vert. Paleontol. 27(3, suppl.):106A.

Lohmann U., Sachs S. 2001. Observations on the postcranial morphology, ontogeny and palaeobiology of *Sclerocephalus haeuseri* (Amphibia: Actinodontidae) from the Lower Permian of Southwest Germany. Mem. Queensland Mus. 46:771–781.

Maxwell E.E., Harrison L.B., Larsson H.C.E. 2010. Assessing the phylogenetic utility of sequence heterochrony: evolution of avian ossification sequences as a case study. Zoology 113:57–66.

Müller H., Oommen O.V., Bartsch P. 2005. Skeletal development of the direct-developing caecilian *Gegeneophis ramaswamii* (Amphibia: Gymnophiona: Caeciliidae). Zoomorphology 124:171–188.

Müller H. 2006. Ontogeny of the skull, lower jaw, and hyobranchial skeleton of *Hypogeophis rostratus* (Amphibia: Gymnophiona: Caeciliidae) revisited. J. Morph. 267:968–986.

de Oliveira C.A., Nogueira J.C., Mahecha G.A.B. 1998. Sequential order of appearance of ossification centers in the opossum Didelphis albiventris (Didelphidae) skeleton during development in the marsupium. Ann. Anat. 180:113–121. [Lack of italics in the original.]

Olori, J.C. 2013. Ontogenetic sequence reconstruction and sequence polymorphism in extinct taxa: an example using early tetrapods (Tetrapoda: Lepospondyli). Paleobiology 39:400–428.

Púgener L.A., Maglia A.M. 1997. Osteology and skeletal development of *Discoglossus sardus* (Anura: Discoglossidae). J. Morphol. 233:267–286.

Reilly S.M. 1986. Ontogeny of cranial ossification in the eastern newt, *Notophthalmus viridescens* (Caudata: Salamandridae), and its relationship to metamorphosis and neoteny. J. Morphol. 188:315–326.

Reilly S.M. 1987. Ontogeny of the hyobranchial apparatus in the salamanders *Ambystoma talpoideum* (Ambystomatidae) and *Notophthalmus viridescens* (Salamandridae): the ecological morphology of two neotenic strategies. J. Morphol. 191:205–214.

Reilly S.M., Altig R. 1996. Cranial ontogeny in *Siren intermedia* (Amphibia: Sirenidae): Paedomorphic, metamorphic, and novel patterns of heterochrony. Copeia 1996:29–41.

Rieppel O. 1990. The structure and development of the jaw adductor musculature in the turtle *Chelydra serpentina*. Zool. J. Linn. Soc. 98:27–62.

Rieppel O. 1993a. Studies on skeleton formation in reptiles. V. Patterns of ossification in the skeleton of *Alligator mississippiensis* Daudin (Reptilia, Crocodylia). Zool. J. Linn. Soc. 109:301–325.

Rieppel O. 1993b. Studies on skeleton formation in reptiles: Patterns of ossification in the skeleton of *Chelydra serpentina* (Reptilia, Testudines). J. Zool. (Lond.) 231:487–509.

Sánchez-Villagra M., Goswami A.,Weisbecker V., Mock O., Kuratani S. 2008. Conserved relative timing of cranial ossification patterns in early mammalian evolution. Evol. Dev. 10:519–530.

de Sá R.O. 1988. Chondrocranium and ossification sequence of *Hyla lanciformis*. J. Morphol. 195:345–355.

de Sá R.O., Swart C.C. 1999. Development of the suprarostral plate of pipoid frogs. J. Morphol. 240:143–153.

Schoch R.R. 2004. Skeleton formation in the Branchiosauridae: a case study in comparing ontogenetic trajectories. J. Vert. Paleontol. 24:309–319.

Schoch R.R. 2003. Early larval ontogeny of the Permo-Carboniferous temnospondyl *Sclerocephalus*. Palaeontology 46:1055–1072.

Schoch R.R. 2009. Evolution of life cycles in early amphibians. Annu. Rev. Earth Planet. Sci. 37:125–162.

Schoch R.R., Witzmann F. 2009. Osteology and relationships of the temnospondyl genus *Sclerocephalus*. Zool. J. Linn. Soc. 157:135–168.

Trueb L. 1966. Morphology and development of the skull in the frog *Hyla septentrionalis*. Copeia 1966:562–573.

Trueb L., Pugener L.A., Maglia A.M. 2000. Ontogeny of the bizarre: An osteological description of *Pipa pipa* (Anura: Pipidae), with an account of skeletal development in the species. J. Morphol. 243:75–104.

Wake T.A., Wake D.B., Wake M.H. 1983. The ossification sequence of *Aneides lugubris*, with comments on heterochrony. J. Herpetol. 17:10–22.

Weisbecker V., Goswami A., Wroe S., Sánchez-Villagra M.R. 2008. Ossification heterochrony in the therian postcranial skeleton and the marsupial–placental dichotomy. Evolution 62:2027–2041.

Weisbecker V. 2011. Monotreme ossification sequences and the riddle of mammalian skeletal development. Evolution 65:1323–1335.

Werneburg R. 2018 (for 2017). Earliest 'nursery ground' of temnospondyl amphibians in the Permian. Semana 32:3–42.

Werneburg I., Tzika A.C., Hautier L., Asher R.J., Milinkovitch M.C., Sánchez-Villagra M.R. 2013. Development and embryonic staging in non-model organisms: the case of an afrotherian mammal: Development and embryonic staging. J. Anat. 222:2–18.

Wiens J.J. 1989. Ontogeny of the skeleton of *Spea bombifrons* (Anura: Pelobatidae). J. Morphol. 202:29–51.

Witzmann F. 2006. Developmental patterns and ossification sequence in the Permo-Carboniferous temnospondyl *Archegosaurus decheni* (Saar-Nahe Basin, Germany). J. Vert. Paleontol. 26:7–17.

Witzmann F., Pfretzschner H.-U. 2003. Larval ontogeny of *Micromelerpeton credneri* (Temnospondyli, Dissorophoidea). J. Vert. Paleontol. 23:750–768.

Yeh J. 2002. The evolution of development: two portraits of skull ossification in pipoid frogs. Evolution 56:2484–2498.

Yukawa M., Hayashi N., Takagi K., Mochizuki K. 1999. The normal development of Mongolian Gerbil foetuses and, in particular, the timing and sequence of the appearance of ossification centres. Anat. Histol. Embryol. 28:319–324.