A phylogenomic approach to reconstruct interrelationships of main clupeocephalan lineages with a critical discussion of morphological apomorphies

Nicolas Straube1,2*, Chenhong Li3, Matthias Mertzen1,4, Hao Yuan3 and Timo Moritz1,4

Abstract

Background: Previous molecular studies on the phylogeny and classification of clupeocephalan fishes revealed numerous new taxonomic entities. For re-analysing these taxa, we perform target gene capturing and subsequent next generation sequencing of putative ortholog exons of major clupeocephalan lineages. Sequence information for the RNA bait design was derived from publicly available genomes of bony fishes. Newly acquired sequence data comprising > 800 exon sequences was subsequently used for phylogenetic reconstructions.

Results: Our results support monophyletic Otomorpha comprising Alepocephaliformes. Within Ostariophysi, Gonorynchiformes are sister to a clade comprising Cypriniformes, Characiformes, Siluriformes and Gymnotiformes, where the interrelationships of Characiformes, Siluriformes and Gymnotiformes remain enigmatic. Euteleosts comprise four major clades: Lepidogalaxiiformes, Protacanthopterygii, Stomiatii, and Galaxiiformes plus Neoteleostei. The monotypic Lepidogalaxiiformes form the sister-group to all remaining euteleosts. Protacanthopterygii, comprising Argentinii-, Esoci- and Salmoniformes, is sister to Stomiatii (Osmeriformes and Stomiatifomes) and Galaxiiformes plus Neoteleostei.

Conclusions: Several proposed monophyla defined by morphological apomorphies within the Clupeocephalan phylogeny are confirmed by the phylogenetic estimates presented herein. However, other morphologically described groups cannot be reconciled with molecular phylogenies. Thus, numerous morphological apomorphies of supposed monophyla are called into question. The interpretation of suggested morphological synapomorphies of otomorph fishes is strongly affected by the inclusion of deep-sea inhabiting, and to that effect morphologically adapted Alepocephaliformes. Our revision of these potential synapomorphies, in the context that Alepocephaliformes are otomorph fishes, reveals that only a single character of the total nine characters proposed as synapomorphic for the group is clearly valid for all otomorphs. Three further characters remain possible apomorphies since their status remains unclear in the deep-sea adapted Alepocephaliformes showing developmental lag and lacking a swim bladder. Further, our analysis places Galaxiiformes as sister group to neoteleosts, which contradicts some previous molecular phylogenetic studies. This needs further investigation from a morphological perspective, as suggested synapomorphies for several euteleostean lineages are challenged or still lacking. For the verification of results presented herein, a denser phylogenomic-level taxon sampling should be applied.

Keywords: Otomorph characters, Euteleostei, Alepocephaliformes, Synapomorphies, Target gene capture, Next generation sequencing, Evolmarkers

* Correspondence: straube@znm.de
1Institut für Zoologie & Evolutionsbiologie, Friedrich-Schiller-Universität Jena, Erbertstraße 1, 07743 Jena, Germany
2Zoologische Staatssammlung München, Staatliche Naturwissenschaftliche Sammlungen Bayerns, Münchhausenstraße 21, 81247 Munich, Germany
Full list of author information is available at the end of the article

© The Author(s). 2018 Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The Creative Commons Public Domain Dedication waiver (http://creativecommons.org/publicdomain/zero/1.0/) applies to the data made available in this article, unless otherwise stated.
Background
With approximately 32,000 species, the teleost fishes comprise about half of the vertebrate species, with representatives in almost any aquatic environment from montane habitats to the deep-sea. With the progress in gene sequencing technologies, several phylogenetic hypotheses have been published for Teleostei in recent years advancing from single gene alignments to mitochondrial genomes and multi-locus approaches increasing in the taxonomic diversity analyzed. However, major discrepancies are evident between morphology and molecular phylogenetics [1–6] and between different DNA sequence based datasets. Further, several deep phylogenetic nodes remain enigmatic [2]. In the course of this paper, we are referring to the classification suggested in [2], if not indicated otherwise.

Phylogenetic analyses suggest that Teleostei comprises three main lineages: Osteoglossomorpha, Elopomorpha and Clupeocephala with the latter being the largest by far. The monophyly of the supercohort Clupeocephala is evidenced by both morphological and molecular data [1, 2, 5–7]. Wiley and Johnson [6] suggested Clupeocephala to contain two major lineages, the Otomorpha and Euteleostei (Fig. 1a). Otomorpha are suggested to comprise the Clupei (herrings and allies, also referred to as Clupeomorpha) and Ostariophysi. The diverse subcohort Ostariophysi includes the most species rich and predominantly fresh water inhabiting lineages, the Cypriniformes (carps) as well as the Siluriformes (catfishes). Further, the clade comprises Gonorynchiformes (milkfishes), Characiformes (characins and allies) and Gymnotiformes (neotropical knifefishes) (Fig. 1). With more than 10,000 described species, the Ostariophysi pose a noteworthy part of today’s Clupeocephalan overall diversity. The situation gets complicated by results from molecular phylogenetic studies providing strong evidence that Alepocephaliformes are Otomorpha [1, 2, 5, 8–10] (Fig. 1b,c). Contrasting all other otomorph fishes, Alepocephaliformes represent a group of deep-sea fishes, which show extensive morphological adaptations to their habitat making a morphological comparison to other taxa within Otomorpha difficult. The first detailed morphological investigation on the systematic position of alepocephaliformes [11] placed them as sistergroup to the Argentinioidei (marine smelts and allies), which was adopted in the classification of [6], placing the Alepocephaloidei in the order Argentiniformes among euteleost fishes. Therefore, molecular phylogenetic analyses challenge proposed morphological synapomorphies on at least five phylogenetic levels rendering higher-level taxa Euteleostei, Otomorpha, Protacanthopterygii, Argentinioidei and Alepocephaloidei either poly- or paraphyletic.

---

**Fig. 1** Summary of previous phylogenetic estimates and classifications of Clupeocephalan fishes. Colours indicate taxa with variable phylogenetic positions.
A sister-group relation of clupeomorphs and ostariophysians was first proposed by Lecointre and Nelson [12]. The respective taxonomic unit, the Otomorpha [6], also named Otocephala [13] or Ostarioclupeomorpha [14], was soon established, but morphological evidence supporting the monophyly of the group remained scarce [6, 12, 14, 15]. The following nine characters have been discussed as possible apomorphies for otomorphs: (1) ossification of autopalatine early in ontogeny- within the palatoquadrate cartilage three bones autogenously ossify. The ancestral teleost condition is an autopalatine, which forms clearly later than metapterygoid and quadrate [16]. Arratia and Schultz [16] found that the autopalatine in Denticeps, Dorsosoma and Chanos ossifies at the same ontogenetic stage. In salmonids, it ossifies slightly later compared to the otomorph palatoquadrate ossification sequence and the authors suggested an early ontogeny of an autopalatine as synapomorphy for Clupeocephalans. In more primitive taxa such as Osteoglossomorpha or Elopomorpha, the ossification of the autopalatine takes place significantly later. Subsequent studies proposed that this character was not apomorphic for clupeocephalans but otomorphs [12–15]. Later, Arratia [7, 17] revised this statement and re-defined the character as apomorphic for the Clupeocephalans again. (2) Fusion of medial extrascapulars with parietals- the absence of a separate median sensory canal bone in the supratemporal commisure, resulting in a canal in the parietal and supraoccipital bones, was first reported from clupeomorphs and regarded as apomorphic for this group [18]. Such a condition is usually interpreted as a fusion of the median extrascapular to the parietal [12]. A respective fusion without inclusion of the supraoccipital was also reported from some species of Gonorynchiformes [19], Cypriniformes, and Characiformes [20, 21]. This led to the conclusion that the character is of apomorphic state in Otomorpha [6, 12, 15].

(3) Ossified epicentra- ossified epicentra are documented already in the Elopomorpha [22], which excludes this character as an apomorphy for Otomorpha. (4) Connection of swim bladder and ear - the otophysic connection, i.e. the connection between swim bladder and inner ear, fundamentally differs in Clupeomorph and Ostariophysi [23–25]. The evolutionary sequence between these two states is unknown. Furthermore, gonorynchiform species do not show an otophysic connection [23, 25], but only show adaptations, which can be interpreted as ancestral conditions of an otophysic connection of the Ostariophysi-type: the division of the swim bladder into an anterior and posterior chambers and an enlarged first rib. This leads to the conclusion that this character cannot be regarded as an apomorphy for otomorphs. Alepocephalids lack a swim bladder and (5) Anterior chamber of swim bladder partly or completely covered by silvery peritoneal tunic- a completely or partially silvery peritoneal cover of the anterior chamber of the swim bladder [23] was discussed as possibly apomorphic for Ostariophysi [24]. Clupeomorph swim bladders do not show an anterior-posterior division as in Ostariophysi, however, within their single chambered swim bladder, the anterior part is also covered with a silvery peritoneal tunic. Therefore, this character is regarded as apomorphy for otomorphs [26]. (6) Hemal spines anterior to second preural centrum fused to centra early in ontogeny- the character “hemal spine of preural-centrum 3 (PU3) and anterior vertebrae fused with their respective centra” [26] usually is supplemented with the information that the character is already developed “from a young juvenile stage on” [6, 14, 24]. Fink and Fink [24] found that also clupeomorphs have all hemal spines fused to their centra and interpreted this condition as hint to a "relationship between Clupeomorpha and Ostariophysi". Later, however, the Fink and Fink [24] listed the same character as an ostariophysean apomorphy disregarding the clupeomorph condition [24]. The interpretation of this character in the light of recent phylogenetic hypothesis is difficult as the deep-sea dwelling alepocephalids show reductions and developmental lag in ossifications. (7) Presence of a pleuro-ostyl- a pleurostyl is found in clupeoids and Ostariophysi [27, 28]. Most clupeiforms and Ostariophysi have a pleuro-ostyl. The sister-taxon of all clupeiforms, Denticeps, as well as fossil stem-group representatives of Ostariophysi, however, do not show this character. In herring-like fishes, it is a fusion of the first uroneural to the (first) preural centrum [27, 28]. Respective fusions in the caudal skeleton of Ostariophysi are more comprehensive and include additionally both ural centra and the hypural 2 [24]. (8) Lack of cartilaginous connection between the bases of hypurals 1 and 2- the lack of a cartilaginous connection between bases of the hypurals 1 and 2 at any ontogenetic stage was proposed as apomorphy for otomorphs [6, 14, 15]. (9) Fusion of hypural 2 with compound centrum. A fusion of hypural 2 with the first ural centrum is present in Ostariophysi and Clupeocephora [6, 24, 27, 28]. Such a fusion is, on the other hand, absent from all gonorynchiforms except Gonorynchus in which, however, caudal element fusions are extensive [25].

In summary, morphological evidence supporting the taxon otomorpha is scarce. Further, several of the proposed apomorphic characters are doubtful and have already been critically discussed [6, 7, 12, 13].

With advances in sequencing technology, phylogenetic analyses shifted from analysing morphological matrices to sequence alignments with growing number of genes and taxa included in the analyses. Based on molecular
phylogenetic results, Betancur-R. et al. [2] define the Euteleostei (referred to as Euteleostomorpha in [2]) to comprise several newly defined clades in their new classification of bony fishes, for example the Lepidogalaxiiformes, a taxonomic unit comprising a single extant species only, the West-Australian salamander fish *Lepidogalaxias salamandroides*. Molecular phylogenetic analyses suggest this species to form the sister taxon to all other euteleost lineages [29]. Besides its unexpected phylogenetic position suggested by molecular data, it also displays numerous noteworthy morphological characters indicating strong specialization and, contrasting molecular phylogenetics, a close phylogenetic relationship with Galaxiiformes [30]. Following Betancur-R. et al. [1, 2], the sister clade of Lepidogalaxiiformes is unnamed including three major clades: the Protacanthopterygii as sister to a clade comprising Stomiati and Neoteleostei. The Protacanthopterygii sensu Betancur-R. et al. [1, 2] comprise Argentiniformes, Galaxiiformes, Salmoniformes (salmons) and Esociformes (pikes and mudminnows) (Fig 1c). It should be noted that, contrasting [1, 2], Near et al. [5] found the Galaxiiformes to form a sister group relationship with neoteleosts, i.e. the order was not clustering along with the Protacanthopterygii sensu Betancur-R. et al. [1, 2] (Fig. 1d, f). From a morphological perspective, another noteworthy result from DNA sequence data is the sister group relationship of Osmeriformes (smelts) with Stomiatiiformes (dragonfish) forming the Stomiati (Fig. 1c, d e, f). Morphological studies considered dragonfishes hitherto as neoteleost fishes sharing proposed neoteleost synapomorphies, especially in branchial arch musculature and tooth attachment type [6]. Wiley and Johnson [6] commented that homology of these characters has not been evaluated from an ontogenetic perspective in Stomiatiiformes and such information is still lacking. Hence, the classification in Betancur-R. et al. [1] challenged the morphological monophyly of neoteleosts and calls the suggested synapomorphic characters into question. Results presented in Betancur-R. et al. [1, 2] reproduce neoteleost fishes as sister to the Stomiati/ Protacanthopterygii clade. In their phylogenetic tree reconstruction, however, this split lacks high node support, confirming the split between Stomiati and Neoteleostei. A similar situation is apparent in the sister group relationship of Protacanthopterygii and the Stomiati/ Neoteleost clade. With updating the classification of bony fishes [2], Protacanthopterygii are sister to a clade now comprising Stomiati and Neoteleostei. Among neoteleosts, Ateleopodiformes (jellynose fishes) are sister to all further higher-level taxa referred to as Eurypterygia (Cyclosquamata, Ctenosquamata, Acan tomorphhata, Euacantomorphacea, and Percomorphada).

In a recent and extensive study, Mirande [31] combines both morphological and molecular data to re-infer the major phylogenetic relationships within Acanthopterygii. The parsimony-based phylogenetic estimates [31] are only partially in accordance with studies analyzing solely molecular data, demonstrating the contrarious phylogenetic signals by morphological and molecular data. The final hypothesis of Mirande [31] recovers the Clupeocephala with high Bremer support, while the Otomorpha, including the Alepocephaliformes as sister to the remaining otomorph taxa, appear only weakly supported. The well-supported euteleosts show two major clades, one comprising Lepidogalaxiiformes and Galaxiiformes as sister groups to all remaining euteleost lineages. The monophyly of the Lepidogalaxiform/ Galaxiiform clade is only weakly supported. A sister group relationship of Protacanthopterygii (sensu Betancur-R. et al. [1, 2]) and Stomiati is not well supported, while the monophyly of neoteleosts appears underpinned based on Bremer support (Fig. 1e).

In summary, several taxonomic entities proposed in previous phylogenetic studies analyzing molecular data are not supported by presently available morphological data. In times of next generation sequencing, researchers started to conduct phylogenomic level approaches to resolve difficult phylogenetic questions such as otophyan interrelationships [32, 33].

Here, we apply a targeted gene capture approach using a set of curated RNA baits to attain a phylogenomic-level dataset of potential ortholog loci to test, if we can sequence a sufficient number of genes from non-model organisms to resolve deep phylogenetic nodes on ordinal level within the evolutionary oldest Clupeocephalan lineages. In this study, we test if (1) we can resolve the otomorph phylogeny and (2) discuss previously suggested morphological synapomorphies of Otomorpha in the light of our phylogenomic reconstruction. (3) New insights into the phylogeny of deep phylogenetic nodes of Euteleostei are discussed with regard to previous phylogenetic reconstructions and morphological characters.

**Results**

Sequencing of 52 taxa resulted in an average number of 6.3 million paired-end reads per specimen. After adaptor and quality trimming of reads, per taxon reads were blasted [34] against bait sequences to create gene bins. Reads mapped to target loci were *de-novo* assembled resulting on average in 3400 contigs per specimen. After reciprocally blasting specimen contigs against the reference genome, which was used for bait design, 368 loci were identified as potential paralogs and excluded from further analysis steps. 13,681 loci were available for the cross-contamination check. The highest percentage of potential cross-contamination between pair of taxa among these loci was only 1.44%, which meant there was no cross-contamination among our samples. (Additional file 1: Table S2). The latter were excluded
from further analyses. The search for phylogenetically informative loci applying the Matrix Reduction Software MARE [35, 36] identified 838 most informative loci. Additional checks for orthology [37, 38] identified four further loci flagged as potential paralogs (Additional file 1: Table S3) and were excluded from subsequent analysis. Another five loci were excluded, which contained less than four taxa. These 829 most informative loci were phylogenetically analysed totaling 202,922 possible sites per specimen in the concatenated nucleotide alignment. RAxML found 117,046 distinct alignments patterns and a proportion of gaps and undetermined characters of 60.3%. The alignments are available for download at the Dryad data repository [39]. See Fig. 2 and Additional file 1: Table S1 for a summary of target capture success.

Phylogenetic analyses

The results from Partitionfinder and PartitionfinderProtein [40–42] and best-fitting partitioning schemes for the maximum likelihood analyses of the concatenated datasets are available for download at the Dryad data repository [39]. ESS values derived from the Phylobayes [43] analyses are further listed in Additional file 1. Overall, results were indicating high quality runs, as the comparisons of bipartition frequencies is around 0.1 for both runs (Additional file 1).

Different phylogenetic estimates in this study are widely congruent with few exceptions, which will be subject of the discussion. Please refer to Fig. 3 and Table 1 for tracking the results listed below. After rooting all resulting trees with the two elopomorph taxa Elops senegalensis and Anguilla anguilla, Ponto-don buchholzi, representing the Osteoglossomorpha, forms the sister group of the monophyletic and well-supported Clupeocephala. Clupeocephala are split in two major clades, Otomorpha, comprising Clupeiformes, Alepocephaliformes and Ostariophys (Rutilus and Danio (Cypriniformes), Gonorynchus, Cromeria and Chaños (Gonorynchiformes), Gymnotus (Gymnotiformes), Alestes and Distichodus (Characiformes) and Schilbe (Siluriformes), and euteleosts including Lepidogalaxias (Lepidogalaxiiformes), Galaxias and Lovettia (Galaxiiformes), Esox, Dallia and Umbra (Esociformes), Bathylagus, Nansenia, Argentina and Opisthoproctus (Argentiniformes), Salmo and Coregonus (Salmoniformes), Stomiattii (Osmerus, Mallotus, Salax and Plecoglossus representing Osmeriformes) and Borostomias, Astronesthes and Maurulicus representing the Stomiiformes). Monophyletic Neoteleostei are part of the euteleosts and represented by several species of Aulopiformes, Ateleopodiformes and Trachypterus (Lampridiformes).

Within Otomorpha, we recovered three major clades, Clupeiformes, Alepocephaliformes, and Ostariophys. Clupeiformes are sister to a clade comprising Alepocephaliformes and Ostariophys (Fig. 3). Alepocephaliformes as sister to Ostariophys is well-supported by all analyses based on nucleotide alignments, whereas amino acid based analyses result in Alepocephaliformes as sister to a clade comprising Clupeiformes and Ostariophys (Table 1). Different types of phylogenetic analyses and datasets (i.e. concatenated amino acid versus concatenated DNA alignments and coalescent analyses) partially show weak node support for the phylogenetic placement of Alepocephaliformes as sister to Ostariophys (Fig. 3, Table 1). The ASTRAL [44] species tree computed from maximum likelihood trees based on amino acid single loci alignments result in a topology where Alepocephaliformes are sister to a clade including Clupeiformes and Ostariophys (Additional file 1: Figure S7). In several phylogenetic analyses Gonorynchus forms a distinct lineage as sister-group to all remaining Ostariophys (Table 1; Additional file 1: Figures S6 & S8). In all other analyses (Table 1), monophyletic Gonorynchiformes are clearly supported as the sister group of Otophysi (Table 1; Additional file 1: Figures S1-S5, S7).

All analyses recover Cypriniformes as sister to a clade comprising Characiformes, Gymnotus (Gymnotiformes) and Schilbe (Siluriformes) with high node support (Table 1), however, the relationships within the latter clade differ between analyses. While several analyses result in Gymnotus forming the sister lineage to a clade comprising Alestes, Distichodus and Schilbe (Table 1; Additional file 1: Figures S3, S4, S8), i.e. rendering Characiformes paraphyletic, only the concatenated amino acid dataset recovers interrelationships as suggested from morphology e.g. [24] and a comprehensive recent phylogenomic study [32] (Table 1, Additional file 1: Figures S1–2).

Regarding euteleost lineages, the phylogenetic analyses recover monospecific Lepidogalaxias as sister-group to all remaining euteleost groups (Table 1, Fig. 3, Additional file 1: Figures S1–8). The major euteleost clade comprises protacanthopterygians sensu [1, 2] excluding Galaxiiformes, i.e. Argentiniformes sister to a clade comprising Esociformes and Salmoniformes. This clade is sister to a clade including Stomiattii sensu Betancur-R. et al. [1, 2], Galaxiiformes and Neoteleostei. All analyses performed here support monophyletic Stomiattii comprising Stomiiformes and Osmeriformes. Retropinna is sister to a clade comprising further osmeriforms representing the major families. The Stomiattii are the sister clade to Galaxiiformes and Neoteleostei. The Galaxiform samples (Galaxias and Lovettia) included in this study form the sister taxon to the neoteleost lineage in all analyses with high node support (Table 1) contradicting results presented in Betancur-R. et al. [1, 2].

The following splits are variable in different types of phylogenetic analyses. Protacanthopterygii (excluding
Fig. 2 Summary of target capture success. X-axis: target species; y-axis: number of target loci captured per specimen.

Fig. 3 Phylogenetic reconstruction based on concatenated DNA sequence alignments of 52 taxa (see Additional file 1: Table S1) using RAxML [42] and best partitioning scheme resulting from a Partitionfinder analysis [40, 41]. Numbers at nodes refer to Table 1. Tree re-rooted with Elopomorpha (Elops senegalensis and Anguilla anguilla).
Table 1: Overview of information on node numbers provided in Fig. 3

| Node number | Node description | Bootstrapping support maximum likelihood analysis concatenated amino acid alignments, partitioned | Additional file 1: | Bootstrapping support maximum likelihood analysis concatenated DNA alignments, partitioned | Additional file 1: | Bootstrapping support maximum likelihood analysis concatenated AA alignments, non partitioned | Additional file 1: | Phylobayes analysis based on concatenated DNA alignments | Additional file 1: | Phylobayes analysis based on concatenated AA alignments | Additional file 1: | ASTRAL species tree based on AA alignments | Additional file 1: | ASTRAL species tree based on DNA alignments | IC/CA values plotted on ML tree based on the non-partitioned DNA dataset |
|-------------|------------------|------------------------------------------------------------------------------------------------|-------------------|------------------------------------------------------------------------------------------------|-------------------|------------------------------------------------------------------------------------------------|-------------------|----------------------------------------------------------------------------------------------------------------|-------------------|----------------------------------------------------------------------------------------------------------------|-------------------|----------------------------------------------------------------------------------------------------------------|-------------------|----------------------------------------------------------------------------------------------------------------|-------------------|
| 1           | Split Otomorpha/ Euteleostei | 100 | Additional file 1: | 100 | 100 | 100 | 1 | 1 | 1 | 1 | 0.404/0.404 |
| 2           | Split Clupeiformes/ (Ostariophysi; Alepocephaliformes) | n.a. | n.a. | n.a. | 100 | 100 | n.a. | 1 | n.a. | 1 | 0.621/0.021 |
| 3           | Radiation Clupeiformes | 100 | 100 | 100 | 100 | 1 | 1 | 1 | 0.545/0.054 |
| 4           | Split Alepocephaliformes/ Ostariophysi | 80 (Clupeiformes) | 80 (Clupeiformes) | 100 | 100 | 1 | 0.599/0.054 |
| 5           | Split Gonorynchiformes/ Otophysi | 100 | 100 | 100 | 100 | 1 | 1 | 1 | 0.705/0.755 |
| 6           | Split Cypriniformes/ (Gymnotus, Schilbe, Characiformes) | 100 | 100 | 100 | 100 | 1 | 1 | 1 | 0.710/0.710 |
| 7           | Split (Gymnotus/ Characiformes/ Schilbe) | 100 | 100 | n.a. | n.a. | n.a. | n.a. | n.a. | n.a. | n.a. | 0.316/0.316 |
| 8           | Split (Schilbe (Alestes/ Distichodus)) | 65 | 63 | n.a. | 57 | n.a. | n.a. | n.a. | n.a. | n.a. | 0.025/0.025 |
| 9           | Split (Distichodus/ Schilbe) | 52 | 43 | 63 | 55 | n.a. | 0.77 | 0.48 | n.a. | 0.025/0.025 |
| 10          | Split Lepidogalaxias/ remaining Euteleostei | 100 | 100 | 100 | 100 | 1 | 1 | 1 | 0.603/0.603 |
| 11          | Split "Protacanthopterygii"/ (Stomiatii, Galaxiiformes, Neoteleostei) | 100 | 100 | 100 | 100 | n.a. | 1 | 1 | 1 | 0.661/0.661 |
| 12          | Split Argentiniformes/ (Salmoniformes, Esociformes) | 88 | 95 | 87 | 77 | n.a. | 0.98 | 0.55 | 0.97 | 0.505/0.505 |
| 13          | Split Stomiati/ (Galaxiiformes, Neoteleostei) | 100 | 100 | 100 | 100 | 1 | 1 | 1 | 0.480/0.480 |
| 14          | Split Stomiiformes/ Osmeriformes | 100 | 100 | 100 | 100 | 1 | 1 | 1 | 0.679/0.679 |
| 15          | Split Retropinnidae/ | 100 | 100 | 100 | 100 | 1 | 1 | 1 | 0.864/0.864 |
Table 1 Overview of information on node numbers provided in Fig. 3 (Continued)

| Node number | Node description | Bootstrap support maximum likelihood analysis concatenated AA alignments, partitioned | Bootstrap support maximum likelihood analysis concatenated amino acid alignments, non partitioned | Bootstrap support maximum likelihood analysis concatenated DNA alignments, partitioned | Bootstrap support maximum likelihood analysis concatenated DNA alignments, non partitioned | Phylobayes analysis based on concatenated AA alignments | Phylobayes analysis based on concatenated DNA alignments | ASTRAL species tree based on AA alignments | ASTRAL species tree based on DNA alignments | IC/CA values plotted on ML tree based on the non-partitioned DNA dataset |
|-------------|------------------|-------------------------------------------------------------------------------------|---------------------------------------------------------------------------------|---------------------------------------------------------------------------------|---------------------------------------------------------------------------------|----------------------------------|----------------------------------|----------------------------------|----------------------------------|---------------------------------------------------------------|
|             |                  | Additional file 1: Figure S1 | Additional file 1: Figure S2 | Additional file 1: Figure S3 | Additional file 1: Figure S4 | Additional file 1: Figure S5 | Additional file 1: Figure S6 | Additional file 1: Figure S7 | Additional file 1: Figure S8 | n.a.                                           |
| 16          | Split Galaxiformes/ Neoteleostei | 100 | 100 | 100 | 100 | 1 | 1 | 0.76 | 1 | 0.254/ 0.254 |
| 17          | Split Trachipterus/ (Ateleopoiformes/ Aulopiformes) | 100 | 100 | 100 | 100 | n.a. | n.a. | n.a. | 1 | 0.316/ 0.316 |
| 18          | Split Ateleopoiformes/ Lampridiformes | n.a. | n.a. | 92 | 87 | n.a. | n.a. | n.a. | 0.87 | 0.797/ 0.797 |
| 19          | Split (Synodus/ Harpadon, Hime)/ (Evermannella, Antcozena, Lestrolepis) | 100 | 100 | 100 | 100 | 1 | 1 | n.a. | n.a. | 0.211/ 0.211 |

Additional file 1: Figure S1
Additional file 1: Figure S2
Additional file 1: Figure S3
Additional file 1: Figure S4
Additional file 1: Figure S5
Additional file 1: Figure S6
Additional file 1: Figure S7
Additional file 1: Figure S8
n.a.
remaining Osmeriformes
Galaxiiformes) appears paraphyletic in the analysis of the concatenated amino acid alignment (Additional file 1: Figure S2) using the Bayesian inference (Additional file 1: Figure S5). In these analyses, Esociformes are sister to a clade including Argentiniformes, Osmeriformes, Galaxiiformes, Stomiiformes, and neoteleost fishes, while all other analyses reconstruct a phylogeny as shown in Fig. 3, however, node support values for the split are partially low (Table 1).

Within neoteleosts, our results show two possible cladograms. Either two clades, where Aulopiformes is sister to a clade comprising Ateleopodiformes and Trachypterus (Table 1, Additional file 1: Figures S3, S4, S8) or Trachypterus as sister to remaining neoteleost lineages included in this study (Table 1, Additional file 1: Figures S1, 2, unresolved in S5 and S6).

For the few nodes, bootstrap and posterior node support values are high, i.e. above 95% or 0.95, respectively (Table 1). Computed IC and ICA values [42, 45–47] do not indicate conflicting bipartitions except for the nodes, which will be discussed below and are also marked with low node support values (Table 1). The overall relative tree certainty is 0.48 indicating low incongruence among trees. The AU test performed in CONSEL [48–51] ranks the phylogenetic estimates based on the concatenated nucleotide datasets higher than all other bifurcating trees, where the tree estimate based on the partitioned nucleotide alignment ranks highest (Additional file 1: Table S4).

Discussion
Molecular and morphological evidence for the Otomorpha

All phylogenetic reconstructions performed in this study readily result in well-supported Otomorpha as sister clade to Euteleostei and include Alepocephaliformes, as previously found in molecular studies [1–5, 10]. However, the phylogenetic placement of the alepocephaliforms within Otomorpha remains not completely ascertained as reflected in weak node support and low IC and ICA values indicating incongruence (Table 1). Future studies should include a denser taxon sampling covering the different inter- and intra-otomorph lineages to verify results presented herein.

Our review of hitherto proposed synapomorphic morphological characters in the light of the phylogenetic reconstruction (Fig. 3) indicates that morphological evidence supporting Otomorpha including Alepocephaliformes is presently limited. After taking into account previous reviews of morphological characters [6, 12, 13, 17, 26], nine characters have been discussed as possible apomorphies (see above) for Otomorpha. Three do not withstand thorough investigation, even without considering the inclusion of alepocephaliforms, i.e. the early ossification of the autoplatine which is apomorphic for clupeocephalans [7, 17], the ossified epicentrum, which are plesiomorphic [22], and the otophyic connection, which principally differs in clupeomorphs and ostariophysi [18, 23, 24] in a way that no transitional states seem likely. For the latter, the fossil record and the condition in Gonorynchiformes further contradict a possible synapomorphic state. Therefore, six characters remain as candidates to characterize the Otomorpha without including alepocephaliforms: (1) parietals fused with extrascapulars, (2) anterior part of swim bladder with silvery peritoneum, (3) fusion of haemal spines anterior of preural centrum 2 with their centra from an early stage on, as well as three characters dealing with the caudal fin skeleton: (4) the absence of a united cartilaginous basis of hypurals 1 and 2, (5) fusion of hypural 2 with the compound centrum, and (6) the presence of a pleurostyl.

Discussing the six characters in Alepocephaliformes reveals the difficulties arising when analysing such a morphologically highly specialized taxon. Fig. 4 (A and B) shows the dorsal view on the posterior right part of the neurocranium in two alepocephalid species focusing on medial extrascapulare. The ossified canals of the occipital commisure can be interpreted as the remnants of extrascapulare. They are not fused with the parietals in any analysed alepocephalid species, exemplified in Fig. 4. This challenges the synapomorphic state of character 1. In Lecointre and Nelson [12], a fusion of the extrascapular with the parietal in two alepocephalids, i.e. Leptodermus and Rouleina is described. A more detailed sample of alepocephalid species is necessary to fully evaluate character 1 and clarify, if this poses a convergence in Leptodermus and Rouleina with clupeomorphs and ostariophysi or a reversal in other alepocephaliforms species. Alepocephaliformes lack a swim bladder, which is likely owed to their deep-sea inhabiting lifestyle. Therefore, the details of the swim bladder such as character 2 can generally not be analysed. This does not exclude the possibility, that this character is still a synapomorphy, as it may have been reduced in the course of adapting to deep-sea conditions and could well be present in the common ancestor.

Despite their tendency for a delayed ossification in development, alepocephaliforms show an early fusion of haemal spines with their centra anterior to preural centrum 2 (character 3) (Fig. 5b, c). Therefore, character 3 seems to be a valid apomorphy for otomorphs [6, 14, 24]. On the other hand, maybe due to the long persisting cartilages in alepocephaliform development, a continuous cartilaginous basis of hypurals 1 and 2 is clearly visible in Holtbyrina and Normichthys (Fig. 5a-c) as well as Xenodermichthys and Maulisia. This state likely excludes character 4 as possible apomorphy for otomorphs. This cartilage also separates hypural 2 from the compound centrum avoiding a fusion (character 5). However, if the slow ossification sequence in
alepocephaliforms is interpreted as apomorphic for this group, the situation in alepocephaliforms could be the result of a reversal. In summary, the status of characters 4 and 5 remain questionable. The situation is clearer for the pleurostyl (character 6) which is clearly absent in alepocephaliforms (Fig. 5a-c). It is further absent from *Denticeps* (Clupeiformes, Fig. 5d), several fossil clupeiforms [52–54] and gonorynchiforms [25]. All this indicates that pleurostyles in Clupeoidei and Ostariophysi have convergently evolved [13, 14].

**Molecular and morphological evidence for otophysean interrelationships**

Figure 3 shows Ostariophysi splitting in two major clades where Gonorynchiformes are sister to the Otophysa (Characiformes, Cypriniformes, Gymnotiformes, and Siluriformes). Within Otophysa, Cypriniformes are sister to a clade comprising Characiformes *Alestes* and *Distichodus* as well as *Gymnotus* sp. representing the Gymnotiformes and *Schilbe* as a representative taxon of Siluriformes. Only recently, the complex phylogeny of this clade has been in the focus of phylogenomic level analyses [32, 33, 55]. A major point of discussion of these studies is the monophyly of Characiformes. While [32, 55] present a phylogeny in congruence to morphological data, [33] do not recover Characiformes as monophyletic.

In our study, the analyses of concatenated amino acid data (Table 1, Additional file 1: Figures S1, S2) reflect the interrelationships of *Gymnotus* and *Schilbe* as sister clade to monophyletic Characiformes. In all other analyses these interrelationships are mixed up showing for example a sister group relationship of *Distichodus* (Characiformes) and *Schilbe* (Siluriformes) or remain unresolved (Additional file 1: Figures S3, S4, S5, S6, S7, S8). We conclude that we did not sample enough variation of these lineages capturing insufficient phylogenetic signal for resolving these interrelationships with confidence. Figure 3 shows a phylogeny derived from the maximum likelihood analysis of concatenated nucleotide alignments, which were identified as significantly more likely compared to other phylogenetic estimates by the AU Test [48–51] (Additional file 1: Table S4). However, only the phylogenetic reconstructions based on the concatenated amino acids (Additional file 1: Figure S1 and S2) align with morphological synapomorphies suggested to characterize these clades. Regarding suggested synapomorphic characters, the Weberian apparatus was considered to characterize the Otophysa as taxonomic unit by Rosen and Greenwood [23] as well as Greenwood [56]. Further, four major clades (i.e. Cypriniformes, Characiformes, Gymnotiformes and Siluriformes) and their interrelationships were strongly supported by several morphological characters considered apomorphic [6, 26]. Contrasting, results of previous molecular phylogenetic studies did not support the monophyly of characiforms [33, 57–61], or the sister-group relation of siluriforms and gymnotiforms, which is strongly supported by several morphological apomorphies [24].

**Phylogeny of Euteleostei**

Our analyses clearly recover monophyletic Euteleostei. So far, only three apomorphies were listed in a previous morphological review for the Euteleostei [6]: (1) a stegural, (2) caudal median cartilages, and (3) a unique pattern of supraneural shape and development. The first two characters are challenged by alepocephaliforms clustering among Otomorpha. Alepocephaliforms show a stegural, which is a uroneural 1 with anterodorsal membranous outgrowth [6], and caudal median cartilages (Fig. 5a-c). Presently only the unique supraneural pattern (pattern 2 in Johnson and Patterson [13]) remains as synapomorphic character for
the Euteleostei, as alepocephaliforms do not show the respective character state.

The *Lepidogalaxias* lineage forms the sister taxon to all remaining euteleost fishes. This endemic West-Australian freshwater species unites several unique morphological characters and may actually be of key importance to understand the early evolution and extant diversity of euteleosts. Its unexpected phylogenetic position further calls morphological features into question, which were used to characterize interrelationships of *Lepidogalaxias* and galaxiids [29], as these characters are shared between both taxa, although they are showing no close phylogenetic relationship in any molecular phylogenetic analysis ([1–3, 5, 29, 55, 61], this study). This leads us to conclude that some of these characters are based on convergent evolution, while others may indeed be of plesiomorphic state. The phylogenetic position of galaxiids remains enigmatic. While Betancur-R et al. [1] suggest them to be part of the supergroup Protacanthopterygii, Near et al. [5] suggest Galaxiiformes to be the sister to all neoteleost lineages with high node support. Recently, in their phylogenomic level study, Hughes et al. [55] report on some cases of incongruence of gene trees and conflicting phylogenetic hypothesis. Here, Galaxiiformes form the
sister group to neoteleosts in all analyses with high support. An estimate ICA value of 0.254 for this split suggests less incongruence compared to 0.078 in [55]. Our results are therefore in favour of Hughes et al.’s [55] hypothesis H1, which may be caused by novel sequencing information from the genus Lovettia (Aplochitonidae, respectively Aplochitoniniae) representing the sister group of all other galaxiids [62]. Thus we suggest Galaxiiformes to form the sister group to neoteleost fishes and that thus Protacanthopterygia do not include Galaxiiformes. They are likely of major importance for understanding the evolution of extant Neoteleostei. Development of characters through ontogeny of both lineages are crucial to identify synapomorphies, which may allow for an update for the morphological synapomorphies of neoteleost fishes.

As aforementioned, Protacanthopterygia sensu Betancur-R. et al. [1, 2] cannot be recovered, as Galaxiiformes do not cluster along with other Prothacanthopterygia in any of our analysis. Fig. 3 shows Argentiniformes as sister to a clade comprising Esoci- and Salmoniformes. This sister group relationship cannot always be recovered, as indicated by low node support via bootstrapping as well as low IC values indicating incongruence (Fig. 3, Table 1). The Bayesian inference analysis of the concatenated amino acid alignment contradicts these interrelationships and suggest Argentiniformes as sister to Stomiati, Galaxiiformes and Neoteleosts (Table 1, Additional file 1: Figure S5). Contrasting, all other analyses of both amino acid and DNA based gene trees align with results from the maximum likelihood analysis from the concatenated DNA and alignment shown in Fig. 3 (Table 1, Additional file 1: Figures S1, S2, S3, S4, S6, S7, S8). Betancur-R. et al. [2] and Hughes et al. [55] discuss the difficult situation for Protacanthopterygia and characterize their classification as sedis mutabilis [2]. Due to the very different phylogenetic hypothesis published for the group, morphological evidence supporting the group is virtually lacking. Candidate characters are cartilaginous epicentrals, and simple (not forked) epineurals and epipleurals [6], which are subject of ongoing studies.

Stomiati sensu Betancur-R. et al. [1, 2] are recovered in all our analyses. The sister group relationship of a mostly coastal marine and coastal freshwater lineage, the Osmeriformes, and an exclusively marine and comparably highly diverse deep-sea lineage, the Stomiati, is noteworthy and somewhat parallels the relationship of Alepocephaliformes and clupeomorphs. Stomiatiiformes share morphological characters of neoteleost fishes, prompting synapomorphies for the latter group. These characters mainly refer to the highly derived branchial musculature including a new muscle, the retractor dorsalis, present in Stomiatiiformes and neoteleosts [13, 63–65]. Morphological evidence for a relationship of Osmeriformes and Stomiatiiformes is presently still lacking. The family Retropinnidae contains several freshwater and brackish water species. Our phylogenetic reconstruction shows that Retropinna is sister to all other osmeriforms (Fig. 3). All phylogenetic analyses performed in this study result in a well-supported sister group relationship of Stomiati to the galaxiform-neoteleost clade (Table 1, Fig. 3, Additional file 1: Figure S1–8). Morphological synapomorphies for this clade are still lacking.

However, our resolution of some deep phylogenetic nodes within the Clupeocephalan phylogeny will help allowing for reviewing morphological apomorphies and identifying candidate characters for the description and subsequent classification in the context of clades presented in this study.

Conclusions
Molecular phylogenies including the tree reconstruction presented herein have called numerous morphological apomorphies of clades into question, as topologies derived from morphology and molecular data differ significantly. The interpretation of suggested morphological synapomorphies of otomorph fishes is strongly affected by the inclusion of deep-sea inhabiting Alepocephaliformes. Our revision of these potential synapomorphies reveals that only a single character of nine characters in total can be flagged as synapomorphy valid for otomorphs. Three further characters remain possible apomorphies since their status cannot be evaluated without ambiguity in Alepocephaliformes.

Our phylogenetic estimate of Euteleost lineages shows that Protacanthopterygia, sensu Betancur-R. et al. [2] comprising Esoci-, Salmoni-, Argenti-, and Galaxiiformes, cannot be recovered, as Galaxiiformes appear to be the sister to all neoteleost fishes and further, a common ancestor of Salmoni-, Esoci and Argentinifomines is not well supported. It is noteworthy that another study contemporaneously aiming to solve the actinopterygian phylogeny with a very similar approach as the analysis presented herein, results in the same challenging nodes in their phylogenetic estimate [55]. Some splits remain uncertain, as e.g. the composition and phylogenetic placement of Protacanthopterygia or the phylogenetic placement of Alepocephaliformes, a group not included in [55]. Our efforts to reconstruct deep phylogenetic nodes based on a phylogenomic level dataset of clupeocephalan fishes still reveals problematic divergence estimates and calls for the application of phylogenomic methods on datasets with an enhanced taxon sampling, which can strengthen some of the phylogenetic hypothesis presented herein. Regarding morphology, data on possible characters for several nodes in the present
clupeocephalan tree are scarce. Therefore, subsequent morphological studies are required to understand character evolution, evolutionary driving forces and origin of species diversity in extant clupeocephalans.

**Methods**

**Material**
The major part of samples was collected during field trips and comprise muscle tissue or fin clips, respectively. The sampling aims for covering the major teleost lineages with a focus on Clupeocephala where neoteleosts, elopomorphs and osteoglossomorphs serve as an outgroup. Please see Additional file 1: Table S1 for an overview of samples analysed.

**Methods**

**Bait design**
To retrieve blueprints for bait sequences, we used the online resource Evolmarkers [66, 67] to search for putative ortholog exon loci in publicly available reference genomes. In a first step, we searched the genome of the zebra fish, *Danio rerio* (Cypriniformes), for single-copy loci using standalone BLAST [34]. In a second step, the results were subsequently BLASTed [34] against further available bony fish genomes, at the time of bait design comprising *Anguilla anguilla*, *Oryzias latipes*, *Tetraodon nigroviridis*, *Lepisosteus oculatus*, *Gadus morhua*, *Gasterosteus aculeatus* and *Oreochromis niloticus*. Finally, only exon sequences with a single BLAST [34] hit in all analysed genomes were used for bait design. Custom RNA baits were manufactured by Arbor Biosciences (Ann Arbor, Michigan, USA) with a length of 120 nucleotides and 60 nucleotides overlap after padding sequence lengths totaling 39,049 unfiltered baits with a 2× flexible tiling density. After removing all baits with any soft-masked sequence, 38,318 baits were put into production.

**Library preparation**
Genomic DNA was extracted from tissue samples listed in Additional file 1: Table S1 using the Machery & Nagel blood and tissue kit*. The DNA content of the final eluate was measured using a Qbit® Fluorometer (Life Technologies) applying the broad range kit. Thereafter, 130 μl with a concentration of at least 3 ng/μl DNA were used for shearing the DNA to ~ 500 bp using a Covaris® Sonicator. Shearing success was checked using gel-electrophoresis. The following steps for Illumina (Illumina, Inc., San Diego, CA) library constructions are based on Li et al. [68] and comprise a size selection step for fragments > 500 bp, blunt end repair using polymerase, adaptor ligation, fill-in and a final amplification of libraries using the KAPA® library amplification kit. DNA content of libraries was measured using a Qbit® fluorometer applying the high sensitivity kit and further checked with gel-electrophoresis to check the size distribution of fragments.

**Target capture**
For the performance of interordinal target capture, amplified libraries from step 2.2 served as starting point for hybrid enrichment. All steps follow the protocol provided in Li et al. [68]. Summarizing, library fragments are hybridized to RNA baits, remaining fragments and unintentionally hybridized fragments are washed off. During hybridisation, blocking oligos are preventing adaptor to adaptor ligation, while human cot DNA serves to avoid repetitive elements to cause non-specific binding. We applied a touch down hybridisation with decreasing hybridisation temperature from 65° to 50 °C in steps of 11 h totaling 36 h of hybridisation. The captured library is again amplified, size selected [69] and used as a starting point for a second round of capture, which is shown to increase the number of genes captured [68, 70].

During the final amplification step, individual sequencing indices are implemented to the adaptors allowing for de-multiplexing of reads after sequencing on an Illumina MiSeq® instrumentation. We aimed for an average coverage of 6.6 million paired-end 250 basepair reads per sample.

**Data analysis**
**Recovery of exon sequence alignments of phylogenetically informative loci** Sequencing reads were first checked for quality and low quality reads were excluded from further analysis with a cut-off value of 20. Adaptors were trimmed from reads using Trimgalore vers 03.07 [71, 72]. Thereafter, we followed the analysis pipeline introduced for target capture data in Yuan et al. [37]. Trimmed reads are first searched for replicate sequences, which are subsequently removed. For that, the first 20 bp of both reads are compared, if identical, they are removed. In a next step, the reads are BLASTed [34] against the bait sequences to sort the reads into corresponding gene bins. Next, reads are *de-novo* assembled into contigs using Trinity vers. 2.2.0 [73]. Output contigs are then separated into folders containing one or more than one contig. Where Trinity was creating more than a single contig sequence, Geneious® R7 was used to further assemble multiple contigs in an effort to create even longer contig sequences. For retrieving the best sequences of each gene in comparison to the bait sequences (query sequences were the bait sequences derived from the *Danio rerio* genome), we predicted the frame of each query sequence using a custom python script (*predict.frame.py*), which is available for download at the Dryad data
depository [39], and trimmed stop codons from it. Subsequently, contigs were translated into amino acids. All contigs were reciprocally blasted against the query sequences to check for homologs, i.e. contigs showing the best blast hit out of the target region were excluded from further analytical steps. As we are performing target capture on inter-ordinal level, the rate of gene losses or duplications is unknown, therefore single copy genes identified in Evolmarkers [66, 67] are not necessarily single copy in phylogenetically distant taxa. Therefore, contigs, which did not pass the reciprocal blast screening, were excluded from further analyses. Finally, intron inserted sequences were merged and subsequently translated to amino acids. We used customized Perl scripts to batch align each gene bin file containing all captured taxa and the bait sequence using MAFFT [74, 75]. As cross-contamination poses a problem in NGS datasets (e.g. [76]), possible cross-contamination were checked using a custom perl-script [39], which uses p-distances computed from single loci alignments and information on taxon groups assumed to be closely related. These are then compared with more distantly related taxa. Potential cross-contamination is indicated by extremely small p-distances (equal to or smaller than 0.002) between distantly related taxa (Additional file 1: Table S2). Although p-distance between conserved loci can be extremely small between distantly related taxa, conservation cannot be ubiquitous among all loci. Thus, there is no cross-contamination between a pair of taxa, if percentage of potential cross-contamination between them among all loci is extremely low. Subsequently, cleaned single loci were concatenated using Geneious® R7. The concatenated sequences were checked for phylogenetically most informative loci using the software Matrix Reduction [35], which is based on the treelikeness computed for single loci [36]. After extracting 838 most informative loci [39], we used customized Perl scripts to batch align each gene bin file containing all captured taxa and the bait sequence using MAFFT [74, 75]. As an additional check for orthology, the 838 loci identified with the Matrix reduction software, were analysed using Orthograph vers. 0–6–3-1 [38] and the custom script reblast.pl (Additional file 1: Table S3). Applying the latter approach, another four loci were eliminated from the dataset. Finally, five loci were excluded as they contained less than four sequences.

Phylogenetic analyses

**Concatenated datasets**

Concatenation was performed in Geneious R7 on 829 phylogenetically informative loci suggested from the MARE [35, 36] analysis (Additional file 1). We analysed two datasets, i.e. the amino acid alignments as well as the DNA alignments. For finding best-fitting partitions for the concatenated datasets, Partition finder vers. 2.1.1 [40–42] was used. A phylogenetic analysis based on maximum likelihood was performed in RAxML vers. 8.2.4 [42] incorporating the best-fitting partitioning schemes. RAxML [42] settings were applying the GTR GAMMA substitution model. Bootstrapping was halted automatically [77] using the fast hill-climbing algorithm.

As an alternative, we computed a phylogenetic tree using a Bayesian inference applying the CAT dirichlet process [78, 79] implemented in PhyloBayes vers. 4.1c [43]. Two chains were run in parallel and checked for convergence using the tracecomp and bccomp scripts provided in PhyloBayes.

All abovementioned analyses were performed in CIPRES [80].

**Coalescence analysis**

For a comparison of phylogenies computed from the concatenated dataset and a coalescence-based approach, we further performed maximum likelihood tree searches on single DNA and amino acid loci alignments using RAxML [42] on batch for attaining a collection of gene trees from both amino acid and DNA datasets. Those were subsequently used for estimating the species tree in ASTRAL vers. 4.10.12 [44].

**Computing tree certainty and performing AU test**

We computed internode certainty (IC/ ICA) and tree certainty (TC/ TCA) values [45, 46] from partial gene trees from the gene trees depicted from the coalescence analysis as implemented in RAxML [42, 47] using the best tree resulting from the maximum likelihood analysis of the concatenated and best-partitioned amino acid alignments. This step was used to evaluate incongruence among trees.

For testing for significant differences of species trees and fully bifurcating trees based on concatenated alignments, we performed an AU test in CONSEL [48–51].

**Morphological comparative material**

Cleaned and double stained collection specimen:

Osteoglossiformes. Osteoglossidae: *Osteoglossum biceps* (Cuvier, 1829): DMM IE/11035, 95.5 mm SL.

Elopiformes. Elopidae: *Elops senegalensis* Regan, 1909: DMM IE/11008, 61.3 mm SL.

Clupeiformes. Denticipitidae: 3 *Denticeps clupeoides* Clausen, 1959: DMM IE/11417, IE11420, 29.2–41.1 mm SL. *Clupea harengus* Linnaeus, 1758: DMM IE/ 11,029 83.1 mm SL.

Alepocephaliformes. Alepocephalidae: *Alepocephalus bicolor* Alcock, 1891: DMM IE/9602, 192 mm SL. *Xenodermichthys copei* (Gill, 1884) DMM IE/10190, 110.1 mm SL. *Platyroctidae: 5 Holbymia anomala* Kretz, 1980: DMM IE/10079, IE 10079, IE 6145, IE 4885, 55.99 mm -
144.4 mm SL; Maulisia argipalla Matsui & Rosenblatt, 1979; DMM IE/10459, 115.6 mm SL. Normichthys operatus Parr, 1951, DMM IE/11040, 97.1 mm SL; Searsia koefoedi Parr 1937; DMM IE/10191, 117.6 mm SL.

Gonorynchiformes. Gonorynchidae: Gonorynchus abbreviates Temminck & Schlegel, 1846, DMM IE/11730, 84.2 mm SL; Chanidae: Chanos chanos (Forsskål, 1775); DMM IE/11010, 72.18 mm SL; Kneridae: Kneria stappersii Boulenger, 1915, DMM IE/12025, 26.4 mm SL.

Cypriniformes. Cyprinidae: 2 Dawkinsia tambraparniei (Silas, 1954): DMM IE/12072, 27.8 mm SL, 28.6 mm SL.

Argentiniformes. Argentinidae: Argentina silus (Ascanius, 1775); DMM IE/11033, 103.2 mm SL; Bathylagidae: Bathyergus euryops Goode & Bean, 1896; DMM IE/11034, 96.3 mm SL.

Osmeriformes. Osmeridae: Osmerus eperlanus (Linnaeus, 1758), DMM IE/11090, 36.5 mm SL.

Salmoniformes: Thymallidae: Thymallus thymallus (Linnaeus, 1758) DMM IE/11820, 99.5 mm SL.

The specimens were cleared and double stained following [81, 82]. Specimens were transferred into 98% ethanol. Afterwards, cartilage was stained with Alician blue in 1:4 acetic acid and ethanol-solution for maximally 48 h. Thereafter, the specimens were transferred via a decreasing alcohol concentration in digestion solution with trypsin. As soon as specimens were cleared, the pigmentation of the skin was eliminated by a bleach bath of potassium hydroxide solution and addition of hydrogen peroxide. Subsequently, bones were stained with Alizarin-red. Finally, specimens were transferred into glycerine for increasing the transparency.

The dissected parts of the specimens were photographed with a Canon EOS 5D0 with a Sigma 105 mm lens and the software EOS Utility 3.0 (Canon). Stacking of images for obtaining advanced and extended focus images were done with software Helicon Focus 6. The images were edited in GIMP 2.8 and were compiled in Inkscape 0.92.1.

Additional file

Acknowledgements

We would like to acknowledge several colleagues for their kind help with sampling: Mark Adams (South Australian Museum); Katherine Pearson Maslenikov (Burke Museum), Samuel Iglesias (Museo nacional de Histoire naturelle), Alastair Graham (Commonwealth Scientific and Industrial Research Organisation), Ulrich Schliebeh, Frederic Schedel, and Dirk Neumann (all Bavarian State Collection of Zoology). Further, we would like to express our sincere thanks to Austéllin Teller, Saurabh Pophaly, and Hanna Mäkle (Technical University Munich) for clustering, David Langenberger (EcSeq), Mario Fasold (EcSeq) and Jörg Lesch (Lesch IT) for IT support.

Funding

This work was funded by the VolkswagenStiftung within the focal point program „Forschung in Museen” AZ-88 353 to TM and NS.

Availability of data and materials

The data sets supporting the results of this article are available in the Dryad data repository, doi:https://doi.org/10.5061/dryad.7b1f933; https://datadryad.org/revision?id=105061/dryad/7b1f933.

Authors’ contributions

TM and NS designed the study. Laboratory work was carried out by NS and MW. NS, CL, HY analyzed the data. NS, TM, MM, CL wrote and drafted the manuscript. All authors read and approved the final version of the manuscript.

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Publisher’s Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Author details

1Institut für Zoologie & Evolutionsbiologie, Friedrich-Schiller-Universität Jena, Erbertstraße 1, 07743 Jena, Germany. 2Zoologische Staatssammlung München, Staatliche Naturwissenschaftliche Sammlungen Bayerns, Münchnerhausenstraße 21, 81247 Munich, Germany. 3Key Laboratory of Exploration and Utilization of Aquatic, Genetic Resources, Shanghai Ocean University, Ministry of Education, Shanghai 201306, China. 4Deutsches Meeresmuseum, Karteninberg 14-20, 18439 Stralsund, Germany.

Received: 26 August 2017 Accepted: 26 September 2018

Published online: 23 October 2018

References

1. Betancur-R, R, Broughton RE, Wiley EO, Carpenter K, López JA, Li C, Holcroft NI, Arcila D, Sanciangco M, Cureton J, II, Zhang F, Buset T, Campbell MA, Ballesteros JA, Roa-Varon A, Willis S, Borden WC, Rowley T, Renua PC, Hough DJ, Lu G, Grande T, Arratia G, Ortí G. The tree of life and a new classification of bony fishes. PLOS Currents. 2013. doi:https://doi.org/10.1371/ currents.c7sba26640d8f0caee75b1656c82688.

2. Betancur-R, Wiley EO, Arratia G, Acero A, Bailly N, Mya M, Lecointre G, Ortí G. Phylogenetic classification of bony fishes. BMC Evol Biol. 2017. doi. https://doi.org/10.1186/s12862-017-0958-z.

3. Lavoué S, Mya M, Poulsen JY, Meller PR, Kinda M. Monophyly, phylogenetic position and inter-familial relationships of the Alepocephaliformes (Teleostei) based on whole mitogenome sequences. Mol Phylogenet Evol. 2008;47:1111–21.

4. Lavoué S, Mya M, Inoue JG, Saitoh K, Ishiguro NB, Nishida M. Molecular systematics of the gonorhynchiform fishes (Teleostei) based on whole mitogenome sequences: implications for higher-level relationships within the Otophysa. Mol Phylogenet Evol. 2005;47:1111–21.

5. Near TJ, Eytan RI, Dornburg A, Kuhn KL, Moore JA, Davis MP, Wainwright PC, Friedman M, Leo Smith W. Resolution of ray-finned fish phylogeny and timing of diversification. PNAS. 2012. https://doi.org/10.1073/pnas.1206625109.

6. Wiley EO, Johnson GD. A teleost classification based on monophyletic groups. In: Nelson JS, Schultz HP, MVH W, editors. Origin and phylogenetic interrelationships of Teleosts. München: Verlag Dr. F. Pfeil; 2010. p. 123–82.

7. Arratia G. The Clupeocephala re-visited: analysis of characters and homologies. Rev Biol Mar Oceanog. 2010;45(Suppl 1):s35–57.

8. Ishiguro NB, Mya M, Nishida M. Basal euteleostean relationships: a mitogenomic perspective on the phylogenetic reality of the Protacanthopterygii. Mol Phylogenet Evol. 2003;27:476–88.

9. Lavoué S, Mya M, Kawauchi A, Yoshino T, Nishida M. The phylogenetic position of an undescribed paedomorphic clupeiform taxon: mitogenomic evidence. Ichthyol Res. 2008;55:328–34.
10. Poulsen JY, Møller PR, Lavoué S, Knudsen SW, Nishida M, Miya M. Higher and lower-level relationships of the deep-sea fish order Alepsocephaliformes (Teleostei: Otocephala) inferred from whole mitogenome sequences. Biol J Linn Soc. 2009. https://doi.org/10.1111/j.1095-8312.2009.01235.x.
11. Greenwood PH, Rosen DE. Notes on the structure and relationships of the alepsocephaloid fishes. Am Mus Novit. 1971;2473:1–41.
12. Lecointre G, Nelson G. Clupeomorpha, sister-group of Otariprisci. In: Stsastys MLJ, Parenti LR, Johnson GD, editors. Interrelationships of fishes. San Diego: Academic Press. 1996. p. 193–207.
13. Johnson GD, Patterson C. Relationships of lower euteleostean fishes. In: Stsastys MLJ, Parenti LR, Johnson GD, editors. Interrelationships of fishes. San Diego: Academic Press. 1996. p. 251–332.
14. Arratia G. Basal teleosts and teleostean phylogeny. Palao Ichthyologica. 1997;7:1–168.
15. Arratia G. The monophyly of Teleostei and stem-group teleosts. Consensus and disagreements. In: Arratia G, Schultz HP, editors. Mesozoic fishes 2. Systematics and fossil record. München: Verlag Dr. Friedrich Pfeil. 1999. p. 265–334.
16. Arratia G, Schultz HP. Palatoquadrate and its ossifications: development and homology within osteichthyans. J Morph. 1991;208:1–81.
17. Arratia G. The varasichthyid and other crossognathiform fishes, and the break-up of Pangaea. In: Cavin L, Longbottom A, Richter M, editors., vol. 295. London: Geological society, Special Publications; 2008. p. 71–92.
18. Grande L. Recent and fossil clupeomorph fishes with materials for revision of the subclade of clupeiformes. Bull Am Mus Nat Hist. 1988;181:231–72.
19. Gayet M. Ramallichthys, Gayet du Cénomanien inférieur marin de Ramallah (Judée), une introduction aux relations phylogénétiques des Otariprisci. Mém Mus Natl Hist Nat Ser. 1986;51:21–31.
20. Daget J. Le crâne des télétostéens. Mém Mus Nat Hist Nat Ser A. 1964;31:163–362.
21. Roberts T. Osteology and relationships of characoid fishes, particularly the genus Hesperus, Salminus, Hoplosternum, and Aequistorchus. Proc Calif Acad Sci. 1969;36:391–500.
22. Blot J. Les Apodes fossiles du Monte Bolca. Stud. Ric. Giacimenti Terziari di Judée), une introduction aux relations phylogénétiques des Ostariophysi. Mém Mus Natl Hist Nat Ser. 1985;65:21–51.
23. Lanfear R, Frandsen PB, Wright AM, Senfeld T, Calcott B. PartitionFinder 2: new methods for selecting partitioned models of evolution for molecular and morphological phylogenetic analyses. Mol Biol Evol. 2016. https://doi.org/10.1093/molbev/msw260.
24. Shimodaira H. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics. 2014. https://doi.org/10.1093/bioinformatics/btu333.
25.колошина Р. PhyloBayes home page. 2017. http://phylobayes.org. Accessed 20 Mar 2017.
26. Mirarab S, Reaz R, Bayzid MS, Zimmermann T, Swenson MS, Warnow T. ASTRAL: genome-scale coalescent-based species tree estimation. Bioinformatics. 2014. https://doi.org/10.1093/bioinformatics/btu682.
27. Salichos L, Rokas A. Inferring ancient divergences requires genes with strong phylogenetic signals. Science. 2013;340:327–31.
28. Salichos L, Stamatakis A, Rokas A. Novel information theory-based measures for quantifying incongruence among phylogenetic trees. Mol Biol Evol. 2014;31:1261–71.
29. Kivell T, Salichos L, Rokas A, Stamatakis A. Computing the intermediate certainty and related measures from partial gene trees. Mol Biol Evol. 2016. https://doi.org/10.1093/molbev/msw040.
30. Shimodaira H. Multiple comparisons of log-likelihoods and combining nonnest models with applications to phylogenetic tree selection. Comm. in Statist., Part A - Theory Meth. 2001;30:1751–72.
31. Salichos L, Rokas A, Stamatakis A. Computer programs for combining phylogenetic trees. Mol Biol Evol. 2014;31:1261–71.
32. Kivell T, Salichos L, Rokas A, Stamatakis A. Computing the intermediate certainty and related measures from partial gene trees. Mol Biol Evol. 2016. https://doi.org/10.1093/molbev/msw040.
33. Shimodaira H. A computer package for combining phylogenetic trees. Mol Biol Evol. 2014;31:1261–71.
34. Shimodaira H. A computer package for combining phylogenetic trees. Mol Biol Evol. 2014;31:1261–71.
35. Shimodaira H. A computer package for combining phylogenetic trees. Mol Biol Evol. 2014;31:1261–71.
36. Shimodaira H. A computer package for combining phylogenetic trees. Mol Biol Evol. 2014;31:1261–71.
37. Shimodaira H. A computer package for combining phylogenetic trees. Mol Biol Evol. 2014;31:1261–71.
38. Shimodaira H. A computer package for combining phylogenetic trees. Mol Biol Evol. 2014;31:1261–71.
39. Shimodaira H. A computer package for combining phylogenetic trees. Mol Biol Evol. 2014;31:1261–71.
60. Chen WJ, Lavoué S, Mayden RL. Evolutionary origin and early biogeography of otophysan fishes (Ostariophysi: Teleostei). Evolution. 2013;67:2218–39.

61. Campbell MA, Alfaro ME, Belasco M, Lopez JA. Early-branching euteleost relationships: areas of congruence between concatenation and coalescent model inferences. PeerJ. 2017. https://doi.org/10.7717/peerj.3548.

62. Burridge CP, McDowall RM, Caw O, Wilson MH. Waters JM. Marine dispersal as a pre-requisite for Gondwanan vicariance among elements of the galaxiid fish fauna. J Biogeogr. https://doi.org/10.1111/j.1365-2699.2011.02600.x.

63. Rosen DE. Interrelationships of higher euteleostean fishes. In: Greenwood PH, Miles RS, Patterson C, editors. Interrelationships of fishes. London: Academic Press; 1973. p. 397–513.

64. Johnson GD. Monophyly of the euteleostean clades: Neoteleostei. Eurypterygii and Ctenosquamata Copeia. 1992;1:8–25.

65. Springer VG, Johnson DG. Study of the dorsal gill-arch musculature of teleostome fishes, with special reference to the Actinopterygii. Bull Biol Soc Wash. 2004;11:236–60.

66. Li C, Riethoven J-JM, Ma L. Exon-primed intron-crossing (EPIC) markers for non-model teleost fishes. BMC Evol Biol. 2010. https://doi.org/10.1186/1471-2148-10-90.

67. Li C, Ort G, Zhang G, Guoquing L. A practical approach to phylogenomics: the phylogeny of ray-finned fish (Actinopterygii) as a case study. BMC Evol Biol. 2007. https://doi.org/10.1186/1471-2148-7-44.

68. Li C, Hofreiter M, Straube N, Corrigan S, Naylor GJP. Capturing protein-coding genes across highly divergent species. BioTechniques. 2013;54(6):321–6.

69. Fisher S, Barry A, Abreu J, Minie B, Nolan J, De loreyT M, YounG , FennellT J, Allen A, Ci bulskis K, Friedrich D, Johnson R, John F, Reilly B, Sharrma R, Stalker J, Walsh J, Zimmer A, Zwirko Z, Robert N, Nusbaum C. A scalable, fully automated process for construction of sequence-ready human exome targeted capture libraries. Genome Biol. 2011. https://doi.org/10.1186/gb-2011-12-1-r1.

70. Springer MS, Signore AV, Paijmans JL, Vélez-Juarbe J, Domning DP, Bauer CE, He K, Cesar L, Campos PF, Murphy WJ, Meredith RW, Gatesy J, Willerslev E, MacPhee RD, Springer MS, Hofreiter M, Campbell KL. Interordinal gene capture, the phylogenetic position of Steller’s sea cow based on molecular and morphological data, and the macroevolutionary history of Sirenia. Mol Phylogenetics Evol. 2015;91:178–93.

71. Martin M. Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet J. 2011. https://doi.org/10.14806/ej.17.1.200.

72. Krueger F. (2016) Trimgalore https://www.bioinformatics.babraham.ac.uk/projects/trim_galore/. Accessed 24 April 2018

73. Haas BJ, Papanicolaou A, Yassour M, Grabherr M, Blood PD, Bowden J, Couger MB, Eccles D, Li B, Lieber M, MacManes MD, Ott M, Orvis J, Pochet N, Strozzi F, Weeks N, Henschel R, LeDuc RD, Friedman N, Regev A. De novo transcript sequence reconstruction from RNA-Seq: reference generation and analysis with Trinity. Nature Protoc. 2013. https://doi.org/10.1038/nprot.2013.084.

74. Katoh K, Misawa K, Kuma KI, Miyata T. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. Nuc Ac Res. 2002. https://doi.org/10.1093/nar/gqf436. PMC 135756.

75. Katoh K, Kuma KI, Toh H, Miyata T. MAFFT version 5: improvement in accuracy of multiple sequence alignment. Nuc Ac Res. 2005. https://doi.org/10.1093/nar/gkl001. PMC 548945.

76. Ballenghien M, Faivre N, Galtier N. Patterns of cross-contamination in a multispecies population genomic project: detection, quantification, impact, and solutions. BMC Biol. 2017. 10.1186/s12915-017-0366-6.

77. Pattengale ND, Alipour M, Bininda-Edmonds ORP, Moret BME, Stamatakis A. How many bootstrap replicates are necessary? J Comput Biol. 2010. https://doi.org/10.1089/cmb.2009.0179.

78. Lartillot N, Philippe H. A Bayesian mixture model for across site heterogeneities in the amino-acid replacement process. Mol Biol Evol. 2004;21(6):1095–109.

79. Lartillot N, Brinkmann H, Philippe H. Suppression of long-branch attraction artefacts in the animal phylogeny using a site-heterogeneous model. BMC Evol Biol. 2007. https://doi.org/10.1186/1471-2148-7-54.

80. Miller MA, Pfeiffer W, Schwartz T. Creating the CIPRES Science Gateway for inference of large phylogenetic trees. In: Proceedings of the Gateway Computing Environments Workshop. San Diego; 2010. p. 1–8.

81. Dingerkus G, Uhler LD. Enzyme clearing of Alcian blue stained whole small vertebrates for demonstration of cartilage. Stain Technol. 1977;52:229–32.

82. Taylor WR, CVD VO. Revised procedures for staining and clearing small fishes and other vertebrates for bone and cartilage study. Cybium. 1985;9:107–19.