WILEY-BLACKWELL

Cladistics (2016) 1–18

Cladistics

10.1111/cla.12171

Combined phylogeny of ray-finned fishes (Actinopterygii) and the use of morphological characters in large-scale analyses

Juan Marcos Mirande*

Unidad Ejecutora Lillo (UEL, Fundación Miguel Lillo-CONICET), San Miguel de Tucumán 4000, Argentina

Accepted 30 June 2016

Abstract

This study evaluates the phylogeny of ray-finned fishes (Actinopterygii) combining most available information (44 markers from nuclear and mitochondrial DNA and 274 morphological characters). The molecular partition of the dataset was produced through a pipeline (GB-to-TNT) that allows the fast building of large matrices from GenBank format. The analysed dataset has 8104 species, including representatives of all orders and 95% of the 475 families of Actinopterygii, making it the most diverse phylogenetic dataset analysed to date for this clade of fishes. Analysed morphological characters are features historically considered diagnostic for families or orders, which can be unequivocally coded from the literature. Analyses are by parsimony under several weighting schemes. General results agree with previous classifications, especially for groups with better gene sampling and those long thought (from morphological evidence) to be monophyletic. Many clades have low support and some orders are not recovered as monophyletic. Additional data and synthetic studies of homology are needed to obtain synapomorphies and diagnoses for most clades.

Most recent phylogenetic hypotheses and classifications of the Actinopterygii have been based either on nuclear DNA (e.g. Betancur-R et al., 2013a), molecular DNA (e.g. Miya et al., 2013) or morphology (e.g. Wiley and Johnson, 2010). The aims of this study are to propose a phylogenetic hypothesis from the combination of all those data and to provide some guidelines and discuss problems associated with the combination of different kinds of information. This phylogenetic analysis is the most diverse for the Actinopterygii and most of its orders and families, including more than 8000 analysed species. It is also the first large-scale hypothesis of the Actinopterygii based on both molecular and morphological data. Morphological characters are not frequently used in large-scale phylogenies and especially in analyses containing several hundreds or thousands of species. This practice may be due to the difficulty of managing such a volume of information, in addition to the challenges related to generating and/or compiling so much morphological data.

The Actinopterygii include the vast majority of fishes, with approximately 32 000 extant species (Eschmeyer and Fong, 2016) that represent about half of vertebrate diversity. The systematics of ray-finned fishes has undergone many changes during the last century and there are still many uncertainties about the phylogenetic relationships, even at the ordinal level, of entire families and the monophyly of many highly diverse groups. Wiley and Johnson (2010) synthesized decades of morphological knowledge in the first published classification of the Actinopterygii based on hypothesized monophyletic groups diagnosed by morphological synapomorphies. That classification included a split of the highly diverse Perciformes, whose monophyly had been previously challenged (e.g. Nelson, 2006). There have also been several contributions to the phylogeny of the Actinopterygii based on molecular data, from either mitochondrial (e.g. Miya

*Corresponding author:
E-mail address: mcmirande@gmail.com

et al., 2003) or (mostly) nuclear DNA (Betancur-R et al., 2013a). However, no attempt to combine those sources of information in a large-scale phylogeny of the Actinopterygii has yet been published.

The proposal by Betancur-R et al. (2013a) constitutes the most comprehensive phylogenetic hypothesis of the Actinopterygii published to date. They also completed the largest genome exploration for nuclear DNA markers useful for phylogenetic analyses on the Actinopterygii and sequenced those markers for many species. Betancur-R et al. (2013a) analysed 1416 species and 21 molecular markers (20 nuclear and one mitochondrial), with representatives of 396 of the 504 families of Actinopterygii recognized in that article. Their classification is based only on molecular data and therefore it is unclear for most of the clades which, if any, are the morphological synapomorphies.

Recently, there has been an increased production of new markers and sequences for many species across the Tree of Life, but the use of morphological data on a large scale for phylogenetic analyses has received comparatively less attention, with some noticeable exceptions in other groups of vertebrates (e.g. Livezey and Zusi, 2007; O'Leary et al., 2013). Those exceptions, however, have been focused mainly on increasing the number of characters, rather than the number of species, which is one premise of this analysis. This paper highlights the importance and utility of morphological data in the context of large-scale phylogenies and shows that incorporating well-established morphological knowledge into phylogenetic analyses is a useful exercise.

Material and methods

All sequences have been obtained from GenBank (accession numbers are provided in the dataset; Appendix S1). The data matrix has been built with the program GB-to-TNT (Goloboff and Catalano, 2012), aligned with Muscle (Edgar, 2004) and analysed under parsimony with TNT (Goloboff et al., 2008). The list of analysed genes and details of the phylogenetic analyses are provided below.

GB-to-TNT (Goloboff and Catalano, 2012) software allows us to generate large TNT matrices combining different markers from previously downloaded GenBank files. With the aid of this program, DNA sequences of each marker were extracted from GenBank files containing hundreds or thousands of accessions, depending on the gene. Duplicate sequences for the same taxa, hybrids and samples not identified to species level were filtered and removed, also with GB-to-TNT. This software may call alignment programs and produce a file ready to be analysed with TNT. However, intermediate Fasta files were produced to be aligned with specific parameters of Muscle (Edgar, 2004) and inspected with BioEdit (Hall, 1999), before being compiled again with GB-to-TNT. The complete systematics of each terminal taxon, as provided in GenBank, is included in the dataset together with the species name and accession number for each molecular block. This information embedded in the TNT file permits a quick reference to clades with TNT, something useful for evaluating results and managing morphological characters, as explained below. More details about GB-to-TNT software are provided by Goloboff and Catalano (2012).

The taxonomy of each species, obtained from GenBank, was updated following Eschmeyer (2016), as an approach to the systematic classification currently accepted for the ichthyological community for most families. Possible contamination of sequences has been detected through the evaluation of superficial parsimony searches complemented with blastn (Altschul et al., 1990). A list of sequences deleted from the original dataset due to possible contamination or problems in the species identification is provided as Appendix S2.

Taxa

The premise of this study was to analyze as many species as possible. Although some clades have historically been considered monophyletic, the internal relationships among their species, even if represented by just a few known sequences, are potentially useful for the resolution of the phylogenetic tree. Therefore, all species for which there was information on GenBank for each of the analyzed DNA markers are included in the complete dataset, which contains data for 14 141 species. This dataset was further restricted to species having at least three DNA markers except for representatives of a few families, of which some species with two or, exceptionally, one sequence were also analyzed. The few species analyzed from < 3 known DNA sequences are either representatives of families that had not been previously included in large-scale analyses or species selected to improve the overlapping of genes in families with low taxonomic sampling. Although the reliability of results depends in good measure on the amount of data, missing data itself has been shown (e.g. Goloboff et al., 2009) not to affect large phylogenetic analyses too much when data from different blocks have some degree of overlap. The inclusion of some species with < 3 known DNA sequences is based on that premise. Only 48 species (about 0.59%) are analyzed from < 3 DNA sequences. Those species are members of the families Amphiliidae (1 species), Aploactinidae (1), Aspredinidae (4), Barbuccidae (1), Batrachoididae (2), Cyttidae (1), Dinopercidae (1), Erethistidae (13), Gnathanacanthidae (1),

Goodeidae (4), Grammicolepididae (1), Heptapteridae (9), Heterenchelyidae (1), Normanichthyidae (1), Pinguipedidae (1), Plectrogeniidae (1), Psettodidae (1), Serpenticobitidae (1), Tetrabrachiidae (1), Tetrarogidae (1), and Valenciidae (1). Among them, the cypriniforms *Barbucca diabolica* (Barbuccidae) and *Serpenticobitis zonata* (Serpenticobitidae) are analyzed from only one DNA sequence each. In the final dataset, only 18 families of Actinopterygii (out of 475) are not represented by at least one sequence of one species (Table 2).

Characters and dataset management. The data matrix includes approximately equal proportions of nuclear and mitochondrial DNA data, with the addition of 274 morphological characters. The dataset consists of 45 blocks, one of morphology and 44 of DNA. Molecular data include 41 coding genes: 13 mitochondrial (*atp6*, *atp8*, *cox1*, *cox2*, *cox3*, *cytb*, *nd1*, *nd2*, *nd3*, *nd4*, *nd4 l*, *nd5* and *nd6*) and 28 nuclear (*egr1*, *egr2b*, *egr3*, *enc1*, *ficd*, *glyt*, *h3*, *irbp*, *kiaa1239*, *mll4*, *myh6*, *panx2*, *plagl2*, *ptr*, *rag1*, *rag2*, *rh*, *ripk4*, *rnf213*, *sh3px3*, *sidkey*, *sreb2*, *svep1*, *tbr1*, *tmo-4c4*, *ube3a*, *vcpip* and *zic1*) and three ribosomal genes (the mitochondrial *12s* and *16s* and the nuclear *28s*). Most alignments are trivial, with relatively few gaps. Only the ribosomal blocks had problems of alignment in some regions and these have been removed from the dataset. The analysed dataset has 30 970 characters. Sequences for most markers are unknown for many species and therefore the matrix has a large proportion (79%) of missing entries. Four species have as much information as 38 molecular blocks, while the average is about 9.12 DNA sequences per terminal taxon. The number of analysed sequences is 73 916, out of a total of what would be 356 576 if all species had information for all DNA markers. A file with the number of sequences analysed for each species is provided as Appendix S3. The dataset is available online at MorphoBank (O'Leary and Kaufman, 2011, 2012).

Most analysed nuclear DNA sequences are the same used by Betancur-R et al. (2013a,b), but there are many additional available sequences, especially from mitochondrial DNA, that are also included. As examples, complete mitochondrial sequences for more than 1400 species of Actinopterygii and a very large number of *cox1* and *cytb* sequences (of 9285 and 7306 species, respectively) are included in the complete dataset.

Morphological characters are defined and coded based on more than 200 papers including phylogenetic analyses, systematic revisions, morphological surveys and species descriptions (see Appendices S4 and S5 for details). These characters represent only a small sample of the morphological variability known to be present among the Actinopterygii, but the lack of comprehensive morphological phylogenies for many groups and the relatively few studies of homologies across orders of fishes prevent a more comprehensive use of data from the literature. Most of the characters analysed comprise features known for decades or centuries to be diagnostic of groups [(e.g. the presence of a suprabranchial organ in the Anabantidae, described by Cuvier and Valenciennes (1831)]. Such characters have been selected for their stability across families and orders or because exceptions in some taxa are well documented. Many morphological characters have been previously synthesized by Nelson (2006), although most of the characters herein analysed have been traced to the papers that served as sources for that author. The presence or absence of complete structures are analysed as phylogenetic characters, regardless of whether these features are considered diagnostic of some clades in the literature. For example, the loss of basisphenoid was considered as a synapomorphy of the Ostariophysi (Fink and Fink, 1981), but that absence was reported for many other fishes for which no phylogenetic inferences have been derived [e.g. the monotypic Luvariidae (Tyler et al., 1989)] and such information was scored in the analysis. Morphological characters dealing with form of structures that are difficult to compare in fishes of different clades or having high intrafamilial or some intrageneric variability have been excluded from this analysis. As a general rule, it was always preferred to leave some potentially useful morphological variation out of the analysis rather than including dubious information. All morphological characters are defined as binary, but some of them represent binary coding of multistate additive characters.

As the assignment of morphological states in a dataset with this number of species is almost impossible to perform in reasonable times, a TNT script (modified from Goloboff et al., 2009; provided as Appendix S5) was used to include information from the literature to the corresponding species. This script takes into account taxonomic groups instead of individual species, allowing us to code whole genera or suprageneric groups and consider the exceptions known in every case. It was not possible to score every morphological character directly for every terminal given the large number of terminals in this study. Thus, the following explicit assumption was made in many cases and was carried out by the TNT script: if reliable sources in the literature described a feature as characterizing a given clade, all terminal members of that clade were coded as having the feature. This was not validated by direct observation due to the large number of species terminals. Also, inapplicable cases have been coded with this script. Assigning states in this way makes it possible to create large blocks of morphology for many species, after a detailed examination of specimens and/or

review of the literature. Assigning states to morphological characters based on data from the literature and explicitly assuming that the presence of particular states in a sample of species means the states represent the condition of all species of a taxonomic group may introduce mistakes if the literature is incomplete or wrong.

The complete dataset, with 14 141 terminal taxa, containing GenBank accession numbers, and a dataset containing only the block of morphology are provided as Appendices S1 and S6.

Phylogenetic analysis. Analyses were done by parsimony with TNT (Goloboff et al., 2008), under equal weighting and extended implied weighting using different parameters, as explained below and in Table 1 (Goloboff, 1993, 2014). For comparative purposes, an analysis based only on molecular data was also done. Various analyses were also conducted to evaluate the influence of the morphological characters and weighting schemes and to have some estimation of clade stability, as a kind of sensitivity analysis. Results and discussions are based on the analyses under implied weighting.

Extended implied weighting (Goloboff, 1993, 2014) overcomes two potential problems when analysing combined datasets by differential weighting: the interdependence of each site with the remaining sequence, which is an argument against separately weighting each site, and the frequent lack of whole blocks of data in many species, which produce artificially high weights in sequences with many missing entries (see Goloboff, 2014). Extended implied weighting weights against the average homoplasy of either an entire gene or regions within each gene and takes into account the non-informative sites in the weight implied for variable (informative) characters (Goloboff, 2014). The configuration and extent of the regions used to calculate weights can be adjusted in TNT from one column (=site) to the entire sequence included in each block (=gene).

The rationale to group columns forming weighting sets is based on the hypothesis that nucleotides are comparatively more dependent on each other than morphological characters, due to their arrangement in sequences, and that the homology across columns is assessed only by a positional criterion, after the use of an aligning method. However, there are regions of each marker that can be considered as homologous between species with more confidence than individual columns. The same happens, naturally, with entire genes (discarding the problem of paralogues). This justifies some grouping of columns while calculating homoplasy to assign weights. For example, when grouping columns to form weighting sets, the comparatively higher homoplasy of third positions will not be reflected in the weights if the weighting sets are formed by successive columns. Therefore, if weighting sets are formed by contiguous positions, third columns will have higher weights than implied by their homoplasy, while first and second columns will have artificially low weights. TNT has the option of grouping first, second and third positions (after alignment) of each block as separate sets. This option may be considered the most reasonable, but it requires either a perfect alignment (without gaps) or the presence of gaps introduced exclusively in multiples of three positions (one codon). Such alignments may be the case for some genes or regions but are not equally applicable to all regions of all markers. It may be deduced that the more there is deviation from a "perfect" alignment, the worse the results such a weighting scheme will be. More details on extended implied weighting are provided by Goloboff (2014).

Six different analyses under extended implied weighting are included, weighting each molecular character by separate (SEP), grouping three (GR3), nine (GR9) and 27 (G27) columns, using the entire marker to average the homoplasy (GEN), and grouping first, second and third positions of each coding sequence as different weighting sets (POS). The number of columns grouped to collectively weight characters is rather exploratory, given that there are no published precedents for these kinds of analyses, but care was put into grouping multiples of three columns to represent

Table 1
Summary of the different searches done. Column groupings refer to the number of columns (characters) grouped to average their homoplasy and collectively weight them. Total weighting sets are the number of molecular sets of columns that are each collectively weighted

Search	Dataset	Weighting	Columns grouping	Total weighting sets
SEP	Combined	Implied	None	30 970
GR3	Combined	Implied	3	10 247
GR9	Combined	Implied	9	3431
G27	Combined	Implied	27	1160
GEN	Combined	Implied	Entire genes	44
POS	Combined	Implied	Codon positions for each coding gene	126
EQW	Combined	Equal	Not applicable	Not applicable
MOL	Molecular	Implied	3	10 247

different numbers of codons (one, three and nine, respectively). Only one analysis was done with the molecules-only dataset (MOL), using the parameters that better performed in the analyses from the complete matrix. Also, an analysis with equal weighting of the combined dataset was done (EQW). The different searches are summarized in Table 1.

Searches under extended implied weighting were made under a reference concavity constant (K) of 200 (Goloboff, 2014). With this value of K, a character with an average number of homoplastic steps (about 47 in an optimal tree under equal weights for this dataset) has approximately 80% of the weight of a character with no homoplastic steps. In previous experiences (Mirande, 2009; Mirande et al., 2011, 2013) the preferred values of K were those in which an average character had 62–74% the weight of a character free of homoplasy. It was decided herein to use a value of K less than in prior work given that coding of the morphological characters may have some problems derived from wrong or incomplete data in the literature. With this lower value of K, the weight of morphological characters would receive comparatively less weight than prior work.

Searches started from a relatively low number of initial Wagner trees and TBR swapping, with most of the time invested in sectorial searches and tree fusing (Goloboff, 1999), which are the most efficient methods available to analyse large matrices (Goloboff et al., 2009). Support was evaluated through Bremer support (Bremer, 1994), but given the size of the dataset and the limitations to obtaining and processing an appropriate number of trees, suboptimal trees used to calculate supports were restricted to a group of approximately 500 trees, composed as following: most parsimonious hypotheses obtained under all the parameters explored, trees thought to be optimal during searches under different parameters but eventually discarded as suboptimal, 100 Wagner trees, and those Wagner trees after TBR swapping interrupted after 5 and 10 min. This tree sample produced a rough approximation of trees with various degrees of suboptimality. Advantages of Bremer support relative to resampling methods are the lower computational demand and that all groups from the most parsimonious trees will have some support, even if it is very low in some cases. Because this is a highly approximate estimation of Bremer supports, the values obtained may be slightly to moderately overestimated.

Results

The complete dataset has 30 970 characters, of which the first 274 are morphological. There are 21 416 informative characters. Most parsimonious trees have between 2 629 862 (EQW) and 2 647 426 steps (SEP) (Tables 1 and 2). Most parsimonious trees from different analyses differ in some basal relationships, but there are 4274 clades (out of 8102 possible ones) that persist under all the conditions analysed, 4332 clades are shared by all the analyses of the combined dataset, while 4790 are shared by the analyses of the combined dataset under extended implied weighting.

The highest number of monophyletic families was obtained in GR3. In that analysis, 284 of the 376 families with at least two species analysed were recovered. The remaining analyses under extended implied weighting of the combined data set recovered 279 or 280 families as monophyletic (Table 1). Only 273 families were recovered as monophyletic in EQW, while MOL recovered 260. As shown, the combined data set recovered 24 more families than MOL, using the same parameters. Also, the analyses of the combined dataset under extended implied weighting recovered between six and nine more monophyletic families than EQW.

Optimal trees in the final hypothesis (GR3) have a fit (used as optimality criterion under extended implied weighting) of 4620.20103 and 2 639 520 steps. Most

Table 2
Results of different searches. Name of searches are those described in Table 1

Search	Fit	Fit under GR3	Families	Total steps	Morphological steps
SEP	3157.58185	4625.02817	277	2647 426	1064
GR3	4620.20103	4620.20103	284	2639 520	1082
GR9	4758.99827	4620.77455	280	2639 476	1086
G27	4827.49437	4621.13474	279	2639 826	1090
GEN	4902.70047	4621.40184	280	2639 497	1090
POS	4044.83082	4625.72560	280	2642 422	1073
EQW	Not applicable	4628.89289	274	2629 862	1137
MOL	4616.26673	4620.50085	260	2639 474	1202

For each search the fit, total number of steps and number of steps of morphological characters of the most parsimonious trees are denoted. Fit under GR3 is the fit of the most parsimonious trees of every search optimized under the parameters herein preferred (grouping each three columns to collectively weight them). Families is the number of monophyletic families recovered as monophyletic in each search.

parsimonious trees obtained with the remaining analyses of the combined dataset, and even with the molecule-only dataset, have as much as 8.7 units of fit more than the optimal when optimized under the same parameters (i.e. grouping every three columns). All the most parsimonious trees obtained with different parameters of extended implied weighting have a difference in fit from the most parsimonious trees of about 5.5 (or fewer) units of fit. Therefore, clades with an estimated Bremer support of 9.0 (or more) are obtained as monophyletic under all the conditions explored and clades with a support of 6.0 (or more) are monophyletic in all the analyses under implied weighting. The consistency index (CI) of the morphological characters in the different analyses of the combined dataset varies from 0.237 to 0.253. Although this is more than 10 times higher than the average CI for the dataset (0.019–0.020), it is clear that the morphological characters analysed are not perfectly hierarchical synapomorphies added to a molecule-based tree, but instead display a large number of parallelisms and reversals.

Most parsimonious trees in the analysis of MOL have only 0.300 points of fit more than the most parsimonious trees with the combined dataset (using the same weighting scheme). This difference in fit is smaller than between GR3 and any other of the explored parameters (0.574 with GR9). This result shows that the small number of morphological characters has little impact on the global optimality of the most parsimonious trees, being less influential than slight variations in the weighting schemes. However, the morphological characters have an important impact on the topology of the most parsimonious trees and, when analysed, the shortest trees include 24 more families as monophyletic than when morphological data are excluded.

Results and Bremer support found for clades at the ordinal level are presented in Figs 1–3, while further discussions and complete cladograms from the different searches, showing supports and mapping morphological synapomorphies, are provided as Appendices S7–S21. A list of the orders and families with information about their monophyly in the final hypothesis is provided as Table 2.

General topology and comparison with prior results

All the analyses show the Actinopterygii as monophyletic, with the Osteoglossomorpha as the sister group of the Elopomorpha and Clupeocephala (Fig. 1). The Polypteriformes are the sister group of the remaining Actinopterygii. Under extended implied weighting, both with the complete dataset and with MOL, the Chondrostei are the sister group of the Holostei and Teleostei (Fig. 1a,b). However, in EQW, the Chondrostei are the sister group of the Holostei (Fig. 1c). The Osteoglossomorpha are monophyletic in all the analyses, but the Osteoglossiformes are paraphyletic in EQW, with the Pantodontidae as the sister group of the Hiodontiformes (Fig. 1c).

The Elopomorpha are also monophyletic, as the sister group of the Clupeocephala. Among the Elopomorpha, the Notacanthiformes are the sister group of the Albuliformes in the analyses under extended implied weighting (as in Greenwood, 1977), while in EQW, they form a clade with the Anguilliformes, as in the hypothesis by Betancur-R et al. (2013a).

Under extended implied weighting, the Otomorpha are recovered as monophyletic, while in EQW, the Alepocephaliformes are the sister group of the Euteleosteomorpha. Differing from most previous hypotheses (e.g. Calcagnotto et al., 2005; Betancur-R et al., 2013a), the Characiformes are not herein obtained as monophyletic and the Citharinoidei are recovered as the sister group of the Characoidei plus Siluriformes in all the analyses.

The Protacanthopterygii, composed of the Argentiniformes, Esociformes, Galaxiiformes (excluding Lepidogalaxiidae) and Salmoniformes (i.e. as defined by Betancur-R et al., 2013a), are not strictly recovered as monophyletic, but the same group of taxa except the Galaxiiformes, is obtained. In most analyses, the Galaxiiformes are recovered as the sister group of *Lepidogalaxias salamandroides*, the single known member of the Lepidogalaxiiformes. Only in SEP, *L. salamandroides* is the sister group of all the remaining Euteleosteomorpha, as obtained also by Betancur-R et al. (2013a). The clade of Argentiniformes, Esociformes and Salmoniformes is monophyletic in all the analyses.

All the combined analyses agree on the monophyly of the Neoteleostei, composed of the Acanthomorphata, Ateleopodiformes, Aulopiformes and Myctophiformes, represented by 5196 species in this analysis. All the analyses also agree in the monophyly of the Euacanthomorphacea (4936 species analysed), Euteleosteomorpha (5421), Clupeocephala (7868), Teleostei (8057) and Actinopteri (8087). Details of variations in the results from different analyses are provided as Appendices S7–S20.

Discussions on phylogenetic results below are restricted to orders or supraordinal clades differing from the current classifications or whose monophyly has been debated recently in the literature. Comparisons of the results obtained at subordinal or subfamilial levels of some of the most diverse orders are also included. More extensive discussions to family level and details of the final hypothesis are provided in Appendices S8, S11, S12 and S20. A summary of the results obtained in GR3 are provided as Table 3.

All basal clades are well supported and in concordance with previous hypotheses (e.g. Wiley and

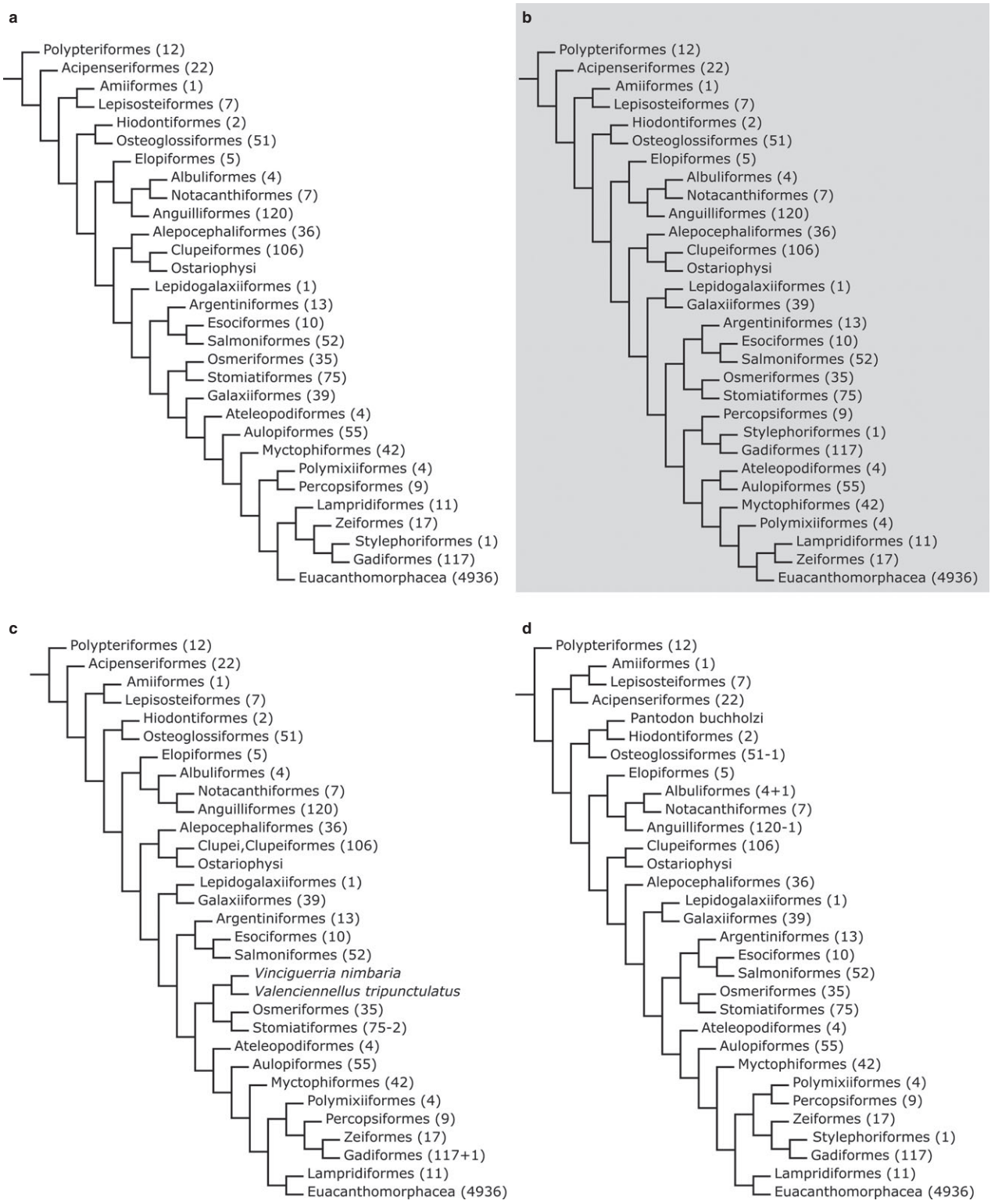

Fig. 1. General topology of the consensus of most parsimonious trees obtained from: (a) combined dataset weighting each character separately (SEP); (b) grouping contiguous characters to weight them collectively (GR3, GR9, G27, GEN, MOL); (c) combined dataset grouping by positions (MOL); (d) combined dataset with equal weights (EQW). The final hypothesis herein proposed is shaded. The number of analysed species for each category is shown in parentheses.

Fig. 2. General topology of the final hypothesis (GR3) showing relationships between orders (Fit = 4619.23957; Length = 2634 502 steps). Subtrees of the Ovalentaria and Carangiaria are shown separately. The subtree of the Eupercaria is shown in Fig. 3. More details are provided in Table 3 and in the Supplementary Online Material.

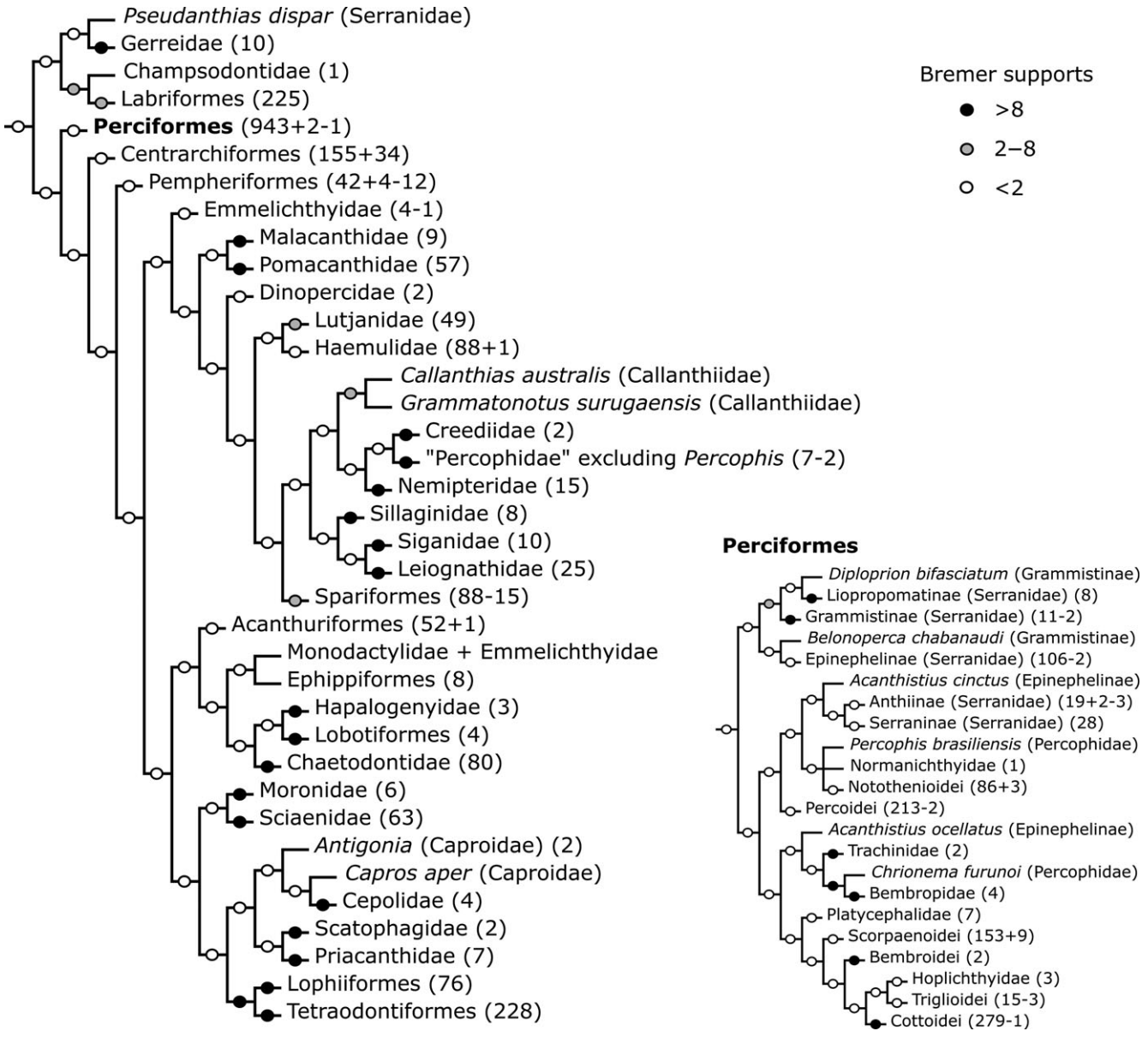

Fig. 3. General topology of the final hypothesis (GR3) showing relationships within the Eupercaria (Fit = 4619.23957; Length = 2634 502 steps). The subtree of the Perciformes is shown separately. More details are provided in the Table 3 and in the Supplementary Online Material.

Johnson, 2010; Betancur-R et al., 2013a). The Osteoglossomorpha are the sister group of the remaining Teleostei in the present hypothesis (Figs 1 and 2), contrary to the monophyly of the Osteoglossocephalai (*sensu* Arratia, 1999), composed of all the Teleostei except the Elopomorpha. Published hypotheses of relationships of the Osteoglossomorpha and the Elopomorpha conflict. Some authors have proposed that those groups form a clade (e.g. Hoegg et al., 2004; Broughton et al., 2013) or are successive sister-groups of the remaining Teleostei, with the Elopomorpha (Betancur-R et al., 2013a; Chen et al., 2014) or the Osteoglossomorpha (Inoue et al., 2001a, 2003, 2004; this analysis) as the sister group of the remaining Teleostei (Clupeocephala). Therefore, the monophyly or paraphyly of the Osteoglossocephalai is still an open question, depending on the dataset analysed and methods used.

The Otomorpha, including the Alepocephaliformes, Clupeiformes and Ostariophysi, are obtained here as monophyletic. The Alepocephaliformes are the sister group of the Ostarioclupeomorpha (*sensu* Wiley and Johnson, 2010), composed of the Clupeiformes and Ostariophysi. Betancur-R et al. (2013a) obtained, instead, the Alepocephaliformes as the sister group of the Ostariophysi. This analysis does not support the

Table 3
List of supraordinal clades (including more than one order), orders and families according to the final hypothesis (GR3)

Supraordinal clades

Actinopterygii (8099). **Cladistia** (12). **Actinopteri** (8087). **Chondrostei** (22). **Neopterygii** (8057). **Holostei** (8). **Teleostei** (8057). **Elopocephalai** (136). **Osteoglossocephalai** (7921,+136): Paraphyletic in terms of the Elopocephalai. **Clupeocephala** (7868). **Otomorpha** (2447). **Ostariophysi** (2305). **Euteleosteomorpha** (5421). **Protacanthopterygii** (114,NF): Diphyletic; clade of 75 taxa composed of Argentiniformes, Esociformes, and Salmoniformes goes with Stomiati; clade of 39 taxa composed of Galaxiiformes, goes with Lepidogalaxiiformes. **Stomiati** (110). **Neoteleostei** (5196). **Ateleopodia** (4). **Eurypterygia** (5192,+4): added Ateleopodiformes with Aulopiformes. **Ctenosquamata** (5137,+59): added Ateleopodiformes and Aulopiformes with Zeiformes, Myctophiformes, Lampridiformes, and Polymixiiformes. **Acanthomorphata** (5095,+101): added Ateleopodiformes, Aulopiformes, and Myctophiformes. **Paracanthomorphacea** (144,−17): Zeiformes goes with Lampridiformes. **Zeiogadaria** (135,−17): Zeiformes goes with Lampridiformes. **Euacanthomorphacea** (4936). **Percomorphaceae** (4857). **Gobiaria** (504,+3): added Trichonotidae with Gobiiformes. **Anabantaria** (95). **Carangiaria** (258). **Ovalentaria** (1447). **Eupercaria** (2294,+2,−3): added Scombropidae (2) with Epigonidae; Trichonotidae (3) goes with Gobiiformes.

Orders

Acipenseriformes (22): Acipenseridae (20); Polyodontidae (2).
Amiiformes, Amiidae (1)
Lepisosteiformes, Lepisosteidae (7)
Elopiformes (5): Elopidae (3); Megalopidae (2)
Albuliformes, Albulidae (4)
Notacanthiformes (7): Halosauridae (3,NF) – Paraphyletic in terms of the Notacanthidae –; Notacanthidae (4)
Hiodontiformes, Hiodontidae (2)
Osteoglossiformes (51): Arapaimidae (2); Gymnarchidae (1); Mormyridae (34); Notopteridae (8); Osteoglossidae (5); Pantodontidae (1)
Anguilliformes (120): Anguillidae (17); Chlopsidae (4); Congridae (16,−8) – Paraphyletic; clade of *Heteroconger* and *Paraconger* goes with Muraenesocidae, Ophichthidae, and clade of *Ariosoma* and *Parabathymyrus* (Congridae); clade with *Rhynchoconger* and *Macrocephenchelys* goes with *Nettastoma* (Nettastomatidae) –; Cyematidae (1); Derichthyidae (2,+1) – paraphyletic in terms of the Colocongridae (1) –; Eurypharyngidae (1); Heterenchelyidae (2); Monognathidae (1); Moringuidae (4); Muraenesocidae (4,−1) – *Gavialiceps* goes with *Macrocephenchelys*, *Rhynchoconger* (Congridae), and *Nettastoma* (Nettastomatidae) –; Muraenidae (35); Myrocongridae (1); Nemichthyidae (3); Nettastomatidae (5,NF) – paraphyletic in terms of *Gavialiceps* (Muraenesocidae), clade of *Rhynchoconger* and *Macrocephenchelys* (Congridae), and clade of *Bathyuroconger*, *Bathycongrus*, *Conger*, and *Uroconger* (Congridae) –; Ophichthidae (10); Protanguillidae (1); Saccopharyngidae (2); Serrivomeridae (4); Synaphobranchidae (6)
Alepocephaliformes (36): Alepocephalidae (29,+7) – paraphyletic in terms of the Leptochilichthyidae (1) and Platytroctidae (6) –
Clupeiformes (106): Clupeidae (46,+8) – added Dussumieriidae (2) and Chirocentridae (1) with Spratelloidini, Pristigasteridae (4) with *Clupea* and *Sprattus*, and Sundasalangidae (1) with *Clupeichthys* –; Denticipitidae (1); Engraulidae (51)
Gonorynchiformes (13): Chanidae (1); Phractolaemidae (1); Gonorynchidae (3); Kneriidae (8)
Cypriniformes (1088): Balitoridae (22); Barbuccidae (1); Botiidae (19); Catostomidae (62); Cobitidae (46); Cyprinidae (912,+2,−1) – added Psilorhynchidae (2) with a large clade of cyprinids including Cyprininae and Labioninae; *Sundadanio* goes as sister group of all remaining Cypriniformes –; Ellopostomatidae (1); Gyrinocheilidae (2); Nemacheilidae (19); Serpenticobitidae (1); Vaillantellidae (1)
Gymnotiformes (36): Apteronotidae (1); Gymnotidae (18); Hypopomidae (8); Rhamphichthyidae (6); Sternopygidae (2)
Citharinifformes (21): Citharinidae (2); Distichodontidae (19)
Characiformes (380): Alestidae (56,+1,−3) – added Hepsetidae (1) with *Arnoldichthys*; *Chalceus* goes as sister group of remaining Characoidei –; Anostomidae (18); Characidae (220,+6,−3) – added Gasteropelecidae (6) with *Brycon*, *Henochilus*, and *Salminus*; Cynodontinae goes with Hemiodontidae –; Chilodontidae (8); Ctenoluciidae (3); Curimatidae (6); Erythrinidae (5); Hemiodontidae (5); Lebiasinidae (3); Parodontidae (2); Prochilodontidae (13); Serrasalmidae (29)
Siluriformes (767): Akysidae (3); Amblycipitidae (13,NF) – clade of *Liobagrus* and *Xiurenbagrus* goes with Akysidae –; Anchariidae (1); Ariidae (101); Auchenipteridae (7); Aspredinidae (8); Bagridae (38,+2,−1) – added *Pachypterus* (Horabagridae) with *Bagrus*, *Sperata*, and some *Hemibagrus*; *Rita* goes with Chacidae and Plotosidae –; Callichthyidae (139); Cetopsidae (3); Chacidae (1); Clariidae (6); Claroteidae
(20,−2) – *Auchenoglanis* and *Parauchenoglanis* goes with *Parailia*, *Pareutropius*, and *Schilbe* (Schilbeidae) –; Cranoglanididae (1); Diplomystidae (6); Doradidae (33); Heptapteridae (12,−2) – *Phreatobius* goes with Pseudopimelodidae –; Heteropneustidae (1); Horabagridae (6,−2) – *Pachypterus* goes with *Bagrus*, *Sperata*, and some *Hemibagrus* (Bagridae) –; Ictaluridae (22); Lacantuniidae (1); Loricariidae (111); Malapteruridae (2); Mochokidae (61); Nematogenyidae (1); Pangasiidae (15,+1) – added *Eutropiichthys* (Schilbeidae) with *Pangasianodon* and some *Pangasius* –; Pimelodidae (42); Plotosidae (6); Pseudopimelodidae (5); Schilbeidae (10,NF) – *Ailia*, *Clupisoma*, and *Laides* go with *Horabagrus* and *Pseudeutropius* (Horabagridae); *Eutropiichthys* goes with *Pangasianodon* and some *Pangasius* (Pangasiidae) –; Scoloplacidae (1); Siluridae (10); Sisoridae (47,+14,−1) – added Erethistidae with *Bagarius*; *Pseudecheneis* goes with *Amblyceps* (Amblycipitidae) –; Trichomycteridae (13). Not analyzed: Astroblepidae; Austroglanidae
Lepidogalaxiiformes, Lepidogalaxiidae (1)
Galaxiiformes, Galaxiidae (39)
Argentiniformes (13): Argentinidae (4); Bathylagidae (4,NF) – *Bathylagoides* and *Lipolagus* go with Microstomatidae –; Microstomatidae (3); Opisthoproctidae (2)
Esociformes (10): Esocidae (5); Umbridae (5,−2) – *Dallia* and *Novumbra* go with Esocidae –
Salmoniformes, Salmonidae (52)
Osmeriformes (35): Osmeridae (14); Plecoglossidae (1); Retropinnidae (5); Salangidae (15)

Table 3
(Continued)

Stomiatiformes (75): Gonostomatidae (16,NF) – Paraphyletic in terms of all other families of the order; *Gonostoma* paraphyletic in terms of *Sigmops* –; Phosichthyidae (6,NF) – *Ichthyococcus*, *Pollichthys*, and *Phosichthys* go within Stomiidae; *Polymetme* and *Yarrella* go with Sternoptychidae; *Vinciguerria* goes with *Valenciennellus* (Sternoptychidae) and *Margrethia* (Gonostomatidae) –; Sternoptychidae (13,+1,−1) – added *Triplophos* (Gonostomatidae) with *Sternoptyx*; *Valencienellus* goes with *Vinciguerria* (Phosichthyidae) and *Margrethia* (Gonostomatidae) –; Stomiidae (40,+3,−3) – added *Pollichthys* (Phosichthyidae) with *Melanostomias* and *Stomias*; added *Phosichthys* and *Ichthyococcus* (Phosichthyidae) with most stomiids; *Chauliodus* goes with *Bonapartia* (Gonostomatidae) –
Percopsiformes (9): Amblyopsidae (6); Aphredoderidae (1); Percopsidae (2)
Stylephoriformes, Stylephoridae (1)
Gadiformes (117): Euclichthyidae (1); Gadidae (18,−1) – *Raniceps* goes with Euclichthyidae –; Lotidae (7,NF) – paraphyletic in terms of Gadidae –; Macrouridae (54,+4,−3) – added Bregmacerotidae (2) with Bathygadinae and *Macruronus* and *Steindachneria* (Merluciidae) with Macrourinae; *Squalogadus* and *Trachirincus* go with Muraenolepididae –; Melanonidae (2); Merluccidae (14,−2) – *Macruronus* and *Steindachneria* go with Macrourinae –; Moridae (11); Muraenolepididae (2); Phycidae (6)
Ateleopodiformes, Ateleopodidae (4)
Aulopiformes (55): Alepisauridae (2); Aulopidae (2); Bathysauridae (1); Chlorophthalmidae (4) – added *Scopelosaurus hoedti* (Notosudidae) with *Chlorophthalmus* –; Evermannellidae (3); Giganturidae (2); Ipnopidae (5); Notosudidae (4,−1) – *Scopelosaurus hoedti* goes with Chlorophthalmidae –; Paralepididae (12,NF) – paraphyletic in terms of the Alepisauridae (2), Anotopteridae (1), Evermannellidae (3), and Omosudidae (1) –; Paraulopidae (1); Pseudotrichonotidae (1); Scopelarchidae (5); Synodontidae (11,NF) – *Harpadon* and *Saurida* are sister group of all members of the order excepting Paraulopidae, Pseudotrichonotidae and remaining synodontids –. Not analyzed: Bathysauroididae
Myctophiformes (42): Myctophidae (39); Neoscopelidae (3)
Polymixiiformes, Polymixiidae (4)
Lampridiformes (11): Lamprididae (2); Lophotidae (1); Regalecidae (2); Trachipteridae (4); Veliferidae (2). Not analyzed: Radiicephalidae
Zeiformes (17): Cyttidae (2,NF) – *Cyttus traversi* goes with *Xenolepidichthys* (Grammicolepididae) –; Grammicolepididae (2,NF) – *Xenolepidichthys* goes with *Cyttus traversi* (Cyttidae) –; Oreosomatidae (5); Parazenidae (2); Zeidae (4); Zeniontidae (2,NF) – *Cyttomimus* as sister group of remaining zeiforms – .
Trachichthyiformes (14): Anomalopidae (2); Diretmidae (3); Monocentridae (1); Trachichthyidae (7,+1) – added Anoplogastridae (1) with *Aulotrachichthys* and *Paratrachichthys* – .
Beryciformes (25): Berycidae (4); Cetomimidae (9); Melamphaidae (8); Rondeletiidae (2,NF) – forming a tetrachotomy with Barbourisiidae (1) and Stephanoberycidae (1) –. Not analyzed: Gibberichthyidae; Hispidoberycidae.
Holocentriformes, Holocentridae (40)
Ophidiiformes (37): Bythitidae (13,+1) – added Aphyonidae (1) with *Bidenichthys* and *Cataetyx* –; Ophidiidae (20,+3) – added Carapidae (3) with *Chilara*, *Genypterus*, *Lepophidium*, and *Ophidion* –. Not analyzed: Parabrotulidae
Batrachoidiformes, Batrachoididae (9)
Kurtiformes (39): Apogonidae (37); Kurtidae (2)
Gobiiformes (465): Eleotridae (52,−14) – Butinae (12) goes with Thalasseleotrididae and Gobiidae and Milyeringinae (2) goes as sister group of remaining Gobiiformes –; Gobiidae (380,+20) – added Kraemeriidae (2) with *Pleurosicya*, Schindleriidae (2) with *Gobiopterus semivestitus*, and Microdesmidae with *Callogobius sclateri* –; Microdesmidae (16,+1) – added *Coryphopterus hyalinus* with *Cerdale* and *Microdesmus* –; Rhyacichthyidae (3); Odontobutidae (8); Thalasseleotrididae (2). Not analyzed: Xenisthmidae [clade with Trichonotidae (3)]
Syngnathiformes (95): Aulostomidae (2); Callionymidae (10); Centriscidae (3); Dactylopteridae (5); Draconettidae (1); Fistulariidae (2); Mullidae (13); Pegasidae (2); Solenostomidae (2); Syngnathidae (55)
Scombriformes (118,−2) – *Scombrops* goes within Pempheriformes –: Ariommatidae (4); Arripidae (3); Bramidae (9); Caristiidae (2); Centrolophidae (8); Chiasmodontidae (5); Gempylidae (16,−1) – *Lepidocybium* goes with Trichiuridae –; Icosteidae (1); Latridae (2); Nomeidae (6); Pomatomidae (1); Scombridae (38); Scombrolabracidae (1); Stromateidae (9); Tetragonuridae (2); Trichiuridae (11). Not analyzed: Amarsipidae
Synbranchiformes (14): Indostomidae (2); Mastacembelidae (9); Synbranchidae (3). Not analyzed: Chaudhuriidae
Anabantiformes (81): Anabantidae (9); Badidae (8); Channidae (14); Helostomatidae (1); Nandidae (4); Osphronemidae (43); Pristolepididae (2)
"Sphyraenid clade": Centropomidae (6); Lactariidae (1); Latidae (5); Psettodidae (3); Sphyraenidae (7)
Istiophoriformes (10): Istiophoridae (9); Xiphiidae (1) [clade with Leptobramidae (1); Menidae (1); Nematistiidae (1); Polynemidae (9); Toxotidae (2)]
Carangiformes (64,−1) – Nematistiidae goes with Toxotidae –: Carangidae (52,−6) – *Oligoplites*, *Scomberoides*, and *Trachinotus* go to clade with Coryphaenidae, Echeneidae, and Rachycentridae –; Coryphaenidae (2); Echeneidae (8); Rachycentridae (1)
Pleuronectiformes (152,−3) – Psettodidae (3) goes with Lactariidae and Latidae –: Achiridae (5); Bothidae (14,−1) – *Grammatobothus* goes within *Pseudorhombus* (Paralichthyidae), within Soleidae –; Citharidae (4,−1) – *Citharoides macrolepidotus* goes within *Poecilopsetta* (Pleuronectidae); Cynoglossidae (16); Paralichthyidae (20,+1,−8) – added *Grammatobothus* within *Pseudorhombus*; clade with *Citharichthys*, *Cyclopsetta*, *Etropus*, and *Syacium* goes with Bothidae –; Pleuronectidae (52,−9) – Rhombosoleinae goes with Achiropsettidae; *Poeciliopsetta* goes with *Citharoides macrolepidotus* within Soleidae –; Samaridae (5); Scophthalmidae (6); Soleidae (25,+3) – added *Citharoides macrolepidotus* (Citharidae) and *Poecilopsetta* (Pleuronectidae) with *Austroglossus* and *Synaptura* –
"Polycentrid clade": Polycentridae (4)
Mugiliformes, Mugilidae (57) [clade with Ambassidae (6) and Congrogadidae (3)]
Pholidichthyiformes, Pholidichthyidae (1)
Cichliformes, Cichlidae (416)

Table 3
(Continued)

Blenniiformes (248): Blenniidae (126); Chaenopsidae (39,−5) – clade of *Mccoskerichthys* and *Neoclinus* goes with *Stathmonotus*, *Paraclinus* (Labrisomidae), Dactyloscopidae and remaining Chaenopsidae; *Stathmonotus* goes with *Paraclinus* (Labrisomidae) –; Clinidae (15); Dactyloscopidae (5); Gobiesocidae (7); Labrisomidae (37,NF) – paraphyletic in terms of the Chaenopsidae and Dactyloscopidae –; Tripterygiidae (19) [clade with Opisthognathidae (4), Grammatidae (3,NF), Pomacentridae (177), and Pseudochromidae (8)]
"Embiotocid clade": Embiotocidae (16), Plesiopidae (2,NF) – *Assessor* goes to a trichotomy with *Lipogramma* (Grammatidae), Pholidichthyiformes, and Cichliformes –.
Beloniformes (77): Adrianichthyidae (8); Belonidae (29,+3) – added Scomberesocidae (3) with *Belone* and *Petalichthys* –; Exocoetidae (19); Hemiramphidae (13,NF) – paraphyletic in terms of Exocoetidae –; Zenarchopteridae (5,NF) – paraphyletic in terms of Belonidae –
Atheriniformes (172): Atherinidae (22,NF) – *Atherinason*, *Atherinosoma*, *Craterocephalus*, and *Kestratherina* sister group of remaining Atherinidae plus Melanotaenioidei –; Atherinopsidae (57,+1) – added Notocheiridae (1) with Menidiinae –; Bedotiidae (11); Isonidae (3); Melanotaeniidae (70,−1) – *Cairnsichthys* goes with Bedotiidae – .
Cyprinodontiformes (253): Anablepidae (5); Aplocheilidae (4); Cyprinodontidae (15); Fundulidae (23); Goodeidae (16); Nothobranchiidae (33); Poeciliidae (84,−2) – *Fluviphylax* goes as sister group of *Aplocheilichthys*, Anablepidae, Cyprinodontidae, Fundulidae, Goodeidae, Poeciliidae, Profundulidae, and Valenciidae; *Aplocheilichthys* goes as sister group of Valenciidae –; Profundulidae (2); Rivulidae (69); Valenciidae (2)
Labriformes (225): Labridae (219,+6) – added Odacidae (6) with *Choerodon* – [clade with Champsodontidae (1); Gerreidae (10); *Pseudanthias dispar* (Serranidae)]
Perciformes (943,+2,1) – added *Chrionema* and *Percophis* (Percophidae); *Pseudanthias dispar* goes with Gerreidae –: Agonidae (13); Anarhichadidae (5); Anoplopomatidae (2); Aploactinidae (2); Artedidraconidae (13); Aulorhynchidae (1); Bathydraconidae (14); Bathymasteridae (6,−1) – *Rathbunella* goes as sister group of remaining Zoarcales –; Bembridae (1); Bembropidae (4); Bovichtidae (3); Channichthyidae (16); Congiopodidae (2); Cottidae (93,+24) – added Agonidae and Hemitripteridae with *Scorpaenichthys*, Psychrolutidae with *Artediellus*, and *Nautichthys* (Hemitripteridae) and Rhamphocottidae (1) as successive sister groups of a diverse clade –; Cryptacanthodidae (3); Cyclopteridae (5); Eleginopsidae (1); Gasterosteidae (8); Gnathanacanthidae (1); Harpagiferidae (2); Hexagrammidae (12); Hoplichthyidae (3); Hypoptychidae (2); Liparidae (19); Neosebastidae (2,NF) – paraphyletic in terms of Scorpaenoidei –; Niphonidae (1); Normanichthyidae (1); Nototheniidae (36,NF) – paraphyletic in terms of remaining Notothenioidei except Bovichtidae, Eleginopsidae, and Pseudaphritidae –; Parabembridae (1); Pataecidae (1); Percidae (210); Percophidae (9, NF) – *Percophis and Chrionema* found into Perciformes and remaining genera in Spariformes –; Peristediidae (6); Pholidae (10;−1) – *Xererpes* goes with *Xiphister* (Stichaeidae) –; Platycephalidae (7); Plectrogeniidae (1); Pseudaphritidae (1); Psychrolutidae (5); Ptilichthyidae (1); Rhamphocottidae (1); Scorpaenidae (35,+1) – added Setarchidae with *Neomerinthe* and *Pontinus* –; Sebastidae (110); Serranidae (172,NF) – clade with Epinephelinae, Grammistinae, and Liopropomatinae go as sister group of remaining Perciformes; *Acanthistius ocellatus* goes with Bembropidae, Trachinidae, and *Chrionema* (Percophidae) –; Setarchidae (1); Stichaeidae (32,NF) – paraphyletic in terms of remaining Zoarcales excepting Bathymasteridae –; Synanceiidae (3,NF) – paraphyletic in terms of Aploactinidae, Gnathanacanthidae, and Pataecidae –; Tetrarogidae (4,NF) – paraphyletic in terms of Aploactinidae, Gnathanacanthidae, Pataecidae, and Synanceiidae –; Trachinidae (2); Trichodontidae (2); Triglidae (9,−3) – *Chelidonichthys* and *Lepidotrigla* go with *Pagothenia* (Nototheniidae) –; Zaproridae (1); Zoarcidae (52,−3) – *Eulophias* goes with *Alectrias* (Stichaeidae); *Neozoarces* and *Zoarchias* go with Anarhichadidae –. Not analyzed: Apistidae; Ereuniidae; Eschmeyeridae; Perryenidae; Scytalinidae.
Centrarchiformes (165,+24) – added Dichistiidae and *Malakichthys griseus* (Acropomatidae) with Terapontoidei; added Uranoscopiformes with Percalatidae; added Leptoscopidae within Uranoscopiformes –: Ammodytidae (6); Aplodactylidae (2); Centrarchidae (33); Centrogenyidae (1); Cheilodactylidae (8,−1) – *Cheilodactylus fasciatus* goes with Chironemidae –; Cheimarrichthyidae (1); Chironemidae (3); Cirrhitidae (11); Dichistiidae (1); Elassomatidae (5); Enoplosidae (1); Girellidae (6); Kuhliidae (12); Kyphosidae (16,NF) – paraphyletic in terms of Girellidae, Kuhliidae, Oplegnathidae, and Terapontidae –; Leptoscopidae (2); Oplegnathidae (3); Percalatidae (2); Percichthyidae (16); Pinguipedidae (5,NF) – paraphyletic in terms of Cheimarrichthyidae and Leptoscopidae –; Sinipercidae (10); Terapontidae (35); Uranoscopidae (7)
Pempheriformes (42,+4,−12) – added Bathyclupeidae with Lateolabracidae, Glaucosomatidae, Pempheridae, and *Malakichthys elegans* (Acropomatidae); added Symphysanodontidae with Howellidae and *Synagrops* (Acropomatidae); added Scombropidae with Epigonidae; Creediidae and most Percophidae go with Nemipteridae; *Malakichthys griseus* (Acropomatidae) goes with Dichistiidae; *Percophis* (Percophidae) goes with Normanichthyidae and Notothenioidei; *Chrionema* (Percophidae) goes with Bembropidae –: Acropomatidae (7,NF) – paraphyletic in terms of most other Pempheriformes –; Banjosidae (1); Bathyclupeidae (1); Callanthiidae (3,NF) – *Callanthias ruber* goes into Haemulidae –; Epigonidae (2); Glaucosomatidae (4); Howellidae (2); Lateolabracidae (1); Ostracoberycidae (1); Pempheridae (5); Pentacerotidae (5); Polyprionidae (3); Scombropidae (2); Symphysanodontidae (1). Not analyzed: Parascorpididae
Spariformes (87,−15) – Nemipteridae (15) goes with Crediidae and Percophidae –: Centracanthidae (5,NF) – four clades within Sparidae –; Lethrinidae (15); Sparidae (53,+5) – added clade of *Spicara maena* and *S. smaris* (Centracanthidae) with *Spondyliosoma*; added *Spicara nigricauda* (Centracanthidae) with *Diplodus*; added *Spicara alta* (Centracanthidae) with *Dentex tumifrons*; added *Centracanthus* (Centracanthidae) with *Pagellus bogaraveo* – [clade with Callanthiidae (3,NF) – *Callanthias ruber* goes to Haemulidae –; Creediidae (2); Dinopercidae (2); "Emmelichthyidae" (4,−1): – *Emmelichthys* (nominal genus) goes to Ephipphiformes –; Haemulidae (88,+1) – added *Callanthias ruber* with *Parapristipoma* and *Plectorhinchus mediterraneus* –; Leiognathidae (25); Lutjanidae (49); Malacanthidae (9); Nemipteridae (15); "Percophidae" (7,−2) – *Chrionema* and *Percophis* (nominal genus) go to Perciformes –; Pomacanthidae (57); Siganidae (10); Sillaginidae (8)]
Acanthuriformes (52,+1) – added Dinolestidae with Luvaridae –: Acanthuridae (50); Dinolestidae (1); Luvaridae (1); Zanclidae (1)
Ephippiformes (8): Drepaneidae (3); Ephippidae (5) [clade with Monodactylidae (2,NF) – trichotomy with *Emmelichthys* –; Emmelichthyidae (4,NF) – *Erythrocles* go to clade including Spariformes]
Lobotiformes (4): Datnioididae (2); Lobotidae (2) [clade with Hapalogenyidae (3); Chaetodontidae (80)]
"Sciaenid clade" (69): Moronidae (6); Sciaenidae (63)
"Cepolid clade" (16): Caproidae (3,NF) – paraphyletic in terms of the Cepolidae –; Cepolidae (4); Priacanthidae (7); Scatophagidae (2)

Table 3
(Continued)

Lophiiformes (76): Antennariidae (26,+2) – added Brachionichthyidae (1) with *Tathicarpus* and Tetrabrachiidae (1) as sister group of a diverse clade –; Caulophrynidae (2); Ceratiidae (3); Chaunacidae (6); Diceratiidae (2); Gigantactinidae (3); Himantolophidae (3); Linophrynidae (3); Lophiidae (11); Melanocetidae (3); Neoceratiidae (1); Ogcocephalidae (6); Oneirodidae (4); Thaumatichthyidae (1). Not analyzed: Lophichthyidae

Tetraodontiformes (228): Aracanidae (9); Balistidae (32); Diodontidae (9); Molidae (3); Monacanthidae (46); Ostraciidae (18); Tetraodontidae (100); Triacanthidae (4); Triacanthodidae (6); Triodontidae (1)

Numbers in parentheses after each taxa represent the species analysed. Monophyletic or monotypic taxa (at least in this analysis) are followed only by the number of species in parentheses. Symbols "–" and "+" indicate the number of species that should be removed and/or added to render the corresponding group monophyletic. NF means "not found (as monophyletic)" and the group cannot be redefined as monophyletic by the addition or removal of a few taxa. Supraordinal clades in the first section of the table and orders in the second section are indicated in bold. Note that discussions and cladograms to familial level are provided as Online Supporting Information.

monophyly of the Characiformes, as traditionally considered, and the Citharinoidei (or "Citharinifformes" in Fig. 2) and Characoidei (or "true" Characiformes) are, instead, successive sister groups of the Siluriformes. The non-monophyly of the Characiformes has been previously proposed by Nakatani et al. (2011), contrary to most other hypotheses (as Betancur-R et al., 2013a) (Figs 1–3). The Cypriniformes are monophyletic and composed of the Cobitioidea and Cyprinoidea (Fig. 3) except for *Sundadanio axelrodi* (Cyprinidae) which is obtained as the sister group of all the remaining cypriniforms, in a odd result that should be further tested. The highly diverse Cyprinidae are paraphyletic in terms of the Psilorhynchidae. Among the Characiformes, the Gasteropelecidae as proposed by Mirande (2009, 2010) from morphological data are obtained as monophyletic (although included in the Characidae) and the Hepsetidae are included in the Alestidae, implying the existence of only one African clade in the Characoidei (or "true" Characiformes), instead of the two usually proposed (e.g. Vari, 1979; Calcagnotto et al., 2005; Zanata and Vari, 2005). The Siluriformes are monophyletic and well supported, with the Diplomystoidei, Siluroidei and Loricarioidei (*sensu* Sullivan et al., 2006) supported as clades.

The Protacanthopterygii are not monophyletic herein. A clade composed of the Galaxiiformes and the monotypic Lepidogalaxiiformes are the sister group of the other Euteleosteomorpha (*sensu* Betancur-R et al., 2014). The remaining Protacanthopterygii are the sister group of the Stomiati (Osmeriformes and Stomiatiformes) and this clade is the sister group of the Neoteleostei. The Paracanthomorphacea according to Grande et al. (2013) and Betancur-R et al. (2013a) are composed of the Percopsaria (Percopsiformes) and Zeiogadaria (Gadiformes, Stylephoriformes and Zeiformes). In the present analysis, the Paracanthomorphacea are paraphyletic, with the Zeiformes as the sister group of the Lampridiformes, instead of being included in this clade. Both the Paracanthomorphacea and the clade obtained here, excluding the Zeiformes, have low support in the analysis by Betancur-R et al. (2013a) and in the present analysis, respectively.

The Euacanthomorphacea (*sensu* Johnson and Patterson, 1993) are monophyletic. According to Betancur-R et al. (2014), this division is composed of the Berycomorphaceae and Percomorphaceae. In this analysis, both the Berycomorphaceae and the Beryciformes, either including the Stephanoberyciformes (as in Betancur-R et al., 2013a) or not (e.g. Wiley and Johnson, 2010), are paraphyletic. The Berycomorphaceae form three clades that are successive sister groups of the Percomorphaceae, corresponding to the Trachichthyiformes (*sensu* Moore, 1993), a clade of Beryciformes *sensu stricto*, and the Holocentriformes.

The Percomorphaceae are recovered as monophyletic (Figs 2 and 3). Within this clade, the Batrachodiaria, Syngnatharia, Ovalentaria, Anabantaria and Carangiaria are monophyletic, while the Gobiaria include the Trichonotidae and the Pelagiaria exclude the Scombropidae. The Trichonotidae were not analysed by Betancur-R et al. (2013a,b), who classified them tentatively as *incertae sedis* within the Eupercaria. The Scombropidae are recovered in the Pempheriformes.

The Ovalentaria (Fig. 3) have been proposed as monophyletic by Smith and Near in Wainwright et al. (2012), including the Atherinomorpha, Blenniidae and Cichlidae, among others. Members of this group share the presence of a suite of features associated with demersal spawning, such as adhesive chorionic filaments on eggs (Smith and Wheeler, 2004; Smith and Craig, 2007). The Ovalentaria are monophyletic in all the analyses. The Anabantaria are monophyletic and composed of the Anabantiformes and Synbranchiformes. As proposed by Miya et al. (2003) and differing from all hypotheses based exclusively on morphology, the Indostomidae are the sister group of the Synbranchidae, in the Synbranchiformes. The Carangiaria are composed of the Carangiformes, Istiophoriformes, Pleuronectiformes and several families considered as *incertae sedis* by Betancur-R et al. (2014).

The Pleuronectiformes are not recovered as monophyletic, with the Psettodoidei forming a clade with the Lactariidae and Latidae. The molecular support of the Pleuronectiformes was subject to recent debate.

Betancur-R et al. (2013a) found the Pleuronectiformes paraphyletic in terms of the Centropomidae, with the Psettodoidei and Pleuronectoidei forming monophyletic groups. Betancur-R and Ortí (2014) concluded that, considering some details of the modelling for maximum likelihood approaches, the Pleuronectiformes have molecular support, while Campbell et al. (2014) maintained that the available molecular information neither clearly supports nor contradicts the monophyly of this clade. In the present analysis, the Pleuronectiformes are consistently found as diphyletic, with the Psettodoidei and Pleuronectoidei as separate clades (Fig. 3). Further discussions about the molecular support of this clade are given in Betancur-R et al. (2013b), Betancur-R and Ortí (2014), and Campbell et al. (2013, 2014).

The Eupercaria form a very diverse clade of fishes including most of the former Perciformes and related groups (Fig. 3). In the definition by Betancur-R et al. (2013a), the Eupercaria are not monophyletic, but may be redefined as such with small taxonomic changes. This paper includes several families that have not previously been analysed in global phylogenies of the Actinopterygii but, in general, their relationships are weakly supported and further studies including additional data are needed. Most of the internal relationships in the Eupercaria are poorly supported (but read details in the Supporting Information).

The Chaetodontiformes, composed of the Chaetodontidae and Leiognathidae, were proposed as a new order in Betancur-R et al. (2013a,b) even recognizing their low support but based on a relatively good stability (considering the congruence with the hypotheses by Near et al., 2012, 2013). In the present analysis both families of this order are monophyletic, but distantly related to each other, rendering this order diphyletic (Fig. 3). The Perciformes, even after their redefinition as a monophyletic unit, still constitute one of the most diverse clades of Actinopterygii, including the former Scorpaeniformes and a number of families that have been traditionally included in this order.

Morphological synapomorphies

This analysis includes some characters considered in the literature to be synapomorphic for families or orders, which usually have low levels of homoplasy. An example of that is the presence of a kinethmoid bone, which is synapomorphic for the Cypriniformes. However, there are also characters (mainly loss of bones or complete girdles) that are hypothesized to be much more variable and considered as highly homoplastic within the Actinopterygii, but not previously evaluated in a global phylogeny. In the most parsimonious trees, the loss of a bony basisphenoid is optimized for 35 clades or species, with four reversals to presence; the loss of the intercalar has 29 steps, with seven reversals; and the loss of free parietals has 22 steps with one reversal, among the structures more frequently lost within the Actinopterygii. The complete loss of the pelvic fin and girdle is optimized 10 times in the obtained phylogeny of the Actinopterygii, while the loss of the pectoral girdle was found for eight clades or species. Some of the most homoplastic characters, however, are synapomorphies of large clades, with no or just a few reversals. As an example, the loss of the orbitosphenoid is synapomorphic for the Anguilliformes plus Notacanthiformes, the Galaxiiformes plus the Lepidogalaxiiformes and, in parallel, for the Esociformes, Gonorynchiformes, Osmeriformes and some smaller clades. Another example is the loss of a bony basisphenoid, which is a parallel synapomorphy for the Ostariophysi, Cyprinodontiformes, Galaxiiformes, Gobiiformes, Lepisosteiformes, Osmeriformes, Tetraodontiformes, Pleuronectoidei and many other smaller clades. The complete list of morphological characters is provided in Appendix S1 and the list of synapomorphies for each node is provided as Appendix S17.

Discussion

This is the first large-scale analysis in which extended implied weighting (Goloboff, 2014) is used and explorations of the parameters of this method have been performed. With the criterion used herein to select between different parameters (i.e. number of families recovered as monophyletic), the one that best performed was grouping every three positions to weight characters (GR3) (the three characters of each codon are therefore given the same weight according to their average homoplasy). Under this weighting scheme, the third positions, often considered more homoplastic than the first two, are given a weight that might be considered artificially high compared with the homoplasy they may have, while the inverse situation occurs with the first two positions.

Morphological characters are frequently not analysed in phylogenetic analyses and especially in studies including thousands of species (e.g. Hackett et al., 2008; Pyron et al., 2013). If considered at all, phenotypic features are often optimized on trees derived from molecular data (e.g. Helmstetter et al., 2016). Also, molecular analyses are often considered refutative over morphology-based hypotheses and, in practice, taken as the single valid source for classifications (e.g. Thomaz et al., 2015). However, most parsimonious trees based only on molecular data are usually suboptimal if morphological characters are added to the analysis. In the present analysis, the morphological characters optimized onto the molecule-only

hypothesis have 1206 steps, versus 1092 steps in the final hypothesis based on combined data. Also, the molecule-only dataset recovers 24 families fewer than the preferred hypothesis based on all data. Therefore, this study underlines why it is problematic to map morphological characters onto a molecule-only hypothesis and, in general, to derive conclusions from this kind of analysis when morphological data are also available for tree building.

The use of morphological characters allowed us to recover as monophyletic 14–24 more families than with DNA data alone. Also, the combined analyses provide morphological synapomorphies for some clades that had been proposed only by molecular analyses, corroborate some of the synapomorphies already proposed, and allow studies of the evolution of structures that had been historically recognized to be repeatedly lost in the Actinopterygii (e.g. the pelvic girdle or the basisphenoid).

The most important factor preventing the inclusion of morphological characters in large-scale phylogenies is the scarcity of this kind information for many clades. The relatively scarce morphological information for many groups and the absence, in many of them, of morphology-based phylogenetic analyses based on a good sample of characters and species became evident during the search of morphological characters for this paper. This is more evident in diverse clades, such as large families and orders. A paradigmatic case is the Cypriniformes that are, in some sense, among the best known fishes in the world. However, the available morphological information is relatively scarce compared with their diversity and there are no published morphology-based phylogenetic hypotheses for that order.

Finally, it must be pointed out that this analysis (and all the preceding ones based on molecular data) excludes information of many species and even entire orders of extinct fishes that have well-known morphology and have been the subject of many phylogenetic studies (e.g. López-Arbarello, 2012; Arratia, 2013; Arratia and Schultze, 2013; Sferco et al., 2015). Those clades and species are important to include if we are to have a better understanding of the phylogeny of the Actinopterygii, but their inclusion in a global phylogenetic analysis of the Actinopterygii requires a better knowledge of the comparative morphology of extant groups.

This analysis corroborates many of the clades proposed by Wiley and Johnson (2010) and/or Betancur-R et al. (2013a,b), but it also challenges the monophyly of some recently proposed groups. Some clades considered to be monophyletic by Betancur-R et al. (2014), such as the Chaetodontiformes, are not obtained as monophyletic in any of the analyses, and others, such as the Centrarchiformes, Pempheriformes or Scombriformes, would need some changes in their compositions to be monophyletic, and even so, they are weakly supported. Also, there are *incertae sedis* families of Perciformes whose low supported relationships preclude their inclusion in existing or new orders. Some differences with previous hypotheses may arise from the use of different analytical methods, but others may result from the inclusion of many taxa that had not been considered in previous analyses.

This analysis corroborates some hypotheses obtained from molecular phylogenies that differ from classifications based only on morphology (e.g. Nelson, 2006; Wiley and Johnson, 2010), such as the inclusion of the Indostomidae in the Synbranchiformes (supported by three morphological synapomorphies, in addition to molecular data). Taxonomic conclusions extracted from the present analysis and especially from those based on more restricted sets of data should be made only in cases with good support or, at least, support from several sources of data.

The aim of this paper is not to provide an alternative classification of the Actinopterygii, but instead to test the current classification in the light of as much available information as possible and to provide general guidelines on how this can be done. In this analysis, the combination of molecular and morphological data and the use of extended implied weighting produced results that are more compatible with the families long recognized as natural groups, compared with equal weighting and/or molecular data alone. Even with 79% of missing entries and a morphological block that is just a small sample of the information that could be included, this analysis recovers 284 of the 376 families having at least two representatives in the analysis.

Acknowledgements

This paper can be seen as a consequence of the phylogenetic analysis of Eukaryota, published with Pablo Goloboff, Claudia Szumik, Santiago Catalano, Salvador Arias, Steve Farris and Mari Källersjö, and I am very grateful to them. I had many interesting discussions, advice, encouragement and suggestions from Pablo Goloboff concerning this paper and in general. I would also thank Ricardo Betancur-R and Guillermo Ortí, who kindly shared with me some experiences in their very important contributions to the phylogeny of the Actinopterygii. I am always indebted to the many ichthyologists who I have studied under, especially Mercedes Azpelicueta, and I am grateful to my closest colleagues Gastón Aguilera, Cristina Butí, Guillermo Terán, Anyelo Vanegas, Felipe Alonso, Fabiana Cancino, Luis Fernández and Baltazar Bugeau. This paper

was almost exclusively done on a Linux platform and I thank Manjaro, Gnome, OpenOffice and Inkscape communities. TNT was provided free by the Willi Hennig Society. This paper was entirely funded by the Argentinian State, through subsidies PICT-2011-0992 (FONCyT) and PIP-0301 (CONICET). I also thank Fundación Miguel Lillo and CONICET for permanent support.

References

Altschul, S.F., Gish, W., Miller, W., Myers, E.W., Lipman, D.J., 1990. Basic local alignment search tool. J. Mol. Biol. 215, 403–410.

Arratia, G., 1999. The monophyly of Teleostei and stem-group teleosts. Consensus and disagreements. In: Arratia, G., Schultze, H.-P. (Eds.), Mesozoic Fishes 2—Systematics and Fossil Record. Verlag Dr. F. Pfeil, München, Germany, pp. 265–334.

Arratia, G., 2013. Morphology, taxonomy, and phylogeny of Triassic pholidophorid fishes (Actinopterygii, Teleostei). J. Vertebr. Paleontol. 33, 1–138.

Arratia, G., Schultze, H.-P., 2013. Outstanding features of a new Late Jurassic on current understanding of pachycormiforms. In: Arratia, G., Schultze, H.-P., Wilson, M.V.H. (Eds.), Mesozoic Fishes 5—Global Diversity and Evolution. Verlag Dr. F. Pfeil, München, Germany, pp. 87–120.

Betancur-R, R., Ortí, G., 2014. Molecular evidence for the monophyly of flatfishes (Carangimorpharia: Pleuronectiformes). Mol. Phylogenet. Evol. 73, 18–22.

Betancur-R, R., Broughton, R.E., Wiley, E.O., Carpenter, K., López, J.A., Li, C., Holcroft, N.I., Arcila, D., Sanciangco, M., Cureton, J.C. II, Zhang, F., Buser, T., Campbell, M.A., Ballesteros, J.A., Roa-Varon, A., Willis, S., Borden, W.C., Rowley, T., Reneau, P.C., Hough, D.J., Lu, G., Grande, T., Arratia, G., Ortí, G., 2013a. The tree of life and a new classification of bony fishes. PLOS Curr. TOL, 1–45. Doi: 10.1371/currents.tol.53ba26640df0ccaee75bb165c8c26288.

Betancur-R, R., Li, C., Munroe, T.A., Balllesteros, J.A., Ortí, G., 2013b. Addressing gene tree discordance and non-stationarity to resolve a multi-locus phylogeny of the flatfishes (Teleostei: Pleuronectiformes). Syst. Biol. 62, 763–785.

Betancur-R, R., Wiley, E.O., Bailly, N., Miya, M., Lecointre, G., Ortí, G., 2014. Phylogenetic Classification of Bony Fishes - Version 3 (http://www.deepfin.org/Classification_v3.htm).

Bremer, K., 1994. Branch support and tree stability. Cladistics 10, 295–304.

Broughton, R.E., Betancur-R, R., Li, C., Arratia, G., Ortí, G., 2013. Multi-locus phylogenetic analysis reveals the pattern and tempo of bony fish evolution. PLOS Curr. TOL, 1–33. doi: 10.1371/currents.tol.2ca8041495ffafd0c92756e75247483e.

Calcagnotto, D., Schaefer, S.A., DeSalle, R., 2005. Relationships among characiform fishes inferred from analysis of nuclear and mitochondrial gene sequences. Mol. Phyl. Evol. 36, 135–153.

Campbell, M.A., Chen, W.J., López, J.A., 2013. Are flatfishes (Pleuronectiformes) monophyletic? Mol. Phyl. Evol. 69, 664–673.

Campbell, M.A., Chen, W.J., López, J.A., 2014. Molecular data do not provide unambiguous support for the monophyly of flatfishes (Pleuronectiformes): a reply to Betancur-R and Ortí. Mol. Phyl. Evol. 75, 149–153.

Chen, J.N., López, J.A., Lavoué, S., Miya, M., Chen, W.J., 2014. Phylogeny of the Elopomorpha (Teleostei): evidence from six nuclear and mitochondrial markers. Mol. Phyl. Evol. 70, 152–161.

Cuvier, G., Valenciennes, A., 1831. Histoire Naturelle des Poissons. Tome septième. Livre huitième. Des Poissons à Pharyngiens Labyrinthiformes, Paris-Strasbourg, France.

Edgar, R.C., 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res. 32, 1792–1797.

Eschmeyer, W.N., 2016. Catalog of fishes: Genera, species, references. (http://research.calacademy.org/research/ichthyology/catalog/fishcatmain.asp). Electronic version.

Eschmeyer, W.N., Fong, J.D., 2016. Species by family/subfamily. (http://researcharchive.calacademy.org/research/ichthyology/catalog/SpeciesByFamily.asp). Electronic version.

Fink, S.V., Fink, W.L., 1981. Interrelationships of ostariophysan fishes (Teleostei). Zool. J. Linn. Soc. 72, 297–353.

Goloboff, P.A., 1993. Estimating character weights during tree search. Cladistics 9, 83–91.

Goloboff, P.A., 1999. Analyzing large data sets in reasonable times: solutions for composite optima. Cladistics 15, 415–428.

Goloboff, P.A., 2014. Extended implied weighting. Cladistics 30, 260–272.

Goloboff, P.A., Catalano, S.A., 2012. GB-to-TNT: facilitating creation of matrices from GenBank and diagnosis of results in TNT. Cladistics 28, 503–513.

Goloboff, P.A., Farris, J.S., Nixon, K.C., 2008. TNT, a free program for phylogenetic analysis. Cladistics 24, 774–786.

Goloboff, P.A., Catalano, S.A., Mirande, J.M., Szumik, C.A., Arias, J.S., Källersjö, M., Farris, J.S., 2009. Phylogenetic analysis of 73 060 taxa corroborates major eukaryotic groups. Cladistics 25, 211–230.

Grande, T., Borden, W.C., Smith, W.L., 2013. Limits and relationships of Paracanthopterygii: a molecular framework for evaluating past morphological hypotheses. In: Arratia, G., Schultze, H.-P., Wilson, M.V.H. (Eds.), Mesozoic Fishes 5—Global Diversity and Evolution. Verlag Dr. F. Pfeil, München, Germany, pp. 385–418.

Greenwood, P.H., 1977. Notes on the anatomy and classification of elopomorph fishes. Bull. Brit. Mus. Nat. Hist. (Zool.) 32, 65–102.

Hackett, S.J., Kimball, R.T., Reddy, S., Bowie, R.C.K., Braun, E.L., Braun, M.J., Chojnowsky, J.L., Cox, W.A., Han, K.L., Harshman, J., Huddleston, C.J., Marks, B.D., Miglia, K.J., Moore, W.S., Sheldon, F.H., Steadman, D.W., Witt, C.C., Yuri, T., 2008. A Phylogenomic study of birds reveals their evolutionary history. Science 320, 1763–1768.

Hall, T.A., 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucl. Acids Symp. Ser. 41, 95–98.

Helmstetter, A.J., Papadopulos, A.S.T., Igea, J., Van Dooren, T.J.M., Leroi, A.M., Savolainen, V., 2016. Viviparity stimulates diversification in an order of fish. Nat. Commun. 7, 11271.

Hoegg, S., Brinkmann, H., Taylor, J.S., Meyer, A., 2004. Phylogenetic timing of the fish-specific genome duplication correlates with the diversification of teleost fish. J. Mol. Evol. 59, 190–203.

Inoue, J.G., Miya, M., Tsukamoto, K., Nishida, M., 2001a. A mitogenomic perspective on the basal teleostean phylogeny: resolving higher-level relationships with longer DNA sequences. Mol. Phyl. Evol. 20, 275–285.

Inoue, J.G., Miya, M., Tsukamoto, K., Nishida, M., 2003. Basal actinopterygian relationships: a mitogenomic perspective on the phylogeny of the "ancient fish". Mol. Phyl. Evol. 26, 110–120.

Inoue, J.G., Miya, M., Tsukamoto, K., Nishida, M., 2004. Mitogenomic evidence for the monophyly of elopomorph fishes (Teleostei) and the evolutionary origin of the leptocephalus larva. Mol. Phyl. Evol. 32, 274–286.

Johnson, D.G., Patterson, C., 1993. Percomorph phylogeny: a survey of acanthomorphs and a new proposal. Bull. Mar. Sci. 52, 554–626.

Livezey, B.C., Zusi, R.L., 2007. Higher-order phylogeny of modern birds (Theropoda, Aves: Neornithes) based on comparative anatomy. II. Analysis and discussion. Zool. J. Linn. Soc. 149, 1–95.

López-Arbarello, A., 2012. Phylogenetic interrelationships of ginglymodian fishes (Actinopterygii: Neopterygii). PLoS ONE 7, e39370.

Mirande, J.M., 2009. Weighted parsimony phylogeny of the family Characidae (Teleostei: Characiformes). Cladistics 25, 574–613.

Mirande, J.M., 2010. Phylogeny of the family Characidae (Teleostei: Characiformes): from characters to taxonomy. Neotrop. Ichthyol. 8, 385–568.

Mirande, J.M., Aguilera, G., Azpelicueta, M.M., 2011. A threatened new species of *Oligosarcus* and its phylogenetic relationships, with comments on *Astyanacinus* (Teleostei: Characidae). Zootaxa 2994, 1–20.

Mirande, J.M., Jerep, F.C., Vanegas-Ríos, J.A., 2013. Phylogenetic relationships of the enigmatic *Carlastyanax aurocaudatus* (Eigenmann) with remarks on the phylogeny of the Stevardiinae (Teleostei: Characidae). Neotrop. Ichthyol. 11, 747–766.

Miya, M., Takeshima, H., Endo, H., Ishiguro, N.B., Inoue, J.G., Mukai, T., Satoh, T.P., Yamaguchi, M., Kawaguchi, A., Mabuchi, K., Shirai, S.M., Nishida, M., 2003. Major pattern of higher teleostean phylogenies: a new perspective based on 100 complete mitochondrial DNA sequences. Mol. Phyl. Evol. 26, 121–138.

Miya, M., Friedman, M., Satoh, T.P., Takeshima, H., Sado, T., Iwasaki, W., Yamanoue, Y., Nakatani, M., Mabuchi, K., Inoue, J.G., Poulsen, J.Y., Fukunaga, T., Sato, Y., Nishida, M., 2013. Evolutionary origin of the Scombridae (tunas and mackerels): members of a paleogene adaptive radiation with 14 other pelagic fish families. PLoS ONE 8, e73535. doi:10.1371/journal.pone.0073535.

Moore, J.A., 1993. Phylogeny of the Trachichthyiformes (Teleostei: Percomorpha). Bull. Mar. Sci. 52, 114–136.

Nakatani, M., Miya, M., Mabuchi, K., Saitoh, K., Nishida, M., 2011. Evolutionary history of Otophysi (Teleostei), a major clade of the modern freshwater fishes: Pangaean origin and Mesozoic radiation. BMC Evol. Biol. 11, 177. doi:10.1186/1471-2148-11-177.

Near, T.J., Eytan, R.I., Dornburg, A., Kuhn, K.L., Moore, J.A., Davis, M.P., Wainwright, P.C., Friedman, M., Smith, W.L., 2012. Resolution of ray-finned fish phylogeny and timing of diversification. Proc. Natl Acad. Sci. USA 109, 13698–13703.

Near, T.J., Dornburg, A., Eytan, R.I., Keck, B.P., Smith, W.L., Kuhn, K.L., Moore, J.A., Price, S.A., Burbrink, F.T., Friedman, M., Wainwright, P.C., 2013. Phylogeny and tempo of diversification in the superradiation of spiny-rayed fishes. Proc. Natl Acad. Sci. USA 101, 12738–12743.

Nelson, J.S., 2006. Fishes of the World, 4th edn. John Wiley & Sons, Inc., New York.

O'Leary, M.A., Kaufman, S.G., 2011. MorphoBank: phylophenomics in the 'cloud'. Cladistics 27, 1–9.

O'Leary, M.A., Kaufman, S.G., 2012. MorphoBank 3.0: Web application for morphological phylogenetics and taxonomy. http://www.morphobank.org.

O'Leary, M.A., Bloch, J.I., Flynn, J.J., Gaudin, T.J., Giallombardo, A., Giannini, N.P., Goldberg, S.L., Kraatz, B.P., Luo, Z.X., Meng, L., Ni, X., Novacek, M.J., Perini, F.A., Randall, Z.S., Rougier, G.W., Sargis, E.J., Silcox, M.T., Simmons, N.B., Spaulding, M., Velazco, P.M., Weksler, M., Wible, J.R., Cirranello, A.L., 2013. The placental mammal ancestor and the post–K-Pg radiation of placentals. Science 339, 662–667.

Pyron, R.A., Burbrink, F.T., Wiens, J.J., 2013. A phylogeny and revised classification of Squamata, including 4161 species of lizards and snakes. BMC Evol. Biol. 13, 93.

Sferco, E., López-Arbarello, A., Báez, A.M., 2015. Anatomical description and taxonomy of †*Luisiella feruglioi* (Bordas) new combination, a freshwater teleost (Actinopterygii, Teleostei) from the Upper Jurassic of Patagonia. BMC Evol. Biol. 15, 268.

Smith, W.L., Craig, M.T., 2007. Casting the percomorph net widely: the importance of broad taxonomic sampling in the search for the placement of serranid and percid fishes. Copeia 2007, 35–55.

Smith, W.L., Wheeler, W.C., 2004. Polyphyly of the mail-cheeked fishes (Teleostei: Scorpaeniformes): evidence from mitochondrial and nuclear sequence data. Mol. Phyl. Evol. 32, 627–646.

Sullivan, J.P., Lundberg, J.G., Hardman, M., 2006. A phylogenetic analysis of the major groups of catfishes (Teleostei: Siluriformes) using rag1 and rag2 nuclear gene sequences. Mol. Phyl. Evol. 41, 636–662.

Thomaz, A.T., Arcila, D., Ortí, G., Malabarba, L.R., 2015. Molecular phylogeny of the subfamily Stevardiinae Gill, 1858 (Characiformes: Characidae): classification and the evolution of reproductive traits. BMC Evol. Biol. 15, 146.

Tyler, J.C., Johnson, G.D., Nakamura, I., Collette, B.B., 1989. Morphology of *Luvarus imperialis* (Luvaridae), with a phylogenetic analysis of the Acanthuroidei. Smith. Contr. Zool. 485, 1–78.

Vari, R.P., 1979. Anatomy, relationships and classification of the families Citharinidae and Distichodontidae (Pisces, Characoidea). Bull. Brit. Mus. Nat. Hist. (Zool.) 36, 261–344.

Wainwright, P.C., Smith, W.L., Price, S.A., Tang, K.L., Sparks, J.S., Ferry, L.A., Kuhn, K.L., Eytan, R.I., Near, T.J., 2012. The evolution of pharyngognathy: a phylogenetic and functional appraisal of the pharyngeal jaw key innovation in labroid fishes and beyond. Syst. Biol. 61, 1001–1027.

Wiley, E.O., Johnson, G.D., 2010. A teleost classification based on monophyletic groups. In: Nelson, J.S., Schultze, H.-P., Wilson, M.V.H. (Eds.). Origin and Phylogenetic Interrelationships of Teleosts. Verlag Dr. Friedrich Pfeil, München, Germany, pp. 123–182.

Zanata, A.M., Vari, R.P., 2005. The family Alestidae (Ostariophysi, Characiformes): a phylogenetic analysis of a trans-Atlantic clade. Zool. J. Linn. Soc. 145, 1–144.

Supporting Information

Additional Supporting Information may be found in the online version of this article:

Appendix S1. Dataset, including the 14 141 species having at least one DNA sequence and their accession numbers.

Appendix S2. List of DNA sequences removed from the final dataset for possible contamination or identification problems (S2_removed.txt).

Appendix S3. List of the number of sequences analysed by each species (S3_in.txt).

Appendix S4. List of morphological characters (S4_morpho.doc).

Appendix S5. TNT script used to assign states from the literature to morphological characters (S5_morphology.tnt).

Appendix S6. Dataset of 8104 species restricted to the first block of data, including the taxonomy of species (S6_names.zip; extract with Winzip or File Roller and execute with TNT).

Appendix S7. Consensus trees for analyses under different parameters in TNT format.

Appendix S8. General topology of the final hypothesis (consensus of GR3) showing relationships between families (Fit = 4620.20103; Length = 2639 520 steps).

Appendix S9. Agreement subtree (2500 spp.) showing only the relationships shared by most parsimonious trees obtained in all the analyses, prunning taxa or clades floating with different parameters of search (S9_agreement.zip; extract with Winzip or File Roller and open with a web browser).

Appendix S10. Agreement subtree (3710 spp.) showing only the relationships shared by most parsimonious trees obtained in the analyses of the combined dataset without grouping or grouping contiguous columns to collectively weight them against their average homoplasy (SEP, GR3, GR9, G27 and GEN)

(S10_agreementIW.zip; extract with Winzip or File Roller and open with a web browser).

Appendix S11. General topology of the final hypothesis (consensus of GR3) (Fit = 4620.20103; Length =2639 520 steps).

Appendix S12. General topology of the final hypothesis (consensus of GR3) (Fit = 4620.20103; Length =2639 520 steps) showing morphological synapomorphies.

Appendix S13. General topology of the consensus of most parsimonious trees under SEP (Fit = 3157.58185; Length = 2647 426 steps) (S13_SEP.zip; extract with Winzip or File Roller and open with a web browser).

Appendix S14. General topology of the consensus of most parsimonious trees under GR9 (Fit = 4758.99827; Length = 2639 476 steps) (S14_GR9.zip; extract with Winzip or File Roller and open with a web browser).

Appendix S15. General topology of the consensus of most parsimonious trees under G27 (Fit = 4827.49437; Length = 2639 826 steps) (S15_G27.zip; extract with Winzip or File Roller and open with a web browser).

Appendix S16. General topology of the consensus of most parsimonious trees under GEN (Fit = 4902.70047; Length = 2639 497 steps) (S16_GEN.zip; extract with Winzip or File Roller and open with a web browser).

Appendix S17. General topology of the consensus of most parsimonious trees under POS (Fit = 4044.83082; Length = 2642 422 steps) (S17_POS.zip; extract with Winzip or File Roller and open with a web browser).

Appendix S18. General topology of the consensus of most parsimonious trees under equal weights (EQW) (Length = 2629 862 steps) (S18_EQW.zip; extract with Winzip or File Roller and open with a web browser).

Appendix S19. General topology of the consensus of most parsimonious trees in the molecule-only analysis (MOL) (Fit = 4616.26673; Length = 2639 474 steps) (S19_MOL.zip; extract with Winzip or File Roller and open with a web browser).

Appendix S20. Systematic results and discussions for each clade to familial level (S20_discussion.doc).

Appendix S21. List of autapomorphies and synapomorphies obtained under GR3. Node numbers correspond to those of Appendix S10 (apom.txt.7z).