



ACADÉMIE
DES SCIENCES
INSTITUT DE FRANCE

Comptes Rendus

Biologies


Elise Parey, Camille Berthelot, Hugues Roest Crolius and Yann Guiguen

Solving an enigma in the tree of life, at the origins of teleost fishes

Volume 347 (2024), p. 1-8

Online since: 5 March 2024

<https://doi.org/10.5802/crbiol.150>

 This article is licensed under the
CREATIVE COMMONS ATTRIBUTION 4.0 INTERNATIONAL LICENSE.
<http://creativecommons.org/licenses/by/4.0/>



*The Comptes Rendus. Biologies are a member of the
Mersenne Center for open scientific publishing*
www.centre-mersenne.org — e-ISSN : 1768-3238



Review article

Solving an enigma in the tree of life, at the origins of teleost fishes

Elise Parey^{Ⓢ,*,a,b}, Camille Berthelot^{Ⓢ,a,c}, Hugues Roest Crollius^{Ⓢ,a} and Yann Guiguen^{Ⓢ,b}

^a Institut de Biologie de l'ENS (IBENS), Département de Biologie, École Normale Supérieure, CNRS, INSERM, Université PSL, 75005 Paris, France

^b INRAE, LPGP, Rennes, France

^c Institut Pasteur, Université Paris Cité, CNRS UMR 3525, INSERM UA12, Comparative Functional Genomics Group, F-75015 Paris, France

Current address: Centre for Life's Origins & Evolution, Department of Genetics, Evolution & Environment, University College London, London, UK (E. Parey)

E-mails: e.parey@ucl.ac.uk (E. Parey), camille.berthelot@pasteur.fr (C. Berthelot), hrc@bio.ens.psl.eu (H. Roest Crollius), yann.guiguen@inrae.fr (Y. Guiguen)

Abstract. Tracing the phylogenetic relationships between species is one of the fundamental objectives of evolutionary biology. Since Charles Darwin's seminal work in the 19th century, considerable progress has been made towards establishing a tree of life that summarises the evolutionary history of species. Nevertheless, substantial uncertainties still remain. Specifically, the relationships at the origins of teleost fishes have been the subject of extensive debate over the last 50 years. This question has major implications for various research fields: there are almost 30,000 species in the teleost group, which includes invaluable model organisms for biomedical, evolutionary and ecological studies. Here, we present the work in which we solved this enigma. We demonstrated that eels are more closely related to bony-tongued fishes than to the rest of teleost fishes. We achieved this by taking advantage of new genomic data and leveraging innovative phylogenetic markers. Notably, in addition to traditional molecular phylogeny methods based on the evolution of gene sequences, we also considered the evolution of gene order along the DNA molecule. We discuss the challenges and opportunities that these new markers represent for the field of molecular phylogeny, and in particular the possibilities they offer for re-examining other controversial branches in the tree of life.

Keywords. Evolution, Phylogeny, Teleost fishes, Synteny, Genomics.

Note. Elise Parey received the Grandes Avancées en Biologie 2023 award.

Funding. Agence Nationale de la Recherche, France (ANR) (Grant agreement no. ANR-16-CE12-003), European Union Horizon 2020 research and innovation program (Grant agreement no. 817923 (AQUA-FAANG)), Newton International Fellowship from the Royal Society (NIF/R1/222125).

Published online: 5 March 2024

1. Reconstructing the tree of life: an ambitious goal initiated by Charles Darwin

The ambitious aim of reconstructing the tree of life consists in representing all 8 million extant species [1] on a single tree that would recapitulate

their evolutionary—or phylogenetic—relationships. In this tree, leaves represent extant species and nodes their common ancestors. Establishing a phylogenetic tree of life is first and foremost a way of listing and classifying the diversity of life on Earth. The tree of life is particularly valuable to conservation efforts, because it rigorously documents the diversity of living organisms and thus informs on the impacts of environmental upheavals. The distribution of

* Corresponding author.

endangered species along the tree of life is informative to predict which branches are the most likely to collapse and lead to the greatest losses in terms of evolutionary diversity [2].

The first sketch of a species tree is attributed to Jean-Baptiste de Lamarck, in 1809 [3]. It was only 50 years later that Charles Darwin definitively popularised the concept of the tree of life, in his seminal work *The Origin of Species* [4]. We also owe Ernst Haeckel many of the terms still in use in the field of evolutionary biology; inspired by Darwin's writings, he proposed the term “phylogeny” for the first time [5]. Early phylogenetic studies used anatomical similarities between species to predict their evolutionary proximity. A century later, the DNA molecule was recognised as a “document of evolutionary history”: a powerful lever for reconstructing evolutionary relationships between species [6]. It became apparent that the mechanisms through which DNA is modified and transmitted imply that the similarity of DNA sequences between species is an indicator of their relatedness. This was followed by the advent of molecular phylogeny, and a continuous improvement of sequence evolution models, which made it possible to infer the history of life on Earth ever more completely and accurately.

Despite the advent of genomics and the refinement of DNA sequencing techniques, several nodes of the animal phylogeny remain debated within the scientific community [7, 8]. One of the most controversial nodes is the one at the base of the animal tree: which of sponges or ctenophores diverged first from the rest of animal species? This question is hotly debated, partly because its resolution will shed light on the evolution of neurons and muscles. Although morphologically distinct [9], neuronal and muscle cells are present in ctenophores and most of the rest of the animals, but are notably absent from sponges. The different phylogenies proposed at the base of the animal tree imply different evolutionary scenarios for the origin of neurons and muscles. For example, the “ctenophores-first” hypothesis implies either a common origin in the ancestor of animals with a secondary loss in sponges, or independent evolution in each of the two groups (ctenophores and the rest of the animals). Thus, resolving this node will provide the necessary framework to evaluate these evolutionary scenarios. A second challenge concerns the position of the enigmatic marine worm *Xenoturbella*,

nicknamed the “deep-sea purple sock” because of its simplified morphology (no eye, no digestive system, no brain). The latter was alternatively placed with molluscs, as the sister group to all bilaterian animals, or close to the sea stars group. Finally, because of their impressive diversity, considerable work remains to be done to draw a complete picture of the evolution of ray-finned fishes (actinopterygians) [10]. Our work has focused on resolving a debated node at the base of the largest fish clade: the origin of teleost fishes.

2. Fifty years of debates to resolve the origin of teleost fishes

The teleost group encompasses more than 96% of all fish species. With a total of over 30,000 recorded species, it includes as many species as the Tetrapods (Amphibians, Mammals, Birds and Reptiles) (Figure 1A). Teleost fishes are subdivided into three groups: the Elopomorpha (tarpon, moray eel, eel), the Osteoglossomorpha or “bony-tongued fish” (arowana, mormyrid) and the Clupeocephala, which make up the majority of teleosts (zebrafish, tetraodon, stickleback, cod, pike, etc.). It is estimated that the last common ancestor of these three groups dates back to the Triassic period, around 250 million years ago [11].

Resolving the evolutionary relationships between these three groups represents a major challenge, one which the community has been grappling with for over 50 years. The first studies were based on anatomical criteria and the analysis of fossils. They first proposed to group the Elopomorpha together with the Clupeocephala (Ref. [15]; Figure 1B), and later the Osteoglossomorpha with the Clupeocephala (Ref. [13]; Figure 1C). This question was subsequently revisited with the advent of molecular phylogeny methods which, although considering increasingly large quantities of data, have alternately supported each of the three possible groupings (Refs. [11, 14, 16–19]; Figure 1B,C,D).

Several hypotheses can be put forward to explain such incongruences across studies. A first argument proposes that the instability of the Osteoglossomorpha and Elopomorpha positions in the reconstructed phylogenies might be due to their large underrepresentation (on average 14 times fewer species included than Clupeocephala in the previously cited

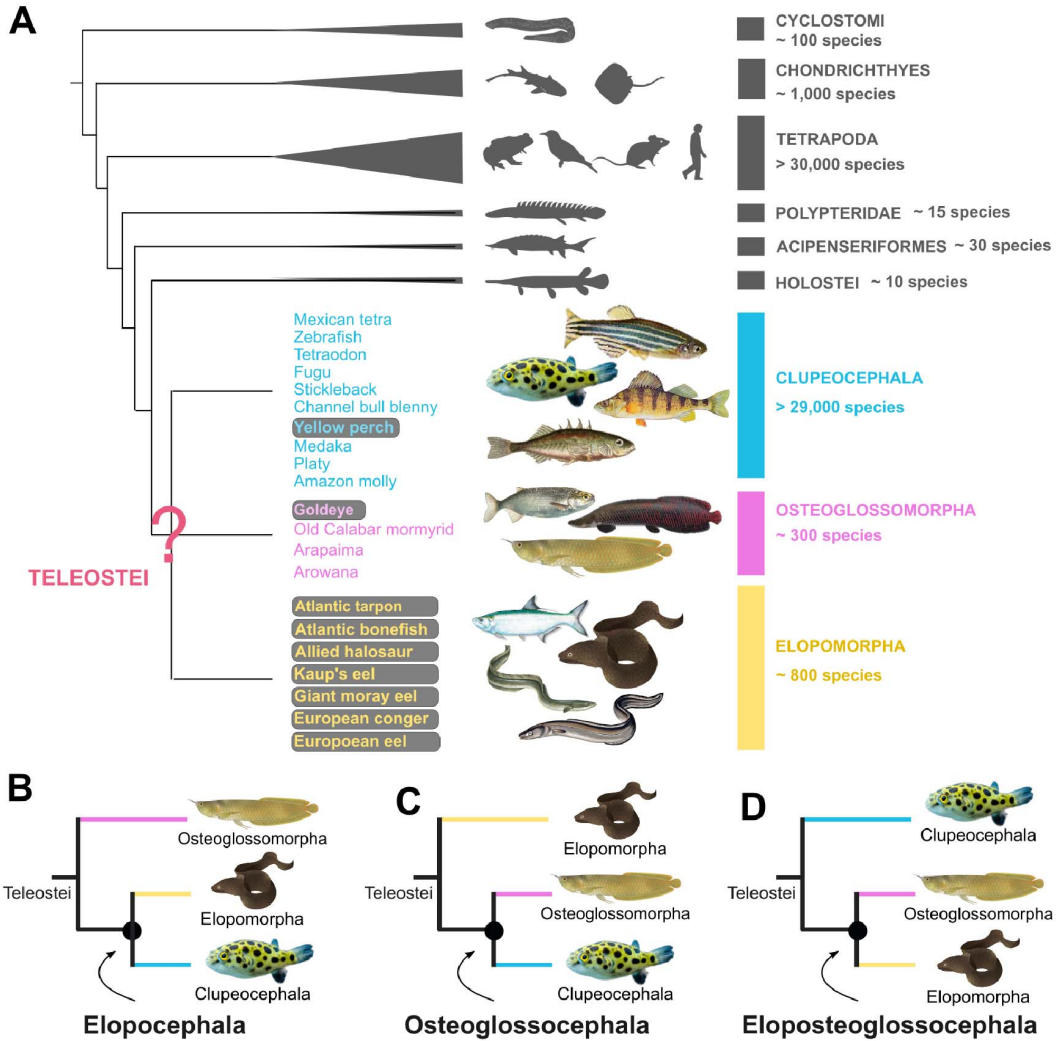


Figure 1. Evolutionary relationships at the origins of teleost fishes. (A) Phylogenetic position of teleost fishes within vertebrates. Species listed to illustrate each of the three major groups (Clupeocephala, Osteoglossomorpha, Elopomorpha) correspond to species considered in our study, including those for which we generated new genomes (grey background). See the original study [12] for the full scientific names of included species. (B–D) Different topologies proposed at the base of the teleost tree: Elopoccephala topology [13–15], Osteoglossocephala [11, 16] and Eloposteoglossocephala [17–19]. Figure adapted from [12].

studies). Other methodological considerations point to the diversity in the methods, and the relevance of technical choices made to model the evolution of sequences. Briefly, two main families of approaches are generally used: (i) the “concatenation” method, which considers all of the DNA sequences in a single block to directly reconstruct the phylogeny of species, and (ii) the “consensus” method, which

considers genes separately to reconstruct a tree for each and deduce the phylogeny. The advantage of the concatenation method is that the quantity of data considered maximises the phylogenetic signal, but the disadvantage is that it is based on the assumption that all sequences share the same evolutionary history, a simplification rarely verified in practice. In contrast, gene trees of the consensus method are

more prone to estimation errors due to lower statistical power, but this method accounts for distinct evolutionary histories across different parts of genomes. A phenomenon known as “incomplete lineage sorting” is most often responsible for these discordant evolutionary histories within the same genome. Incomplete lineage sorting occurs when pre-existing genetic differences within an ancestral population (different alleles of a gene, for example) are retained differentially in descendant species. For instance, if two alleles A and B of the same gene existed in the ancestral population of teleost fish and each of the three major groups randomly retained either allele A or allele B, then the evolutionary history of the sequence of this gene will not necessarily follow the phylogeny of the species. The effect of incomplete lineage sorting is all the more important when several groups of species diverge from each other in a short period of time, as was the case at the origins of teleost fishes. Introgression is another biological phenomenon that can blur the phylogenetic signal: the occurrence of hybridisation between ancient populations of the three major groups could have led to exchanges of DNA between lineages after they had separated. Consensus methods often model incomplete lineage sorting, but rarely introgression.

Finally, one of the main challenges in reconstructing the phylogenies of teleost fishes is tied to their complex evolutionary history: all extant teleosts descend from an ancestor that underwent a whole-genome duplication event [20]. As a result, many genes still exist in two copies within teleost genomes, making it difficult to identify “marker” genes, i.e. genes that are comparable between species and can be leveraged for phylogenetic analyses. Specifically, the challenge consists in distinguishing orthologous genes (genes that descend from the same copy of the ancestral gene) from paralogous genes (genes that descend from duplicated copies). The duplication event that separates paralogous genes from one another introduces a discrepancy between the evolutionary history of the gene and of the species. Orthologous gene sequences, on the other hand, directly reflect the evolutionary history of species and are therefore good markers for reconstructing their phylogeny. Although the difficulties encountered in phylogenetic reconstructions can be explained by a combination of several of the factors mentioned above, the confounding effect of genome duplication

has been formally demonstrated [21] and has probably largely contributed to the incongruities reported at the origins of teleost fishes. A parallel can be drawn with the rapid diversification of the three main families of Salmonidae whose evolutionary relationships have long been ambiguous [22], linked to a whole-genome duplication event in their ancestral lineage.

3. A reanalysis in the light of new genomic data

In order to resolve the phylogenetic relationships at the origin of teleost fishes, we performed new analyses, designed to mitigate the limitations of previous studies. As part of our study, we generated new genomic resources, in particular for the Elopomorpha for which we sequenced the genomes of 7 species (Figure 1A). In total, we considered 25 genomes (Figure 1A), carefully selected to avoid significantly over-representing any one group over the others: 7 Elopomorpha, 4 Osteoglossomorpha, 10 Clupeocephala and 4 non-teleost vertebrates.

The reconstruction of a molecular phylogeny involves three main steps: (i) the identification of marker genes, present and confidently identifiable across all considered genomes (orthologous genes), (ii) the alignment of these gene sequences across species, to reveal positions that have changed during evolution (iii) the inference of a phylogenetic tree based on the observed sequence changes in the sequences. To identify marker genes across teleost fishes, a task made complicated by their ancient genome duplication event, we built upon our previous work [23,24]. We had developed methods specifically tailored to the specificities of teleost genomes, enabling to establish a set of marker genes both more complete and more robust than leveraged in previous studies. This ortholog identification method is based on a signature left in the genomes following whole-genome duplication events. Initially, all of the chromosomes are duplicated, but, subsequently, duplicated chromosomes evolve independently and accumulate distinct gene losses and genomic rearrangements, making it possible to differentiate and identify them across species. We identify orthologous genes on the basis of their sequence conservation, but also on the conservation of their local genomic environment, which reflects their common chromosomal origin [23, 24]. Using this approach, we consider a set of 955 marker genes, representing

around 5% of the complete gene repertoire, for a total size of aligned gene sequences of 2,328,657 nucleotides. In comparison, the alignments analysed in the most comprehensive previous studies [11, 18, 19] comprised between 500,000 and 1,000,000 nucleotides, i.e. approximately 2 to 5 times less.

We also took advantage of a wide range of different methodologies: we reconstructed a total of 16 phylogenetic trees, through both concatenation (direct inference of a single tree based on the entire gene set) and consensus (reconstruction of one tree per marker gene, subsequently reconciled into a single phylogeny) approaches. All these phylogenetic analyses converged to the same topology: the Eloposteoglossocephala phylogeny (Figure 1D). Although proposed by several previous studies, this is the first time that this topology has been supported regardless of the methodology employed.

4. Resolving the origins of teleost fishes through innovative methods

These new phylogenetic analyses represent an important step towards shedding light on the evolutionary relationships at the origins of teleost fishes. Nevertheless, the main contribution of our work lies in the large size of the dataset considered (number of genes), an advantage highlighted by each of the previous studies compared with their predecessors. With the aim of breaking apart from these previous debates, we also set out to take advantage of innovative molecular phylogeny methods based on novel markers.

Traditional molecular phylogeny methods leverage changes observed in the DNA sequences of genes. However, sequences are not the only genomic feature that accumulates changes over time. Genome structures are also dynamic, and involve, over the course of generations, changes in the ordering of genes along the chromosomes. These modifications accumulate more slowly than changes affecting sequences, and therefore offer a complementary perspective to study genomes evolution. Moreover, gene order evolution is potentially less affected by introgression [25]. Although the relevance of structural genomic changes for evolutionary studies was demonstrated almost a century ago [26], it is only very recently that they have started to be used to reconstruct species phylogenies [12, 25, 27, 28].

The mechanisms governing the evolution of gene organisation on chromosomes—the evolution of synteny—remain less characterised than those leading to sequence changes. As a result, there is no well-established probabilistic model to describe the evolution of synteny. In the absence of models, we applied methods that estimate evolutionary distances between pairs of genomes, and subsequently reconstruct a tree that best reflects these distances. Here, we estimated evolutionary distances by quantifying the degree of gene reordering between two genomes under consideration (Figure 2A,B), and used the Neighbor-Joining algorithm [29] to reconstruct a tree from this distance matrix (Figure 2C). The Neighbor-Joining algorithm was previously widely used in the field of molecular phylogeny, before the advent of probabilistic methods. We also applied a similar method (PhyChro, [27]), which offers an improvement on the Neighbor-Joining method, to adapt it to genome organisation markers. In particular, in the PhyChro algorithm, the distance calculation between a pair of genomes also examines all the other genomes included in the analysis, in order to specifically consider the genomic rearrangements that are the most informative to reconstruct the phylogeny.

We reconstructed a total of five teleost phylogenies based on the organisation of their genomes, examined at different scales (in particular: conserved adjacencies between marker genes, ordering of gene blocks and organisation of entire chromosomes). In each of the five reconstructed phylogenies, we again recover the Eloposteoglossocephala topology, which groups the Elopomorpha with the Osteoglossomorpha (Figure 2C). We note that these different approaches use marker gene sets identified by different strategies (see [12] for more details), implying that the result is robust to the differences across considered gene sets. In conclusion, through a wide range of complementary phylogenetic analyses, we have been able to resolve the evolutionary relationships at the origin of teleost fishes and thus demonstrate that eels are more closely related to bony-tongued fishes than to the rest of the teleost fishes.

5. Towards a more systematic use of synteny to solve the tree of life

One fundamental prerequisite has to be met in order to use genome structures as a lever to trace their

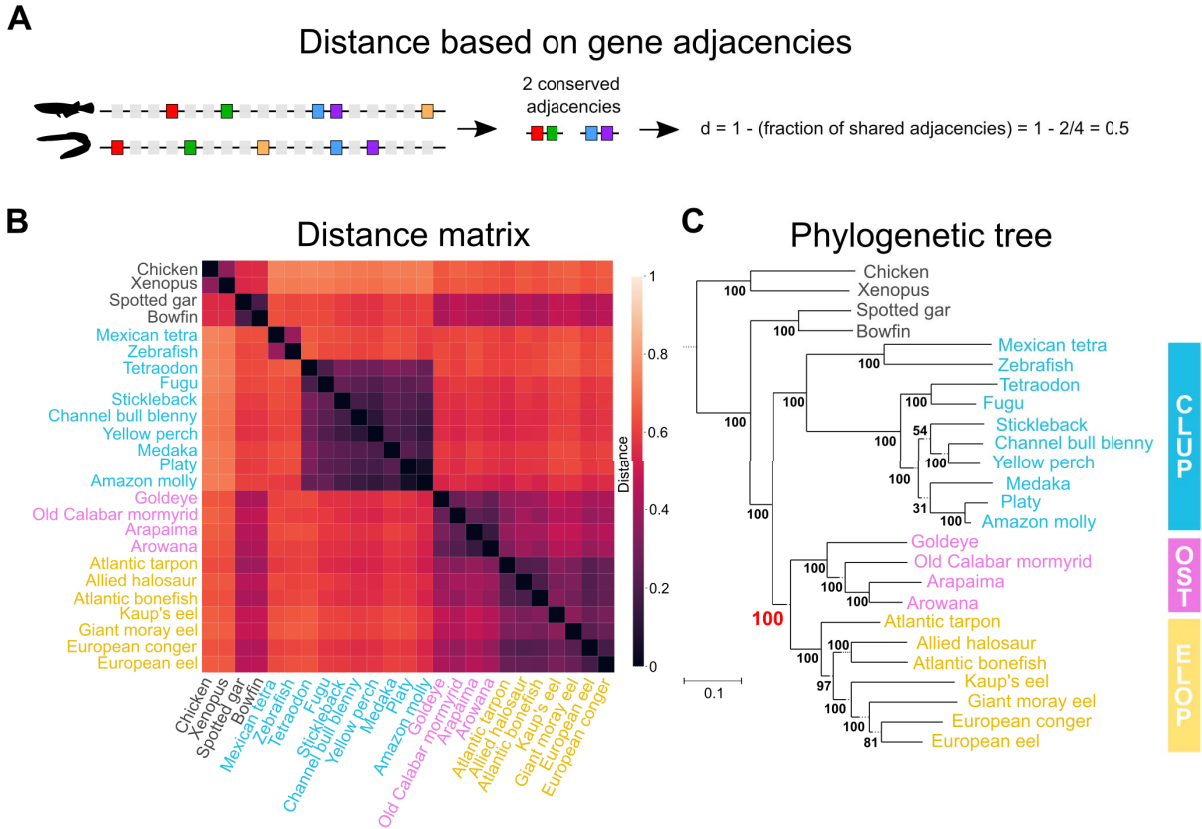


Figure 2. Teleost fishes phylogeny based on gene order evolution. (A) Definition of a distance between two genomes based on the conservation of gene adjacencies. Genes are represented by rectangles along a DNA segment, with “marker” genes (i.e. genes identifiable in all considered genomes) shown in colour. A measure of similarity is calculated on the basis of the proportion of conserved adjacencies between marker genes in the two genomes being compared, then transformed into a distance. (B) Pairwise distance matrix obtained for the 25 genomes considered and 3041 marker genes, based on the d distance introduced in (A). The end-to-end distance covered by the markers considered corresponds on average to 86% of the genome. (C) Phylogenetic tree reconstructed from the distance matrix using the Neighbor-Joining algorithm [29]. Osteoglossomorpha are grouped with Elopomorpha (Eloposteoglossocephala topology, Figure 1D), with strong support (100% bootstrap).

evolutionary relationships: the genomes considered must be of sufficient quality to provide a complete and accurate representation of gene orderings along chromosomes. These high-quality genomic resources are already available for many species, but their quantity is set to explode in the coming years, thanks in particular to many large-scale biodiversity sequencing projects (African BioGenome Project, ATLASa, Darwin Tree of Life, Earth Biogenome Project, European Reference Genome Atlas). In this context, a more systematic use of synteny to reconstruct

phylogenies is becoming a feasible and promising next step, although this still presents many methodological challenges [30]. In particular, deep nodes remain intrinsically difficult to resolve due to the substantial amount of elapsed evolutionary time, which erodes phylogenetic signal both in DNA sequences and genome structures. For the biological signal provided by synteny to be usable in phylogenetic analyses, gene order must be variable between study species without being completely shuffled. Defining indicators that quantify synteny conservation and

degradation limits will represent a crucial step towards defining the optimum range of application for these new methods.

Recently, a study examined synteny evolution patterns to investigate the evolutionary relationships at the origin of animals [28], a node dated around 650 million years ago. This work evidenced similarities between the organisation of the genomes of sponge and other animals, thus supporting a phylogeny in which ctenophores constitute the sister group of a clade bringing together sponges and the rest of animals. This surprising result, which contradicts the most recent molecular phylogeny analyses [8, 31, 32], is galvanising research efforts to better understand the methodological and/or biological reasons underlying these incongruences. The thousands high-quality genomes that will become available in the near future offer unprecedented opportunities to understand the mechanisms that govern the evolution of genome organisation, and re-examine many controversial branches of the tree of life.

Declaration of interests

The authors do not work for, advise, own shares in, or receive funds from any organization that could benefit from this article, and have declared no affiliations other than their research organizations.

Funding

This work was supported by the Agence Nationale de la Recherche, France (ANR) on the GenoFish project, 2016–2021 (grant No. ANR-16-CE12-003) and the European Union Horizon 2020 research and innovation program under Grant Agreement No 817923 (AQUA-FAANG). EP is currently supported by a Newton International Fellowship from the Royal Society (NIF/R1/222125).

Acknowledgements

We thank all of our co-authors from the original study, for their contribution to the previous work that we cover in this review article: Alexandra Louis, Jerome Montfort, Olivier Bouchez, Céline Roques, Carole Iampietro, Jerome Lluch, Adrien Castinel, Cécile Donnadiou, Thomas Desvignes, Christabel

Floi Bucuo, Elodie Jouanno, Ming Wen, Sahar Mejri, Ron Dirks, Hans Jansen, Christiaan Henkel, Wei-Jen Chen, Margot Zahm, Cédric Cabau, Christophe Klopp, Andrew W. Thompson, Marc Robinson-Rechavi, Ingo Braasch, Guillaume Lecointre, Julien Bobe and John H. Postlethwait.

References

- [1] C. Mora, D. P. Tittensor, S. Adl, A. G. B. Simpson, B. Worm, “How many species are there on earth and in the ocean?”, *PLoS Biol.* **9** (2011), no. 8, article no. e1001127.
- [2] W. Thuiller, S. Lavergne, C. Roquet, I. Boulangeat, B. Lafourcade, M. B. Araujo, “Consequences of climate change on the tree of life in Europe”, *Nature* **470** (2011), p. 531-534.
- [3] J.-B. Lamarck, *Philosophie Zoologique*, Baillière, Paris, 1809.
- [4] C. Darwin, *On the Origin of Species by Means of Natural Selection, or, The Preservation of Favoured Races in the Struggle for Life*, J. Murray, London, 1859.
- [5] E. Haeckel, *Generelle Morphologie der Organismen : allgemeine Grundzüge der organischen Formen-Wissenschaft, mechanisch begründet durch die von Charles Darwin reformirte Descendenz-Theorie*, G. Reimer, Berlin, 1866.
- [6] E. Zuckerkandl, L. Pauling, “Molecules as documents of evolutionary history”, *J. Theor. Biol.* **8** (1965), p. 357-366.
- [7] G. Giribet, “Genomics and the animal tree of life: conflicts and future prospects”, *Zool. Scr.* **45** (2016), p. 14-21.
- [8] P. Kapli, M. J. Telford, “Topology-dependent asymmetry in systematic errors affects phylogenetic placement of Ctenophora and Xenacoelomorpha”, *Sci. Adv.* **6** (2020), article no. eabc5162.
- [9] C. W. Dunn, S. P. Leys, S. H. Haddock, “The hidden biology of sponges and ctenophores”, *Trends Ecol. Evol.* **30** (2015), p. 282-291.
- [10] A. Dornburg, T. J. Near, “The emerging phylogenetic perspective on the evolution of actinopterygian fishes”, *Annu. Rev. Ecol. Evol. Syst.* **52** (2021), p. 427-452.
- [11] L. C. Hughes, G. Ortí, Y. Huang, Y. Sun, C. C. Baldwin, A. W. Thompson, D. Arcila, R. Betancur-R, C. Li, L. Becker, N. Bellora, X. Zhao, X. Li, M. Wang, C. Fang, B. Xie, Z. Zhou, H. Huang, S. Chen, B. Venkatesh, Q. Shi, “Comprehensive phylogeny of ray-finned fishes (Actinopterygii) based on transcriptomic and genomic data”, *Proc. Natl. Acad. Sci.* **115** (2018), p. 6249-6254.
- [12] E. Parey, A. Louis, J. Montfort, O. Bouchez, C. Roques, C. Iampietro, J. Lluch, A. Castinel, C. Donnadiou, T. Desvignes, C. Floi Bucuo, E. Jouanno, M. Wen, S. Mejri, R. Dirks, H. Jansen, C. Henkel, W.-J. Chen, M. Zahm, C. Cabau, C. Klopp, A. W. Thompson, M. Robinson-Rechavi, I. Braasch, G. Lecointre, J. Bobe, J. H. Postlethwait, C. Berthelot, H. Roest Crolius, Y. Guiguen, “Genome structures resolve the early diversification of teleost fishes”, *Science* **379** (2023), p. 572-575.
- [13] G. Arratia, “Basal teleosts and teleostean phylogeny”, *Paleo Ichtyol.* **7** (1997), p. 1-168.

- [14] C. M. Austin, M. H. Tan, L. J. Croft, M. P. Hammer, H. M. Gan, "Whole genome sequencing of the asian arowana (*Scleropages formosus*) provides insights into the evolution of ray-finned fishes", *Genome Biol. Evol.* **7** (2015), p. 2885-2895.
- [15] C. Patterson, D. E. Rosen, "Review of ichthyodectiform and other Mesozoic teleost fishes, and the theory and practice of classifying fossils", *Bull. Am. Mus. Nat. Hist.* **158** (1977), article no. 2.
- [16] T. J. Near, R. I. Eytan, A. Dornburg, K. L. Kuhn, J. A. Moore, M. P. Davis, P. C. Wainwright, M. Friedman, W. L. Smith, "Resolution of ray-finned fish phylogeny and timing of diversification", *Proc. Natl. Acad. Sci. USA* **109** (2012), p. 13698-13703.
- [17] C. Bian, Y. Hu, V. Ravi, I. S. Kuznetsova, X. Shen, X. Mu, Y. Sun, X. You, J. Li, X. Li, Y. Qiu, B.-H. Tay, N. M. Thevasagayam, A. S. Komissarov, V. Trifonov, M. Kabilov, A. Tupikin, J. Luo, Y. Liu, H. Song, C. Liu, X. Wang, D. Gu, Y. Yang, W. Li, G. Polgar, G. Fan, P. Zeng, H. Zhang, Z. Xiong, Z. Tang, C. Peng, Z. Ruan, H. Yu, J. Chen, M. Fan, Y. Huang, M. Wang, X. Zhao, G. Hu, H. Yang, J. Wang, J. Wang, X. Xu, L. Song, G. Xu, P. Xu, J. Xu, S. J. O'Brien, L. Orbán, B. Venkatesh, Q. Shi, "The asian arowana (*Scleropages formosus*) genome provides new insights into the evolution of an early lineage of teleosts", *Sci. Rep.* **6** (2016), article no. 24501.
- [18] M.-Y. Chen, D. Liang, P. Zhang, "Selecting question-specific genes to reduce incongruence in phylogenomics: a case study of jawed vertebrate backbone phylogeny", *Syst. Biol.* **64** (2015), p. 1104-1120.
- [19] R. A. Vialle, J. E. S. de Souza, K. de P. Lopes, D. G. Teixeira, P. de A. Alves Sobrinho, A. M. Ribeiro-dos Santos, C. Furtado, T. Sakamoto, F. A. Oliveira Silva, E. Herculano Corrêa de Oliveira, I. G. Hamoy, P. P. Assumpção, Â. Ribeiro-dos Santos, J. P. M. Santos Lima, H. N. Seuánez, S. J. de Souza, S. Santos, "Whole genome sequencing of the pirarucu (*arapaima gigas*) supports independent emergence of major teleost clades", *Genome Biol. Evol.* **10** (2018), p. 2366-2379.
- [20] O. Jaillon, J.-M. Aury, F. Brunet, J.-L. Petit, N. Stange-Thomann, E. Mauceli, L. Bouneau, C. Fischer, C. Ozouf-Costaz, A. Bernot, S. Nicaud, D. Jaffe, S. Fisher, G. Lutfalla, C. Dossat, B. Segurens, C. Dasilva, M. Salanoubat, M. Levy, N. Boudet, S. Castellano, V. Anthouard, C. Jubin, V. Castelli, M. Katinka, B. Vacherie, C. Biémont, Z. Skalli, L. Cattolico, J. Poulain, V. de Berardinis, C. Cruaud, S. Duprat, P. Brottier, J.-P. Coutanceau, J. Gouzy, G. Parra, G. Lardier, C. Chapple, K. J. McKernan, P. McEwan, S. Bosak, M. Kellis, J.-N. Volff, R. Guigó, M. C. Zody, J. Mesirov, K. Lindblad-Toh, B. Birren, C. Nusbaum, D. Kahn, M. Robinson-Rechavi, V. Laudet, V. Schachter, F. Qué-tier, W. Saurin, C. Scarpelli, P. Wincker, E. S. Lander, J. Weissenbach, H. Roest Crollius, "Genome duplication in the teleost fish *tetraodon nigroviridis* reveals the early vertebrate proto-karyotype", *Nature* **431** (2004), p. 946-957.
- [21] H. Xiong, D. Wang, C. Shao, X. Yang, J. Yang, T. Ma, C. C. Davis, L. Liu, Z. Xi, "Species tree estimation and the impact of gene loss following whole-genome duplication", *Syst. Biol.* **71** (2022), p. 1348-1361.
- [22] D. J. Macqueen, I. A. Johnston, "A well-constrained estimate for the timing of the salmonid whole genome duplication reveals major decoupling from species diversification", *Proc. R. Soc. Lond. B* **281** (2014), article no. 20132881.
- [23] E. Parey, A. Louis, C. Cabau, Y. Guiguen, H. Roest Crollius, C. Berthelot, "Synteny-guided resolution of gene trees clarifies the functional impact of whole-genome duplications", *Mol. Biol. Evol.* **37** (2020), p. 3324-3337.
- [24] E. Parey, A. Louis, J. Montfort, Y. Guiguen, H. Roest Crollius, C. Berthelot, "An atlas of fish genome evolution reveals delayed rediploidization following the teleost whole-genome duplication", *Genome Res.* **32** (2022), p. 1685-1697.
- [25] T. Zhao, A. Zwaenepoel, J.-Y. Xue, S.-M. Kao, Z. Li, M. E. Schranz, Y. Van de Peer, "Whole-genome microsynteny-based phylogeny of angiosperms", *Nat. Commun.* **12** (2021), article no. 3498.
- [26] A. H. Sturtevant, Th. Dobzhansky, "Inversions in the third chromosome of wild races of *Drosophila pseudoobscura*, and their use in the study of the history of the species", *Proc. Natl. Acad. Sci. USA* **22** (1936), p. 448-450.
- [27] G. Drillon, R. Champeimont, F. Oteri, G. Fischer, A. Carbone, "Phylogenetic reconstruction based on synteny block and gene adjacencies", *Mol. Biol. Evol.* **37** (2020), p. 2747-2762.
- [28] D. T. Schultz, S. H. D. Haddock, J. V. Bredeson, R. E. Green, O. Simakov, D. S. Rokhsar, "Ancient gene linkages support ctenophores as sister to other animals", *Nature* **618** (2023), p. 110-117.
- [29] N. Saitou, M. Nei, "The neighbor-joining method: a new method for reconstructing phylogenetic trees", *Mol. Biol. Evol.* **4** (1987), p. 406-425.
- [30] J. Steenwyk, N. King, "From genes to genomes: Opportunities, challenges, and a roadmap for synteny-based phylogenomics", 2023, preprint, <https://doi.org/10.20944/preprints202309.0495.v3>.
- [31] C. G. P. McCarthy, P. O. Mulhair, K. Siu-Ting, C. J. Creevey, M. J. O'Connell, "Improving orthologous signal and model fit in datasets addressing the root of the animal phylogeny", *Mol. Biol. Evol.* **40** (2023), article no. msac276.
- [32] A. K. Redmond, A. McLysaght, "Evidence for sponges as sister to all other animals from partitioned phylogenomics with mixture models and recoding", *Nat. Commun.* **12** (2021), article no. 1783.