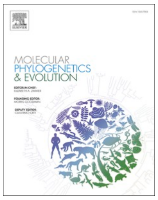
Molecular Phylogenetics and Evolution 189 (2023) 107942

Contents lists available at ScienceDirect

Molecular Phylogenetics and Evolution

journal homepage: www.elsevier.com/locate/ympev



Phylogenomic investigation of lampreys (Petromyzontiformes)

Brianna Smith, Alexandra Walling, Rachel Schwartz [*]

Department of Biological Sciences, College of the Environment and Life Sciences, University of Rhode Island, 120 Flagg Road, Kingston, RI 02881, United States

A R T I C L E I N F O

Keywords:
Lamprey
Petromyzontiformes

A B S T R A C T

The history of lamprey evolution has been contentious due to limited morphological differentiation and limited genetic data. Available data has produced inconsistent results, including in the relationship among northern and southern species and the monophyly of putative clades. Here we use whole genome sequence data sourced from a public database to identify orthologs for 11 lamprey species from across the globe and build phylogenies. The phylogeny showed a clear separation between northern and southern lamprey species, which contrasts with some prior work. We also find that the phylogenetic relationships of our samples of two genera, *Lethenteron* and *Eudontomyzon,* deviate from the taxonomic classification of these species, suggesting that they require reclassification.

1. Introduction

Lamprey are one of two surviving clades of jawless fishes that date as far back as 360 million years (Gess et al., 2006). In that time they have incurred surprisingly few morphological changes (Green and Bronner, 2014), thus their evolutionary history is of great interest due to both their highly conserved body plans and potential as a model for understanding the origins of vertebrate evolution (Xu et al., 2016). However, understanding the morphological and genomic evolution of lamprey first requires an understanding of relationships among these species.

There are three major families of lamprey with antitropical distribution: the northern clade *Petromyzontidae*, and the southern, species-poor families *Geotriidae* and *Mordaciidae*. It is generally agreed that the northern clade is monophyletic and includes the following genera (the number of species for each genus is given based on Brownstein and Near (2023)): *Capsiomyzon* (3 species)*, Entosphenus* (5 species)*, Eudontomyzon* (4 species)*, Ichthyomyzon* (6 species)*, Lampetra* (14 species)*, Lethenteron* (7 species)*, Petromyzon* (1 species), and *Tetrapleurodon* (1 species) (Green and Bronner, 2014). The southern families *Geotriidae* and *Mordaciidae* contain only one genus each, *Geotria* and *Mordacia* respectively. *Geotria* contains two species: *G. australis* and *G. macrostoma*. *Mordacia* contains three species: *M. mordax*, *M. praecox*, and *M. lapicida.*

Multiple conflicting hypotheses have been proposed for relationships within lampreys and with their close relatives. Phylogenies that rely on morphological or molecular data often produce conflicting results, particularly in groups of this age with minimal morphological differentiation, few characters for phylogenetics, and thus a high potential for homoplasy. For example, there is debate around whether lamprey are sister to hagfish or gnathostomes (jawed-fishes). Morphological data pairs lamprey sister to gnathostomes; however, molecular data alone, as well as studies that use combined morphological and molecular data, pair lamprey with hagfish in the Cyclostomata (Heimberg et al., 2010; Miller et al., 2021).

Relationships among southern lamprey species are also not well understood. Prior work has found an ambiguous relationship among the three families (Gill et al., 2003), that *Geotriidae* is more closely related to the northern *Petromyzontidae* than to other southern species (Lang et al., 2009; Riva-Rossi et al., 2020), and that the southern species are monophyletic (Brownstein and Near, 2023). However, this research has been limited to morphology and few molecular markers, and support (including in the latter case) is low. Just recently, *G. macrostoma* was reinstated as a separate species after being mistakenly classified as *G. australis* (Riva-Rossi et al., 2020). In particular, the lack of knowledge about the developmental forms of lampreys commonly causes researchers to confuse larval forms of species, and accidently split or lump species as in this case (Riva-Rossi et al., 2020).

Here we examine the relationships among 11 species of lamprey (from 8 of 10 extant genera) for which there are publicly available genomic sequence data. These include the northern species *I. fossor*, *I. castaneus*, *P. marinus*, *La. fluviatilis*, *La. planeri*, *Le. reissneri*, *Le. camtschaticum*, *En. tridentatus,* and *Eu. morii,* and southern species *G. australis* and *M. mordax*. Using a genome-wide dataset, we provide additional data and information to elucidate some previously ambiguous and

---

[*] Corresponding author.
E-mail address: rsschwartz@uri.edu (R. Schwartz).

https://doi.org/10.1016/j.ympev.2023.107942
Received 31 July 2023; Received in revised form 2 October 2023; Accepted 4 October 2023

contentious relationships.

2. Methods

2.1. Raw read sequence data

We used all publicly available whole genome sequence data for lampreys including species from both southern genera and 6 of the 8 northern genera. We obtained raw reads from paired-end Illumina sequences from the European Nucleotide Archive. Information detailing accession numbers and number of base pairs per taxon can be found in Table 1. The outgroup, *Salmo salar,* was chosen due to the availability of data for this species. Given that the relationships among lampreys, hagfish, and bony fish/tetrapods is not well established, any extant individual of the latter two groups is expected to be equally and substantially distant in evolutionary time from any extant lamprey.

2.2. Obtaining orthologous data

We used the SISRS v2.0 pipeline (Literman and Schwartz, 2021; Schwartz et al., 2015) to identify orthologous loci from next-generation whole-genome sequence data for lampreys. Briefly, this includes assessing the quality of the raw reads using fastQC v0.11.5 (Andrews, 2010), then trimming low-quality reads using BBtools v.37.31 (Bushnell, 2014). With only the high-quality reads remaining, we subsampled to a coverage depth of ~10x total for all lamprey species assuming an average lamprey genome size of 2.31 Gb (Smith et al., 2010) in order to generate data that could be combined to generate a "composite genome" comprised of likely orthologous loci. This genome was produced using Ray v2.3.1 with default parameters including a k-mer length of 31. The trimmed reads for each of the 11 taxa plus the outgroup were mapped to the composite genome using Bowtie v1.3.1 (Langmead et al., 2009) to determine the sequence of each contig for each species. Sites were called as 'N' for a particular species if they were (1) covered by less than three reads, or (2) included variation among reads. This latter step serves to remove sites that (1) may be variable within the population, and thus evolving too rapidly for accurate inference of synapomorphies, and (2) data from multiple paralogs aligning to a single contig that would result in inaccurate inference of gene trees. This step particularly aids in removing erroneous data for *S. salar*, which is a tetraploid derivative and thus has an abundance of paralogs. Sites that were not parsimony informative across species, or included insertions/deletions, were removed to create a final alignment. The final alignment was then filtered for sites with three different levels of missing data due to the potential for both dataset size and completeness (amount of missing data) to affect phylogenetic estimates (Wiens and Morrill, 2011 and references therein).

2.3. Phylogeny construction

The maximum likelihood approach of RAxML v8.2.12 (Stamatakis, 2014) was used with the GTR + GAMMA model and Lewis ascertainment bias correction (Lewis, 2001) to create phylogenies for each of the three concatenated datasets. 100 bootstrap replicates were used to assess support. Additionally, we generated a phylogeny for each dataset using a multispecies coalescent approach implemented in SVDquartets (Chifman and Kubatko, 2014) with default parameters of 1000 bootstrap replicates. This approach, in contrast, treats each site as having its own unique gene tree that is a product of the underlying species tree. We also examined support for each node in each unique topology using site concordance factors (sCF) implemented in IQ-Tree v2.1.2 (Minh et al., 2020) with parameters of 100 quartets as suggested by the manual.

Table 1
All Lamprey Accession Numbers and sequence information.

Taxon	Accession Numbers	Base Pairs
Entosphenus tridentatus	SRR11364986	59827952466
Eudontomyzon morii	SRR14574376	2156547800
Eudontomyzon morii	SRR14574377	2480544800
Eudontomyzon morii	SRR14574378	2749907800
Eudontomyzon morii	SRR14574379	2697822400
Eudontomyzon morii	SRR14574380	2856233800
Eudontomyzon morii	SRR14574381	2860383600
Ichthyomyzon castaneus	SRR13530976	10050002371
Ichthyomyzon castaneus	SRR13530977	10655975865
Ichthyomyzon fossor	SRR13530970	9651363172
Ichthyomyzon fossor	SRR13530971	11284277200
Lampetra fluviatilis	SRR3477525	435221855
Lampetra fluviatilis	SRR3477526	405730209
Lampetra fluviatilis	SRR3477527	577173917
Lampetra fluviatilis	SRR3477534	456105434
Lampetra fluviatilis	SRR3477538	335404401
Lampetra fluviatilis	SRR3477539	186242751
Lampetra fluviatilis	SRR3477540	215066145
Lampetra fluviatilis	SRR3477541	183934191
Lampetra fluviatilis	SRR3477542	158703865
Lampetra fluviatilis	SRR3477543	171115940
Lampetra fluviatilis	SRR3477544	242127942
Lampetra fluviatilis	SRR3477545	172657839
Lampetra fluviatilis	SRR3477546	199409702
Lampetra fluviatilis	SRR3477547	170172394
Lampetra fluviatilis	SRR3477548	229332375
Lampetra fluviatilis	SRR3477549	173965390
Lampetra fluviatilis	SRR3477550	178231810
Lampetra fluviatilis	SRR3477551	259720686
Lampetra fluviatilis	SRR3477552	315290343
Lampetra fluviatilis	SRR3477553	253444818
Lampetra fluviatilis	SRR3477554	359762355
Lampetra fluviatilis	SRR3477555	233867702
Lampetra fluviatilis	SRR3477556	614469995
Lampetra fluviatilis	SRR3477557	379809847
Lampetra fluviatilis	SRR3477558	473014155
Lampetra fluviatilis	SRR3477559	591688556
Lampetra fluviatilis	SRR3477560	498427230
Lampetra fluviatilis	SRR3477561	165549826
Lampetra fluviatilis	SRR3477562	452294264
Lampetra fluviatilis	SRR3477563	369184977
Lampetra fluviatilis	SRR3477564	370205209
Lampetra fluviatilis	SRR3477565	512588010
Lampetra fluviatilis	SRR3477566	377347487
Lampetra fluviatilis	SRR3477567	465984235
Lampetra fluviatilis	SRR3477568	674793696
Lampetra fluviatilis	SRR3477569	583587023
Lampetra fluviatilis	SRR3477570	666713641
Lampetra fluviatilis	SRR3477571	533237377
Lampetra fluviatilis	SRR3477572	672441691
Lampetra fluviatilis	SRR3477573	507449536
Lampetra fluviatilis	SRR3477574	544559525
Lampetra fluviatilis	SRR3477575	273232731
Lampetra fluviatilis	SRR3477576	238009002
Lampetra fluviatilis	SRR3477577	331133670
Lampetra fluviatilis	SRR3477578	329403163
Lampetra fluviatilis	SRR3477579	228179455
Lampetra fluviatilis	SRR3477580	78464722
Lampetra fluviatilis	SRR3477581	493655114
Lampetra fluviatilis	SRR3477582	250681104
Lampetra fluviatilis	SRR3477583	211834651
Lampetra fluviatilis	SRR3477584	261214951
Lampetra fluviatilis	SRR3477585	401624254
Lampetra planeri	SRR3477787	367503926
Lampetra planeri	SRR3477788	521676856
Lampetra planeri	SRR3477789	434710725
Lampetra planeri	SRR3477790	312409415
Lampetra planeri	SRR3477791	563368440
Lampetra planeri	SRR3477792	445830604
Lampetra planeri	SRR3477793	205516460
Lampetra planeri	SRR3477794	539215828
Lampetra planeri	SRR3477795	443443331
Lampetra planeri	SRR3477796	216578243
Lampetra planeri	SRR3477797	606677780

(*continued on next page*)

Table 1 (*continued*)

Taxon	Accession Numbers	Base Pairs
Lampetra planeri	SRR3477798	526519176
Lampetra planeri	SRR3477799	488557289
Lampetra planeri	SRR3477800	273409151
Lampetra planeri	SRR3477801	534452321
Lampetra planeri	SRR3477802	217434216
Lampetra planeri	SRR3477803	284539443
Lampetra planeri	SRR3477804	221482974
Lampetra planeri	SRR3477805	320466397
Lampetra planeri	SRR3477806	140182187
Lampetra planeri	SRR3477807	431045176
Lampetra planeri	SRR3477808	264152136
Lampetra planeri	SRR3477809	308234198
Lampetra planeri	SRR3477813	270922295
Lampetra planeri	SRR3477815	331619312
Lampetra planeri	SRR3477818	374531581
Lampetra planeri	SRR3477819	310937223
Lampetra planeri	SRR3477820	640548356
Lampetra planeri	SRR3477821	574476569
Lampetra planeri	SRR3477822	481118253
Lampetra planeri	SRR3477823	322239167
Lampetra planeri	SRR3477824	454625941
Lampetra planeri	SRR3477829	407696703
Lampetra planeri	SRR3477831	224280051
Lampetra planeri	SRR3477839	370768220
Lampetra planeri	SRR3477840	284956172
Lampetra planeri	SRR5230929	4217632800
Lethenteron camtschaticum	DRR150092	8653514600
Lethenteron camtschaticum	SRR1695991	11361029328
Lethenteron reissneri	DRR016690	7834368808
Lethenteron reissneri	SRR10202883	9056498700
Mordacia mordax	SRR2146922	4778470590
Petromyzon marinus	SRR12778999	7433726427
Petromyzon marinus	SRR12779000	7130928146
Petromyzon marinus	SRR12779001	2383473706
Geotriia Australis	SRR2146917	5230111078
Geotriia Australis	SRR2146918	5275638848
Geotriia Australis	SRR2146919	7458448828

3. Results

3.1. Data

On average, there are 17.04 Gb per species in our dataset (~7.4x coverage), with a minimum of 4.78 Gb (*M. mordax*; ~2x coverage) and a maximum of 20.94 Gb (*I. fossor*; ~9x coverage). The composite genome included 527,345 contigs >= 100nt with a total length of 123,918,125 and N50 of 237. 4,301,130 variable sites were obtained for the initial output alignment; of these 3,222,618 sites were singletons resulting in an alignment excluding singletons (i.e. including only parsimony informative sites) of 1,078,512 variable sites.

Due to the tradeoffs between the size of the dataset and its completeness (Schwartz et al., 2015; Wiens and Morrill, 2011), we initially filtered the SISRS output alignment (which contained all variable parsimony-informative sites with information for at least three species) for sites missing no more than 5 species. This dataset contained 579,195 sites. For comparison, we filtered out data to produce a dataset missing no more than 3 species at each site, which contained 166,202 sites, as well as a dataset with no missing information, which contained 158 sites. Notably, while this latter dataset appears small, because it contains only variable, parsimony-informative sites, it is likely comparable to a much larger dataset of loci containing many invariable or singleton sites.

3.2. Phylogeny

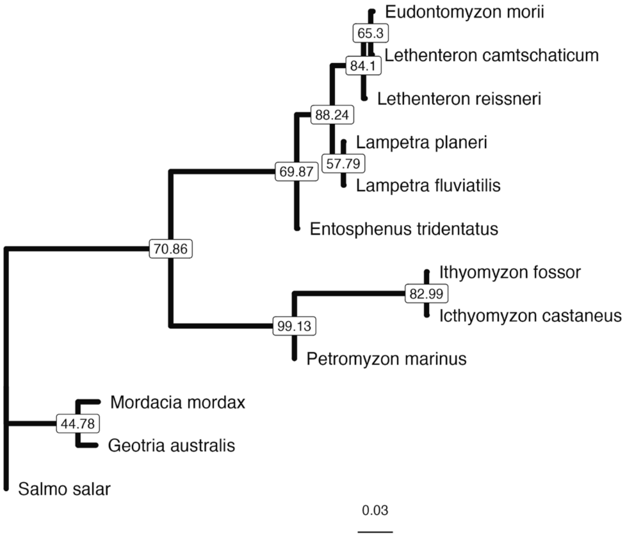


Fig. 1. A phylogenetic tree of lampreys produced using maximum-likelihood analysis and 579,195 sites from across the genome. The tree confirms the paraphyly of *Lethenteron* with high support, and supports monophyly of the southern lamprey families. Putatively orthologous loci were obtained for 11 lamprey species using SISRS with third-generation short-read genome sequence data. This tree was generated in RAxML and rooted using *Salmo salar* as the outgroup. Notably, trees that were identical topologically were also produced using SVDQuartets and smaller datasets filtered for greater completeness. The numbers at the nodes indicate sCF support - though not shown here, bootstraps for the indicated nodes were all 100. Branch lengths reflect substitutions-per-site for the sites in the alignment, which are highly conserved due to selection of loci that were alignable across the deep evolutionary distance of all lamprey families.

The species relationships generated from RAxML and SVDquartets for all three datasets were identical. Consequently, only the tree inferred using RAxML for the largest dataset is shown here (Fig. 1). All subsequent descriptions of the phylogeny and support values are for the dataset with up to five species missing per site. For all phylogenies, when rooting the tree using *Salmo salar*, there is an early branch that clearly differentiates the northern and southern clades. Most of the genera are monophyletic with the exception that *Lethenteron camtschaticum* is sister to *Eudontomyzon morii* rather than *Lethenteron reissneri.*

For all relationships, the bootstrap values were 100, while the sCF values varied between 44.8 and 99.1 - the lowest values of which were found between *Mordaciia* and *Geotriia*, and between *Lampetra* species. The lowest value 44.8 is associated with the node between *Mordaciia* and *Geotriia.* However, the associated site discordance factor values of 23.9 and 31.3 suggest that the observed relationship estimated by maximum likelihood is the most well supported. For the node supporting the monophyly of *Lampetra,* where the sCF was 57.8, the corresponding sDF are 18.5 and 23.7. The values at the node supporting *Le. camtschaticum* and *E. morii* as sister species have an sCF value of 65.3, and sDF scores 17.2 and 17.5.

4. Discussion

Our dataset revealed strong and consistent support across methods for the observed relationships. Our results strongly confirm the monophyly of northern hemisphere lamprey and indicate that there is a need for taxonomic reclassification of *Lethenteron* and *Eudontomyzon*. Additionally, we find that the species from the two southern hemisphere families form a monophyletic clade, resolving prior ambiguous and conflicting findings regarding the relationships among groups.

While all nodes had 100% bootstrap support, this level of support could be due to the large size of the datasets. However, most nodes contained high sCF support as well. The exceptions were (1) between southern species *G. australis* and *M. mordax* (44.8), and (2) *La. planeri* and *La. fluviatalis* (57.8). Despite these low sCF scores, sDF support confirms the resolution of these nodes as the most likely topologies. While in the former case, additional sequence data for *M. mordax* and/or sequencing additional species in the two southern genera could further clarify these relationship, the finding that smaller and more complete filtered datasets also strongly supported these relationships provides some suggestion that data limitations did not severely affect phylogenetic estimates. All other nodes, including that of the grouping between *Le. camtschaticum* and *Eu. morii* (65.3) are much more highly supported. While support for this node is certainly on the lower side compared to those of other branches, sDF scores 17.2 and 17.5 reveal that this resolution is more likely than one where *Le. camtschaticum* is sister to *Le. reissneri*.

4.1. Paraphyly of Lethenteron

Similar to results found in prior work, we find that *Le. camtschaticum* is placed sister to *Eu. morii* with a high degree of support rather than the other member of genus *Lethenteron* observed in this study, *Le. reissneri* (Pereira et al., 2021; Potter et al., 2015). In older work, *Eu. morii* is classified as *Lethenteron* (Lang et al., 2009) and recent work has supported this relationship as well (Brownstein and Near, 2023); however, this classification does not appear elsewhere or in the NCBI Taxonomy. Importantly, we have dramatically expanded the dataset beyond two mitochondrial markers (*cytochrome b* and *c*), and 2 nuclear internal transcribed spacers (*ITS*) loci as recommended due to the limitations of these loci (Miller et al., 2021). This result also contradicts morphological observations suggesting that the genera *Eudontomyzon* and *Lampetra* should be considered monophyletic as presently described (Potter et al., 2015).

4.2. Relationships among families

Prior work suggests that the southern clades are paraphyletic, with *Geotriidae* related to *Petromyzontidae* (Lang et al., 2009; Potter et al., 2015). However, this relationship has generally been considered uncertain (Gill et al., 2003; Miller et al., 2021). Here we find support for the monophyly of the southern clades. Only one prior study (Brownstein and Near, 2023) has suggested this relationship; however, their limited dataset (cytochrome *b* and morphological data) produced limited to mixed support.

4.3. Comparison with morphological results

We confirm much of the work of Gill et al. (2003) with genome-scale molecular data. These results defined the monophyly of many clades in our study. In particular, we confirm that *Icthyomyzon* + *Petromyzon* (who share absence of median velar tentacles), as well as *Entosphenus* + *Lethenteron* + *Eudontomyzon* + *Lampetra* (who share loss of lateral teeth alongside the oral disc), and *Lethenteron* + *Eudontomyzon* + *Lampetra* (who share tooth and mouth synapomorphies such as placement of lateral teeth on both sides of the oral disc and a median tooth atop 'transverse lingual lamina') all form monophyletic clades (Gill et al., 2003). It is important to note that *Entosphenus* + *Lethenteron* + *Eudontomyzon* + *Lampetra* are considered monophyletic with tooth and mouth synapomorphies as evidence, however, *Eudontomyzon* does have teeth present alongside the oral disc unlike the other members of this group, suggesting that there was a reversal of this character to its ancestral state (Gill et al., 2003).

However, we cannot confirm the result of Gill et al. that *Eudontomyzon* (who share a reversal to the presence of lateral teeth alongside the oral disc) are monophyletic due to having data for a single representative of the genus. Li (2014) found that *Eudontomyzon* were monophyletic with the exception of the placement of *E. morii* as a member of *Lethenteron*. Brownstein and Near (2023) show a similar classification. This could be consistent with our placement of *E. morii* as a member of *Lethenteron*. While we find that *Eudontomyzon* + *Lampetra* (who share certain oral synapomorphies) do not form a clade, due to our single sample of *Eudontomyzon*, this grouping may hold if *E. morii* is reclassified as *Lethenteron*.

5. Conclusion

The lamprey phylogeny continues to pose challenges, particularly when it comes to resolving molecular and morphological estimates of relationships and understanding how phenotypes have evolved. In this study, we expanded genetic sampling from a few genes to whole genomes to address two of the many questions surrounding this lineage. First, we find support for prior work suggesting the need for reclassification of one or more species of *Eudontomyzon* with *Lethenteron*. Second, we find support for monophyly of the southern families of lampreys. This latter result in particular has implications for our understanding of lamprey evolution.

CRediT authorship contribution statement

Brianna Smith: Conceptualization, Formal analysis, Investigation, Writing – original draft, Visualization. **Alexandra Walling:** Visualization, Supervision. **Rachel Schwartz:** Supervision, Writing – review & editing, Resources.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

We thank members of the Schwartz lab for valuable comments and suggestions. The High-Performance Research Computing (HPC) facility at the University of Rhode Island provided the computational resources for the analyses in this manuscript and Dr. Kevin Bryan, the HPC systems manager, provided additional assistance. This work is supported by the USDA National Institute of Food and Agriculture, Hatch project accession no. 1017848, as well as grant 1942273 from the National Science Foundation (USA) to RS.

Appendix A. Supplementary data

Supplementary Table 1. Support Values (sCF, sDF) for the phylogeny estimated from 579,195 sites, allowing up to 5 species missing per site. Supplementary data to this article can be found online at https://doi.org/10.1016/j.ympev.2023.107942.

References

Andrews, S., 2010. FastQC A Quality control tool for high throughput sequence data [WWW Document]. URL https://www.bioinformatics.babraham.ac.uk/projects/fastqc/.

Brownstein, C.D., Near, T.J., 2023. Phylogenetics and the cenozoic radiation of lampreys. Curr. Biol. 33, 397–404.e3. https://doi.org/10.1016/j.cub.2022.12.018.

Bushnell, B., 2014. BBMap. https://sourceforge.net/projects/bbmap/.

Chifman, J., Kubatko, L., 2014. Quartet inference from SNP data under the coalescent model. Bioinform. 30, 3317–3324. https://doi.org/10.1093/bioinformatics/btu530.

Gess, R.W., Coates, M.I., Rubidge, B.S., 2006. A lamprey from the Devonian period of South Africa. Nature 443, 981–984. https://doi.org/10.1038/nature05150.

Gill, H.S., Renaud, C.B., Chapleau, F., Mayden, R.L., Potter, I.C., 2003. Phylogeny of living p lampreys (Petromyzontiformes) based on morphological data. Copeia 2003, 687–703. https://doi.org/10.1643/ia02-085.1.

Green, S.A., Bronner, M.E., 2014. The Lamprey: A jawless vertebrate model system for examining origin of the neural crest and other vertebrate traits. Differentiation 87, 44. https://doi.org/10.1016/j.diff.2014.02.001.

Heimberg, A.M., Cowper-Sal-lari, R., Sémon, M., Donoghue, P.C.J., Peterson, K.J., 2010. microRNAs reveal the interrelationships of hagfish, lampreys, and gnathostomes and the nature of the ancestral vertebrate. Proc. Natl. Acad. Sci. USA 107, 19379–19383. https://doi.org/10.1073/pnas.1010350107.

Lang, N.J., Roe, K.J., Renaud, C.B., Gill, H.S., Potter, I.C., Freyhof, J., Naseka, A.M., Cochran, P., Perez, H.E., Habit, E.M., Kuhajda, B.R., Neely, D.A., Reshetnikov, Y.S., Salnikov, V.B., Stoumboudi, M.T., Mayden, R.L., 2009. Novel relationships among lampreys (Petromyzontiformes) revealed by a taxonomically comprehensive molecular data set, in: Brown, L.R., Chase, S.D., Mesa, M.G., Beamish, R.J., Moyle, P. B. (Eds.), Biology, Management, and Conservation of Lampreys in North America. American Fisheries Society. Doi: 10.47886/9781934874134.ch2.

Langmead, B., Trapnell, C., Pop, M., Salzberg, S.L., 2009. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biol. 10, 1–10. https://doi.org/10.1186/gb-2009-10-3-r25.

Lewis, P.O., 2001. A likelihood approach to estimating phylogeny from discrete morphological character data. Syst. Biol. 50, 913–925. https://doi.org/10.1080/106351501753462876.

Li, Y., 2014. Phylogeny of the lamprey genus *Lethenteron* Creaser and Hubbs 1922 and closely related genera using the mitochondrial cytochrome b gene and nuclear gene introns. The University of Manitoba.

Literman, R., Schwartz, R., 2021. Genome-scale profiling reveals noncoding loci carry higher proportions of concordant data. Mol. Biol. Evol. 38, 2306–2318. https://doi.org/10.1093/molbev/msab026.

Miller, A.K., Baker, C., Kitson, J.C., Yick, J.L., Manquel, P.E.I., Alexander, A., Gemmell, N.J., 2021. The southern hemisphere lampreys (Geotriidae and Mordaciidae). Rev. Fish Biol. Fish. 31, 201–232. https://doi.org/10.1007/s11160-021-09639-x.

Minh, B.Q., Hahn, M.W., Lanfear, R., 2020. New methods to calculate concordance factors for phylogenomic datasets. Mol. Biol. Evol. 37, 2727–2733. https://doi.org/10.1093/molbev/msaa106.

Pereira, A.M., Levy, A., Vukić, J., Šanda, R., Levin, B.A., Freyhof, J., Geiger, M., Choleva, L., Francisco, S.M., Robalo, J.I., 2021. Putting European lampreys into perspective: A global-scale multilocus phylogeny with a proposal for a generic structure of the Petromyzontidae. J. Zool. Syst. Evol. Res. 59, 1982–1993. https://doi.org/10.1111/jzs.12522.

Potter, I.C., Gill, H.S., Renaud, C.B., Haoucher, D., 2015. The taxonomy, phylogeny, and distribution of lampreys. In: Lampreys: Biology. Conservation and Control. Springer, Netherlands, pp. 35–73.

Riva-Rossi, C., Barrasso, D.A., Baker, C., Quiroga, A.P., Baigún, C., Basso, N.G., 2020. Revalidation of the Argentinian pouched lamprey *Geotria macrostoma* (Burmeister, 1868) with molecular and morphological evidence. PLoS One 15, e0233792.

Schwartz, R.S., Harkins, K.M., Stone, A.C., Cartwright, R.A., 2015. A composite genome approach to identify phylogenetically informative data from next-generation sequencing. BMC Bioinf. 16, 193. https://doi.org/10.1186/s12859-015-0632-y.

Smith, J.J., Stuart, A.B., Sauka-Spengler, T., Clifton, S.W., Amemiya, C.T., 2010. Development and analysis of a germline BAC resource for the sea lamprey, a vertebrate that undergoes substantial chromatin diminution. Chromosoma 119, 381–389. https://doi.org/10.1007/s00412-010-0263-z.

Stamatakis, A., 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinform. 30, 1312–1313. https://doi.org/10.1093/bioinformatics/btu033.

Wiens, J.J., Morrill, M.C., 2011. Missing data in phylogenetic analysis: reconciling results from simulations and empirical data. Syst. Biol. 60, 719–731. https://doi.org/10.1093/sysbio/syr025.

Xu, Y., Zhu, S.-W., Li, Q.-W., 2016. Lamprey: a model for vertebrate evolutionary research. Zool. Res. 37, 263–269. https://doi.org/10.13918/j.issn.2095-8137.2016.5.263.