Organisms Diversity & Evolution (2023) 23:621–629
https://doi.org/10.1007/s13127-022-00587-3

ORGANISMS DIVERSITY & EVOLUTION

METHODS PAPER

ELAV Intron 8: a single-copy sequence marker for shallow to deep phylogeny in *Eupulmonata* Hasprunar & Huber, 1990 and *Hygrophila* Férussac, 1822 (Gastropoda: Mollusca)

Jeffrey C. Nekola[1] · Markéta Nováková[1] · Michal Horsák[1] · Coen M. Adema[2]

Received: 17 June 2022 / Accepted: 30 September 2022 / Published online: 4 November 2022

Abstract
Progress in eupulmonate gastropod taxonomy is limited by absence of single- or low-copy nuclear DNA markers that provide species-scale insights. We detail here the single-copy intron 8 of the *embryonic lethality and abnormal visual system* gene (*ELAVI8*). High sequence conservation within flanking exons 8 and 9 allowed design of non-redundant primers that enable PCR amplification across a broad phylogenetic extent. Across the Eupulmonata and Hygrophila ELAVI8 product ranged from 602 to 802 base pairs in length. A multiple sequence alignment across 62 taxa, representing 36 genera and 12 major clades, showed 1296 sites with 865 variable positions. Typically, 2–10 base pair differences were noted between closely related species within a genus, with 74–200 bases differing between major infra-order clades. ELAVI8 maximum-likelihood trees showed essentially identical topologies at high levels of support compared to those created from concatenated ITS1 + 2 (within a genus) or 28S (across Eupulmonata and Hygrophila). ELAVI8 is superior to these two phylogenetic marker regions, however, because it (1) resolves phylogenetic pattern from species- to sub-cohort scales in a single amplicon, (2) was better able to resolve family- and deeper-clades at high support, (3) allows for outgroup rooting between genera, and (4) possesses no intercopy haplotype variability.

Keywords PCR · nDNA · Intron · Phylogenetic marker · Gastropods

Introduction

Because of introgression, incomplete sorting, horizontal gene transfer, and differential heritability mechanisms, signal emanating from only a single DNA marker may not reflect true taxonomic status or evolutionary relationships (e.g., Bryja et al., 2018; Haase et al., 2021; Nekola et al., 2015). The identification of robust pattern in molecular-based taxonomy and phylogeny is thus contingent upon recognition of consensus across multiple differentially sorting units. For the ~ 35,000 species of eupulmonate gastropods (Barker, 2001), species-scale signal from multiple areas within the mitochondrial genome is possible via polymer chain reaction (PCR) combined with Sanger sequencing (e.g., cytochrome oxidase subunit 1 – COI, cytochrome oxidase subunit II – COII, cytochrome B – CytB, 16S ribosomal RNA– 16S). Many fewer options exist in nuclear DNA (nDNA). The most commonly used nDNA markers in molluscan phylogenetics are 18S and 28S of the rRNA gene cassette (e.g., Saadi et al., 2021; Wade et al., 2006). However, these represent highly conserved areas that are only suitable for robust documentation of deep (e.g., genus or greater) evolutionary relationships. Likewise, the protein-encoding sequences for adenine nucleotide translocase – ANT, calmodulin – Cal, cyclophylin A – CycA (Audzijonyte & Vrijenhoek, 2010), and the non-transcribed spacer sequence between H3/H4 in the histone gene cluster (Armbruster et al., 2005) lack the variability required for robust node resolution between closely related species. Taxon-dependent variability in flanking sequence also prevents effective universal primer design to amplify the Cal and CycA marker outside of limited phylogenetic groups (Audzijonyte & Vrijenhoek, 2010).

✉ Jeffrey C. Nekola
nekola@sci.muni.cz

[1] Department of Botany and Zoology, Faculty of Science, Masaryk University, Kotlářská 2, 61137 Brno, Czech Republic

[2] Department of Biology, University of New Mexico, Albuquerque MSC03 2020N M87131, USA

To date, easily obtainable species-scale nDNA signal in eupulmonate and related hygrophilid gastropods (Jörger et al., 2010) has been limited to the internal transcribed spacer 1 – ITS1 (Armbruster & Bernhard, 2000) and internal transcribed spacer 2 – ITS2 (Wade & Mordan, 2000) of the rRNA gene cassette. Amplification of these two markers is facilitated by existence of universal primers placed within ultra conserved elements (Bejerano et al., 2004) in the flanking 18S, 5.8S, and 28S regions. While concatenated ITS1 + 2 provides roughly the same number of variable bases as mitochondrial 16S — thereby allowing for excellent species-level resolution and support (e.g., Nekola et al., 2018) — it also has severe limitations: (1) neither ITS1 nor ITS2 can be aligned between genera (Nekola & Coles, 2016; Wade & Mordan, 2000), thus preventing tree rooting and deeper level phylogenetic reconstruction; (2) nucleotide variability is occasionally too low for species-scale node resolution and support (Horsáková et al., 2020; Nekola et al., 2015); and (3) haplotype heterogeneity may exist among the tandemly arranged multiple genomic copies of the rRNA gene cassette. This latter issue occurs in multiple genera, and often limits nDNA phylogenetic reconstruction. Within California *Haplotrema* Ancey, 1881 land snails, for instance, this problem occurs to some degree in almost half of analyzed individuals and becomes so profound in 10% that their ITS1 + 2 sequence traces cannot be interpreted because of mixed nucleotide signals contributed by heterogenic amplicons. Subcloning of seventeen ITS1 amplicons from one of these problematic samples revealed the presence of ten unique haplotypes whose fragment lengths varied by 7–41 bases while possessing 9–23 variable base pairs (*author's unpublished data*).

It is thus essential for additional species-scale nDNA markers to be developed for the eupulmonate and hygrophilid phylogenetics toolkit. We detail here a novel nDNA marker that provides phylogenetic resolution from species- to superorder-scales: intron 8 of the *embryonic lethality and abnormal visual system* gene (ELAVI8). We describe its location in the Eupulmonata and Hygrophila genome, methods for PCR amplification, and levels of base pair variability across various taxonomic scales. We then illustrate the ability of ELAVI8 to resolve phylogenetic pattern at both shallow (within a genus) and deep (up to superorder) levels and compare the ELAVI8 phylogenetic trees to those generated from ITS1 + 2 or 28S.

Methods

Target region

We focused on intronic regions because they represent obvious targets for species level and lower phylogenetic nDNA signal (Côrte-Real et al., 1994; Jennings & Etter, 2011; Lessa, 1992). We searched the mapped and annotated genome assemblies of *Achatina* (*Lissachatina*) *fulica* (Bowdich, 1822; Guo et al., 2019) — a stylommatophoran terrestrial gastropod, and *Biomphalaria glabrata* (Say, 1818) (BglaB1; Adema et al., 2017) — a hygrophilid freshwater gastropod, and identified ELAVI8 as a likely target because it was (1) roughly 800 bp in length and thus short enough to allow complete Sanger sequencing in a single PCR product; (2) flanked by highly conserved sequence; and (3) a single-copy gene. In these two assembled gastropod genomes, the ELAV gene region was approximately 20 K base pairs long and divided into 11 exons (Fig. 1). The ELAV gene family codes for RNA-binding proteins, and while they can occur in gonads most appear to function in neuron development where they serve as gene-specific regulators of alternative splicing. ELAV proteins also appear to be involved in

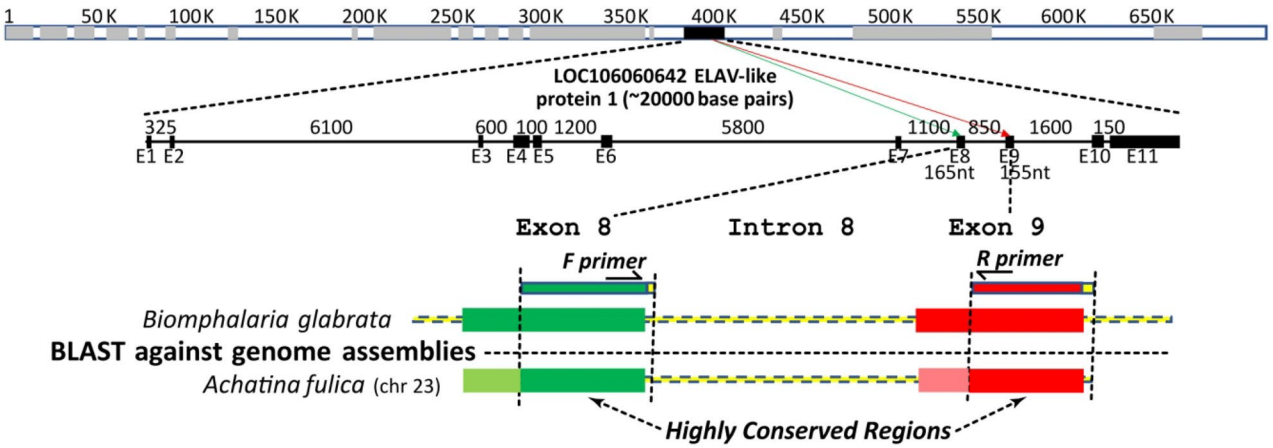

Fig. 1 *Biomphalaria glabrata* genomic scaffold ASM4573v1 LG4_Random_Scaffold46, whole genome shotgun sequence, with the location of Exons 8 and 9, Intron 8, and primers used for ELAVI8 PCR amplification

a number of other post-transcriptional processes including mRNA stability, translation, and transport (Soller & White, 2004).

PCR optimization

Because flanking sequences in exons 8 and 9 are conserved between distantly-related *A. fulica* and *B. glabrata*, we designed both non-degenerate standard and nested primer sets targeting these regions to allow PCR amplification of the entire intervening intron 8 (Table 1, Fig. 1). Initial assessment of primer ability to amplify ELAVI8 was done by performing PCR with genomic DNA extractions across seven terrestrial gastropod genera (see Appendix 1 in Supporting Information) representing a cross section of the land snail phylogeny (Wade et al., 2006). Freshwater *Gyraulus parvus* was additionally used to represent the Hygrophila. The DNA extractions used for *Columella* Westerlund, 1878, *Haplotrema* and *Vertigo* O.F. Müller, 1773 had previously all produced unresolveable ITS1 Sanger sequence traces due to intercopy haplotype variability. PCR polymerases, thermal cycling protocols, and primer sets (including nested primers) were then optimized both in terms of target region yield and minimization of off-target products.

Testing ELAVI8 at shallow and deep phylogenetic pattern levels

To assess the ability of ELAVI8 to recover phylogenetic relationships across multiple scales, we employed two specimen sets (see Appendix 1 in Supporting Information for specimen and taxon authorship data):

1. A shallow scale set consisting of 21 individuals from 10 California *Haplotrema* species (*H. alameda*, *H. catalinense*, *H. duranti*, *H. keepi*, *H. minimum*, *H. transfuga*, *H. voyanum,* and 3 putative new species), plus six outgroup samples from across the Stylommatophora (*Cochlicopa lubrica*, *Columella edentula*, *Discus ruderatus*, *Euconulus alderi*, *Nesovitrea* n.sp., *Vertigo ronnebyensis*). *Haplotrema* was chosen as the focal genus because of its high rates of intercopy ITS1 and ITS2 haplotype variability.
2. A deep scale set including 50 species from 36 genera representing the range of major eupulmonate clades from Wade et al. (2006), with *Physa acuta* within the Hygrophila being used for outgroup comparison.

PCR was performed using a variety of commercially available Taq polymerase kits, with the most effective being used for remaining reactions. All ELAVI8 PCR amplicons were sequenced on both strands using Perkin Elmer ABI Big Dye termination and standard protocols. Forward and reverse sequence-traces from each dye-terminated product were checked by eye for quality and then assembled into a single contig using Sequencher 4.7. Primer ends were removed. The contigs were then aligned via MAFFT v7.0 (Katoh et al., 2017), with the resulting Multiple Sequence Alignment (MSA) being optimized by eye. See Supporting Information for the final MSA (ELAVI8_panpul.fas). Phylogenetic reconstruction was then conducted using Maximum-Likelihood (ML) in MEGA 11.0.11 with gamma distributed variation, invariant sites, and nearest-neighbor interchange. jModelTest v. 2.1.10 (Darriba et al., 2012) was used to verify that the Tamura-Nei substitution model was appropriate for

Table 1 PCR conditions for ELAVI8 amplification

A. primer sets

Standard	ELAVI8-F	5′–GAATACTTTATGATCAAACCAC–3′
	ELAVI8-R	5′–GGCAGCCCACTTACATACAAG–3′
Nested	ELAVI8-F_Outer	5′–GCTAACTTATATATTAGTGGCTTGC–3′
	ELAVI8-R_Outer	5′–GTGTGTGTATCATAAAGAATTCG–3′
	ELAVI8-F_Inner	5′–ATGACACAGTTGGACTTGGAG–3′
	ELAVI8-R_Inner	5′–GAGAACAACTTCTCCAAATCCT–3′

B. template: 1–2 µL of 3–30 ng/µL genomic extraction/25 µL PCR reaction

C. cycling conditions: *TopBio LA Hot Start Master Mix*

Standard			***Modified touchdown***		
Initial denaturation	96 °C	10 min	Initial denaturation	96 °C	1 min
Cycle denaturation	94 °C	30 s	Cycle A denaturation	94 °C	30 s
Cycle anneal	50 °C	45 s	Cycle A anneal	55 °C	45 s
Cycle elongation	72 °C	90 s	Cycle A elongation	72 °C	90 s
Number of cycles	**40**		Number of A cycles	**10**	
Final elongation	72 °C	10 min	Cycle B denaturation	94 °C	30 s
			Cycle B anneal	50 °C	45 s
			Cycle B elongation	72 °C	90 s
			Number of B Cycles	**30**	
			Final elongation	72 °C	

these data. Support values were calculated based on 1000 bootstrap replicates. ML was used because it reliably provides the most conservative support values (Douady et al., 2003; Rannala & Yang, 1996). Because the MSA included considerable insertions/gaps, we chose the "use all sites" option in MEGA.

The ML trees based on ELAVI8 sequences were then compared to ML trees based on ITS1 + 2 (shallow scale) or 28S (deep scale). These sequences were either generated de novo by PCR from the same extractions used for ELAVI8 — using methods provided in Nekola and Coles (2016) and Nekola et al. (2018) — or were obtained from archived GenBank sequence from identical or closely related species (See Appendix 1 in Supporting Information). MAFFT was again used to generate initial sequence alignments, with the MSA being optimized by eye. MEGA 11.0.11 was again used to reconstruct ML trees using the above parameters.

Results

PCR amplification and optimization

While a number of commercial Taq preparations amplified ELAVI8, TopBio LA Hot Start Master Mix afforded the most reliable PCR results. Other commercial kits which functioned well included Promega GoTaq G2 Hot Start Green and Qiagen All Taq PCR Core kit. Although PCR was successful across the Eupulmonata and Hygrophila, the ELAVI8 primers failed to amplify product in the Heterobranchia (*Valvata* O.F. Müller, 1773), Neritomorpha (*Hydrocena* Küster, 1844; *Theodoxus* Montfort, 1810), and Caenogastropoda (*Bithynia* Linnaeus, 1758; *Bythinella* Moquin-Tandon, 1856; *Platyla* Moquin-Tandon, 1856; *Truncatella* Risso, 1826). PCR reactions that produced single ELAVI8 amplicon agarose gel bands generated unambiguous chromatograms. For some DNA extractions, the standard PCR protocol also co-amplified additional products ranging from 200–1400 bp in length (Fig. 2A). An annealing temperature gradient experiment using *Nesovitrea* n.sp. template — which had the largest number of co-amplified products — led to development of a modified step-down PCR profile for optimum target-to-non-target yield and elimination of smaller off-target amplicons (Fig. 2B; Table 1). While in some cases, long off-target PCR products could not be eliminated, the higher yield/concentration of target- to non-target product was sufficient to allow unequivocal determination of ELAVI8 sequence from resultant chromatograms. Only in the case of *Gyraulus parvus* did the similar size and yield of off-target amplicons make ELAVI8 sequence determination impossible from chromatograms. This problem could have been overcome through manual excision of the target band from the agarose gel. Recalcitrant DNA samples that failed to amplify with the standard primer set often yielded ELAVI8 product through use of nested primers, whereby PCR was initially conducted with outer primers and the touchdown thermal cycling profile. This product was then used as template for a second PCR using the inner primer set under standard cycling conditions (Table 1). Amplicons from such nested PCR were sequenced using inner primers.

Genetic variability of eupulmonate and hygrophilid ELAVI8

Haplotrema ELAVI8 amplicons from individual samples ranged from 649 to 665 bases in length (Table 2), with 0–6 variable bases within a species, 9–15 variable bases between closely related species, and 17–103 variable bases between distantly related species (Table 2). Sequence inspection across nine other genera (see MSA in Supporting

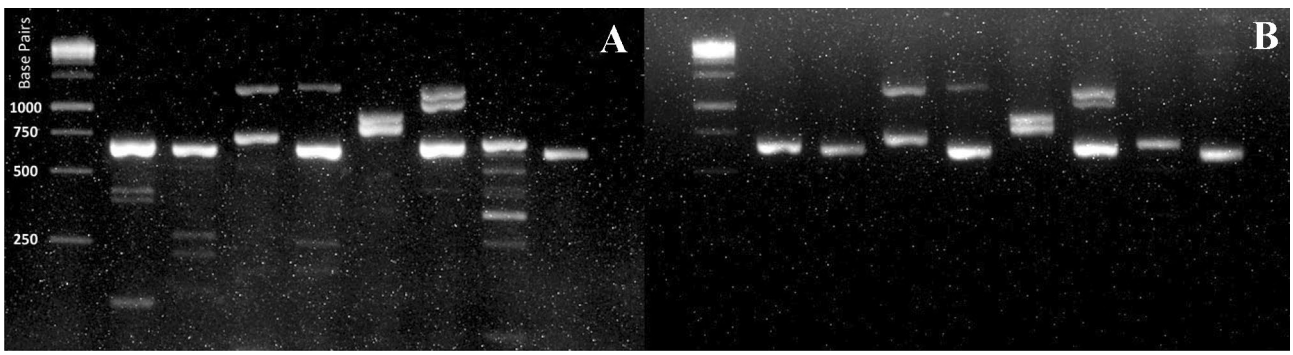

Fig. 2 ELAVI8 PCR yield and stringency across using a 1.1% Agarose gel in 1 × TBE buffer with GoldView Dye and 3 µl of PCR product from each reaction. Panel **A** represents 50° C annealing temperature for 40 cycles while **B** represents the modified touchdown procedure of 54° C anneal for 10 cycles followed by 50° C for 30 cycles. The order of specimens in both panels (left to right) is: *Cochlicopa lubrica* [DNA sample H29], *Columella edentula, Discus ruderatus* [H23], *Euconulus alderi*, *Gyraulus parvus, Haplotrema keepi* [HAP16], *Nesovitrea* n.sp. [H49], *Vertigo ronnebyensis*

Table 2 Species-scale ELAVI8 vs. ITS1/2 genetic variability

A. California *Haplotrema* in Fig. 3

		Base pair variability		
	Amplicon length	Within species	Between species	
			Closely related	Distant
ELAVI8	649–665	0–6	9 – 15	17–102.7
ITS1/2	1217–1263	1.2–34.7	10.5–26.5	39–92

B. Other genera in Fig. 4

			Base pair variability			
	Amplicon length		Within species		Between species	
	ELAVI8	ITS1/2	ELAVI8	ITS1/2	ELAVI8	ITS1/2
Cochlicopa	633–643	–	0–0.7	–	2.3–4.7	–
Columella	605–608	1521–1628	0–2	5.5–13.5	1.1–9.8	31.6–76.9
Discus	691–699	–	0–2	–	3–25	–
Euconulus	619	1514–1546	–	0.3–1.4	0–3	4.5–6.6
Nesovitrea	675–681	1384–1420	0–1.4	0.3–4.7	2.3–8	8.2–45.2
Paralaoma	689–690	–	0–3	–	7.5–9.5	–
Punctum	667–672	–	0–1.3	–	1.6–11.5	–
Vallonia	621–622	1361–1368	0–0.7	0–1.5	1.1–8.2	8.0–45.9
Zoogenetes	628	1474–1478	0.2	0.7	–	–

Information) showed that ELAVI8 amplicon length varied by 1–10 bases within a genus. Base pair variability was generally < 2 within a species. In general, base-pair variability between closely related species was double that observed within those species. ELAVI8 base pair variation rates between species tended to be ~ 50% less than ITS1 + 2 (Table 2). Some genera demonstrated inadequate sequence variation for robust species-level designation, including *Euconulus* O.W.H. Reinhardt, 1883 (no base pair differences between some closely related species) and *Vallonia* Risso, 1826 (< 2 bases).

Across the Eupulmonata and Hygrophila, ELAVI8 amplicon length ranged from 602 to 802 base pairs (Table 3). The MSA contained 1297 sites with 865 variable positions. Indels were relatively short and less frequent in the first half of the MSA, but longer and more common over the final 400 aligned positions (Fig. 3). Although this variability in the second half of the ELAVI8 MSA challenged alignment, the upstream portion often contained enough diagnostic base pair differences to allow species demarcation at high support. Total ELAVI8 base pair variability within a genus ranged from 3 to 215, with most falling between 10 and 30. Approximately 150–400 variable bases were observed within, and 75–200 bases between major taxonomic groups (Table 3). By comparison, 28S amplicons in these taxa ranged from 773 to 827 bases and exhibited only 389 total variable bases across its 945 aligned sites. On average 28S possessed only 2–35 variable sites within genera. Because ELAVI8 possessed 50–110 variable bases within and 34–97 between these same major taxonomic groups, it exhibited roughly twice the amount of phylogenetic information as compared to 28S at deeper taxonomic scales.

Phylogenetic reconstructions

Comparison of ELAVI8- and ITS1 + 2-based ML trees for *Haplotrema* demonstrated essentially identical topologies with high support for all species level clades (Fig. 4). While the ITS1 + 2 tree showed somewhat greater resolution of species-level clades with slightly higher support, its coverage was limited due to exclusion of some individuals because of intercopy haplotype variation. This issue was especially concerning in *H. keepi* where 2/3 of specimens had unreadable chromatograms. While this issue was also present in *H. duranti*, sequence

Table 3 ELAVI8 vs. 28S genetic variability in Eupulmonata and Hygrophila

			Base pair variability			
	Amplicon length	Aligned length	Total	Within genera	Major groups	Between major clades
ELAVI8	602–802	1296	865	3–215	~ 150–380	74–200
28S	773–827	945	389	2–35	~ 50–109	34–97

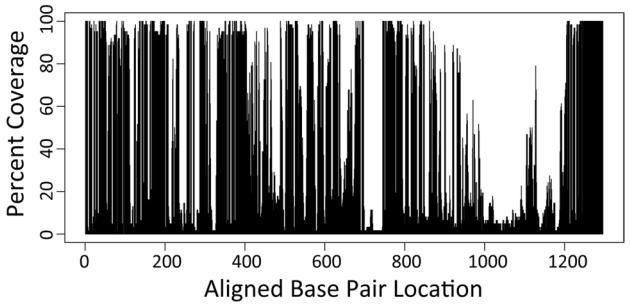

Fig. 3 Base pair coverage across the 1296 aligned sites in the ELAVI8 MSA (see ELAVI8_panpul.fas in the Supporting Information). The y-axis represents the percentage of sites that are represented at that location across all specimens

extrapolation was possible because fewer haplotypes were present, allowing for base pairs from the most common haplotype to be read. Even here, though, final sequences must be seen as approximate because some individual base calls were difficult and could be determined only though comparison with other sequences. At deeper phylogenetic scales, ELAVI8 replicates the topology of 28S from genus to superorder scales but at higher resolution and better node support (Fig. 5). For instance, the 28S ML tree assigned the Limacoidea into two non-coalescing subclades (*Nesovitrea*/*Aegopinella* Lindhold, 1927 vs. *Oxychilus* Fitzinger, 1833/*Euconulus*) while ELAVI8 shows them as members of a single, highly-supported clade.

Discussion

These results demonstrate that ELAVI8 can be amplified in the Eupulmonata and Hygrophila using universal non-degenerate primers and generates effective nDNA phylogenetic signal. This ELAVI8 sequence generates trees that generally replicate topologies displayed by ITS1 + 2 at the species scale and 28S from genus to superorder scales. ELAVI8 represents an improvement over ITS1 + 2 at the species scale by being: (1) a single copy gene that cannot express intercopy haplotype variability. This allowed for unambiguous sequence determination from all California *Haplotrema* specimens; (2) alignable across the breadth of the Eupulmonata and Hygrophila. This allowed for inclusion of outgroups and rooting of trees outside of a focal genus; (3) amplifiable in a single PCR reaction yielding a product of appropriate length for effective Sanger sequencing. This allowed for expedient and cost-effective generation of

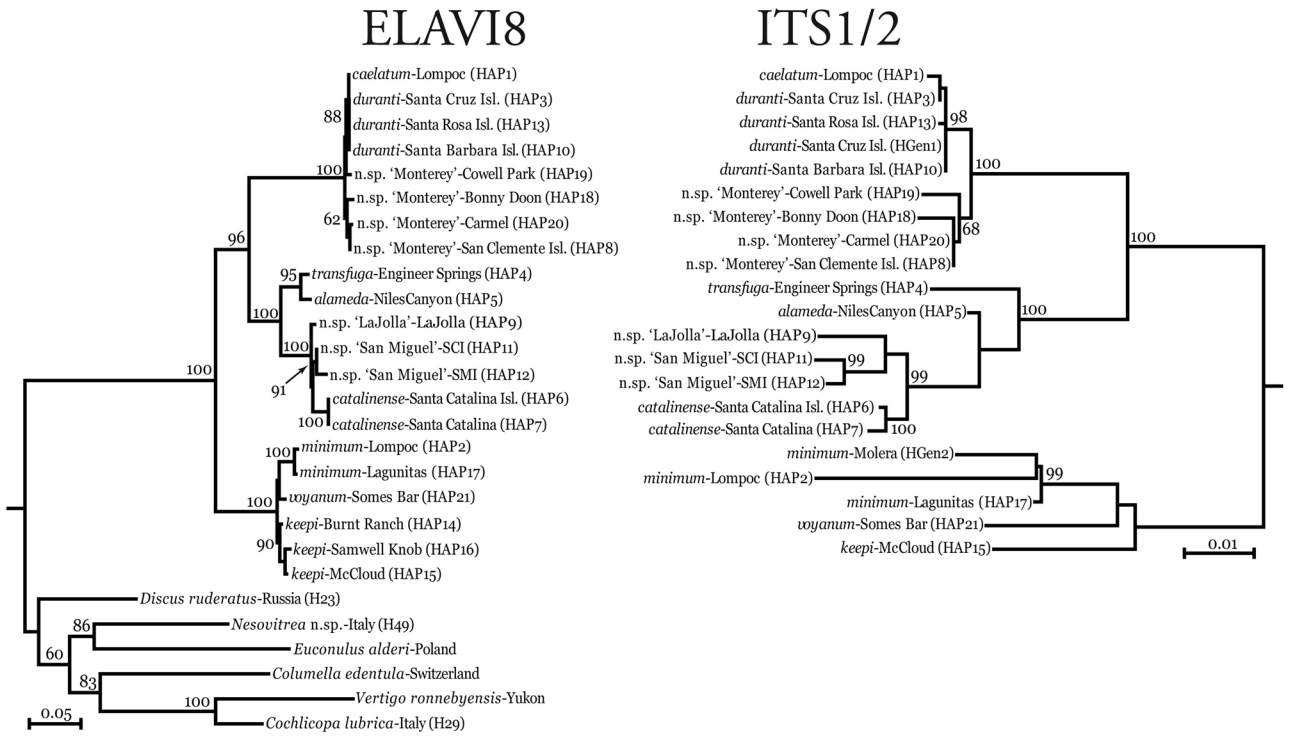

Fig. 4 Comparison of ML phylogenetic reconstructions based on ELAVI8 and concatentated ITS1 + 2 sequence from 21 California, USA, *Haplotrema* specimens representing 9 described and 3 undescribed species. The ELAVI8 *Haplotrema* tree was rooted on a Discidae, a chondrinid, two Orthurethra and two Limacoidea species. This was not possible in ITS1 + 2 because of profound sequence differences between genera. Node support values are based on 1000 bootstrap replicates. Extraction codes are provided for taxa that have multiple DNA sequences for a given amplicon in Appendix I

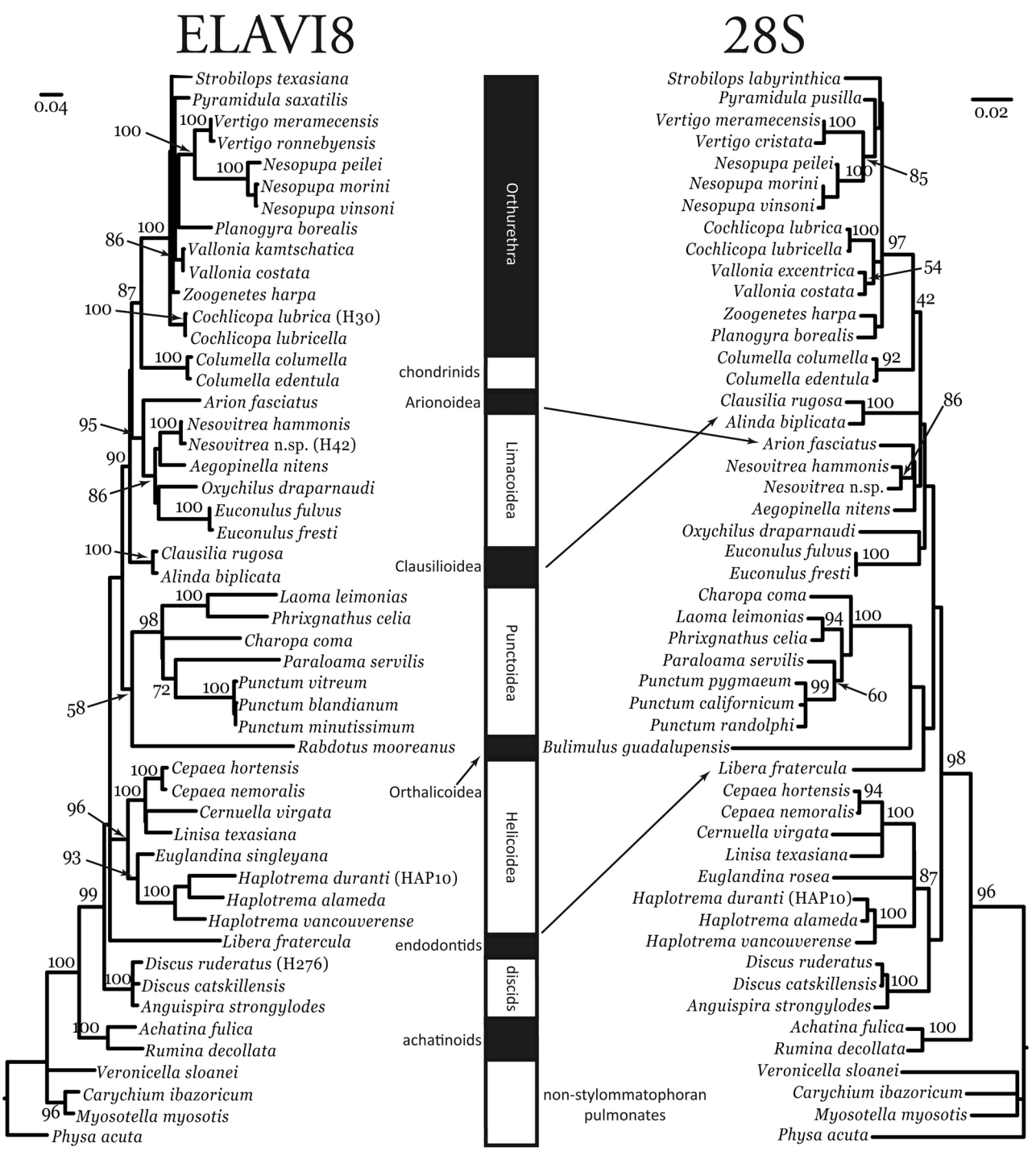

Fig. 5 Comparison of ML phylogenetic reconstructions based on ELAVI8 and 28S sequence from 50 specimens/37 genera representing all of the major panpulmonate land snail clades identified by Wade et al. (2006). Node support values are based on 1000 bootstrap replicates. Extraction codes are provided for taxa that have multiple DNA sequences for a given amplicon in Appendix I

sequence data. While the ITS1 + 2 marker theoretically provides greater resolution and support for species-level clades due to its twofold greater amount of within-genus base pair variability, in practice high rates of intercopy haplotype variation frequently limits the number of specimens that can be analyzed to only the subset which possess a single haplotype across all copies. While the ability of ELAVI8 to resolve species-scale clustering at high support is limited in some genera due to modest levels of base pair variation (e.g., *Vallonia, Euconulus*), this same issue also exists in

ITS1 + 2 (e.g., *Pupilla* Fleming, 1828—Nekola et al., 2015; *Euconulus*—Horsáková et al., 2020). It is possible that in such cases concatenation of ELAVI8 and ITS1 + 2 may provide the statistical power needed to generate higher support, provided that both signals can be shown to display topological concordance.

ELAVI8 also offers improved resolution and support as compared to 28S across deep phylogenetic scales because of its two-fold higher amount of base pair variability. Not only did ELAVI8 generate high node support for genus-level clades (e.g., *Columella, Haplotrema, Nesovitrea*) but it was also able to reproduce the same family-, superfamily-, and superorder-scale groupings as 28S but often at 50% or greater node support. In addition, it demonstrated better node resolution: for example (Fig. 5), ELAVI8 was able to (1) place the chondrinids within the Orthurethra as suggested by Saadi et al. (2021); (2) resolve monophyly within the Limacoidea as suggested by Wade et al. (2006); and (3) demonstrate stronger linkage between *Haplotrema* and *Euglandina* Crosse & Fischer, 1870, and their association with the Helicoidea, than was documented by Wade et al. (2006).

While ELAVI8 appears to provide excellent multiscale phylogenetic signal within the Eunpulmonata and Hygrophila, PCR failed outside of the Panpulmonata. BLAST searches within Genbank-accessible genomic assemblies from Caenogastropoda (e.g., *Batillaria attramentaria*, JACVVK010000098) suggest that their ELAV architecture may differ, with Intron 8 being absent. While ELAVI8 thus will not function as phylogenetic marker across all gastropods, let alone Mollusca, it remains possible that it could function across the entirety of the Panpulmonata.

Conclusions

While advancements in NextGen Sequencing (NGS) might suggest that the development of new PCR amplification targets represents an antiquated activity, we posit that Sanger-based approaches will continue to provide important insights. Even though costs for NGS processing have dramatically fallen, they remain at least an order of magnitude higher than PCR on a per-specimen basis. PCR thus allows for substantially larger sample sizes for a given budget. This is crucial for allowing potential replication across an entity's geographic and ecological range. Higher NGS analysis costs may also block researcher access in poor (and biodiversity-rich) nations from the empirical methods required for species-level hypothesis testing. Because the Eupulmonata and Hygrophila represent approximately one-third of all molluscan species and are world's second most diverse terrestrial invertebrate group (Zhang, 2011), the results from this study indicate considerable potential for ELAVI8 to aid taxonomic revision and phylogenetic reconstruction for this important facet of the Earth's biodiversity.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s13127-022-00587-3.

Acknowledgements Brian Coles, Owen Griffiths, Bill Hoyer, and Ben Hutchins provided some of the specimens used in analysis. Additionally, Tim Pearce provided the *Haplotrema vancouverense* DNA extraction and Rodrigo Salvador provided ELAVI8 sequence for a number of punctids/charopids. Field work to secure California *Haplotrema* and to subclone ITS1 products from a single *H. keepi* individual was supported by Naval Base Ventura County, United States Department of Defense.

Funding The remainder was funded through the Czech Science Foundation (no. 22-23005S).

Data accessibility All sequence data used in this study is stored in GenBank, for accession numbers see Appendix 1. Additionally, the MSA used for all ELAVI8 analysis (ELAVI8_panpul.fas) is provided in the supplementary on-line materials.

Declarations

Competing interests The authors declare no competing interests.

References

Adema, C. M., Hillier, L., Jones, C. S., Loker, E. S., Knight, M., et al. (2017). Whole genome analysis of a schistosomiasis-transmitting freshwater snail. *Nature Communications, 8*, 15451.

Armbruster, G. F. J., & Bernhard, D. (2000). Taxonomic significance of ribosomal ITS-1 sequence markers in self-fertilizing land snails of *Cochlicopa* (Stylommatophora, Cochlicopidae). *Zoosystematics & Evolution, 76*, 11–18.

Armbruster, G. F. J., Böhme, M., Bernhard, D., & Schlegel, M. (2005). The H3/H4 histone gene cluster of land snails (Gastropoda: Stylommatopora): TS/TV ratio, GC3 drive and signals in Stylommatophoran phylogeny. *Journal of Molluscan Studies, 71*, 339–348.

Audzijonyte, A., & Vrijenhoek, R. C. (2010). Three nuclear genes for phylogenetic, SNP and population genetic studies of molluscs and other invertebrates. *Molecular Ecology Resources, 10*, 200–204.

Barker, G. M. (2001). Gastropods on land: Phylogeny, diversity, and adaptive morphology. In G. M. Barker (Ed.), *The Biology of Terrestrial Molluscs* (pp. 1–146). CABI Publishing.

Bejerano, G., Pheasant, M., Makunin, I., Stephen, S., Kent, W. J., Mattick, J. S., & Haussler, D. (2004). Untraconserved elements in the human genome. *Science, 304*, 1321–1325.

Bryja, J., Kostin, D., Meheretu, Y., Šumbera, R., Bryjová, A., et al. (2018). Reticulate Pleistocene evolution of Ethiopian rodent genus along remarkable altitudinal gradient. *Molecular Phylogenetics and Evolution, 118*, 75–87.

Côrte-Real, H. B. S. M., Dixon, D. R., & Holland, P. W. H. (1994). Intron-targeted PCR: A new approach to survey neutral DNA polymorphism in bivalve populations. *Marine Biology, 120*, 407–413.

Darriba, D., Taboada, G. L., Doallo, R., & Posada, D. (2012). jModelTest 2: More models, new heuristics and parallel computing. *Nature Methods, 9*, 772.

Douady, C. J., Delsuc, F., Boucher, Y., Doolittle, W. F., et al. (2003). Comparison of Bayesian and maximum likelihood bootstrap

measures of phylogenetic reliability. *Molecular Biology and Evolution, 20*, 248–254.

Guo, Y., Zhang, Y., Liu, Q, Huang, Y., Mai, G., et al. (2019). A chromosome-level gene assembly for the giant African snail *Achatina fulica*. *Gigascience, 8*(10), giz124.

Haase, M., Meng, S., & Horsák, M. (2021). Tracking parallel adaptation of shell morphology through geological times in the land snail genus *Pupilla* (Gastropoda: Stylommatophora: Pupillidae). *Zoological Journal of the Linnean Society, 191*, 720–747.

Horsáková, V., Nekola, J. C., & Horsák, M. (2020). Integrative taxonomic consideration of the Holarctic *Euconulus fulvus* group of land snails (Gastropoda, Stylommatophora). *Systematics and Biodiversity, 18*, 142–160.

Jennings, R. M., & Etter, R. J. (2011). Exon-primed, intron-crossing (EPIC) loci for five nuclear genes in deep-sea protobranch bivalves: Primer design, PCR protocols and locus utility. *Molecular Ecology Resources, 11*, 1102–1112.

Jörger, K. M., Stöger, I., Kano, Y., Fukuda, H., Knebelsberger, T., & Schrödl, M. (2010). On the origin of Acochlidia and other enigmatic euthyneuran gastropods, with implications for the systematics of Heterobranchia. *BMC Evolutionary Biology, 10*, 323.

Katoh, K., Rozewicki, J., & Yamada, K. D. (2017). MAFFT online service: Multiple sequence alignment, interactive sequence choice and visualization. *Briefings in Bioinformatics, 20*, 1160–1166.

Lessa, E. P. (1992). Rapid surveying of DNA sequence variation in natural populations. *Molecular Biology and Evolution, 9*, 323–330.

Nekola, J. C., Chiba, S., Coles, B. F., Drost, C. A., vonProschwitz, T., & Horsák, M. (2018). A phylogenetic overview of the genus *Vertigo* O. F. Müller, 1773 (Gastropoda: Pulmonata: Pupillidae: Vertigininae). *Malacologia, 62*, 21–161.

Nekola, J. C., & Coles, B. F. (2016). Supraspecific taxonomy in the Vertiginidae (Gastropoda, Stylommatophora). *Journal of Molluscan Studies, 82*, 208–212.

Nekola, J. C., Coles, B. F., & Horsák, M. (2015). Species assignment in *Pupilla* (Gastropoda: Pulmonata: Pupillidae): Integration of DNA-sequence data and conchology. *Journal of Molluscan Studies, 81*, 196–216.

Rannala, B., & Yang, Z. (1996). Probability distribution of molecular evolutionary trees: A new method of phylogenetic inference. *Journal of Molecular Evolution, 43*, 304–311.

Saadi, A. J., Mordan, P. B., & Wade, C. M. (2021). Molecular phylogeny of the Orthurethra (Panpulmonata: Styommatophora). *Zoological Journal of the Linnean Society, 193*, 1126–1140.

Soller, M., & White, K. (2004). ELAV. *Current Biology, 14*, R53.

Wade, C. M., & Mordan, P. B. (2000). Evolution within the gastropod molluscs; using the ribosomal RNA gene-clustre as an indicator of phylogenetic relationships. *Journal of Molluscan Studies, 66*, 565–570.

Wade, C. M., Mordan, P. B., & Naggs, F. (2006). Evolutionary relationships among the Pulmonate land snails and slugs (Pulmonata, Stylommatophora). *Biological Journal of the Linnean Society, 87*, 593–610.

Zhang, Z. Q. (2011). Animal biodiversity: An outline of higher-level classification and survey of taxonomic richness. *Zootaxa, 3148*, 1–237.