The complete mitogenome of *Arion vulgaris* Moquin-Tandon, 1855 (Gastropoda: Stylommatophora): Mitochondrial genome architecture, evolution and phylogenetic considerations within Stylommatophora

Özgül Doğan [Corresp., 1, 2], **Michael Schrödl** [2, 3, 4], **Zeyuan Chen** [2, 3]

[1] Department of Molecular Biology and Genetics, Faculty of Science, Sivas Cumhuriyet University, Sivas, Turkey

[2] SNSB-Bavarian State Collection of Zoology, Munich, Germany

[3] Department Biology II, Ludwig-Maximilians-Universität, Munich, Germany

[4] GeoBio-Center LMU, Munich, Germany

Corresponding Author: Özgül Doğan
Email address: odogan@cumhuriyet.edu.tr

Stylommatophora is one of the most speciose orders of Gastropoda, including terrestrial snails and slugs, some of which are economically important as human food, agricultural pests, vectors of parasites or due to invasiveness. Despite their great diversity and relevance, the internal phylogeny of Stylommatophora has been debated. To date, only 34 stylommatophoran mitogenomes were sequenced. Here, the complete mitogenome of an invasive pest slug, *Arion vulgaris* Moquin-Tandon, 1855 (Stylommatophora: Arionidae), was sequenced using next generation sequencing, analysed and compared with other stylommatophorans. The mitogenome of *A. vulgaris* measures 14,547 bp and contains 13 protein-coding, two rRNA, 22 tRNA genes, and one control region, with an A + T content of 70.20%. All protein coding genes (PCGs) are initiated with ATN codons except for *COX1, ND5* and *ATP8* and all are ended with TAR or T− stop codons. All tRNAs were folded into a clover-leaf secondary structure except for *trnC* and *trnS1* (AGN). Phylogenetic analyses confirmed the position of *A. vulgaris* within the superfamily Arionoidea, recovered a sister group relationship between Arionoidea and Orthalicoidea, and supported monophyly of all currently recognized superfamilies within Stylommatophora. Initial diversification time of the Stylommatophora was estimated as 138.55 million years ago corresponding to Early Cretaceous. The divergence time of *A. vulgaris* and *Arion rufus* (Linnaeus, 1758) was estimated as 15.24 million years ago corresponding to one of Earth's most recent, global warming events, the Mid-Miocene Climatic Optimum. Furthermore, selection analyses were performed to investigate the role of different selective forces shaping stylommatophoran mitogenomes. Although purifying selection is the predominant selective force shaping stylommatophoran mitogenomes, six genes (*ATP8*, *COX1*, *COX3*, *ND3*, *ND4* and *ND6*)

detected by the branch-specific aBSREL approach and three genes (*ATP8*, *CYTB* and *ND4L*) detected by codon-based BEB, FUBAR and MEME approaches were exposed to diversifying selection. The positively selected substitutions at the mitochondrial PCGs of stylommatophoran species seems to be adaptive to environmental conditions and affecting mitochondrial ATP production or protection from reactive oxygen species effects. Comparative analysis of stylommatophoran mitogenome rearrangements using MLGO revealed conservatism in Stylommatophora; exceptions refer to potential apomorphies for several clades including rearranged orders of *trnW-trnY* and of *trnE-trnQ-rrnS-trnM-trnL2-ATP8-trnN-ATP6-trnR* clusters for the genus *Arion*. Generally, tRNA genes tend to be rearranged and tandem duplication random loss, transitions and inversions are the most basic mechanisms shaping stylommatophoran mitogenomes.

The complete mitogenome of *Arion vulgaris* Moquin-Tandon, 1855 (Gastropoda: Stylommatophora): Mitochondrial genome architecture, evolution and phylogenetic considerations within Stylommatophora

Özgül Doğan[1,2*], Michael Schrödl[2,3,4], Zeyuan Chen[2,3]

[1] Department of Molecular Biology and Genetics, Faculty of Science, Sivas Cumhuriyet University, 58140 Sivas, Turkey.

[2] SNSB-Bavarian State Collection of Zoology, 81247 Munich, Germany.

[3] Department Biology II, Ludwig-Maximilians-Universität, 82152 Planegg-Martinsried, Munich, Germany.

[4] GeoBio-Center LMU, 80333 Munich, Germany.

[*]Corresponding Author:

Özgül Doğan[1,2]

Department of Molecular Biology and Genetics, Faculty of Science, Sivas Cumhuriyet University, 58140 Sivas, Turkey.

Email address: odogan@cumhuriyet.edu.tr

ABSTRACT

Stylommatophora is one of the most speciose orders of Gastropoda, including terrestrial snails and slugs, some of which are economically important as human food, agricultural pests, vectors of parasites or due to invasiveness. Despite their great diversity and relevance, the internal phylogeny of Stylommatophora has been debated. To date, only 34 stylommatophoran mitogenomes were sequenced. Here, the complete mitogenome of an invasive pest slug, *Arion vulgaris* Moquin-Tandon, 1855 (Stylommatophora: Arionidae), was sequenced using next generation sequencing, analysed and compared with other stylommatophorans. The mitogenome of *A. vulgaris* measures 14,547 bp and contains 13 protein-coding, two rRNA, 22 tRNA genes, and one control region, with an A + T content of 70.20%. All protein coding genes (PCGs) are initiated with ATN codons except for *COX1, ND5* and *ATP8* and all are ended with TAR or T− stop codons. All tRNAs were folded into a clover-leaf secondary structure except for *trnC* and *trnS1* (AGN). Phylogenetic analyses confirmed the position of *A. vulgaris* within the superfamily Arionoidea, recovered a sister group relationship between Arionoidea and Orthalicoidea, and supported monophyly of all currently recognized superfamilies within Stylommatophora. Initial diversification time of the Stylommatophora was estimated as 138.55 million years ago corresponding to Early Cretaceous. The divergence time of *A. vulgaris* and *Arion rufus* (Linnaeus, 1758) was estimated as 15.24 million years ago corresponding to one of Earth's most recent, global warming events, the Mid-Miocene Climatic Optimum. Furthermore, selection analyses were performed to investigate the role of different selective forces shaping stylommatophoran mitogenomes. Although purifying selection is the predominant selective force shaping stylommatophoran mitogenomes, six genes (*ATP8*, *COX1*, *COX3*, *ND3*, *ND4* and *ND6*) detected by the branch-specific aBSREL approach and three genes (*ATP8*, *CYTB* and *ND4L*) detected by codon-based BEB, FUBAR and MEME approaches were exposed to diversifying

selection. The positively selected substitutions at the mitochondrial PCGs of stylommatophoran species seems to be adaptive to environmental conditions and affecting mitochondrial ATP production or protection from reactive oxygen species effects. Comparative analysis of stylommatophoran mitogenome rearrangements using MLGO revealed conservatism in Stylommatophora; exceptions refer to potential apomorphies for several clades including rearranged orders of *trnW-trnY* and of *trnE-trnQ-rrnS-trnM-trnL2-ATP8-trnN-ATP6-trnR* clusters for the genus *Arion*. Generally, tRNA genes tend to be rearranged and tandem duplication random loss, transitions and inversions are the most basic mechanisms shaping stylommatophoran mitogenomes.

INTRODUCTION

Gastropoda is the most speciose class of Mollusca, including snails and slugs with very diverse feeding habits and a wide range of habitats (Barker, 2009). The about 63,000 gastropod species represent 476 families (Bouchet et al., 2017) and radiated in marine, freshwater and terrestrial ecosystems with detritivorous, herbivorous, carnivorous, predatory or parasitic life styles (Ponder & Lindberg, 1997). Most of the terrestrial gastropods are stylommatophoran pulmonates, with approximately 30,000 species distributed from polar to tropical regions (Mordan & Wade, 2008). Stylommatophorans are economically important as human food and because of their status of being major agricultural pests and/or vectors of parasites and invasiveness (Barker, 2009). The origin of Stylommatophora is within panpulmonate heterobranchs (Jörger et al., 2010) and the monophyly of the order is undisputed. Internal phylogenetic relationships of stylommatophorans were poorly resolved based on morphology but then investigated molecularly in different sampling sets of taxa with various methods and basically relatively short sequences. Tillier, Masselmot & Tillirt, (1996) used the D2 region of *28S rRNA* to explore the phylogenetic relationships of pulmonates including a few stylommatophoran species, however they reported that these short sequences would not have sufficient resolving power for investigating the relationships owing to the probable rapid radiation of pulmonate species. Wade, Mordan & Clarke, (2001) and Wade, Mordan & Naggs, (2006) presented more comprehensive molecular phylogenies based on relatively longer sequence information of the rRNA gene-cluster using 104 species (Wade, Mordan & Clarke, 2001) and 160 species (Wade, Mordan & Naggs, 2006) from Stylommatophora. Although these phylogenetic reconstructions accurately supported the monophyly of achatinoid and non-achatinoid clades, some clades of families that traditionally have been assumed to be

78 monophyletic and some of the morphological groups based on excretory system, in particular
79 monophyly of some families and morphological groups were not supported.

80 The emergence and divergence time of Stylommatophora is also doubtful due to the fragmentary
81 fossil records. The earliest land snails identified as stylommatophoran species are from upper
82 Carboniferous and Permian but their classification has still been controversial (Solem &
83 Yochelson, 1979; Hausdorf, 2000). Bandel, (1991) and Roth et al., (1996) suggested the oldest
84 known fossil records from late Jurassic and Early Cretaceous (*Cheruscicola*) and Early
85 Cretaceous (Pupilloidea). Tillier, Masselmot & Tillirt, (1996) inferred that the Stylommatophora
86 emerged in the transition between late Cretaceous and Paleocene (65-55 Ma) congruent with
87 fossil records, based on the molecular data. However, all of the previous molecular dating
88 analyses on Stylommatophora have been performed either with limited numbers of taxa or
89 molecular markers (Tillier, Masselmot & Tillirt, 1996; Jörger et al., 2010; Dinapoli &
90 Klussmann-Kolb, 2010; Zapata et al., 2014), therefore there is a need for further investigations in
91 a more comprehensive sampling using more markers for better understanding of the phylogeny
92 and timing of evolution of Stylommatophora.

93 In recent years, there is a rapid increase in the number of sequenced mitochondrial genomes
94 (mitogenomes) in parallel to revolution on high throughput DNA sequencing technology and
95 data mining, providing a powerful tool for phylogenetic analysis (Moritz, Dowling & Brown,
96 1987; Boore, 1999; Bernt et al., 2013a). Animal mitogenomes are double-stranded circular
97 molecules which are ~16 kb in length and contain 13 protein coding genes (PCGs) forming the
98 respiratory chain complexes: Complex I or NADH: ubiquinone oxidoreductase contains seven
99 subunits of NADH dehydrogenase (*ND1–6* and *ND4L*), complex III or ubiquinol: cytochrome *c*
100 oxidoreductase consists of cytochrome b (*CYTB*), complex IV or cytochrome *c* oxidase

101 comprises three subunits of cytochrome c oxidase (*COX1–COX3*) and complex V or ATP
102 synthase includes two subunits of the ATPase (*ATP6* and *ATP8*). The mitochondrial PCGs have
103 generally been supposed to be evolving under neutral or nearly neutral selection (Ballard &
104 Kreitman, 1995). Although it has been suggested that these genes are likely to be under strong
105 purifying selection considering their functional importance, the selective pressures might vary
106 even among closely related species and be influenced by environmental conditions (Meiklejohn,
107 Montooth & Rand, 2007). The mitogenomes also encode the small and large subunit rRNAs
108 (*rrnL* and *rrnS*) and twenty-two tRNA genes for the translation process of PCGs. In general, they
109 harbour a single large non-coding region containing control elements necessary for replication
110 and transcription (Boore, 1999). Mitogenomes have become widely used tools in recent
111 phylogeny, phylogeography and molecular dating analyses in various taxa, because of their (1)
112 relatively small size, (2) the high copy number, (3) maternal inheritance type and (4) relatively
113 rapid rate of evolutionary change (Moritz, Dowling & Brown, 1987; Gray, 1989). The sequence
114 information of mitogenomes has also been used in reconstructing phylogenies of several
115 taxonomic groups within/including Gastropoda (White et al., 2011; Stöger & Schrödl, 2013;
116 Sevigny et al., 2015; Uribe et al., 2016a,b; Romero, Weigand & Pfenninger, 2016; Yang et al.,
117 2019). Although there have been some criticisms about the usage of mitogenomes in
118 construction of gastropod phylogeny because of long branch attraction, substitution saturation
119 and strand-specific skew bias (Stöger & Schrödl, 2013), within the recently diversified lineages
120 of gastropods, the use of mitogenomes resulted in highly resolved phylogenies (Williams, Foster
121 & Littlewood, 2014; Osca, Templado & Zardoya, 2014). Besides the use of the mitogenome in
122 sequence-based phylogenies, mitogenome rearrangements can also provide phylogenetic signals
123 (Grande, Templado & Zardoya, 2008; Stöger & Schrödl, 2013; Xie et al., 2019b). Although the

124 mitogenome is widely used in phylogeny of many gastropod groups (Arquez, Colgan & Castro,
125 2014; Osca, Templado & Zardoya, 2014; Sevigny et al., 2015; Uribe, Zardoya & Puillandre,
126 2018), there are limited numbers of reported stylommatophoran mitogenomes and phylogenetic
127 studies in Stylommatophora in terms of usage of mitogenome sequence and rearrangement
128 (Romero, Weigand & Pfenninger, 2016; Xie et al., 2019a; Yang et al., 2019). To date, complete
129 or nearly complete mitogenomes have been reported for only 34 stylommatophoran species
130 (NCBI, September, 2019).

131 In this study, we sequenced and annotated the complete mitogenome of *Arion vulgaris* Moquin-
132 Tandon, 1855 (Gastropoda: Stylommatophora), which is considered as a serious invasive pest
133 both in agriculture and private gardens. We compared it with the mitogenome of its congener
134 *Arion rufus* (Linnaeus, 1758), and with all other previously reported stylommatophoran
135 mitogenomes. We also reconstructed a phylogeny from stylommatophoran mitogenomes to
136 estimate the phylogenetic position of *A. vulgaris* and to test the informativeness of mitogenome
137 data in the reconstruction of Stylommatophora phylogeny. In addition, we obtained a dated
138 phylogeny using this mitogenome dataset and fossil calibrations to estimate divergence times
139 within Stylommatophora. Furthermore, selection analyses were performed to investigate the role
140 of different selective forces shaping stylommatophoran mitogenomes. Finally, we compared the
141 mitogenome organisations of stylommatophoran species using a comparative and phylogeny
142 based method and tried to uncover the evolutionary pathways of mitogenome rearrangements.

143

MATERIALS AND METHODS

Specimen collection and DNA extraction

The specimen of *A. vulgaris* was collected from the garden of the Zoologische Staatssammlung München (ZSM), Germany. Total genomic DNA was extracted from mantle tissue using CTAB method (Doyle & Doyle, 1987).

Mitogenome sequencing, annotation and analyses

The whole-genome sequencing was conducted with 150 bp pair-end reads on the Illumina Hiseq4000 Platform (Illumina, San Diego, CA) using 350 bp insert size libraries. Raw reads were processed by removing low quality reads, adapter sequences and possible contaminated reads using Fastp v0.20.0 (Chen et al., 2018) and Lighter v1.0.7 (Song, Florea & Langmead, 2014). In total, about 7.5G high quality base pairs of sequence data were obtained and the mitogenome was assembled using the MitoZ software (Meng et al., 2019), followed by manual curation using Geneious R9 (Kearse et al., 2012).

The annotation of tRNA genes of the *A. vulgaris* mitogenome was performed using MITOS (http://mitos.bioinf.uni-leipzig.de/index.py) (Bernt et al., 2013b) and ARWEN web servers (Laslett & Canbäck, 2008) based on their secondary structures and anticodon sequences. The locations and boundaries of PCGs and rRNA genes were identified manually by comparing with the *A. rufus* (KT626607) homologous gene sequences. The visualization of the secondary structure of tRNA genes was performed using VARNA v3-93 (Darty, Denise & Ponty, 2009) and RNAviz 2.0.3 (De Rijk, Wuyts & De Wachter, 2003). Intergenic spacers and overlapping regions between genes were estimated manually. The largest non-coding region was defined as control region and the Mfold server (Zuker, 2003) was used to predict the secondary structure of

this region. The "palindrome" tool within the European Molecular Biology Open Software Suite (EMBOSS) (Rice, Longden & Bleasby, 2000) was used for searching the palindromic sequences in the control region. Finally, the complete mitogenome of *A. vulgaris* was deposited in GenBank under accession number MN607980. The mitogenome of *A. vulgaris* is visualized using OrganellarGenomeDRAW (OGDRAW) (Greiner, Lehwark & Bock, 2019).

The nucleotide compositions, average nucleotide and amino acid sequence divergences and the relative synonymous codon usages (RSCU) of PCGs were computed using MEGA v7.0 (Kumar, Stecher & Tamura, 2016). The strand asymmetries were calculated according to the following formulas: AT-skew = [A − T] / [A + T] and GC-skew = [G − C] / [G + C] (Perna & Kocher, 1995).

Phylogenetic and comparative analyses

Alignment and model selection

Phylogenetic and comparative analyses were performed using the mitogenome dataset of 35 stylommatophoran species representing 18 families, and using one species from Systellommatophora, one species from Hygrophila, and one species from Ellobioidea as outgroups (Table 1). Each tRNA and rRNA gene was aligned individually using MAFFT (Katoh & Standley, 2013) algorithm in Geneious R9 (Kearse et al., 2012). The alignment of nucleotide sequences of each PCG was performed using MAFFT algorithm and the "translation align" option implemented in Geneious R9. The final alignment files were then concatenated using SequenceMatrix v.1.7.8 (Vaidya, Lohman & Meier, 2011). The optimal partitioning scheme and substitution models were inferred by PartitionFinder v1.1.1 (Lanfear et al., 2012) using the Bayesian Information Criterion (BIC) and the "greedy" algorithm with the option of "unlinked"

branch lengths. The best-fit partitioning scheme and nucleotide substitution models were used in phylogenetic analyses (Table S1).

Assessing the substitution saturation level

The substitution saturation levels in different genes and codon positions were estimated comparing the uncorrected p-distances and the distances calculated by applying the GTR + G + I evolutionary model selected based on the BIC using jModelTest v2.1.7 (Darriba et al., 2012). All genetic distances were computed with PAUP v4.0 b10 (Swofford, 2002).

Phylogenetic reconstruction

Two different datasets were created for phylogenetic analyses to test the influence of saturated genes and codon positions: (1) 13 PCGs including all codon positions plus the 22 tRNAs and two rRNAs (P123RNA) and (2) PCGs excluding the five saturated genes and third codon positions, plus 22 tRNAs and two rRNAs (8P12RNA, Table S2). Maximum likelihood (ML) trees were constructed with RAxML v8.0.9 (Stamatakis, 2014) implemented in Geneious R9 applying the best-fit evolutionary model for each partition under 1000 bootstrap replicates. For Bayesian Inference (BI) analyses, MrBayes v3.2.2 (Ronquist et al., 2012) was employed with two independent runs of 10 million generations with four Markov chains (three cold, one heated), sampling every 1000 generations and a burn-in of 25% trees. The stationarity of the chains was assessed using the program Tracer v1.7 (Rambaut et al., 2018). The consensus phylogenetic trees were visualized using FigTree v1.4.0 (Rambaut, 2012).

Divergence time estimation

MCMCTree program implemented in the Phylogenetic Analysis by Maximum Likelihood (PAML) package v4.9 (Yang, 2007) was used for Bayesian estimation of divergence times of each species. Substitution rate per site was estimated by BASEML and was used to set the prior for the mean substitution rate in the Bayesian analysis. MCMC was run by 50 × 10000 iterations with the REV substitution model. The soft bounds of *Helix pomatia* + *Cornu aspersum* [divergence time between 34 million years ago (Ma) and 42 Ma], *Mastigeulota kiangsinensis* + (*Dolicheulota formosensis* + (*Aegista aubryana* + *Aegista diversifamilia*)) (divergence time between 25 Ma and 51 Ma), and *Camaena cicatricosa* + *Camaena poyuensis* (divergence time between 16 Ma and 39 Ma) were used as external calibrations (Razkin et al., 2015) and the estimated nodal age of Tectipleura [244 Ma (210–279 Ma)] was used for the calibration of the root (Kano et al., 2016).

Selection analyses

The CODEML implemented in PAML was used to estimate the ratio of nonsynonymous/synonymous substitution rate (ω=dN/dS) and to explore the role of different selective constraints working on each PCG under the one-ratio model (Model A: model=0, NSsites=0, fix_omega=0, omega= 1). Gaps and ambiguous sites of sequence alignments were included in the analyses. For each PCG, likelihood ratio tests (LRTs) were used to compare the null neutral model (Model B: model = 2, NSsites = 2, fix_omega = 1, omega = 1) against alternative models of branch-specific positive selection (Model C: model=2, NSsites=2, fix_omega=0, omega= 1.5). The Bayes Empirical Bayes (BEB) algorithm in CODEML was used to detect the positively selected sites. Furthermore, the adaptive branch-site random effects

likelihood (aBSREL) (Smith et al., 2015) implemented in DATAMONKEY webserver (Weaver et al., 2018) was used to search the signatures of episodic positive diversifying selection testing each branch. In addition, mixed effects model of evolution (MEME) (Murrell et al., 2012) was used to detect episodic or diversifying selection at individual sites and a fast, unconstrained Bayesian approximation for inferring selection (FUBAR) (Murrell et al., 2013) was used for providing additional support to the detection of sites evolving under positive or negative selection. Each PCG was also evaluated in terms of properties and magnitude of amino acid changes using TreeSAAP v3.2 (Woolley et al., 2003), which uses 31 properties of amino acids and categorizes the degree of substitutions to eight categories (1−8).

Comparison of mitogenome organizations

Mitogenome organizations and gene rearrangements of stylommatophoran species were analysed via the CREx web server (http://pacosy.informatik.uni-leipzig.de/crex) (Bernt et al., 2007). The gene orders of ancestral nodes were reconstructed using the Maximum Likelihood for Gene Order Analysis (MLGO, http://geneorder.org/) (Hu, Lin & Tang, 2014) with the input tree obtained by phylogenetic approaches, and the orders of the protein coding, rRNA and tRNA genes were compared with the inferred ancestral mitogenomes. A distance matrix was calculated based on number of common intervals, and the output diagram visually examined to identify shared and/or derived gene rearrangements as well as mechanisms of rearrangements.

RESULTS AND DISCUSSION

Mitogenome characteristics and nucleotide composition

The complete mitogenome sequence of *A. vulgaris* was obtained with a length of 14,547 bp (Table 2) and its size was within the range of the those of other reported stylommatophoran mitogenomes, varying between 13,797 bp in *Camaena poyuensis* and 16,879 bp in *Partulina redfieldi* (Price et al., 2018). It includes the entire set of 37 mitochondrial genes: 13 PCGs, 22 tRNAs and two rRNAs. Twenty-four genes were located on the J strand, while the remainings were encoded by the opposite N strand (Table 2, Fig. 1).

The nucleotide composition of *A. vulgaris* mitogenome was distinctly biased towards A and T, with a 70.20% A + T content, and comparable to other reported stylommatophoran mitogenomes, varying between 59.79% A+T in *Cepea nemoralis* (Yamazaki et al., 1997) and 80.07% A + T in *Achatinella mustelina* (Price et al., 2016a) (Tables 3 and S3). A bias towards A and T nucleotides was also observed in PCGs of the *A. vulgaris* mitogenome with a 69.34% A + T content (Table 3). The A + T content of the 3rd codon position (79.64%) was higher than those of the 2nd (64.21%) and 1st codon positions (64.18%). Similar to other reported stylommatophoran mitogenomes (Table S3), the AT− and GC−skews were found slightly negative (−0.0756) and positive (0.0431) in the whole mitogenome of *A. vulgaris*, respectively. A pronounced T and G skew was also observed in all PCGs (−0.1508, 0.0472), PCGs on the majority strand (−0.1447, 0.0596), and tRNA genes (−0.0010, 0.1582) (Table 3). The T− and G−skewed mitogenome of *A. vulgaris* might be explained by the spontaneous deamination of cytosine during replication and transcription processes (Reyes et al., 1998). The PCGs encoded on the minority strand displayed a T− and C− skewed pattern (−0.1783 AT−skew, −0.0065

277 GC−skew), contrary to the expected high rates of Ts and Gs on the minority strand for most of
278 the metazoans (Hassanin, Léger & Deutsch, 2005).

279

280 **Protein coding genes and codon usage**

281 In comparison, the lengths of the PCGs of *A. vulgaris* mitogenome were within the range of
282 those of other stylommatophoran mitochondrial PCGs. The *ND4* gene was the most variable
283 gene in length and has a variability of 53 codons among stylommatophorans (419 codons in
284 *Microceramus pontificus* and 472 codons in *Orcula dolium*). The most conserved gene in length
285 was *COX1* and it exhibits variability with only 16 codons between species of Stylommatophora
286 (501 codons in *Achatinella mustelina* and 517 codons in *Achatina fulica*). Compared with the
287 mitogenome of *A. rufus*, the lengths of PCGs of *A. vulgaris* were distinct except for *COX1*,
288 *COX2*, *CYTB* and *ND1* genes. The *ND6* gene was the most variable gene in length and was
289 longer in the *A. vulgaris* mitogenome by 11 codons. Based on the amino acid identities, the most
290 conserved PCG was *COX1* (56.45%) whereas the least conserved was *ND6* (11.92%) among the
291 stylommatophoran mitogenomes. The most conserved PCG was *COX1* (97.45%) whereas the
292 least conserved was *ATP8* (68.18%) based on the amino acid identities between the two *Arion*
293 mitogenomes (Table S4).

294 In the *A. vulgaris* mitogenome, most of the PCGs initiated with typical ATN start codon, except
295 for *COX1*, *ND5* and *ATP8* genes which use TTG, ACA and GTG triplets as start codons,
296 respectively (Table 2). The TTG and GTG start codons are also accepted as canonical start
297 codons for invertebrate mitogenomes (Yang et al., 2019), however, ACA as start codon for *ND5*
298 gene was reported for the first time for stylommatophoran mitogenomes. Most of the PCGs were
299 inferred to use TAR as termination codon, except for *ND4L* and *COX3* which have an

abbreviated T− termination codon and their products are probably completed via post-transcriptional polyadenylation (Anderson et al., 1981; Ojala, Montoya & Attardi, 1981).

The most frequently used amino acids by the PCGs of the mitogenomes of *A. vulgaris* and *A. rufus* were leucine (16.71% and 15.91% respectively) and serine (10.33% and 10.18% respectively), similar to PCGs of the mitogenome of other stylommatophoran species (Leu 16,60%, Ser 10.21% on average). The codons rich in A and T, such as UUA−Leu, AUU−Ile, UUU−Phe, AUA−Met, UAU−Tyr, were the most frequently used codons in all stylommatophoran mitochondrial PCGs. The codons rich in terms of G and C content, CGC−CGG−Arg, CAG−Gln, UGC−Cys, CUC−Leu and UCG−Ser were rarely used in both *Arion* mitogenomes (Table S5, Fig. 2). CGN−Arg, CCS-Pro, GCS, UCG and UGC codons are seldom used or never used also in the stylommatophoran mitogenomes and reflected a significant relationship between codon usage and nucleotide content (Table S5).

tRNA and rRNA genes

All of the tRNA genes could be folded into a usual clover-leaf secondary structure, except for *trnS1* (AGN) and *trnC* which lacked dihydrouridine (DHU) and TΨC arms, respectively and formed simple loops (Fig. S1). Their lengths ranged between 57 bp (*trnC*) and 78 bp (*trnG*), with an average 72.72% A + T content. 26 mismatched positions were observed in stem regions and all of the mismatches were G-U pairs (Fig. S1).

The exact boundaries of rRNA genes were determined as being bounded by the adjacent tRNA genes. The *rrnL* gene was located between *trnV* and *trnL1* genes, and the *rrnS* gene was located between *trnQ* and *trnM* genes. The length of the *rrnL* gene was 1013 bp, with a 71.17% A + T content, while that of *rrnS* gene was 747 bp, with a 71.75% A + T content. These were

323 comparable in ranges to homologous genes in other reported stylommatophoran species, ranging
324 from 605 to 1215 bp in *rrnL* and from 564 to 857 bp in *rrnS.*

325

326 **Non-coding and overlapping regions**

327 The total length of intergenic regions in the *A. vulgaris* mitogenome was 670 bp in 16 locations
328 ranging between 1 and 370 bp (Table 2). In general, the largest non-coding region in the animal
329 mitogenomes is considered to contain the signals for replication and transcription, and so called
330 as the control region (Wolstenholme, 1992). The possible candidate for the control region in *A.*
331 *vulgaris* mitogenome was the largest non-coding region located between *trnY* and *trnG genes*
332 with 370 bp in length. This sequence did not give BLAST hits with other putative CRs of other
333 molluscan mitogenomes, however a part of the sequence with 67 bp in length displayed 79.11%
334 sequence similarity with the mitochondrial control region of an amphibian species (*Indotyphlus*
335 *maharashtraensis,* KF540157). Nucleotide composition of this region was slightly biased
336 towards A + T with a 69.73% A + T content. The putative control region had a nine bp poly-T
337 stretch and formed a stable secondary structure comprising seven stems and loops (Fig. 3).
338 Furthermore, this sequence also contained a lot of palindromic sequences which are varying
339 between 4 and 8 bp, but tandemly repeated sequences were not found.

340 The second largest non-coding region was found between *trnW* and *trnY* with a length of 91 bp
341 (Table 2). The A + T composition of the sequence was higher than that of whole genome and
342 putative control region with an 86.81% A + T. This non-coding region also contained a seven bp
343 poly-A stretch and was folded into a secondary structure with two stem and loops. This
344 secondary structure forming AT-rich sequence might function as the origin of the second strand
345 (Wolstenholme, 1992).

346 Eleven overlapping regions with a total length of 164 bp were found throughout the mitogenome
347 of *A. vulgaris*. The largest overlapping region was 41 bp in length and located between *ND6* and
348 *ND5* genes, while the second largest was 32 bp and located between *trnL2* and *ATP8* (Table 2).

349

350 **Phylogeny and divergence times of stylommatophoran species**

351 Regression analyses of pairwise distances revealed that the 1st and 2nd codon positions of *ATP8*,
352 *ND2*, *ND3*, *ND4L* and *ND6* genes, as well as the 3rd codon positions of all PCGs were saturated
353 (Table S2). Four phylogenetic reconstruction analyses were performed with combination of
354 inference methods and different data matrices to test the influence of inference methods and
355 saturation level of genes/codon positions on tree topology and nodal support. Three different tree
356 topologies were obtained as the results of these analyses, and topologies were sensitive to both
357 inference methods and exclusion of saturated genes/ codon positions (Figures 4 and S2-4). Nodal
358 support values were always higher in BI trees than ML trees of the corresponding dataset. The
359 usage of all mitochondrial genes and codon positions (P123RNA dataset) under both approaches
360 resulted with identical tree topology (Figures S3 and S4), which were similar to the results of Yang
361 et al (2019) obtained using only amino acid sequences of mitochondrial PCGs. The results of these
362 analyses supported the monophyly of all included superfamilies with high nodal supports except
363 for the superfamily Helicoidea which recovered with low nodal support [Bayesian Posterior
364 Probability (BPP)=0.75, Bootstrap support (BS)=58%] and recovered Arionoidea superfamily as
365 sister group to Urocoptoidea + (Polygyroidea + Helicoidea) clade (BPP=1.00, BS=100%), and
366 Succineoidea + Orthalicoidea clade was recovered as sister group to Arionoidea + (Urocoptoidea
367 + (Polygyroidea + Helicoidea)) (BPP=1.00, BS=50%). The ML and BI analyses performed using
368 the dataset constructed with the removal of the saturated PCGs and codon positions (8P12RNA)

resulted in two different tree topologies (Figures 4 and S2). The phylogenetic tree obtained from ML analysis did not support the monophyly of the superfamily of Helicoidea and the superfamily of Polygyroidea placed within the superfamily of Helicoidea (BS=61%, Fig. S2). A highly resolved tree with higher nodal support values was obtained from the BI approach of the dataset 8P12RNA, and hence considered as most reliable tree for discussion. The results confirmed the taxonomic position of *A. vulgaris* as sister species to *A. rufus* and recovered the monophyly of the Arionoidea superfamily (Arionidae + Philomycidae) with high support values (BPP=1.00). A well-supported sister group relationship between Arionoidea and Orthalicoidea was recovered (BPP=0.98) for the first time. However, previous studies using different datasets and sampling of taxa have proposed different sister groups with Arionoidea superfamily. Wade, Mordan & Naggs, (2006) have found the superfamily Limacoidea as sister group to the superfamily Arionoidea using 160 stylommatophoran species, however they used only 823 nucleotides from rRNA gene-cluster. Holznagel, Colgan & Lydeard, (2010) have proposed a sister group relationship between Arionoidea and Limacoidea + Zonitoidea based on the 28S rRNA sequences using seven species from Stylommatophora. The sister group relationships between Arionoidea and Urocoptoidea + Enoidea + Helicoidea (Jörger et al., 2010) or Limacoidea + (Succineoidea + Helicoidea) (Dayrat el al., 2011) have also been suggested by previous studies using relatively longer DNA sequences, however in both studies, only five stylommatophoran species were included for phylogenetic analyses. Furthermore, Xie et al., (2019b) have proposed sister group relationship between Arionoidea and Succineoidea using only amino acid dataset of mitochondrial PCGs, and stated it might be an artefact of poor taxon sampling.

In the phylogenetic tree obtained from 8P12RNA under BI approach, the monophyly of all included families and superfamilies were also supported with high support values except for the

392 superfamily Helicoidea which supported with a low nodal support (BPP=0.82) (Fig. 4).
393 Arionoidea + Orthalicoidea clade was recovered as sister group to Succineoidea + (Urocoptoidea
394 + (Polygyroidea + Helicoidea)). The tree (Fig. 4) also recovered *Deroceras reticulatum*
395 (Limacoidea: Agriolimacidae) at the most basal placement and did not support the monophyly of
396 the suborder Helicina similar to the tree in Yang et al. (2019).

397 A chronogram for Stylommatophora divergence times based on the obtained tree topology is
398 shown in Figure 5. According to our divergence time analysis, the crown age of
399 stylommatophorans was estimated as 138.55 Ma (180.8−107.4 Ma, 95% CI) corresponding to
400 Early Cretaceous. Our estimated times for initial diversification of Stylommatophora are slightly
401 older but broadly congruent with the fossil records and previous studies (Tillier, Masselmot &
402 Tillirt, 1996; Jörger et al., 2010; Dinapoli & Klussmann-Kolb, 2010). Although Solem &
403 Yochelson, (1979) suggested a Paleozoic origin for Stylommatophora, the widely accepted fossil
404 records with recognizable taxa began from Late Cretaceous (Bandel & Riedel, 1994). The
405 Cretaceous origin of stylommatophoran species was also suggested by sequence studies of 28S
406 rDNA fragments by Tillier, Masselmot & Tillirt, (1996), of combined data of 18S, 28S, 16S
407 rDNA and COI by Dinapoli & Klussmann-Kolb, (2010) and Jörger et al., (2010). The
408 diversification of the stylommatophoran species may have been influenced by the explosive
409 radiation of angiosperms and speciation by host-switching during Cretaceous (Friis, Pedersen &
410 Crane, 2010).

411 The split time of *Achatina fulica* from other stylommatophoran species was inferred as 131.91
412 Ma in Early Cretaceous. The splits of the superfamilies Orthalicoidea and Arionoidea, of
413 Succineoidea from Urocoptoidea + (Polygyroidea + Helicoidea), and of Clausilioidea +
414 (Pupilloidea + Achatinelloidea) were dated to 114.18 Ma (148.7-87.2 Ma, 95% CI), 113.30 Ma

(146.1-87.9 Ma, 95% CI) and 111.88 Ma (148.1-84.2 Ma, 95% CI), respectively, coinciding to the beginning of the Albian (Early Cretaceous). The crown ages of the superfamilies Arionoidea, Urocoptoidea, Helicoidea and Pupilloidea were estimated corresponding to Late Cretaceous (84.82, 75.23, 74.39 and 70.06 Ma, respectively). The split of the two *Arion* species and the crown age of Achatinelloidea species were dated to 15.24 Ma (30.0-7.6 Ma, 95% CI) and 13.01 Ma (19.7-8.2 Ma, 95% CI), respectively, corresponding to the Miocene. The divergence time of *A. vulgaris* and *A. rufus* corresponds to one of Earth's most recent, global warming events, the Mid-Miocene Climatic Optimum (MMCO, 17-14.75 Ma) (Böhme, 2003). The MMCO is thought to have contributed to floristic and faunistic diversity across the world and so to animal-plant interactions, correlating with the rise in temperature (Barnosky & Carrasco, 2002; Vicentini et al., 2008; Tolley, Chase & Forest, 2008). The change of plant diversity, emergence of new host plants and the relative warm period may have triggered the diversification of the two *Arion* species. The divergence time of two polygyroid species was inferred as 0.45 Ma (1.1-0.1 Ma, 95% CI), in the Pleistocene.

Selective pressures on stylommatophoran mitogenomes

The ω value for each of the 13 PCGs was inferred under one-ratio model using PAML and presented in Table 4. All of the ω values were extremely low (ω<1), ranging between 0.0129 for *COX1* and 0.2198 for *ATP8,* reflecting that all genes were under strong purifying selection consistent with the general mitogenome evolution pattern in animals (Rand, 2001; Bazin, Glemin & Galtier, 2006). Although purifying selection is the predominant selective force shaping stylommatophoran mitogenomes, the comparison of the null neutral model and alternative branch-specific positive selection model revealed six of the PCGs (*ATP6*, *COX2*, *COX3*, *ND2*,

438 *ND4* and *ND5*) have variation in ω values along different branches. The variability in ω values
439 indicated different selective forces acting on each gene as well as each branch. A more sensitive
440 branch-site method, aBSREL, providing three states for each branch and allowing each site to
441 evolve under any kind of the value (<1, 1 or >1) (Smith et al., 2015), was used for evaluating and
442 confirming the selective forces across lineages determined by PAML analysis. All of the
443 branches in the stylommatophoran phylogeny were tested with aBSREL analysis for each PCG,
444 and the genes detected as under episodic diversifying selection were different from the results of
445 branch-site model of PAML (Table 5) except for *COX3* and *ND4*. The aBSREL analyses
446 discovered episodic diversifying selection in *ATP8* (at the branch leading to *Microceramus*
447 *pontificus*), *COX1* (at the branch leading to *Achatinella mustelina*), *COX3* (at the branch leading
448 to Arionoidea and the branch leading to *Philomycus bilineatus*), *ND3* (at the branch leading to
449 *Helicella itala*), *ND4* (at the branch leading to *Succinea putris*) and *ND6* (at the branch leading
450 to *Vertigo pusilla*). Due to their important function, mitochondrial genes might have a few
451 positively selected sites and the signatures of purifying selection likely mask those of positive
452 selection (Meiklejohn, Montooth & Rand, 2007; da Fonseca et al., 2008). Therefore, two
453 different methods were used to detect positive selection in addition to BEB analysis: FUBAR
454 which estimates the rates of nonsynonymous and synonymous substitutions at each codon in a
455 phylogeny, and MEME which estimates the probability for a codon to have experienced episodic
456 positive selection and allows the ω ratio to vary across branches and codons. BEB analysis
457 identified eight positively selected codons in total in three genes (*ND2*, *ND4* and *ND5*), whereas
458 FUBAR defined six positively selected codons in five genes (*ATP6*, *ATP8*, *COX2*, *CYTB* and
459 *ND4L*). The MEME analysis found the signals of episodic positive selection at 22 codons in nine
460 genes (*ATP6*, *ATP8*, *CYTB*, *ND2-6*, and *ND4L*). There was not any shared codon determined by

all of the three analyses (Table 6). Only four codons in three genes were shared by the results of FUBAR and MEME analyses: 44th codon in *ATP8* gene, 12th codon in *CYTB* gene, and 13th and 57th codons in *ND4L* gene. Therefore, we focused only on these four codons in the TreeSAAP analyses. The positively selected substitution at codon 44 in *ATP8* gene was the change of TTA (Leu) to ATT (Ile) at branches leading to *M. kiangsinensis*, *Cerion incanum* and *Cerion uva*. This substitution was a radical chemical change with a magnitude category of 8 and had an impact on the increment of the equilibrium constant (ionization of COOH). The change at the codon 12 in *CYTB* gene was a conserved change with a magnitude category of 1 and was a substitution of TTG (Leu) to ATG (Met). The positively selected substitutions in *ND4L* gene were the change of ATT (Ile) to ATA (Met) at branch leading to *H. pomatia*, to GTT (Val) at branch leading to *C. nemoralis* at codon 13, and the change of TTT (Phe) to AAT (Asn) at branch leading to Arionidae family at codon 57. The substitution at the 13th codon was a radical change with a magnitude category of 8 altering the equilibrium constant (ionization of COOH), while that at the 57th codon was a radical change with a magnitude category of 7 and modifying the solvent accessibility of the protein.

Consequently, six positive selected genes (*ATP8*, *COX1*, *COX3*, *ND3*, *ND4* and *ND6*) detected by branch-specific aBSREL approach and three genes (*ATP8*, *CYTB* and *ND4L*) detected by codon-based BEB, FUBAR and MEME approaches were exposed to diversifying selection. Four of these genes (*ND3*, *ND4*, *ND4L* and *ND6*) play an important role in oxidative phosphorylation and are subunits of NADH dehydrogenase (Complex I) which is the most complicated and largest proton pump of the respiratory chain coupling electron transfer from NADH to ubiquinone. In addition to its important role in energy production, it has been shown that complex I is implicated in the regulation of reactive oxygen species (ROS) (Sharma, Lu & Bai,

2009). Substitutions in this complex might have been favoured for increasing the efficiency of proton pumping or regulating the response to ROS depending varying amount of oxygen in the atmosphere and adaptation to conditions in new habitats (temperature, humidity, altitude) and / or hosts. *CYTB* gene encodes only mitogenome derived subunit of Complex III and catalyses reversible electron transfer from ubiquinol to cytochrome c (da Fonseca et al., 2008). The positively selected sites in complexes I and III have been suggested to contribute to environmental adaptation in different groups such as mammals, birds, fishes and insects (da Fonseca et al., 2008; Garvin, Bielawski & Gharrett, 2011; Garvin et al., 2014; Melo-Ferreira et al., 2014; Morales et al., 2015; Li et al., 2018). In the cytochrome *c* oxidase complex (Complex IV), *COX1* protein catalyses electron transfer to the molecular oxygen; *COX2* and *COX3* belong to the catalytic core of the complex may act as a regulator. *ATP8* gene encodes the part of ATP synthase (Complex V) regulating the assembly of complex (da Fonseca et al., 2008). The favoured substitutions in *COX3* and *ATP8* gene might have an impact on assembly of the complexes IV and V. The positively selected substitutions and random accumulation of variation in mitochondrial PCGs of stylommatophoran species thus seem to be adaptive and affecting mitochondrial ATP production or protection from ROS effects, however effects of substitutions should be examined in a larger sample by considering protein folding and three-dimensional structure of complexes.

Gene rearrangements in stylommatophoran mitogenomes

The ancestral mitogenome organisation of each node in the phylogeny was inferred using the maximum likelihood approach. The organisation of the hypothetical ancestral Stylommatophora mitogenome (node: A34, Fig. 4) was identical with that of *Deroceras reticulatum* as well as

those of *Albinaria caerulea*, *Cernuella virgata* and *Helicella itala*. The mitogenome of *Achatina fulica* has only experienced the transposition of *trnP* to the downstream of *trnA* compared to its most recent ancestral mitogenome organisation. The common ancestors of Clausilioidea + (Pupilloidea + Achatinelloidea) (node: A31, Fig. 4), Orthalicoidea + Arionoidea (node: A22, Fig. 4), and Succineoidea + (Urocoptoidea + (Polygyroidea + Helicoidea)) (node: A18, Fig. 4) maintained the same order of hypothetical ancestral stylommatophoran mitogenome. In the mitogenome of the most recent common ancestor (MRCA) of Pupilloidea + Achatinelloidea (node: A30, Fig. 4), the reversal of *trnW*, *trnG* and *trnH* genes occurred individually and were followed by the reversal of the cluster *trnW-trnG-trnH*. In the superfamily Pupilloidea, rearrangements of several tRNA genes were observed: the transposition of the cluster *trnD-trnC* to downstream of *trnW* in *Pupilla muscorum*, transpositions of cluster *trnH-trnG* to downstream of *trnW* and of *trnT* to upstream of *COX3* in *Orcula dolium*, transposition of *trnG* to downstream of *trnW* in *Vertigo pusilla* and the reversal of *trnQ* in *Gastrocopta cristata*. In the mitogenome of the MRCA of the superfamily Achatinelloidea (node: A26, Fig. 4), *trnF-COX2-trnY-trnH-trnG-trnW-trnQ-ATP8-trnN-ATP6-trnR-trnE-rrnS-trnM* gene cluster rearranged as *trnW-trnQ-ATP8-ATP6-trnR-trnE-rrnS-trnM-trnF-COX2-trnY-trnH-trnG-trnN* via tandem duplication random loss (TDRL) mechanism. The organisation of the mitogenomes of achatinelloid species nearly matched with the putative ancestral order, except for *Achatinella sowerbyana* which has a transposed position of *trnK* to downstream of *ATP8* and a second copy of *trnL2*, and for *Partulina redfieldi* which has the inversion of *trnE* and *trnN* genes.

The mitogenome of *Naesiotus nux* has almost the same organisation with its MRCA (node: A22, Fig. 4), except for the second inverted copy of *ND4L* located between *trnL1* and *trnP*. The MRCA of the superfamily Arionoidea (node: A21, Fig. 4) had also identical mitogenome

organisation with the ancestor of Stylommatophora, and the MRCAs of the families Arionidae (node: A19, Fig. 4) and of Philomycidae (node: A20, Fig. 4) were derived from this ancestor. The mitogenome of node A19 had shuffled positions of *trnY* and *trnW*, and also transpositions of *trnE* to downstream of *trnQ* and of *rrnS*-*trnM* to upstream of *trnQ*. Both *Arion* mitogenomes also shared this mitogenome organisation and the rearranged orders of *trnW-trnY* and *trnE-trnQ-rrnS-trnM-trnL2-ATP8-trnN-ATP6-trnR* clusters seem to be synapomorphies of this genus. The mitogenome organisation of the node A20 was quite different from those of other stylommatophoran species, which had rearranged positions of almost all genes between *COX1* and *trnI* via two-steps TDRL, and two Philomycidae species also had identical organisation except for *Philomycus bilineatus* had a second copy of *trnC* located downstream of the original copy.

The mitogenome of *Succinea putris* has experienced the transpositions of *trnF* to upstream of *trnD* and of *trnW* to upstream of *trnY*, and also reverse transposition of the cluster *trnW-trnY* to the upstream of *ND3*. The MRCA of the all urocoptoid species (node: A16, Fig. 4) only had the reversal of *trnQ* gene from minor to major strand and *M. pontificus* has also maintained the identical arrangement. The four step requiring rearranged gene cluster was identified in the mitogenome of the MRCA of the genus *Cerion* (node: A15, Fig. 4): (i) reversal of *trnV-rrnL-trnL1*, (ii) reversal of *trnP*, (iii) reversal of *trnA*, and (iv) reversal of the cluster *trnL1-rrnL-trnV-trnP-trnA*. The mitogenome organisation remained the same in all three *Cerion* species and the rearranged state of *trnA-trnP-trnV-rrnL-trnL1* cluster might be a synapomorphy for this genus.

The MRCAs of the polygroid species (node: A12, Fig. 4) and Camaenidae + Bradybaenidae (node: A5, Fig. 4), as well as the *Polygyra cereolus* and *Praticolella mexicana,* had the transposed position of *trnG-trnH* to the upstream of *trnY*. The rearrangement of this cluster as

553 *trnG-trnH-trnY* could be suggested as a synapomorphy for Polygyroidea, but more sampling is
554 required to confirm its status at superfamily level. In the superfamily Helicoidea, the MRCAs of
555 Geomitridae (node: A9, Fig. 4) and Geomitridae + Helicidae (node: A10, Fig. 4) shared the
556 identical mitogenome organisation with the MRCA of Stylommatophora. Both of the
557 Geomitridae species had also same mitogenome organisation except for *ATP8* in *Cernuella*
558 *virgata,* in which this gene was missing, however it seems to be likely a misannotation. The
559 mitogenome of MRCA of Helicidae species (node: A8, Fig. 4) had experienced the
560 transpositions of *trnP* and cluster *trnT-COX3* to downstream of *ND6* and to upstream of *trnS1*,
561 respectively. The mitogenome organisations of *Helix pomatia*, *Cornu aspersum* and *Cepaea*
562 *nemoralis* have not changed and *trnA-ND6-trnP* and *trnS2-trnT-COX3-trnS1* gene orders might
563 be interpreted as synapomorphic for these three species. However, the individual reversals of
564 *trnA*, *ND6* and *trnP* genes followed by reversal of the cluster *trnA-ND6-trnP*, and reversal of
565 *trnS1* were observed in *Cylindrus obtusus* mitogenome. In the mitogenomes of the species of the
566 family Camaenidae, only the transpositions of *trnD* and *trnY* to downstream of *COX2* and to
567 upstream of *trnG* were found, respectively. The arrangement of the *trnC-trnF-COX2-trnD-trnY-*
568 *trnG* cluster could be considered as a synapomorphy for camaenid species, however the
569 taxonomic level of this synapomorphy need to be evaluated in a wider taxonomic range. In the
570 family of Bradybaenidae, the MRCA mitogenome had experienced only the reversal of *trnW*. In
571 addition to this rearrangement, *Aegista* species also have the transposition of *ND3* gene to the
572 downstream of *trnW* and the rearranged position of *ND3-trnW* cluster appears to be a
573 synapomorphy for the genus.
574

CONCLUSIONS

The sequencing and annotation of the mitogenome of *A. vulgaris* and its comparison with other stylommatophoran mitogenomes allow us to denote several conclusions: (i) the mitogenome characteristics of *A. vulgaris* are mostly consistent with the reported stylommatophoran mitogenomes, (ii) rearrangement events are detected in the *trnW-trnY* and *trnE-trnQ-rrnS-trnM-trnL2-ATP8-trnN-ATP6-trnR* gene clusters which may be apomorphic for the genus *Arion*, but further investigations are necessary, (iii) stylommatophoran mitogenome sequence information without the saturated positions seems to be useful for reconstructing phylogeny and estimating divergence times, and the taxon set used should be expanded, (iv) although purifying selection is the dominant force in shaping the stylommatophoran mitogenomes, in the background, several codons or different branches have experienced diversifying selection suggesting adaptation to new environmental conditions.

ACKNOWLEDGMENTS

This study has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 764840.

REFERENCES

Ahn S-J., Martin R., Rao S., Choi M-Y. 2017. The complete mitochondrial genome of the gray garden slug *Deroceras reticulatum* (Gastropoda: Pulmonata: Stylommatophora). *Mitochondrial DNA Part B* 2:255–256. DOI: 10.1080/23802359.2017.1318677.

Anderson S., Bankier AT., Barrell BG., De Bruijn MHL., Coulson AR., Drouin J., Eperon IC., Nierlich DP., Roe BA., Sanger F., Schreier PH., Smith AJH., Staden R., Young IG.

1981. Sequence and organization of the human mitochondrial genome. *Nature* 290:457–465. DOI: 10.1038/290457a0.

Arquez M., Colgan D., Castro LR. 2014. Sequence and comparison of mitochondrial genomes in the genus *Nerita* (Gastropoda: Neritimorpha: Neritidae) and phylogenetic considerations among gastropods. *Marine Genomics* 15:45–54. DOI: 10.1016/j.margen.2014.04.007.

Ballard JW., Kreitman M. 1995. Is mitochondrial DNA a strictly neutral marker? *Trends in Ecology & Evolution* 10:485–8.

Bandel K. 1991. Gastropods from brackish and fresh water of the Jurassic—Cretaceous transition (a systematic reevaluation). *Berliner Geowissenschaftliche Abhandlungen Reihe A Geologie und Palaeontologie* 134:9–55.

Bandel K., Riedel F. 1994. The Late Cretaceous gastropod fauna from Ajka (Bakony Mountains, Hungary): A revision. *Annalen des Naturhistorischen Museums in Wien.* 1–65.

Barker GM. 2009. Gastropods on land: phylogeny, diversity and adaptive morphology. In: *The biology of terrestrial molluscs*. DOI: 10.1079/9780851993188.0001.

Barnosky AD., Carrasco MA. 2002. Effects of Oligo-Miocene global climate changes on mammalian species richness in the northwestern quarter of the USA. *Evolutionary Ecology Research*.

Bazin E., Glemin S., Galtier N. 2006. Population size does not influence mitochondrial genetic diversity in animals. *Science* 312:570–572. DOI: 10.1126/science.1122033.

618 Bernt M., Bleidorn C., Braband A., Dambach J., Donath A., Fritzsch G., Golombek A., Hadrys
619 H., Jühling F., Meusemann K., Middendorf M., Misof B., Perseke M., Podsiadlowski
620 L., von Reumont B., Schierwater B., Schlegel M., Schrödl M., Simon S., Stadler PF.,
621 Stöger I., Struck TH. 2013a. A comprehensive analysis of bilaterian mitochondrial
622 genomes and phylogeny. *Molecular Phylogenetics and Evolution* 69:352–364. DOI:
623 10.1016/j.ympev.2013.05.002.

624 Bernt M., Donath A., Jühling F., Externbrink F., Florentz C., Fritzsch G., Pütz J., Middendorf
625 M., Stadler PF. 2013b. MITOS: Improved de novo metazoan mitochondrial genome
626 annotation. *Molecular Phylogenetics and Evolution* 69:313–319. DOI:
627 10.1016/j.ympev.2012.08.023.

628 Bernt M., Merkle D., Ramsch K., Fritzsch G., Perseke M., Bernhard D., Schlegel M., Stadler
629 PF., Middendorf M. 2007. CREx: Inferring genomic rearrangements based on common
630 intervals. *Bioinformatics*. DOI: 10.1093/bioinformatics/btm468.

631 Böhme M. 2003. The Miocene Climatic Optimum: Evidence from ectothermic vertebrates of
632 Central Europe. *Palaeogeography, Palaeoclimatology, Palaeoecology*. DOI:
633 10.1016/S0031-0182(03)00367-5.

634 Boore JL. 1999. Animal mitochondrial genomes. *Nucleic acids research* 27:1767–80.

635 Bouchet P., Rocroi J-P., Hausdorf B., Kaim A., Kano Y., Nützel A., Parkhaev P., Schrödl M.,
636 Strong EE. 2017. Revised classification, nomenclator and typification of gastropod and
637 monoplacophoran families. *Malacologia* 61:1–526. DOI: 10.4002/040.061.0201.

638 Chen S., Zhou Y., Chen Y., Gu J. 2018. fastp: an ultra-fast all-in-one FASTQ preprocessor.
639 *Bioinformatics* 34:i884–i890. DOI: 10.1093/bioinformatics/bty560.

640 Darriba D., Taboada GL., Doallo R., Posada D. 2012. jModelTest 2: more models, new
641 heuristics and parallel computing. *Nature Methods* 9:772–772. DOI:
642 10.1038/nmeth.2109.

643 Darty K., Denise A., Ponty Y. 2009. VARNA: Interactive drawing and editing of the RNA
644 secondary structure. *Bioinformatics* 25:1974–1975. DOI:
645 10.1093/bioinformatics/btp250.

646 Dayrat B., Conrad M., Balayan S., White TR., Albrecht C., Golding R., Gomes SR.,
647 Harasewych MG., de Frias Martins AM. 2011. Phylogenetic relationships and evolution
648 of pulmonate gastropods (Mollusca): New insights from increased taxon sampling.
649 *Molecular Phylogenetics and Evolution* 59:425–437. DOI:
650 10.1016/j.ympev.2011.02.014.

651 Deng P-J., Wang W-M., Huang X-C., Wu X-P., Xie G-L., Ouyang S. 2016. The complete
652 mitochondrial genome of Chinese land snail *Mastigeulota kiangsinensis* (Gastropoda:
653 Pulmonata: Bradybaenidae). *Mitochondrial DNA* 27:1441–1442. DOI:
654 10.3109/19401736.2014.953083.

655 Dinapoli A., Klussmann-Kolb A. 2010. The long way to diversity – Phylogeny and evolution of
656 the Heterobranchia (Mollusca: Gastropoda). *Molecular Phylogenetics and Evolution*
657 55:60–76. DOI: 10.1016/j.ympev.2009.09.019.

658 Doyle J., Doyle J. 1987. A rapid isolation procedure for small amounts of leaf tissue.
659 *Phytochemical Bulletin* 19:11–15. DOI: 10.2307/4119796.

660 da Fonseca RR., Johnson WE., O'Brien SJ., Ramos MJ., Antunes A. 2008. The adaptive
661 evolution of the mammalian mitochondrial genome. *BMC Genomics* 9:119. DOI:
662 10.1186/1471-2164-9-119.

663 Friis EM., Pedersen KR., Crane PR. 2010. Diversity in obscurity: fossil flowers and the early
664 history of angiosperms. *Philosophical Transactions of the Royal Society B: Biological*
665 *Sciences* 365:369–382. DOI: 10.1098/rstb.2009.0227.

666 Gaitán-Espitia JD., Nespolo RF., Opazo JC. 2013. The Complete Mitochondrial Genome of the
667 Land Snail *Cornu aspersum* (Helicidae: Mollusca): Intra-Specific Divergence of
668 Protein-Coding Genes and Phylogenetic Considerations within Euthyneura. *PLoS ONE*
669 8. DOI: 10.1371/journal.pone.0067299.

670 Garvin MR., Bielawski JP., Gharrett AJ. 2011. Positive darwinian selection in the piston that
671 powers proton pumps in Complex I of the mitochondria of Pacific salmon. *PLoS ONE*
672 6. DOI: 10.1371/journal.pone.0024127.

673 Garvin MR., Bielawski JP., Sazanov LA., Gharrett AJ. 2014. Review and meta-analysis of
674 natural selection in mitochondrial complex I in metazoans. *Journal of Zoological*
675 *Systematics and Evolutionary Research* 53. DOI: 10.1111/jzs.12079.

676 González VL., Kayal E., Halloran M., Shrestha Y., Harasewych MG. 2016. The complete
677 mitochondrial genome of the land snail *Cerion incanum* (Gastropoda:
678 Stylommatophora) and the phylogenetic relationships of Cerionidae within
679 Panpulmonata. *Journal of Molluscan Studies* 82:525–533. DOI:
680 10.1093/mollus/eyw017.

681 Grande C., Templado J., Zardoya R. 2008. Evolution of gastropod mitochondrial genome
682 arrangements. *BMC Evolutionary Biology* 8:61. DOI: 10.1186/1471-2148-8-61.

683 Gray MW. 1989. Origin and evolution of mitochondrial DNA. *Annual Review of Cell Biology*.
684 DOI: 10.1146/annurev.cb.05.110189.000325.

685 Greiner S., Lehwark P., Bock R. 2019. OrganellarGenomeDRAW (OGDRAW) version 1.3.1:
686 expanded toolkit for the graphical visualization of organellar genomes. *Nucleic Acids*
687 *Research* 47:W59–W64. DOI: 10.1093/nar/gkz238.

688 Groenenberg DSJ., Harl J., Duijm E., Gittenberger E. 2017. The complete mitogenome of
689 *Orcula dolium* (Draparnaud, 1801); ultra-deep sequencing from a single long-range
690 PCR using the Ion-Torrent PGM. *Hereditas* 154:7. DOI: 10.1186/s41065-017-0028-2.

691 Groenenberg DSJ., Pirovano W., Gittenberger E., Schilthuizen M. 2012. The complete
692 mitogenome of *Cylindrus obtusus* (Helicidae, Ariantinae) using Illumina next
693 generation sequencing. *BMC Genomics* 13. DOI: 10.1186/1471-2164-13-114.

694 Hassanin A., Léger N., Deutsch J. 2005. Evidence for multiple reversals of asymmetric
695 mutational constraints during the evolution of the mitochondrial genome of Metazoa,
696 and consequences for phylogenetic inferences. *Systematic Biology* 54:277–298. DOI:
697 10.1080/10635150590947843.

698 Harasewych MG., González VL., Windsor AM., Halloran M. 2017. The complete
699 mitochondrial genome of *Cerion uva uva* (Gastropoda: Panpulmonata:
700 Stylommatophora: Cerionidae). *Mitochondrial DNA Part B* 2:159–160. DOI:
701 10.1080/23802359.2017.1303343.

Hatzoglou E., Rodakis GC., Lecanidou R. 1995. Complete sequence and gene organization of the mitochondrial genome of the land snail *Albinaria coerulea*. *Genetics* 140:1353–1366.

Hausdorf B. 2000. Biogeography of the Limacoidea sensu lato (Gastropoda: Stylommatophora): vicariance events and long-distance dispersal. *Journal of Biogeography* 27:379–390. DOI: 10.1046/j.1365-2699.2000.00403.x.

He Z-P., Dai X-B., Zhang S., Zhi T-T., Lun Z-R., Wu Z-D., Yang T-B. 2016. Complete mitochondrial genome of the giant African snail, *Achatina fulica* (Mollusca: Achatinidae): a novel location of putative control regions (CR) in the mitogenome within Pulmonate species. *Mitochondrial DNA* 27:1084–1085. DOI: 10.3109/19401736.2014.930833.

Holznagel WE., Colgan DJ., Lydeard C. 2010. Pulmonate phylogeny based on 28S rRNA gene sequences: A framework for discussing habitat transitions and character transformation. *Molecular Phylogenetics and Evolution* 57:1017–1025. DOI: 10.1016/j.ympev.2010.09.021.

Hu F., Lin Y., Tang J. 2014. MLGO: phylogeny reconstruction and ancestral inference from gene-order data. *BMC Bioinformatics* 15:354. DOI: 10.1186/s12859-014-0354-6.

Huang C-W., Lin S-M., Wu W-L. 2015. Mitochondrial genome sequences of landsnails *Aegista diversifamilia* and *Dolicheulota formosensis* (Gastropoda: Pulmonata: Stylommatophora). *Mitochondrial DNA*:1–3. DOI: 10.3109/19401736.2015.1053070.

722 Hunter SS., Settles ML., New DD., Parent CE., Gerritsen AT. 2016. Mitochondrial Genome
723 Sequence of the Galápagos Endemic Land Snail *Naesiotus nux*. *Genome*
724 *Announcements* 4. DOI: 10.1128/genomeA.01362-15.

725 Jörger KM., Stöger I., Kano Y., Fukuda H., Knebelsberger T., Schrödl M. 2010. On the origin
726 of Acochlidia and other enigmatic euthyneuran gastropods, with implications for the
727 systematics of Heterobranchia. *BMC Evolutionary Biology*. DOI: 10.1186/1471-2148-
728 10-323.

729 Kano Y., Brenzinger B., Nützel A., Wilson NG., Schrödl M. 2016. Ringiculid bubble snails
730 recovered as the sister group to sea slugs (Nudipleura). *Scientific Reports* 6:30908. DOI:
731 10.1038/srep30908.

732 Katoh K., Standley DM. 2013. MAFFT Multiple Sequence Alignment Software Version 7:
733 Improvements in Performance and Usability. *Molecular Biology and Evolution* 30:772–
734 780. DOI: 10.1093/molbev/mst010.

735 Kearse M., Moir R., Wilson A., Stones-Havas S., Cheung M., Sturrock S., Buxton S., Cooper
736 A., Markowitz S., Duran C., Thierer T., Ashton B., Meintjes P., Drummond A. 2012.
737 Geneious Basic: an integrated and extendable desktop software platform for the
738 organization and analysis of sequence data. *Bioinformatics* (Oxford, England) 28:1647–
739 9. DOI: 10.1093/bioinformatics/bts199.

740 Korábek O., Petrusek A., Rovatsos M. 2019. The complete mitogenome of *Helix pomatia* and
741 the basal phylogeny of Helicinae (Gastropoda, Stylommatophora, Helicidae). *ZooKeys*
742 827:19–30. DOI: 10.3897/zookeys.827.33057.

743 Kumar S., Stecher G., Tamura K. 2016. MEGA7: Molecular Evolutionary Genetics Analysis
744 Version 7.0 for Bigger Datasets. *Molecular Biology and Evolution* 33:1870–4. DOI:
745 10.1093/molbev/msw054.

746 Lanfear R., Calcott B., Ho SYW., Guindon S. 2012. PartitionFinder: Combined selection of
747 partitioning schemes and substitution models for phylogenetic analyses. *Molecular*
748 *Biology and Evolution* 29:1695–1701. DOI: 10.1093/molbev/mss020.

749 Laslett D., Canbäck B. 2008. ARWEN: A program to detect tRNA genes in metazoan
750 mitochondrial nucleotide sequences. *Bioinformatics* 24:172–175. DOI:
751 10.1093/bioinformatics/btm573.

752 Li X-D., Jiang G-F., Yan L-Y., Li R., Mu Y., Deng W-A. 2018. Positive Selection Drove the
753 Adaptation of Mitochondrial Genes to the Demands of Flight and High-Altitude
754 Environments in Grasshoppers. *Frontiers in Genetics* 9. DOI:
755 10.3389/fgene.2018.00605.

756 Lin J-H., Zhou W., Ding H-L., Wang P., Ai H-M. 2016. The mitochondrial genome of the land
757 snail *Cernuella virgata* (Da Costa, 1778): the first complete sequence in the family
758 Hygromiidae (Pulmonata, Stylommatophora). *ZooKeys* 589:55–69. DOI:
759 10.3897/zookeys.589.7637.

760 Liu G-H., Wang S-Y., Huang W-Y., Zhao G-H., Wei S-J., Song H-Q., Xu M-J., Lin R-Q.,
761 Zhou D-H., Zhu X-Q. 2012. The Complete Mitochondrial Genome of *Galba pervia*
762 (Gastropoda: Mollusca), an Intermediate Host Snail of *Fasciola* spp. *PLoS ONE*
763 7:e42172. DOI: 10.1371/journal.pone.0042172.

764 Meiklejohn CD., Montooth KL., Rand DM. 2007. Positive and negative selection on the
765 mitochondrial genome. *Trends in Genetics*: TIG 23:259–63. DOI:
766 10.1016/j.tig.2007.03.008.

767 Melo-Ferreira J., Vilela J., Fonseca MM., Da Fonseca RR., Boursot P., Alves PC. 2014. The
768 elusive nature of adaptive mitochondrial DNA evolution of an arctic lineage prone to
769 frequent introgression. *Genome Biology and Evolution* 6:886–896. DOI:
770 10.1093/gbe/evu059.

771 Meng G., Li Y., Yang C., Liu S. 2019. MitoZ: a toolkit for animal mitochondrial genome
772 assembly, annotation and visualization. *Nucleic Acids Research*. DOI:
773 10.1093/nar/gkz173.

774 Minton RL., Martinez Cruz MA., Farman ML., Perez KE. 2016. Two complete mitochondrial
775 genomes from *Praticolella mexicana* Perez, 2011 (Polygyridae) and gene order
776 evolution in Helicoidea (Mollusca, Gastropoda). *ZooKeys* 626:137–154. DOI:
777 10.3897/zookeys.626.9633.

778 Morales HE., Pavlova A., Joseph L., Sunnucks P. 2015. Positive and purifying selection in
779 mitochondrial genomes of a bird with mitonuclear discordance. *Molecular Ecology*
780 24:2820–2837. DOI: 10.1111/mec.13203.

781 Mordan P., Wade C. 2008. Heterobranchia II: The Pulmonata. In: *Phylogeny and Evolution of the*
782 *Mollusca*. University of California Press, 409–426. DOI:
783 10.1525/california/9780520250925.003.0015.

784 Moritz C., Dowling TE., Brown WM. 1987. Evolution of Animal Mitochondrial DNA:
785 Relevance for Population Biology and Systematics. *Annual Review of Ecology and*
786 *Systematics* 18:269–292. DOI: 10.1146/annurev.es.18.110187.001413.

787 Murrell B., Moola S., Mabona A., Weighill T., Sheward D., Kosakovsky Pond SL., Scheffler
788 K. 2013. FUBAR: A fast, unconstrained Bayesian approximation for inferring selection.
789 *Molecular Biology and Evolutio*n 30:1196–1205. DOI: 10.1093/molbev/mst030.

790 Murrell B., Wertheim JO., Moola S., Weighill T., Scheffler K., Kosakovsky Pond SL. 2012.
791 Detecting individual sites subject to episodic diversifying selection. *PLoS Genetics*
792 8:e1002764. DOI: 10.1371/journal.pgen.1002764.

793 Ojala D., Montoya J., Attardi G. 1981. tRNA punctuation model of RNA processing in human
794 mitochondria. *Nature* 290:470–474. DOI: 10.1038/290470a0.

795 Osca D., Templado J., Zardoya R. 2014. The mitochondrial genome of *Ifremeria nautilei* and
796 the phylogenetic position of the enigmatic deep-sea Abyssochrysoidea (Mollusca:
797 Gastropoda). *Gene* 547:257–266. DOI: 10.1016/j.gene.2014.06.040.

798 Perna NT., Kocher TD. 1995. Patterns of nucleotide composition at fourfold degenerate sites of
799 animal mitochondrial genomes. *Journal of Molecular Evolution* 41:353–358. DOI:
800 10.1007/BF01215182.

801 Ponder WF., Lindberg DR. 1997. Towards a phylogeny of gastropod molluscs: an analysis
802 using morphological characters. *Zoological Journal of the Linnean Society* 119:83–265.
803 DOI: 10.1006/zjls.1996.0066.

804 Price MR., Forsman ZH., Knapp I., Hadfield MG., Toonen RJ. 2016a. The complete
805 mitochondrial genome of *Achatinella mustelina* (Gastropoda: Pulmonata:

806 Stylommatophora). *Mitochondrial DNA Part B* 1:175–177. DOI:
807 10.1080/23802359.2016.1149787.

808 Price MR., Forsman ZH., Knapp IS., Toonen RJ., Hadfield MG. 2016b. The complete
809 mitochondrial genome of *Achatinella sowerbyana* (Gastropoda: Pulmonata:
810 Stylommatophora: Achatinellidae). *Mitochondrial DNA Part B* 1:666–668. DOI:
811 10.1080/23802359.2016.1219631.

812 Price MR., Forsman ZH., Knapp I., Toonen RJ., Hadfield MG. 2018. A comparison of
813 mitochondrial genomes from five species in three genera suggests polyphyly in the
814 subfamily Achatinellinae (Gastropoda: Pulmonata: Stylommatophora: Achatinellidae).
815 *Mitochondrial DNA Part B* 3:611–612. DOI: 10.1080/23802359.2018.1473737.

816 Rambaut A. 2012. FigTree v1. 4.0. A graphical viewer of phylogenetic trees. Institute of
817 Evolutionary Biology University of Edinburgh.

818 Rambaut A., Drummond AJ., Xie D., Baele G., Suchard MA. 2018. Posterior Summarization in
819 Bayesian Phylogenetics Using Tracer 1.7. *Systematic Biology* 67:901–904. DOI:
820 10.1093/sysbio/syy032.

821 Rand DM. 2001. The units of selection on mitochondrial DNA. *Annual Review of Ecology and*
822 *Systematics* 32:415–448. DOI: 10.1146/annurev.ecolsys.32.081501.114109.

823 Razkin O., Gómez-Moliner BJ., Prieto CE., Martínez-Ortí A., Arrébola JR., Muñoz B., Chueca
824 LJ., Madeira MJ. 2015. Molecular phylogeny of the western Palaearctic Helicoidea
825 (Gastropoda, Stylommatophora). *Molecular Phylogenetics and Evolution*. DOI:
826 10.1016/j.ympev.2014.11.014.

827 Reyes A., Gissi C., Pesole G., Saccone C. 1998. Asymmetrical directional mutation pressure in
828 the mitochondrial genome of mammals. *Molecular Biology and Evolution*. DOI:
829 10.1093/oxfordjournals.molbev.a026011.

830 Rice P., Longden L., Bleasby A. 2000. EMBOSS: The European Molecular Biology Open
831 Software Suite. *Trends in Genetics*. DOI: 10.1016/S0168-9525(00)02024-2.

832 De Rijk P., Wuyts J., De Wachter R. 2003. RnaViz 2: an improved representation of RNA
833 secondary structure. *Bioinformatics* (Oxford, England) 19:299–300.

834 Romero PE., Weigand AM., Pfenninger M. 2016. Positive selection on panpulmonate
835 mitogenomes provide new clues on adaptations to terrestrial life. *BMC Evolutionary*
836 *Biology* 16:164. DOI: 10.1186/s12862-016-0735-8.

837 Ronquist F., Teslenko M., van der Mark P., Ayres DL., Darling A., Höhna S., Larget B., Liu
838 L., Suchard MA., Huelsenbeck JP. 2012. MrBayes 3.2: Efficient Bayesian Phylogenetic
839 Inference and Model Choice Across a Large Model Space. *Systematic Biology* 61:539–
840 542. DOI: 10.1093/sysbio/sys029.

841 Roth B., Poinar GO., Acra A., Acra F. 1996. Probable Pupillid land snail of Early Cretaceous
842 (Hauterivian) age in amber from Lebanon. *Veliger* 39:87–88.

843 Sevigny JL., Kirouac LE., Thomas WK., Ramsdell JS., Lawlor KE., Sharifi O., Grewal S.,
844 Baysdorfer C., Curr K., Naimie AA., Okamoto K., Murray JA., Newcomb JM. 2015.
845 The mitochondrial genomes of the Nudibranch mollusks, *Melibe leonina* and *Tritonia*
846 *diomedea*, and their impact on gastropod phylogeny. *PLoS ONE* 10:e0127519. DOI:
847 10.1371/journal.pone.0127519.

848 Sharma L., Lu J., Bai Y. 2009. Mitochondrial Respiratory Complex I: Structure, Function and
849 Implication in Human Diseases. *Current Medicinal Chemistry*. DOI:
850 10.2174/092986709787846578.

851 Smith MD., Wertheim JO., Weaver S., Murrell B., Scheffler K., Kosakovsky Pond SL. 2015.
852 Less is more: An adaptive branch-site random effects model for efficient detection of
853 episodic diversifying selection. *Molecular Biology and Evolution* 32:1342–1353. DOI:
854 10.1093/molbev/msv022.

855 Solem A., Yochelson EL. 1979. North American Paleozoic land snails, with a summary of
856 other Paleozoic non-marine snails. *Geological Survey Professional Paper*.

857 Song L., Florea L., Langmead B. 2014. Lighter: fast and memory-efficient sequencing error
858 correction without counting. *Genome Biology* 15:509. DOI: 10.1186/s13059-014-0509-
859 9.

860 Stamatakis A. 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of
861 large phylogenies. *Bioinformatics* 30:1312–1313. DOI: 10.1093/bioinformatics/btu033.

862 Stöger I., Schrödl M. 2013. Mitogenomics does not resolve deep molluscan relationships
863 (yet?). *Molecular Phylogenetics and Evolution* 69:376–392. DOI:
864 10.1016/j.ympev.2012.11.017.

865 Sun BN., Wei LL., Shen HD., Wu HX., Wang DF. 2016. Phylogenetic analysis of euthyneuran
866 gastropods from sea to land mainly based on comparative mitogenomic of four species
867 of Onchidiidae (Mollusca: Gastropoda: Pulmonata). *Mitochondrial DNA Part A*
868 27:3075–3077. DOI: 10.3109/19401736.2014.1003916.

869 Swofford DL. 2002. Phylogenetic Analysis Using Parsimony (*and Other Methods). Version 4.
870 Sunderland, Massachusetts, USA: Sinauer Associates.

871 Tillier S., Masselmot M., Tillirt A. 1996. Phylogenetic relationships of the Pulmonate
872 Gastropods from rRNA sequences, and tempo and age of the Stylommatophoran
873 radiation. In: Taylor JD ed. Origin and Evolutionary Radiation of the Mollusca. *Oxford:*
874 *Oxford Press*, 267–284.

875 Tolley KA., Chase BM., Forest F. 2008. Speciation and radiations track climate transitions
876 since the Miocene Climatic Optimum: a case study of southern African chameleons.
877 *Journal of Biogeography* 35:1402–1414. DOI: 10.1111/j.1365-2699.2008.01889.x.

878 Uribe JE., Colgan D., Castro LR., Kano Y., Zardoya R. 2016a. Phylogenetic relationships
879 among superfamilies of Neritimorpha (Mollusca: Gastropoda). *Molecular Phylogenetics*
880 *and Evolution* 104:21–31. DOI: 10.1016/j.ympev.2016.07.021.

881 Uribe JE., Kano Y., Templado J., Zardoya R. 2016b. Mitogenomics of Vetigastropoda: insights
882 into the evolution of pallial symmetry. *Zoologica Scripta* 45:145–159. DOI:
883 10.1111/zsc.12146.

884 Uribe JE., Williams ST., Templado J., Abalde S., Zardoya R. 2017. Denser mitogenomic
885 sampling improves resolution of the phylogeny of the superfamily Trochoidea
886 (Gastropoda: Vetigastropoda). *Journal of Molluscan Studies* 83:111–118. DOI:
887 10.1093/mollus/eyw049.

888 Uribe JE., Zardoya R., Puillandre N. 2018. Phylogenetic relationships of the conoidean snails
889 (Gastropoda: Caenogastropoda) based on mitochondrial genomes. *Molecular*
890 *Phylogenetics and Evolution* 127:898–906. DOI: 10.1016/j.ympev.2018.06.037.

891 Vaidya G., Lohman DJ., Meier R. 2011. SequenceMatrix: concatenation software for the fast
892 assembly of multi-gene datasets with character set and codon information. *Cladistics*
893 27:171–180. DOI: 10.1111/j.1096-0031.2010.00329.x.

894 Vicentini A., Barber JC., Aliscioni SS., Giussani LM., Kellogg EA. 2008. The age of the
895 grasses and clusters of origins of C4 photosynthesis. *Global Change Biology*. DOI:
896 10.1111/j.1365-2486.2008.01688.x.

897 Wade CM., Mordan PB., Clarke B. 2001. A phylogeny of the land snails (Gastropoda:
898 Pulmonata). *Proceedings of the Royal Society of London. Series B: Biological Sciences*
899 268:413–422. DOI: 10.1098/rspb.2000.1372.

900 Wade CM., Mordan PB., Naggs F. 2006. Evolutionary relationships among the Pulmonate land
901 snails and slugs (Pulmonata, Stylommatophora). *Biological Journal of the Linnean*
902 *Society*. DOI: 10.1111/j.1095-8312.2006.00596.x.

903 Wang P., Yang H., Zhou W., Hwang C., Zhang W., Qian Z. 2014. The mitochondrial genome
904 of the land snail *Camaena cicatricosa* (Müller, 1774) (Stylommatophora, Camaenidae):
905 the first complete sequence in the family Camaenidae. *ZooKeys* 451:33–48. DOI:
906 10.3897/zookeys.451.8537.

907 Weaver S., Shank SD., Spielman SJ., Li M., Muse S V., Kosakovsky Pond SL. 2018.
908 Datamonkey 2.0: A modern web application for characterizing selective and other
909 evolutionary processes. *Molecular Biology and Evolution* 35:773–777. DOI:
910 10.1093/molbev/msx335.

911 White TR., Conrad MM., Tseng R., Balayan S., Golding R., de Frias Martins AM., Dayrat BA.
912 2011. Ten new complete mitochondrial genomes of pulmonates (Mollusca: Gastropoda)

913 and their impact on phylogenetic relationships. *BMC Evolutionary Biology* 11:295.
914 DOI: 10.1186/1471-2148-11-295.

915 Williams ST., Foster PG., Littlewood DTJ. 2014. The complete mitochondrial genome of a
916 turbinid vetigastropod from MiSeq Illumina sequencing of genomic DNA and steps
917 towards a resolved gastropod phylogeny. *Gene* 533:38–47. DOI:
918 10.1016/j.gene.2013.10.005.

919 Wolstenholme DR. 1992. Animal Mitochondrial DNA: Structure and Evolution. *International*
920 *Review of Cytology* 141:173–216. DOI: 10.1016/S0074-7696(08)62066-5.

921 Woolley S., Johnson J., Smith MJ., Crandall KA., McClellan DA. 2003. TreeSAAP: Selection
922 on amino acid properties using phylogenetic trees. *Bioinformatics* 19:671–672. DOI:
923 10.1093/bioinformatics/btg043.

924 Xie J., Feng J., Guo Y., Ye Y., Li J., Guo B. 2019a. The complete mitochondrial genome and
925 phylogenetic analysis of *Nerita yoldii* (Gastropoda: Neritidae). *Mitochondrial DNA Part*
926 *B* 4:1099–1100. DOI: 10.1080/23802359.2019.1586485.

927 Xie G-L., Köhler F., Huang X-C., Wu R-W., Zhou C-H., Ouyang S., Wu X-P. 2019b. A novel
928 gene arrangement among the Stylommatophora by the complete mitochondrial genome
929 of the terrestrial slug *Meghimatium bilineatum* (Gastropoda, Arionoidea). Molecular
930 Phylogenetics and Evolution 135:177–184. DOI: 10.1016/j.ympev.2019.03.002.

931 Yamazaki N., Ueshima R., Terrett JA., Yokobori SI., Kaifu M., Segawa R., Kobayashi T.,
932 Numachi KI., Ueda T., Nishikawa K., Watanabe K., Thomas RH. 1997. Evolution of
933 pulmonate gastropod mitochondrial genomes: Comparisons of gene organizations of

934 *Euhadra*, *Cepaea* and *Albinaria* and implications of unusual tRNA secondary
935 structures. *Genetics* 145:749–758.

936 Yang Z. 2007. PAML 4: Phylogenetic Analysis by Maximum Likelihood. *Molecular Biology*
937 *and Evolution* 24:1586–1591. DOI: 10.1093/molbev/msm088.

938 Yang X., Xie G-L., Wu X-P., Ouyang S. 2016. The complete mitochondrial genome of Chinese
939 land snail *Aegista aubryana* (Gastropoda: Pulmonata: Bradybaenidae). *Mitochondrial*
940 *DNA Part A* 27:3538–3539. DOI: 10.3109/19401736.2015.1074207.

941 Yang T., Xu G., Gu B., Shi Y., Mzuka HL., Shen H. 2019. The Complete Mitochondrial
942 Genome Sequences of the *Philomycus bilineatus* (Stylommatophora: Philomycidae) and
943 Phylogenetic Analysis. *Genes* 10:198. DOI: 10.3390/genes10030198.

944 Zapata F., Wilson NG., Howison M., Andrade SCS., Jörger KM., Schrödl M., Goetz FE.,
945 Giribet G., Dunn CW. 2014. Phylogenomic analyses of deep gastropod relationships
946 reject Orthogastropoda. *Proceedings of the Royal Society B: Biological Sciences*
947 281:20141739. DOI: 10.1098/rspb.2014.1739.

948 Zuker M. 2003. Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic*
949 *Acids Research*.

950

951

952

953

954 **Figure Legends**

955 **Figure 1. Mitogenome organization of *Arion vulgaris*.** Genes transcribed from the J and N
956 strands are shown outside and inside of the circle, respectively. PCGs coding complex I,
957 complex III, complex IV and complex V components are marked with yellow, purple, pink and
958 green, respectively. rRNA genes are coloured with red and putative control region is coloured
959 with cyan, while tRNA genes are coloured with dark blue and labelled by the single letter
960 amino acid code.

961 **Figure 2. Relative synonymous codon usage (RSCU) of the *A. vulgaris* and *A. rufus***
962 **mitogenomes.** Codon families are provided on the x axis. The stop codons are not given.

963 **Figure 3**. **Predicted secondary structure of putative control region of *A. vulgaris***
964 **mitogenome.** Nucleotides are coloured as follows: Adenine is green, thymine is red, cytosine is
965 blue and guanine is black. The poly-T stretch is labelled with purple.

966 **Figure 4. Stylommatophoran phylogenetic tree constructed under BI using the dataset**
967 **8P12RNA.** *Carychium tridentatum* (Ellobioidea), *Platevindex mortoni* (Systellommatophora)
968 and *Galba pervia* (Hygrophila) were used as outgroup. Nodes are labelled with numbers refer
969 to hypothetical ancestral mitogenome organizations inferred by MLGO.

970 **Figure 5. Dated phylogenetic tree.** The axis on the bottom refers to million years. Letters in
971 the boxes refers to external calibration points. The split between *A. vulgaris* and *A. rufus* was
972 shown with a red bar, while the remainings were shown with blue bars.

Figure 1

Mitogenome organization of *Arion vulgaris*.

Genes transcribed from the J and N strands are shown outside and inside of the circle, respectively. PCGs coding complex I, complex III, complex IV and complex V components are marked with yellow, purple, pink and green, respectively. rRNA genes are coloured with red and putative control region is coloured with cyan, while tRNA genes are coloured with dark blue and labelled by the single letter amino acid code.

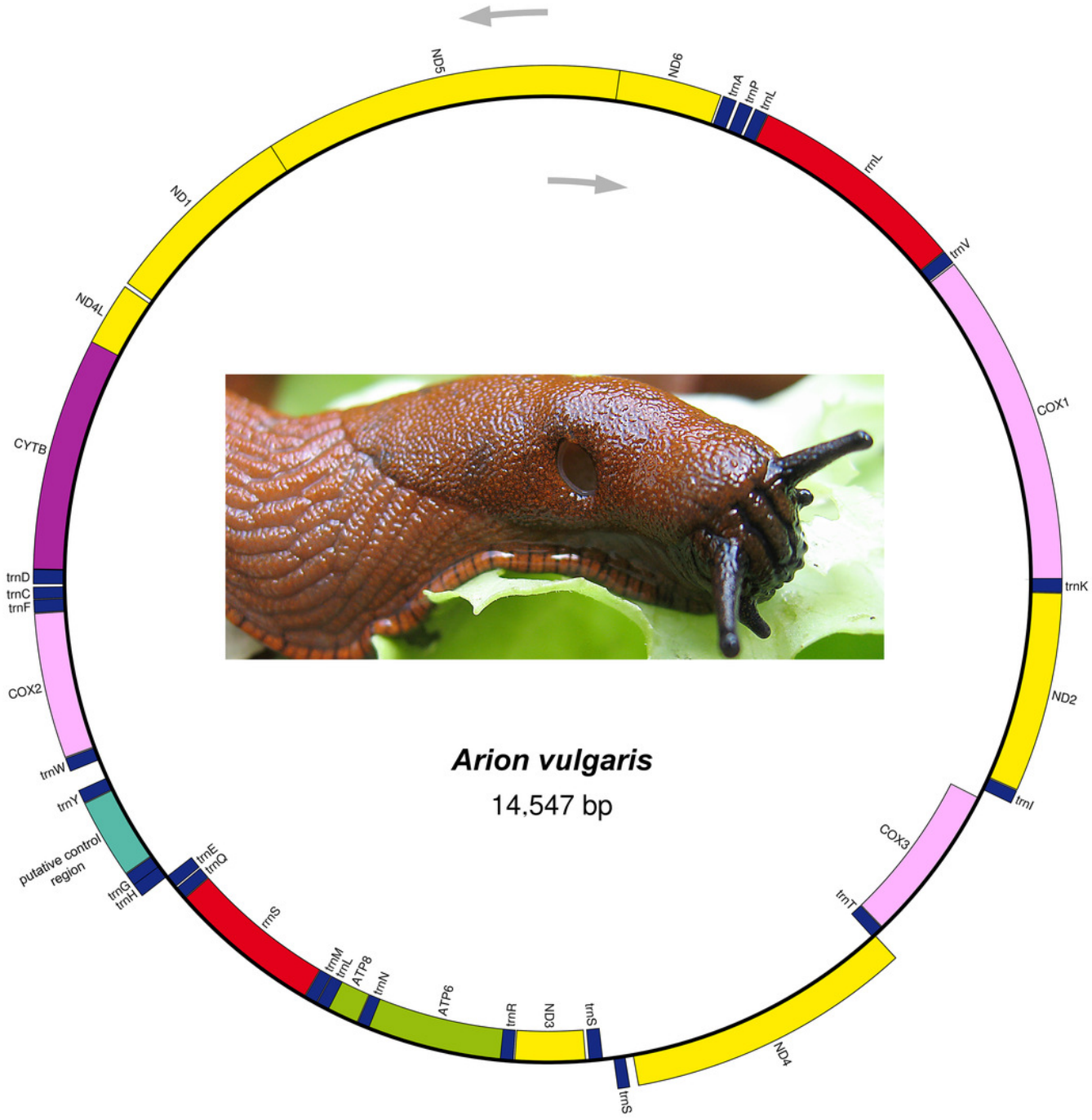
ND5
ND6
trnA
trnP
trnL
rrnL
trnV
COX1
trnK
ND2
trnI
COX3
trnT
ND4
trnS
ND3
trnR
ATP6
ATP8
trnN
trnM
trnL
rrnS
trnE
trnQ
trnG
trnH
putative control region
trnY
trnW
COX2
trnD
trnC
trnF
CYTB
ND4L
ND1
Arion vulgaris
14,547 bp

Figure 2

Relative synonymous codon usage (RSCU) of the *A. vulgaris* and *A. rufus* mitogenomes.

Codon families are provided on the x axis. The stop codons are not given.

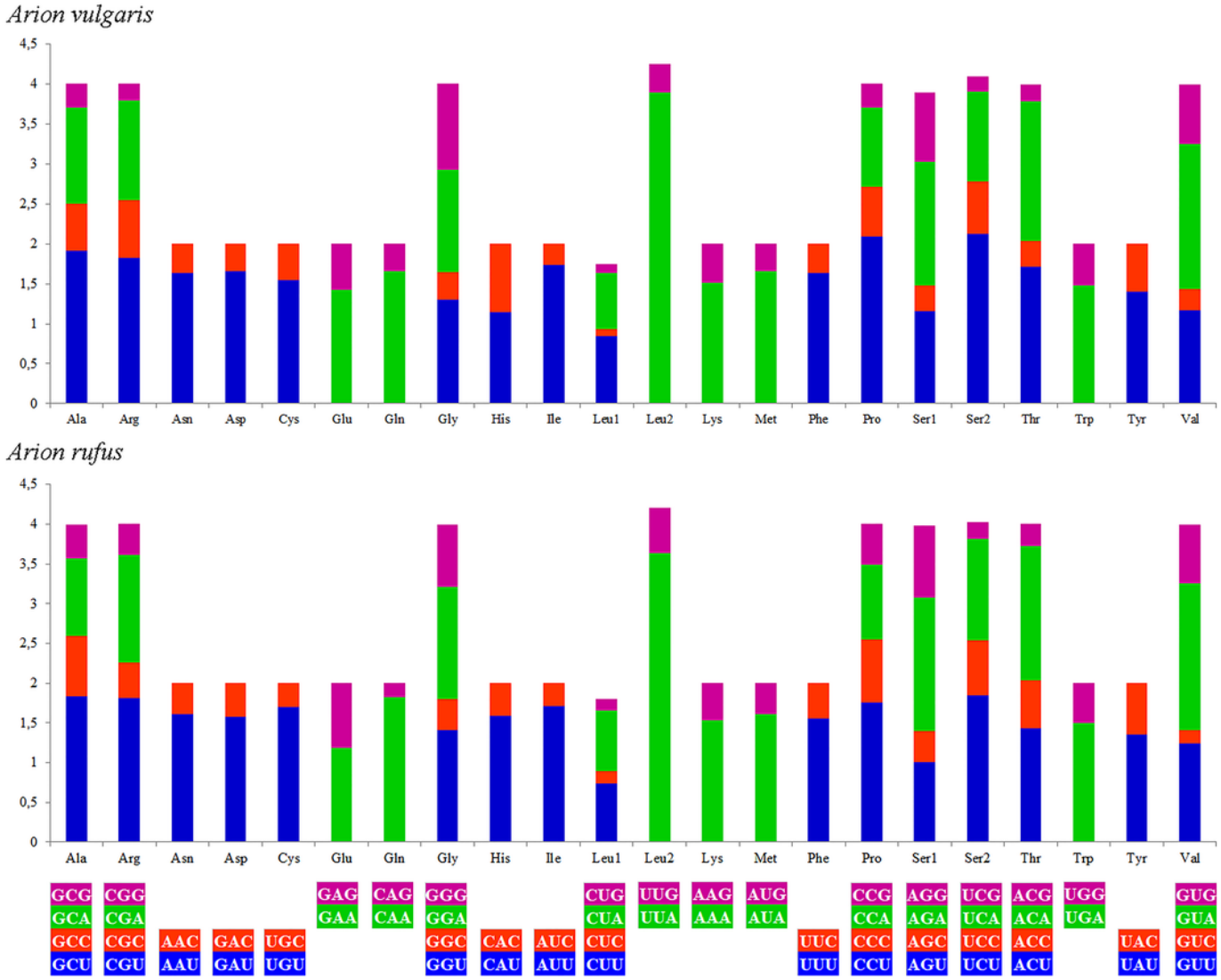

Figure 3

Predicted secondary structure of putative control region of *A. vulgaris* mitogenome.

Nucleotides are coloured as follows: Adenine is green, thymine is red, cytosine is blue and guanine is black. The poly-T stretch is labelled with purple.

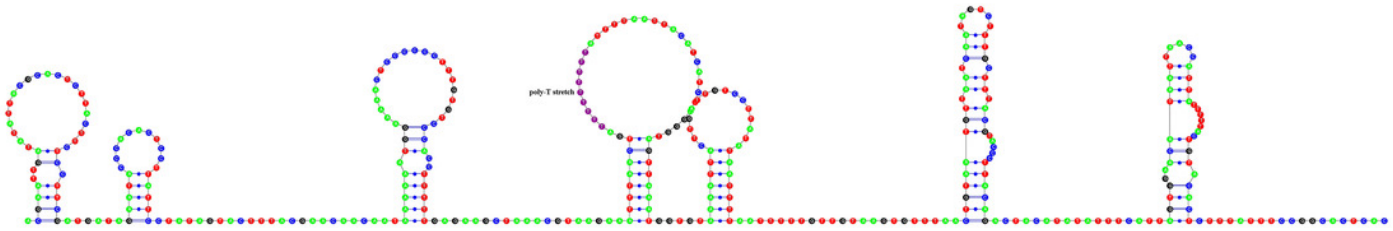

Figure 4

Stylommatophoran phylogenetic tree constructed under BI using the dataset 8P12RNA.

Carychium tridentatum (Ellobioidea), *Platevindex mortoni* (Systellommatophora) and *Galba pervia* (Hygrophila) were used as outgroup. Nodes are labelled with numbers refer to hypothetical ancestral mitogenome organizations inferred by MLGO.

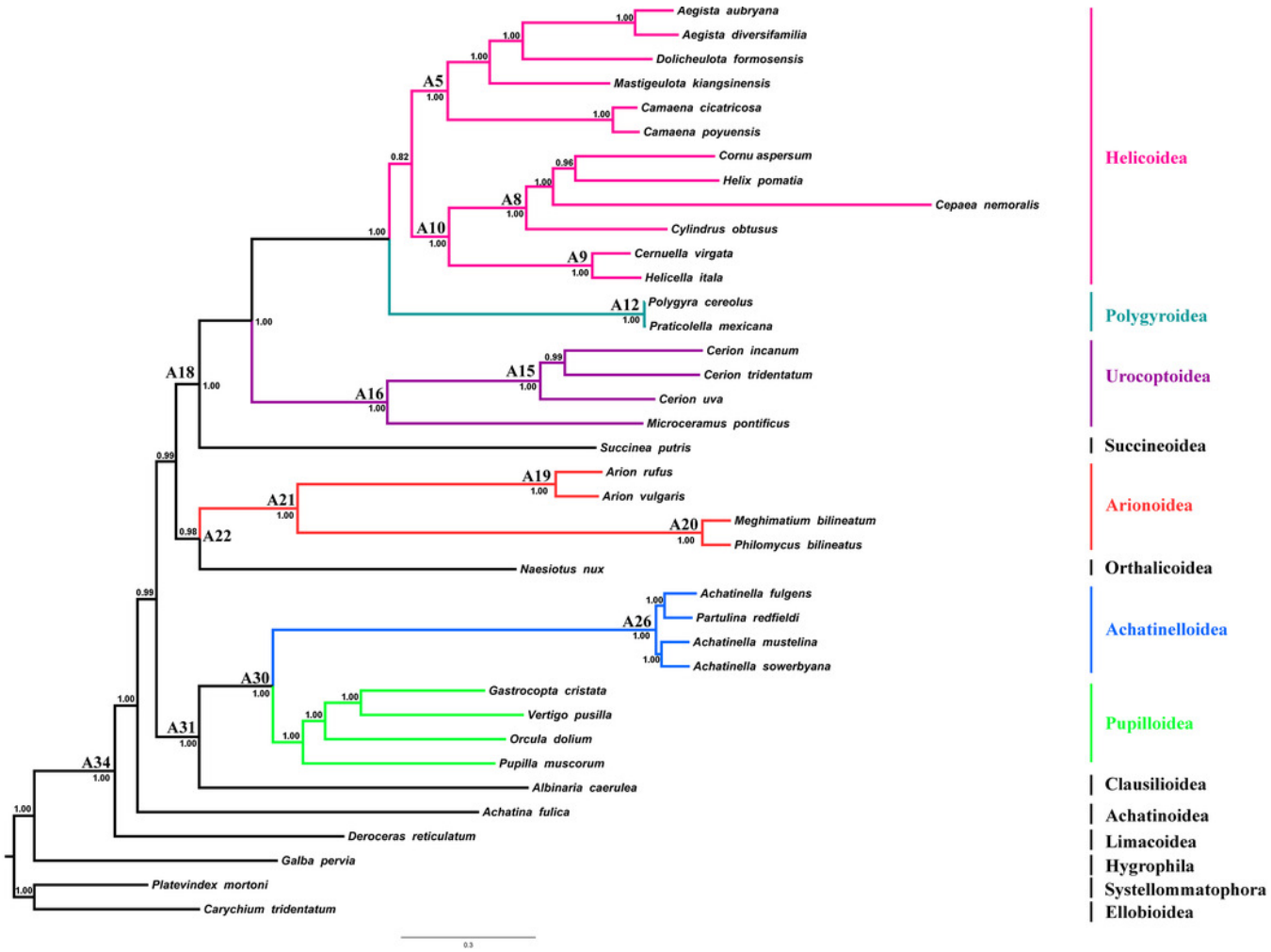

Figure 5

Dated phylogenetic tree.

The axis on the bottom refers to million years. Letters in the boxes refers to external calibration points. The split between *A. vulgaris* and *A. rufus* was shown with a red bar, while the remainings were shown with blue bars.

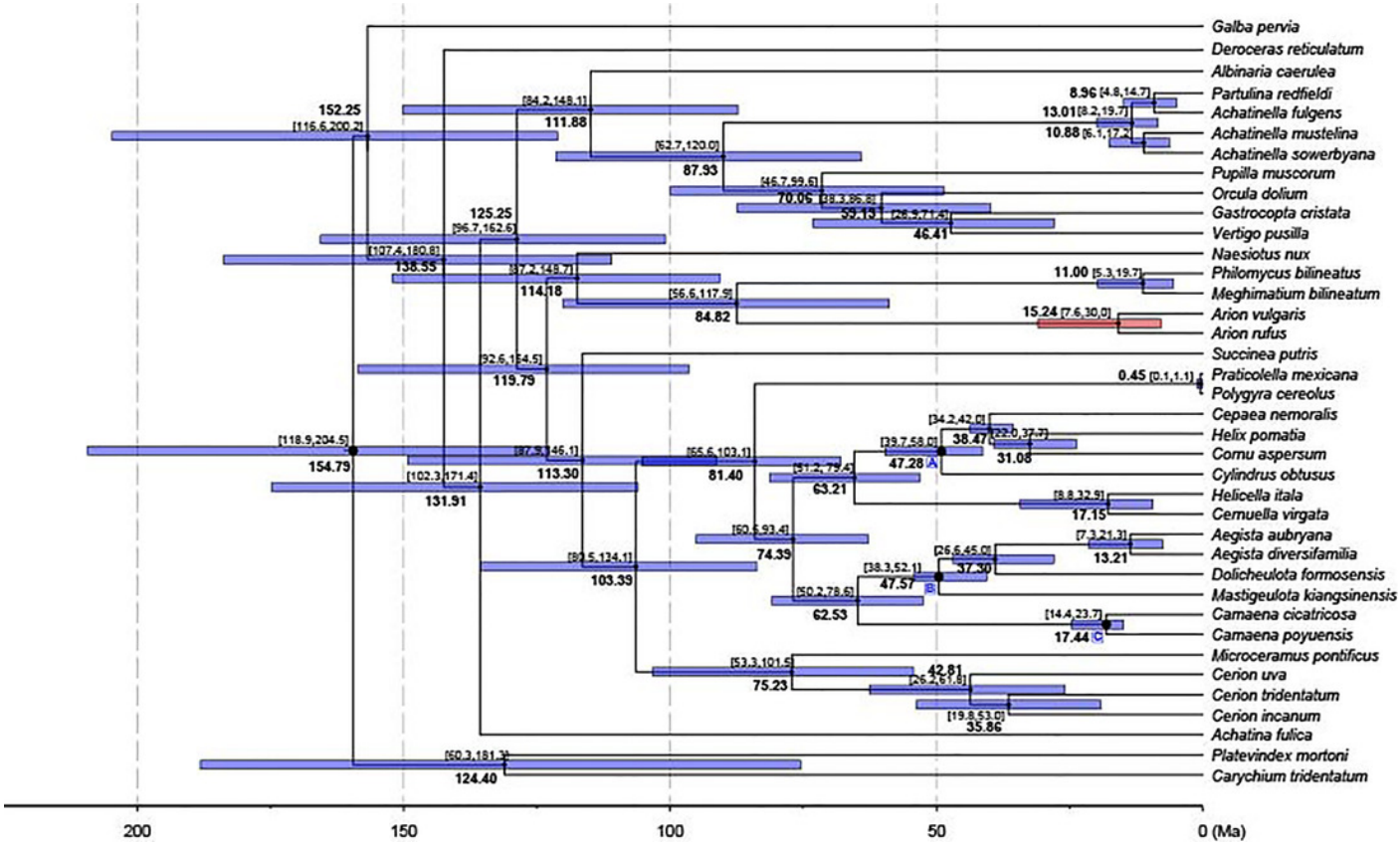

Table 1(on next page)

List of stylommatophoran mitogenomes used in phylogenetic and comparative analyses

Table 1. List of stylommatophoran mitogenomes used in phylogenetic and comparative analyses

	Species	Family	Accession number	References
	Arion vulgaris	Arionidae	MN607980	This study
	Arion rufus	Arionidae	KT626607	Romero et al. (2016)
	Achatinella fulgens	Achatinellidae	MG925058	Price et al., (2018)
	Achatinella mustelina	Achatinellidae	NC030190	Price et al., (2016a)
	Achatinella sowerbyana	Achatinellidae	KX356680	Price et al., (2016b)
	Partulina redfieldi	Achatinellidae	MG925057	Price et al., (2018)
	Achatina fulica	Achatinidae	KM114610	He et al., (2016)
	Deroceras reticulatum	Agriolimacidae	NC035495	Ahn et al., (2017)
	Aegista aubryana	Bradybaenidae	NC029419	Yang et al., (2016)
	Aegista diversifamilia	Bradybaenidae	NC027584	Huang, Lin & Wu, (2015)
	Dolicheulota formosensis	Bradybaenidae	NC027493	Huang, Lin & Wu, (2015)
	Mastigeulota kiangsinensis	Bradybaenidae	NC024935	Deng et al., (2016)
	Camaena cicatricosa	Camaenidae	NC025511	Wang et al., (2014)
	Camaena poyuensis	Camaenidae	KT001074	Unpublished
	Cerion incanum	Cerionidae	NC025645	González et al., (2016)
	Cerion tridentatum costellata	Cerionidae	KY249249	Unpublished
	Cerion uva	Cerionidae	KY124261	Harasewych et al., (2017)
	Albinaria caerulea	Clausiliidae	NC001761	Hatzoglou et al., (1995)
	Gastrocopta cristata	Gastrocoptidae	NC026043	Unpublished
	Cernuella virgata	Geomitridae	NC030723	Lin et al., (2016)
	Helicella itala	Geomitridae	KT696546	Romero et al. (2016)
	Cepaea nemoralis	Helicidae	NC001816	Yamazaki et al., (1997)
	Cylindrus obtusus	Helicidae	NC017872	Groenenberg et al., (2012)
	Cornu aspersum	Helicidae	NC021747	Gaitán-Espitia et al., (2013)
	Helix pomatia	Helicidae	NC041247	Korábek et al., (2019)
	Orcula dolium	Orculidae	NC034782	Groenenberg et al., (2017)
	Naesiotus nux	Orthalicidae	NC028553	Hunter et al., (2016)
	Meghimatium bilineatum	Philomycidae	NC035429	Xie et al., (2019)
	Philomycus bilineatus	Philomycidae	MG722906	Yang et al., (2019)
	Polygyra cereolus	Polygyridae	NC032036	Unpublished
	Praticolella mexicana	Polygyridae	KX240084	Minton et al., (2016)
	Pupilla muscorum	Pupillidae	NC026044	Unpublished
	Succinea putris	Succineidae	NC016190	White et al., (2011)
	Microceramus pontificus	Urocoptidae	NC036381	Unpublished
	Vertigo pusilla	Vertiginidae	NC026045	Unpublished
Ellobioidea	*Carychium tridentatum*	Ellobiidae	KT696545	Romero et al. (2016)
Hygrophila	*Galba pervia*	Lymnaeidae	NC018536	Liu et al., (2012)
Systellommatophora	*Platevindex mortoni*	Onchidiidae	GU475132	Sun et al., (2016)

Table 2(on next page)

Mitogenome summary of *Arion vulgaris*

Table 2. Mitogenome summary of *Arion vulgaris*

Gene	Strand	From	To	Size	Start codon	Stop codon	Anticodon	IGN
COX1	J	1	1530	1530	TTG	TAG		4
tRNA-Val	J	1535	1600	66			UAC	0
16S rRNA	J	1601	2613	1013				0
tRNA-Leu	J	2614	2675	62			UAG	11
tRNA-Pro	J	2687	2752	66			UGG	13
tRNA-Ala	J	2766	2831	66			UGC	7
ND6	J	2839	3312	474	ATG	TAG		-41
ND5	J	3272	4960	1689	ACA	TAA		-10
ND1	J	4951	5853	903	ATG	TAG		15
ND4L	J	5869	6163	295	ATA	T-		-15
CYTB	J	6149	7228	1080	ATG	TAA		-2
tRNA-Asp	J	7227	7296	70			GUC	10
tRNA-Cys	J	7307	7363	57			GCA	0
tRNA-Phe	J	7364	7425	62			GAA	0
COX2	J	7426	8094	669	ATG	TAG		1
tRNA-Trp	J	8096	8160	65			UCA	91
tRNA-Tyr	J	8252	8315	67			GUA	0
Control region	J	8316	8685	370				0
tRNA-Gly	J	8686	8763	78			UCC	-20
tRNA-His	J	8744	8809	66			GUG	-3
tRNA-Glu	N	8807	8873	67			UUC	5
tRNA-Gln	N	8879	8942	64			UUG	0
12S rRNA	N	8943	9689	747				0
tRNA-Met	N	9690	9754	65			CAU	0
tRNA-Leu	N	9755	9820	66			UAA	-32
ATP8	N	9789	9971	183	GTG	TAA		0
tRNA-Asn	N	9972	10033	62			GUU	-8
ATP6	N	10026	10688	663	ATA	TAA		-9
tRNA-Arg	N	10680	10746	67			UCG	3
ND3	N	10750	11094	345	ATG	TAA		13
tRNA-Ser2	N	11108	11176	69			UGA	49
tRNA-Ser1	J	11226	11283	58			GCU	36
ND4	J	11320	12633	1314	ATA	TAG		-18
tRNA-Thr	N	12616	12681	66			UGU	0
COX3	N	12682	13462	781	ATG	T-		41
tRNA-Ile	J	13504	13567	64			GAU	1
ND2	J	13569	14486	918	ATG	TAA		0
tRNA-Lys	J	14487	6	67			UUU	-6

J. major; N. minor; IGN, intergenic nucleotides. Minus indicates overlapping sequences between adjacent genes.

4

Table 3(on next page)

Nucleotide composition of the *Arion vulgaris* mitogenome

Table 3. Nucleotide composition of the *Arion vulgaris* mitogenome

Feature	T%	C%	A%	G%	A+T%	AT-skew	GC-skew
Whole mitogenome	37.75	14.26	32.45	15.54	70.20	-0.076	0.043
Protein coding genes	39.90	14.61	29.44	16.05	69.34	-0.151	0.047
First codon position	33.54	14.01	30.64	21.81	64.18	-0.045	0.218
Second codon position	45.65	19.42	18.55	16.38	64.21	-0.422	-0.085
Third codon position	40.50	10.39	39.14	9.97	79.64	-0.017	-0.021
Protein coding genes-J	39.78	14.34	29.72	16.16	69.50	-0.145	0.060
First codon position-J	32.94	13.93	31.23	21.90	64.17	-0.027	0.222
Second codon position-J	45.91	19.07	18.53	16.49	64.44	-0.425	-0.073
Third codon position-J	40.50	10.01	39.41	10.08	79.90	-0.014	0.003
Protein coding genes-N	40.42	15.80	28.19	15.60	68.60	-0.178	-0.006
First codon position-N	36.24	14.37	27.98	21.41	64.22	-0.129	0.197
Second codon position-N	44.50	20.95	18.65	15.90	63.15	-0.409	-0.137
Third codon position-N	40.52	12.08	37.92	9.48	78.44	-0.033	-0.121
tRNA genes	36.40	11.48	36.33	15.80	72.72	-0.001	0.158
rRNA genes	33.24	13.58	38.18	15.00	71.42	0.069	0.050
Control region	38.92	18.65	30.81	11.62	69.73	-0.116	-0.232

Table 4(on next page)

Likelihood ratios of PAML analysis showing different selective pressures on the mitochondrial PCGs in Stylommatophora

Table 4. Likelihood ratios of PAML analysis showing different selective pressures on the mitochondrial PCGs in Stylommatophora

Models[a]		A		B		C		A-B		B-C	
Gene	ω	lnL[b]	Np[c]	lnL	Np	lnL	Np	LRT	P	LRT	P[d]
ATP6	0.0509	-19658.6890		-19377.7446		-19374.7961		561.8887	0.000	-5.896982	**0.015**
ATP8	0.2198	-6460.1406		-6355.5745		-6355.5745		209.1322	0.000	-0.000006	1.000
COX1	0.0129	-26428.6827		-26169.9332		-26169.9332		517.4990	0.000	0.000000	1.000
COX2	0.0393	-16022.1910		-15851.6973		-15848.6152		340.9874	0.000	-6.164250	**0.013**
COX3	0.0317	-18617.4781		-18345.8456		-18341.7284		543.2650	0.000	-8.234404	**0.004**
CYTB	0.0416	-26893.7004		-26297.6496		-26297.6496		1192.1017	0.000	0.000082	0.992
ND1	0.0430	-24857.6972	76	-24528.7292	78	-24529.0150	79	657.9359	0.000	0.571480	0.450
ND2	0.0670	-31769.5279		-31502.5684		-31498.5036		533.9191	0.000	-8.129668	**0.004**
ND3	0.0607	-11403.3678		-11183.4729		-11183.4729		439.7899	0.000	0.000010	1.000
ND4	0.0511	-40498.0727		-40032.1913		-40026.0861		931.7628	0.000	-12.210500	**0.000**
ND4L	0.0727	-10118.1495		-10076.4001		-10075.2708		83.4989	0.000	-2.258408	0.133
ND5	0.0676	-51970.2413		-51088.0681		-51092.3346		1764.3464	0.000	8.533012	**0.003**
ND6	0.1009	-16916.1956		-16691.7006		-16691.7006		448.9899	0.000	0.000018	1.000

Degrees of freedom = 1.

[a] A: All branches have one ω; B: All branches have same ω=1; C: Each branch has its own ω.

[b] The natural algorithm of the likelihood value.

[c] Number of parameters.

[d] Bold faced figure indicate the statistical significance ($P < 0.05$).

Table 5(on next page)

Genes and branches detected to be exposed episodic diversifying selection using the aBSREL approach.

Table 5. Genes and branches detected to be exposed episodic diversifying selection using the aBSREL approach.

Gene	Number of Selected Branches (P<0.05)	Taxon	ω	Proportion of Codons Under Selection
ATP8	1	*Microceramus pontificus*	288	0.460
COX1	1	*Achatinella mustelina*	2180	0.086
COX3	2	Arionoidea	670	0.053
		Philomycus bilineatus	119	0.092
ND3	1	*Helicella itala*	49.4	0.220
ND4	1	*Succinea putris*	4.18	0.370
ND6	1	*Vertigo pusilla*	15.6	0.370

Table 6(on next page)

Genes/codons under diversifying or positive selection under codon-based models.

Table 6. Genes/codons under diversifying or positive selection under codon-based models.

Gene	BEB	FUBAR	MEME
ATP6	-	4	44
ATP8	-	44	44, 57, 64, 92, 109
COX1	-	-	-
COX2	-	32	-
COX3	-	-	-
CYTB	-	12	12
ND1	-	-	-
ND2	188	-	14, 16, 174
ND3	-	-	27
ND4	109, 170, 192, 301, 386, 427	-	9, 99
ND4L	-	13, 57	13, 57, 109, 111
ND5	451	-	260, 501
ND6	-	-	109, 179, 183