

This Provisional PDF corresponds to the article as it appeared upon acceptance. Fully formatted PDF and full text (HTML) versions will be made available soon.

Evolution of gastropod mitochondrial genome arrangements

BMC Evolutionary Biology 2008, **8**:61 doi:10.1186/1471-2148-8-61

Cristina Grande (grande@berkeley.edu)
Jose Templado (mcnt150@mncn.csic.es)
Rafael Zardoya (rafaz@mncn.csic.es)

ISSN 1471-2148

Article type Research article

Submission date 12 June 2007

Acceptance date 26 February 2008

Publication date 26 February 2008

Article URL <http://www.biomedcentral.com/1471-2148/8/61>

Like all articles in BMC journals, this peer-reviewed article was published immediately upon acceptance. It can be downloaded, printed and distributed freely for any purposes (see copyright notice below).

Articles in BMC journals are listed in PubMed and archived at PubMed Central.

For information about publishing your research in BMC journals or any BioMed Central journal, go to

<http://www.biomedcentral.com/info/authors/>

Evolution of gastropod mitochondrial genome arrangements

Cristina Grande^{1,2}, José Templado¹, Rafael Zardoya¹

¹Departamento de Biodiversidad y Biología Evolutiva, Museo Nacional de Ciencias Naturales, CSIC, José Gutiérrez Abascal, 2, 28006, Madrid, Spain

²Department of Integrative Biology, University of California, Berkeley, CA, USA

Corresponding author:

Cristina Grande

Current address: Dept. of Integrative Biology, University of California, Berkeley, CA, USA

Phone: 510/643-6227

Email address:

CG: grande@berkeley.edu

JT: mcnt150@mncn.csic.es

RZ: rafaz@mncn.csic.es

Abstract

Background: Gastropod mitochondrial genomes exhibit an unusually great variety of gene orders compared to other metazoan mitochondrial genome such as e.g those of vertebrates. Hence, gastropod mitochondrial genomes constitute a good model system to study patterns, rates, and mechanisms of mitochondrial genome rearrangement. However, this kind of evolutionary comparative analysis requires a robust phylogenetic framework of the group under study, which has been elusive so far for gastropods in spite of the efforts carried out during the last two decades. Here, we report the complete nucleotide sequence of five mitochondrial genomes of gastropods (*Pyramidella dolabrata*, *Ascobulla fragilis*, *Siphonaria pectinata*, *Onchidella celtica*, and *Myosotella myosotis*), and we analyze them together with another ten complete mitochondrial genomes of gastropods currently available in molecular databases in order to reconstruct the phylogenetic relationships among the main lineages of gastropods.

Results: Comparative analyses with other mollusk mitochondrial genomes allowed us to describe molecular features and general trends in the evolution of mitochondrial genome organization in gastropods. Phylogenetic reconstruction with commonly used methods of phylogenetic inference (ME, MP, ML, BI) arrived at a single topology, which was used to reconstruct the evolution of mitochondrial gene rearrangements in the group.

Conclusions: Four main lineages were identified within gastropods: Caenogastropoda, Vetigastropoda, Patellogastropoda, and Heterobranchia. Caenogastropoda and Vetigastropoda are sister taxa, as well as, Patellogastropoda and Heterobranchia. This result rejects the validity of the derived clade Apogastropoda (Caenogastropoda + Heterobranchia). The position of Patellogastropoda remains unclear likely due to long-branch attraction biases. Within Heterobranchia, the most heterogeneous group of gastropods, neither Euthyneura (because of the inclusion of *P. dolabrata*) nor Pulmonata (polyphyletic) nor Opisthobranchia (because of the inclusion *S. pectinata*) were recovered as monophyletic groups. The gene order of the Vetigastropoda might represent the ancestral mitochondrial gene order for Gastropoda and we propose that at least three major rearrangements have taken place in the evolution of gastropods:

one in the ancestor of Caenogastropoda, another in the ancestor of Patellogastropoda, and one more in the ancestor of Heterobranchia.

Background

The animal mitochondrial (mt) genome is a circular double-stranded DNA molecule, which typically encodes for two rRNAs, 22 tRNAs, and 13 proteins that are essential for mitochondrial function [1].

The advent of long PCR and automated sequence techniques has recently simplified and accelerated the determination of complete sequences (about 16 Kb) of animal mtDNAs [2], and as of May 2007 there were more than 1000 complete metazoan mtDNAs deposited in GenBank [3].

Most (73%) of these sequences were from vertebrate mtDNAs, which show a relatively fixed genome organization with only few instances of gene rearrangements [e.g. 4, 5]. Instead, changes in gene order are widespread in non-vertebrate mt genomes, and particularly frequent in several phyla such as e.g. nematodes or mollusks [6]. Tandem duplication followed by random loss of redundant genes has been demonstrated to be the main mechanism for gene rearrangement of adjacent tRNAs in vertebrate mt genomes [5]. However, other mechanisms such as illegitimate recombination mediated by e.g. topoisomerases need to be invoked to explain observed gene inversions, as well as transpositions of genes to distant positions in non-vertebrate mt genomes [6]. In order to determine rates and mechanisms of gene rearrangement in metazoans, additional comparative analyses of non-vertebrate mt genome gene orders within an evolutionary framework are largely wanting.

On the other hand, mt genome arrangement comparisons may be useful for phylogenetic reconstruction [6, 7]. For these molecular markers, it is generally assumed that convergence is unlikely because of the great number of potential mt arrangements and the general low rate of rearrangements. Therefore, the presence of rare mtDNA gene orders shared by different taxa can be interpreted as a result of common ancestry. However, structural constraints and the reported existence of hotspots for gene rearrangement (e.g. near the origins of replication in vertebrate mt genomes, [5]) must increase the chances of convergence [5, 6]. Both Parsimony [8] and Bayesian [9, 10] methods are available to reconstruct phylogenies based on genome arrangement data. However, these methods have several limitations [9], and hitherto have not been extensively applied. Further studies based on larger genome arrangement data sets and more sophisticated

methods of inference are expected to confirm the potential of mt genome arrangements as phylogenetic markers, as well as their performance and levels of homoplasy.

Gastropod mollusk mt genomes present high diversity of gene orders [e.g. 11], and offer a suitable model system to study the rates and mechanisms of mt genome rearrangement, as well as the phylogenetic utility of arrangement comparisons. Thus far, ten complete gastropod mt genomes have been reported including *Albinaria coerulea* [12], *Cepaea nemoralis* [13], *Pupa strigosa* [11], *Roboastra europea* [14], *Biomphalaria glabrata* [15], *Haliotis rubra* [16], *Aplysia californica* [17], *Lottia digitalis* and *Ilyanassa obsoleta* [18], and *Lophiotoma cerithiformis* [19] (Fig. 1). In addition, the incomplete mt genomes of *Euhadra herklotsi* [20], *Littorina saxatilis* [21], and *Omalogyra atomus* [22] have been also described (Fig. 1). These mt genomes are all of relatively reduced size (13-15 Kb) except that of *L. digitalis*, which has two large non-coding regions that increase the total length of the mt genome up to 26 Kb [18].

Any meaningful evolutionary comparison between gastropod mt genome arrangements must rely explicitly on a robust phylogeny of these mollusks. However, phylogenetic relationships among extant groups of gastropods are the subject of a long-standing debate that lasts over a century [23-25]. The current classification of gastropods is a consensus of phylogenetic hypotheses proposed by several authors (for example [26-34]) during the last two decades, and generally includes six major groups: Patellogastropoda, Cocculiniformia, Neritopsina, Vetigastropoda, Caenogastropoda and Heterobranchia (Fig. 1). Although the monophyly of these groups is well supported based both on morphological and molecular evidence, phylogenetic relationships among them are still contentious. Morphological [26, 28] but not molecular data [30, 34] support Patellogastropoda (Eogastropoda) as the sister group of the remaining gastropods (Orthogastropoda) (Fig. 1).

Furthermore, Caenogastropoda and Heterobranchia are considered to be sister taxa (Apogastropoda) based both on morphological and molecular data [28, 34, 35] (Fig. 1).

Out of the six major gastropod groups, Heterobranchia seems to be the most heterogeneous one. This clade includes the paraphyletic Heterostropha (*Omalogyroidea*, *Pyramidelloidea*, and other smaller groups) [26, 28], and *Euthyneura*, a monophyletic group that includes the highly diversified *Opisthobranchia* and *Pulmonata* [24] (Fig. 1). The phylogenetic relationships between the

paraphyletic heterostrophan lineages and Euthyneura remain unresolved, as well as the reciprocal monophyly of Opisthobranchia and Pulmonata [e.g. 27, 28, 36, 37]. The absence of a clear morphological distinction between pulmonates (land snails and slugs) and opisthobranchs (sea slugs) has complicated the definition of both groups, with some taxa (eg. Siphonariidae or Onchidiidae) historically included in one or the other group depending on the characters considered [26, 38-44]. Moreover, recent mt sequence data analyses tentatively supported the paraphyly of opisthobranchs and the polyphyly of pulmonates [37, 45, 46].

In this study, we analyze the evolution of mt gene order arrangements within gastropods in order to gain insights on rates and mechanisms of genome rearrangement within the group. We sequenced anew the complete mt genomes of five gastropods species including one representative of Heterostropha (*Pyramidella dolabrata*), one of Opisthobranchia (*Ascobulla fragilis*), and three of Pulmonata (*Siphonaria pectinata*, *Onchidella celtica*, and *Myosotella myosotis*). The newly reported sequences were aligned with all available complete mt genomes of gastropods deposited in GenBank, and subjected to commonly used methods of phylogenetic inference in order to reconstruct a robust phylogeny of gastropods. Genome arrangements of all gastropod mtDNAs were mapped onto the recovered phylogeny in order to determine rearrangement events, and to assess the phylogenetic utility of mt gene order comparisons.

Results

Genome structural features

The main structural features of the five gastropod complete mt genomes that were sequenced anew in this study are described in Table 1. Their total length ranges from 13,856 bp (*P. dolabrata*) to 14,745 bp (*A. fragilis*). Their A+T content varies from 55 % (*M. myosotis*) to 67% (*A. fragilis*). All of them encode a total of 37 (13 protein coding, 2 rRNA, and 22 tRNA) genes. Of these, 13 genes (*trnQ*, *trnL (uur)*, *atp8*, *trnN*, *atp6*, *trnR*, *trnE*, *rrnS*, *trnM*, *nad3*, *trnS (ucn)*, *trnT*, and *cox3*) are encoded in all the mt genomes by the minus strand.

Overlapping of adjacent genes (even between genes encoded by the same strand) is fairly common in the five mt genomes. In almost all cases the overlap involves tRNA genes, although *cox1* gene

overlaps with *trnK* gene in three mt genomes (*O. celtica*, *M. myosotis*, and *S. pectinata*) and with *trnY* gene in that of *P. dolabrata* (Fig. 2). In addition, *nad4* gene overlaps with *trnS* (*agn*) and *trnT* genes in *P. dolabrata*. The total length of intergenic spacers is extremely small for the *P. dolabrata* genome (89 bp), medium for the *S. pectinata*, *O. celtica*, and *M. myosotis* genomes (200-300 bp) and relatively long for the mt genome of *A. fragilis* (644 bp) (Table 1). The longest non-coding region (173 bp) was found in the mt genome of *O. celtica* between *nad6* and *nad5* genes. This region was checked for potential secondary structures (stem loops or tRNA-like structures) and repetitive motives. Putative *trnQ*-like and *trnF*-like structures were found (Fig. 2). The potential origin of replication was located in the five mt genomes in a non-coding sequence between *cox3* and *trnI* genes by comparison with other gastropod genomes. These non-coding sequences (42-53 bp long) have an extremely high A+T content (Table 1).

Initiation and termination codons for the 13 protein coding genes encoded by the five mt genomes are summarized in Table 1. Most protein coding genes start with TTG, although ATG, ATT, ATA and GTG are also common. The ORFs normally end with TAG and TAA stop codons. Incomplete stop codons (T or TA) are common in the two smallest mt genomes (i.e. *P. dolabrata* and *S. pectinata*). Interestingly, *cox3* gene ends with the incomplete stop codon T in all five mt genomes sequenced in this study.

The 22 tRNAs genes of each of the five mt genomes were identified based on the corresponding anticodons, and their typical cloverleaf secondary structure. All *trnS* genes lack the DHU stem. In some tRNAs genes, the acceptor stem is mispaired (Fig. 2). An extreme reduction in length of the TΨC stem was observed in several tRNAs genes in the mt genome of *P. dolabrata* (*trnF*, *trnQ*, *trnG*, *trnH*, and *trnT*) (Fig. 2). Remarkably, nucleotide sequences of the *trnY* and *trnK* of *P. dolabrata* are almost identical (Fig. 2).

Phylogeny of gastropods

The deduced amino-acid sequences of 12 mitochondrial protein-coding genes (all except *atp8*) from 15 gastropods and 3 cephalopods were combined into a single data set that produced an alignment of 3,046 positions. Of these, 714 were invariant, and 1,870 were parsimony-informative. Mean

character distances among ingroup taxa varied between 0.11 (*I. obsoleta* and *L. cerithiformis*) and 0.60 (*L. digitalis* and *C. nemoralis*). The average mean character distance was 0.37 (± 0.07) among Heterobranchia lineages, 0.46 (± 0.02) between Heterobranchia and Caenogastropoda, 0.53 between *L. digitalis* and Caenogastropoda, and 0.57 (± 0.01) between *L. digitalis* and Heterobranchia.

The reconstructed Bayesian 50% majority rule consensus tree (-lnL= 68443.87) based on the concatenated data set is shown in figure 3. All nodes received maximal BPP support. MP (12,933 steps; CI=0.67), ME (score= 3.10), and ML (-lnL=68,436.70) arrived at identical topology. All nodes received high BP support except the sister group relationships of *P. dolabrata* + *O. celtica* and *P. strigosa* + *R. europaea* in the MP and ME analyses, respectively (Fig. 3). In addition, neither MP nor ME analyses supported the clade *A. fragilis* + *S. pectinata*. BI based on an amino acid data set considering each mt gene as a different partition, and using a partitioned mixed model of the combined data recovered an almost identical tree (-lnL= 69881.74). The only differences to the tree shown in Figure 3 were that *S. pectinata* was recovered as sister group of *A. californica* to the exclusion of *A. fragilis* with low BPP support (91%), and that the sister group relationship of *P. dolabrata* + *O. celtica* was not supported. All the other nodes received maximal BPP support (not shown). An identical topology to that shown in Figure 3 was recovered when the polyplacophoran *Katharina tunicata* was used as outgroup (not shown).

Four main lineages were identified within gastropods: Caenogastropoda (including *I. obsoleta* and *L. cerithiformis*), Vetigastropoda (*H. rubra*), Patellogastropoda (*L. digitalis*) and Heterobranchia (all remaining analyzed species). According to the recovered tree, Caenogastropoda and Vetigastropoda are sister group taxa, as well as Patellogastropoda and Heterobranchia (Fig. 3). Within Heterobranchia, neither Euthyneura nor Pulmonata nor Opisthobranchia were recovered as monophyletic groups (Fig. 3). The monophyly of Euthyneura was strongly rejected because the heterostrophan *P. dolabrata* was recovered deep within Pulmonata + Opisthobranchia (Fig 3). The recovered relative positions of Stylommatophora, Ellobiidae, Basommatophora (non-monophyletic), and Systelommatophora render Pulmonata polyphyletic whereas Opisthobranchia is recovered

paraphyletic due to the inclusion within this group of the basommatophoran pulmonate *S. pectinata* (Fig. 3).

The AU, SH and KH tests (Table 2) were performed to evaluate whether alternative morphology-based hypotheses could be rejected based on the analyzed mt data set. The AU and KH test values for the sister group relationship of Eogastropoda and Orthogastropoda, and the monophyly of Apogastropoda, Euthyneura, Pulmonata, and Basommatophora were ≤ 0.01 in all cases. For the SH test, all values for alternative hypotheses were ≤ 0.01 except for the sister group relationship of Eogastropoda and Orthogastropoda, and the monophyly of Apogastropoda. Overall, the performed tests showed that alternative morphological hypotheses could be statistically rejected based on our data set.

Evolution of gastropod mt genome arrangements

Each of the five newly determined gastropod mt genomes exhibits a different gene order (Fig. 4). Moreover, among all reported gastropod mt genomes only two sets of species (*I. obsoleta* + *L. cerithiformis*, and *A. californiaca* + *P. strigosa*) have the same gene order (Fig. 4). Despite the observed arrangement diversity, several general patterns may be inferred. The gene order of the *L. digitalis* mt genome is the most divergent among all gastropod mtDNAs sequenced thus far. The vetigastropoda *H. rubra* mt genome has the same gene order of the cephalopod *O. vulgaris* mt genome except for the relative position of three tRNA (*trnC*, *trnD* and *trnN*) genes (Fig. 4). The identical gene order exhibit by the two caenogastropod mt genomes only differs from that of *O. vulgaris* in one translocation and one inversion (Fig. 4). Despite several autapomorphies, all heterobranch mt genomes analyzed in this study share a general conserved gene order, which substantially differs from that of other gastropods (Fig. 4). *S. pectinata* and *P. dolabrata* show several autapomorphies that involve changes not only in the relative position of tRNAs but also of some protein coding genes (*cox2*, *nad4L*, and *atp6*) (Fig. 4). *M. myosotis* and *C. nemoralis* have autapomorphic relative positions for *nad4L* and *cox3* genes, respectively (Fig. 4). Changes in the remaining analyzed heterobranch mt genomes affect the arrangement of *trnW*, *trnY*, *trnC*, *trnP*, and *trnT* genes (Fig. 4).

Genome arrangements were mapped onto the phylogenetic hypothesis presented in figure 3 in order to gain insights on the patterns, rates, and mechanisms of gastropod mt genome rearrangements. The evolutionary comparative analyses showed that most gene rearrangements occurred in the lineages leading to Patellogastropoda and Heterobranchia, whereas Vetigastropoda and Caenogastropoda mostly retain the ancestral gene arrangement (as represented in the analysis by the cephalopod mt genome arrangement). Within Heterobranchia, rates of rearrangement seem to have significantly slowed down, and accelerated again in the lineages leading to *P. dolabrata* and *S. pectinata*. Most inferred rearrangements involve tRNA genes rather than protein coding genes. Furthermore, changes in gene order are normally related with translocations, most moving to proximal regions but also some to more distant regions. Among all inferred rearrangements, only one instance of gene inversion between *H. rubra* and the ancestor of the Caenogastropoda was found. In most cases, rearrangements involve one or few genes. However, the described inversion involved a genome fragment with 16 genes.

Attempts to recover phylogenetic relationships among gastropods based on genome arrangement information and using parsimony- or Bayesian-based methods of phylogenetic inference rendered highly unresolved trees (the only sister-group relationships that were confidently recovered were *A. californica* + *P. strigosa* and *I. obsoleta* + *L. cerithiformis*, which each have identical gene order) (data not shown).

Discussion

Evolutionary trends in the structural features of gastropod mt genomes

The mt genomes of the gastropods sequenced so far contain the 37 genes described for the majority of mt genomes within Metazoa [47]. This is not the case for other groups of mollusks like bivalves, which lack some genes (e.g. *atp8*, [48], but see [49]) or both bivalves and cephalopods, which show duplicated genes (e.g. *rrnS*, *cox1*, *cox3*, *atp6*, and *atp8*) [50-52]. This observation suggests that changes in gene content and number might be exclusive features characterizing bivalves and cephalopods (unless new data from other groups of mollusks points otherwise), and that gastropods maintain the ancestral number of genes with no translocations of genes to the

nucleus and/or any signal of duplication events (but see below for striking exceptions to this general rule).

The compact organization, gene order, and molecular features of the five heterobranch mt genomes sequenced anew in this study fit well within the general description of the gastropod mt genomes that have been sequenced so far [11-20]. The small lengths of the genes and intergenic spacers, as well as the high number of overlaps between adjacent genes render gastropod mt genomes amongst the smallest in size within Metazoa. The only exception to this general trend is the relatively larger size of the mt genome of the patellogastropodan *L. digitalis*, which is due to the presence of several non-coding tandem repeat units.

Several hypotheses have been proposed to explain overlapping of adjacent mt genes that are transcribed from the same strand, including the existence of multiple promoters, differential cleavage to generate diverse RNAs or post-transcriptional editing of the RNA [48]. All overlapping events that can be observed in the five mt genomes described in this study are between adjacent tRNA genes or between adjacent tRNA and protein coding genes and could be explained by tRNA post-transcriptional mechanisms as described by Yokobori and Pääbo [53]. Observed overlaps in the five gastropod mt genomes are restricted to the acceptor stem, where the 5' end of the tRNAs overlaps with the downstream gene. The nucleotides that overlap seem to be part of the downstream gene while the complementary nucleotides needed in the acceptor stem of the tRNA genes could be added by polyadenylation [53]. The presence of truncated TΨC arms in many tRNA genes in *P. dolabrata* could be the result of the extreme reduction in length of the genes in the mt genome of this species.

Two tRNA-like structures (*Q*-like and *F*-like) were found in the mt genome of *O. celtica*, in a non-coding region between the protein coding genes *nad6* and *nad5*. These tRNA-like structures could result from duplication events in the mt genome of *O. celtica* and may have been able to remain as pseudogenes because they were located in a non-coding region. It has been suggested that mt genes are expressed as a polycistron, and that tRNA-like structures might be related to the processing of the primary transcripts liberating the flanking gene specific mRNAs [20]. Due to the relevance of this hypothetical function, one would expect to find these structures in high frequencies

in mt genomes. However, many protein coding genes in the mt genomes of gastropods abut directly (without tRNA genes or tRNA-like structures between them) and only in very few cases some secondary structures have been described [20]. Therefore, the genome organization data available so far do not support the above-mentioned hypothesis as a general mechanism for processing mRNAs in mt genomes of gastropods, although further studies based on the transcription and processing of mt genes in this group are needed.

There is no general pattern in the use of initiation and termination codons in the 13 protein coding genes of the five mt genomes analyzed in this study. Although a strand-specific pattern of termination codon usage has been described for the gastropod *H. rubra* [16], this does not seem to be a general feature for gastropod mt genomes. The use of incomplete termination codons (T, TA) is especially high in the two mt genomes with smallest sizes (i.e. *P. dolabrata* and *S. pectinata*). Post-transcriptional poly-adenylation could generate the complete termination codon [15, 20]. Our results suggest that *cox3* might be a very conserved gene both in length and in codon usage within Gastropoda.

New insights into the phylogeny of Gastropoda

The reconstructed gastropod phylogeny based on 12 mt protein-coding gene sequence data supports four natural groups within Gastropoda: Vetigastropoda, Caenogastropoda, Patellogastropoda, and Heterobranchia. The systematic validity of these groups is in full agreement with most recent phylogenetic analyses based on morphological and molecular data [26, 28, 34]. Morphological analyses also distinguish two additional lineages within gastropods including Neritopsina [28] and the enigmatic Cocculiniformia [26]. Mt genomes of representatives of these two groups need to be sequenced in the future to test their phylogenetic position within gastropods. Although many efforts have been made in the last years to resolve phylogenetic relationships among the main lineages within Gastropoda, many questions remain open. One of the most controversial issues refers to the position of Patellogastropoda. Morphological studies consistently place patellogastropods as the sister taxon to the rest of gastropods [25, 26, 28, 42, 54]. However, Colgan et al. 2003 [34] proposed a reevaluation of the morphological characters defining

Patellogastropoda and its relationships with other gastropods. The morphological dataset was updated with new data about buccal cartilages and the fine structure of the cephalic tentacles, and the validity of previous polarization of some characters like the absence of the hypobranchial gland or the flexoglosste condition of the radula was questioned [34]. Moreover, molecular studies produced conflicting results, recovering unstable positions for Patellogastropoda [30, 34]. In our phylogenetic analyses, *L. digitalis* is recovered as the sister group to Heterobranchia with statistical support when trees are rooted with both cephalopods and polyplacophorans (a sister group relationship between Eogastropoda and Orthogastropoda was rejected by the AU and KH (but not SH) tests; Table 2). The mt genome of *L. digitalis* exhibits a unique mt gene order, and in the phylogenetic analyses, this taxon has a long branch compared to that of other gastropods. A number of studies have suggested that nucleotide substitution and gene rearrangement rates may be correlated in mitochondrial genomes (see [55] and references therein). This may be the case in *L. digitalis*, and long-branch attraction phenomena could likely be biasing the inference of the phylogenetic position of this species, particularly taking into account that any outgroup taxa that could be included in the phylogenetic analyses are all distantly related to gastropods. The complete mt genomes of more representatives of Patellogastropoda need to be analyzed in order to resolve the phylogenetic position of Patellogastropoda and to discern among competing hypotheses.

Previous morphological and molecular studies based on nuclear markers recovered Caenogastropoda and Heterobranchia as sister taxa, conforming the Apogastropoda [26, 28, 31, 35, 56]. However, all the analyses performed in this study support a strikingly close relationship between Vetigastropoda and Caenogastropoda. The AU and KH (but not SH) tests consistently rejected the validity of Apogastropoda (Table 2). Morphological synapomorphies described for Apogastropoda should be revisited taking into account our results, and perhaps new synapomorphies for Vetigastropoda and Caenogastropoda may be found.

The monophyly of Heterobranchia is well supported by several morphological synapomorphies like the presence of pigmented mantle organs, longitudinal rows of cilia in the mantle cavity, a chalaze in the egg masses, heterostrophy, and simultaneous hermaphroditism [57]. Phylogenetic analyses based on mt genome sequence data support the monophyly of heterobranchs (Fig. 3). However,

the monophyly and phylogenetic relationships between the different groups included within Heterobranchia (i.e. Heterostropha and Euthyneura (Pulmonata and Opisthobranchia)) have been the subject of controversy for many years [25, 26, 58-60]. Heterostropha was defined as a paraphyletic group including species with typically ancestral and derived characters mixed together in the same forms [26, 28, 36]. The unique representative of Heterostropha included in this study (*P. dolabrata*) belongs to Pyramidelloidea, a group of gastropods that has been excluded of Heterobranchia by some authors [23, 25, 61], placed as a basal heterobranch with respect to Euthyneura [58], or included in Opisthobranchia by others [39, 62-65] depending on the characters considered. All our analyses recover *P. dolabrata* as closely related to systelommatophoran pulmonates and opisthobranchs, and confirm our previous studies [37].

Pulmonata includes marine, freshwater and terrestrial gastropods with very different body plans. The monophyly of Pulmonata has been accepted by many authors based on some morphological characters like the streptoneuran innervation of the cephalic tentacles, and the lack of rhinophoric nerve (present in opisthobranchs and pyramidellids). However, the essential, traditionally accepted morphological synapomorphy of Pulmonata is the presence of a special neurosecretory system comprising procerebrum and cerebral gland [26, 66-68]. The procerebrum is formed of small and large neuronal cells, and because it links the peripheral tentacular structures with the central nervous system, an olfactory function has been assumed. The cerebral gland is a neuronal structure associated with the cerebral ganglia. New molecular data reject Pulmonata as a natural group based on both nuclear and mitochondrial data [32, 37, 45]. In this study, we have included representatives of all major lineages within pulmonates (Systelommatophora, Basommatophora, Ellobiidae, and Stylommatophora). All these lineages independently reject the definition of pulmonates as a natural group in all the performed analyses (Fig. 3; Table 2). Stylommatophora (land snails) is a monophyletic group, in agreement with previous morphological studies [69], and it is recovered as the sister group to all other heterobranchs studied (Fig. 3). Our results provide new insights into land colonization by heterobranch gastropods. The transition to a land lifestyle was accompanied by a variety of refined morphological and physiological modifications. As a result, land snails and slugs constitute a well-defined group of pulmonates with several morphological

synapomorphies in the cephalic tentacles, kidney, and central nervous system, as well as in several aspects of their ontogeny. Previous phylogenetic hypotheses had suggested that the transition to land was a rather derived event in the history of pulmonates. Our molecular phylogeny instead supports a different scenario in which gastropod land colonization, and subsequent radiation was an early and significant event in the evolution of Heterobranchia.

The marine Basommatophora *S. pectinata* is more closely related to opisthobranchs than to the freshwater basommatophoran *B. glabrata* or to any other group of pulmonates considered in this study (Fig. 3). The phylogenetic affinities of *S. pectinata* have been under debate for many years (see discussion in [37]). All new data and analyses presented here support *S. pectinata* as an opisthobranch. To test the monophyly of freshwater basommatophorans, mt genome sequence data of more representatives of this group should be included in future analyses.

The monophyly of pulmonates is also rejected by the location of a representative of Systelommatophora (*O. celtica*) as an independent lineage more closely related to *P. dolabrata* and to opisthobranchs than to any other Pulmonate (Fig. 3), which corroborates previous morphological hypotheses [58, 70].

The robust results described in this study provide new insights on the systematics of gastropods. The majority of our evolutionary hypotheses are in agreement with traditional morphological hypotheses although there are some cases of strong discrepancy (i.e. the phylogenetic position of Patellogastropoda, the sister group relationship between Vetigastropoda and Caenogastropoda, and the polyphyly of Pulmonata). Among these results, the polyphyly of pulmonates is perhaps the most remarkable, and warns against only relying on one or few morphological characters (even if they seem to be free of convergent evolution) to define deep phylogenetic relationships. In any case, the results here presented should be interpreted as a working phylogenetic hypothesis, which needs to be further confirmed with a larger taxon sampling of the studied groups, and the addition to the phylogenetic analyses of new taxa representing not previously included major lineages of gastropods.

Mitochondrial genome rearrangements during the evolution of Gastropoda.

Gene order rearrangements in mt genomes are relatively rare, and if shared derived by two taxa can be considered molecular synapomorphies and may provide useful data for phylogenetic reconstruction. In this study, we have mapped gene orders of gastropod mt genomes onto the gastropod phylogeny and tentatively reconstructed the evolutionary history of mt gene order rearrangements in gastropods. The Vetigastropoda *H. rubra* and the cephalopod *O. vulgaris* have the same gene order (with only three changes in *trnC*, *trnD* and *trnN*) suggesting that *H. rubra* may retain the ancestral mt gene order of Gastropoda. The relative placement of *trnD* and *trnN* in *H. rubra* might constitute autapomorphies in this species since these two tRNAs show the same location in Caenogastropoda and the cephalopod *O. vulgaris*.

Considering all the data available so far, three major rearrangements have taken place in the evolutionary history of gastropods: one in the ancestor of Caenogastropoda, another in the ancestor of Patellogastropoda, and another one in the ancestor of Heterobranchia.

The two complete mt genomes of caenogastropods sequenced so far and the incomplete mt genome of the caenogastropod *L. saxatilis* have identical gene arrangements (data not shown). Two rearrangements (one inversion and one translocation) separate the hypothetical ancestral state of gastropods from the gene order found in Caenogastropoda (Fig. 4). Long inversion events rendering new gene order configurations have been already described for other closely related groups [71].

The mt gene order of *L. digitalis* is very distinctive compared to those other mollusk mt genomes sequenced so far. The remarkable number of gene rearrangement events that need to be invoked to explain the mt gene order of *L. digitalis* requires considering special mechanisms such as intramolecular recombination facilitated by tandem repeat sequences, as described for other metazoans [72]. Sequencing of other patellogastropodan mt genomes should determine whether the striking gene order of *L. digitalis* is unique to this species or a more general feature of Patellogastropoda.

The mt gene order of the heterobranchs sequenced so far, including those of the incomplete mt genomes of the Stylommatophoran *E. herklotsi* and the Heterostrophan *O. atomus* (data not shown), has only few gene boundaries in common with the hypothetical ancestral mt gene order of

gastropods. Considering all the data available so far, it is not possible to determine the precise mechanism responsible for rearrangements in this transition (tandem duplication-random loss [73], inversion [71], transposition [74], and/or intramolecular recombination [72]). The compact organization of the mt genomes of Heterobranchia (with very few and short non coding sequences) suggests strong selection against maintaining remnants of duplication events. However, the *trnF*-like structure described in this study, might represent one remnant of a putative duplication event, and would be in support of tandem duplication and random loss as a mechanism acting in heterobranch mt genome rearrangements.

The new data presented here suggest that the gene order among heterobranchs is not as well conserved as previously thought [11, 14, 17, 20, 22]. Most part of these mt genomes shows a rather conserved gene order, being gene rearrangements concentrated between *nad1* and *trnE* genes (Fig. 4). Several gene order autapomorphies (not only in tRNAs but also in protein coding genes) can be detected in heterobranchs, especially in *S. pectinata* and *P. dolabrata*.

Two additional aspects about gene order in heterobranchs can be highlighted. First, the gene order *trnY-trnW-nad4L* supports the close relationship between opisthobranchs and *S. pectinata*.

Second, *trnY* bounds with *cox1* gene in the mt genome of *P. dolabrata* (instead of *trnK* gene as in all other heterobranchs). However, the sequences of the *trnY* and *K* genes in *P. dolabrata* are nearly identical (Fig. 2), suggesting that two mutation events might have occurred in the anticodon triplet producing changes in the identity of these two tRNAs. This event, called tRNA recruitment, has also been previously described in other taxa [75].

The phylogenetic inferences based on genome arrangement data rendered inconclusive results. Despite the relatively high number of different gene orders found in gastropod mt genomes, few of them are shared and derived, and thus both parsimony and Bayesian inference arrived at rather unresolved trees. Our results point out that the comparative study of gene arrangement in gastropods may provide valuable phylogenetic information, but that current methods of phylogenetic inference based on gene arrangements still need further development.

Conclusions

According to our results, the validity of Apogastropoda should be questioned, Pyramidelloidea should be included within Euthyneura, *S. pectinata* should be recognized as a member of Opisthobranchia, and Pulmonata should not be considered a natural group. Our results stress the need of a thorough re-evaluation of the morphological characters that have been used to analyze relationships within Gastropoda, and in particular those that supported the monophyly of pulmonates. Although the number of complete mt genomes is increasing rapidly in the last years, still more genomic data is needed to further understand gastropod systematics. The phylogenetic affinities of Neritopsina and Cocculiniformia with respect to other gastropods, the position of Patellogastropoda among gastropods, and the delimitation of Heterostropha and Euthyneura are still open questions that could be tackled in the future with the approach proposed in this study. The recovered phylogeny based on complete mt genome data provides a first instance of robust phylogenetic framework for comparative studies in gastropods, and will allow a better understanding of evolutionary trends within this group. In particular, it seems to be quite useful and promising for determining the rates and mechanisms of gene rearrangement in gastropod mt genomes.

Methods

Taxon sampling

Each of the five gastropod complete mt genomes sequenced anew in this study was obtained from a single specimen. *A. fragilis* was collected in Murcia (southeastern Iberian Peninsula) and preserved frozen at -20°C . *S. pectinata* and *O. celtica* were sampled in Ceuta (northern Africa) and preserved in EtOH 100%. *P. dolabrata* was collected in Annobon Island (western Africa) and preserved in EtOH 100%. *M. myosotis* was collected in Vigo (northwestern Iberian Peninsula) and preserved frozen at -20°C . All specimens were sampled between 2000 and 2002.

DNA extraction, PCR amplification, cloning and sequencing

Total cellular DNA was purified following a standard phenol/chloroform extraction [76]. Universal primers were used to amplify by polymerase chain reaction (PCR) fragments of the mitochondrial *cox1* (LCO-1490 and HCO-2198, [77]), *rnrL* (16Sar-L and 16Sbr-H, [78]), and *rnrS* (H1478 and L1091, [79]) genes. Standard PCR reactions containing 67 mM Tris-HCl, pH 8.3, 1.5 mM MgCl₂, 0.4 mM of each dNTP, 2.5 μM of each primer, template DNA (10-100 ng), and Taq DNA polymerase (1 unit, Biotools) in a final volume of 25 μl were subjected to 30 cycles of denaturing at 94°C for 60 s, annealing at 42°C for 60 s, and extending at 72°C for 90 s. The PCR amplified fragments were sequenced with the BigDye Deoxy Terminator cycle-sequencing kit (Perkin Elmer Biosystems) in an automated DNA sequencer (ABI PRISM 3100) using the PCR primers, and following manufacturer's instructions.

The sequences of these fragments were used to design three sets of specific primers for each species that amplified, by long PCR, three fragments that covered the remaining mt genome. Long PCRs containing 60 mM Tris-SO₄ (pH 9.1), 18 mM (NH₄)₂ SO₄, 1-2 mM MgSO₄, 0.2 mM of each dNTP, 0.4 μM of each primer, and Takara enzyme (1 unit; Life Technologies) in a final volume of 50 μl were subjected to 40 cycles of denaturing at 94°C for 30 s, annealing at 52°C for 30 s, and extending at 68°C for 7 min. Long PCR products in some cases and total cellular DNA extractions in others were used as DNA templates to amplify by standard PCR reactions (see conditions above) overlapping fragments that covered the complete mt genomes. These overlapping PCRs were performed using degenerated primers (designed based on published mt genome sequences of gastropods) and/ or specific walking primers for each species. The sequences of all these primers are available from the authors upon request. PCR products were cloned into the pGEM-T vector (Promega), and sequenced using M13 universal primers in an automated sequencer (see above).

Molecular and Phylogenetic analyses

Gene annotation was performed using BLAST [80] comparisons against published gastropod mtDNAs. Sequence data were handled with MacClade version 4.05 OSX [81] and PAUP* version 4.0b10 [82]. Protein coding genes were recognized by inferring open reading frames (ORFs), and

by delimiting start and stop codons. Cloverleaf secondary structures of all tRNA genes were reconstructed by hand upon localization of the specific anticodons. The sequences of the complete mt genomes reported in this paper have been deposited at the EMBL/GenBank data libraries under accession numbers AY345054 (*P. dolabrata*), AY098929 (*A. fragilis*), AY345049 (*S. pectinata*), AY345048 (*O. celtica*), and AY345053 (*M. myosotis*).

Phylogenetic analyses included the five newly determined mtDNA sequences, as well as all gastropod complete mt genome sequences available in GenBank. In addition, the corresponding sequences of the cephalopod species *Octopus vulgaris*, *Todarodes pacificus* [52], and *Loligo bleekeri* [83], and the polyplacophoran *Katharina tunicata* [84] were used as outgroups. The deduced amino-acid sequences of each mt protein-coding gene (except *atp8*) were aligned independently using Clustal X version 1.62b [85] followed by refinement by eye in an effort to maximize positional homology. Ambiguous alignments and gaps were discarded from further phylogenetic analyses. The aligned amino acid sequences were concatenated into a single data set, which was subjected to maximum-parsimony (MP), minimum evolution (ME), maximum likelihood (ML), and Bayesian inference (BI) methods of phylogenetic reconstruction. MP analyses were performed with PAUP* using heuristic searches (TBR branch swapping; MulTrees option in effect) with 10 random additions of taxa. ME analyses [86] were carried out with PAUP* using mean character distances. Robustness of the resulting MP and ME trees was evaluated with non-parametric bootstrap proportions (BPs, [87]) as implemented in PAUP* with 1,000 pseudoreplicates. ProtTest version 1.3 [88] was used to estimate the evolutionary model that best fit the amino-acid data set. The Akaike information criterion (AIC) implemented in ProtTest selected the WAG+I+G [89] evolutionary model. ML analyses using the WAG+I+G were performed with PHYML version [90], and robustness of the resulting ML tree was evaluated by bootstrapping with 500 pseudoreplicates. BI analyses were performed with MrBayes 3.12 [91] by simulating a Metropolis-coupled Markov chain Monte Carlo (MCMCMC) with four simultaneous chains, each of 10^6 generations (sampled every 100 generations) under the WAG+I+G model. Trees sampled before the cold chain reached stationarity (as judged by plots of ML scores) were discarded as “burn-in”. Runs were repeated twice. Robustness of the resulting BI tree was evaluated using Bayesian

posterior probabilities (BPPs). In addition, BI was also applied to a data set that included the aligned amino acid sequence of each gene as independent partitions. The AIC implemented in ProtTest was used to select the substitution models for each gene: *atp6* (RtREV+G+F), *cox1* (WAG+G+F), *cox2* (RtREV+I+G+F), *cox3* (cpREV+G+F), *cob* (cpREV+I+G+F), *nad1* (RtREV+G+F), *nad2* (WAG+I+G+F), *nad3* (RtREV+G+F), *nad4* (WAG+I+G+F), *nad4L* (MtREV+G+F), *nad5* (RtREV+I+G+F), *nad6* (Dayhoff+G+F). This data set was subjected to the same searching parameters for BI described above plus the “set partition” and “unlink” options. Alternative phylogenetic hypotheses were tested using the approximately unbiased (AU), Shimodaira- Hasegawa (SH), and Kishino-Hasegawa (KH) tests [92] as implemented in CONSEL version 0.1 [93] using default settings. The alternative hypotheses tested were *a priori* morphology-based hypotheses.

Phylogenetic relationships among gastropods were also reconstructed based on genome arrangement data using parsimony and Bayesian inferences with the GRAPPA [8] and Badger [94] programs.

Abbreviations

BI: Bayesian inference; ME: minimum evolution; ML: maximum likelihood; MP: maximum-parsimony; mt: mitochondrial. BPP: Bayesian posterior probabilities; BP: Bootstrap proportions;

Authors' contributions

JT collected some specimens for this study. CG and RZ gathered the sequences, and performed phylogenetic analyses. All authors wrote the manuscript, read and approved the final version of the manuscript.

Acknowledgements

We are grateful to E. Rolán for providing the samples of *P. dolabrata* and *M. myosotis*. Diego San Mauro helped with the AU test. CG was sponsored by a postdoctoral fellowship (2005-0511) of the Ministerio de Educación y Ciencia. This work received financial support from projects of the Ministerio de Educación y Ciencia to RZ (CGL2007-60954) and to JT (CTM2004-01769).

References

1. Wolstenholme DR: **Animal mitochondrial genome: structure and evolution.** *Int Rev Cytol* 1992, 141: 172-216.
2. Boore JL, Macey JR, Medina M: **Sequencing and comparing whole mitochondrial genomes of animals.** *Methods Enzymol* 2005, 395: 311-348.
3. Website title [http://www.ncbi.nlm.nih.gov/genomes/static/euk_o.html]
4. Desjardins P, Morais R: **Sequence and gene organization of the chicken mitochondrial genome - a novel gene order in higher vertebrates.** *J Mol Biol* 1990, 212: 599-634.
5. San Mauro D, Gower DJ, Zardoya R, Wilkinson M: **A hotspot of gene order rearrangement by tandem duplication and random loss in the vertebrate mitochondrial genome.** *Mol Biol Evol* 2006, 23: 227-234.
6. Boore JL, Brown WM: **Big trees from little genomes: mitochondrial gene order as a phylogenetic tool.** *Curr Opin Genet Dev* 1998, 8: 668-674.
7. Rokas A, Holland PWH: **Rare genomic changes as a tool for phylogenetics.** *TREE* 2000, 15: 454-459.
8. Moret BME, Bader DA, Warnow T: **High-performance algorithm engineering for computational phylogenetics.** *J Supercomput* 2002, 22: 99-110.
9. Larget B, Kadane JB, Simon DL: **A Bayesian approach to the estimation of ancestral genome arrangements.** *Mol Phylogenet Evol* 2005, 36: 214-223.
10. Larget B, Simon DL, Kadane JB, Sweet D: **A Bayesian analysis of metazoan mitochondrial genome arrangements.** *Mol Biol Evol* 2005, 22: 486-495.
11. Kurabayashi A, Ueshima R: **Complete sequence of the mitochondrial DNA of the primitive opisthobranch gastropod *Pupa strigosa*: systematics implication of the genome organization.** *Mol Biol Evol* 2000, 17: 266-277.
12. Hatzoglou E, Rodakis GC, Lecanidou R: **Complete sequence and gene organization of the mitochondrial genome of the land snail *Albinaria coerulea*.** *Genetics* 1995, 140: 1353-1366.
13. Terrett JA, Miles S, Thomas RH: **Complete DNA Sequence of the mitochondrial genome of *Cepaea nemoralis* (Gastropoda: Pulmonata).** *J Mol Evol* 1996, 42: 160-168.
14. Grande C, Templado J, Cervera JL, Zardoya R: **The complete mitochondrial genome of the nudibranch *Roboastra europaea* (Mollusca: Gastropoda) supports the monophyly of opisthobranchs.** *Mol Biol Evol* 2002, 19: 1672-1685.
15. DeJong RJ, Emery AM, Adema CM: **The mitochondrial genome of *Biomphalaria glabrata* (Gastropoda: Basommatophora), intermediate host of *Schistosoma mansoni*.** *J Parasitol* 2004, 90: 991-997.

16. Maynard BT, Kerr LJ, McKiernan JM, Jansen ES, Hanna PJ: **Mitochondrial DNA sequence and gene organization in Australian backup abalone *Haliotis rubra* (leach).** *Mar Biotechnol (NY)* 2005, 7: 645-658.
17. Knudsen B, Kohn AB, Nahir B, McFadden CS, Moroz LL: **Complete DNA sequence of the mitochondrial genome of the sea-slug, *Aplysia californica*: conservation of the gene order in Euthyneura.** *Mol Phylogenet Evol* 2006, 38: 459-469.
18. Simison WB, Lindberg DR, Boore JL: **Rolling circle amplification of metazoan mitochondrial genomes.** *Mol Phylogenet Evol* 2006, 39: 562-567.
19. Bandyopadhyay PK, Stevenson BJ, Cady MT, Olivera BM, Wolstenholme DR: **Complete mitochondrial DNA sequence of a Conoidean gastropod, *Lophiotoma (Xenuroturrus) cerithiformis*: Gene order and gastropod phylogeny.** *Toxicon* 2006, 48: 29-43.
20. Yamazaki N, Ueshima R, Terret JA, Yokobori S, Kaifu M, Segawa R, Kobayashi T, Numachi K, Ueda T, Nishikawa K, Watanabe K, Thomas RH: **Evolution of pulmonate gastropod mitochondrial genomes: Comparisons of gene organizations of *Euhadra*, *Cepaea* and *Albinaria* and implications of unusual tRNA secondary structures.** *Genetics* 1997, 145: 749-758.
21. Wilding CS, Mill PJ, Grahame J: **Partial sequence of the mitochondrial genome of *Littorina saxatilis*: relevance to gastropod phylogenetics.** *J Mol Evol* 1999, 48: 348-359.
22. Kurabayashi A, Ueshima R: **Partial mitochondrial genome organization of the heterostrophan gastropod *Omalogyra atomus* and its systematic significance.** *Venus* 2000, 59: 7-18.
23. Thiele J: *Handbuch der Systematischen Weichtierkunde*, vol. 1. Jena 1931.
24. Spengel J: **Die Geruchsorgane und das Nervensystem der Mollusken. Zeitschrift für Wissenschaftliche Zoologie.** *Leipzig* 1881, 35: 333-383.
25. Golikov AN, Starobogatov YI: **Systematics of prosobranch gastropods.** *Malacologia* 1975, 15: 185-232.
26. Haszprunar G: **On the origin and evolution of major gastropods group, with special reference to the streptoneura.** *J Moll Stud* 1988, 54: 367-441.
27. Salvini-Plawen L, Steiner G: **Synapomorphies and plesiomorphies in higher classification of Mollusca.** In: *Origin and evolutionary radiation of the Mollusca*. Edited by Taylor J. London, The Malacological Society of London 1996, 29-51.
28. Ponder WF, Lindberg DR: **Towards a phylogeny of gastropod molluscs: Analysis using morphological characters.** *Zool J Linn Soc* 1997, 119: 83-265.
29. Winnepenninckx B, Steiner G, Backeljau T, de Wachter R: **Details of gastropod phylogeny inferred from 18s rRNA sequences.** *Mol Phylogenet Evol* 1998, 9: 55-63.

30. Colgan DJ, Ponder WF, Egger PE: **Gastropod evolutionary rates and phylogenetic relationships assessed using partial 28S rDNA and histone H3 sequences.** *Zoologica Scripta* 2000, 29: 29-63.
31. Harasewych MG, McArthur AG: **A molecular phylogeny of the Patellogastropoda (Mollusca: Gastropoda).** *Mar Biol* 2000, 137: 183-194.
32. Yoon SH, Kim W: **Phylogeny of some gastropod mollusks derived from 18s rRNA sequences with emphasis on the Euthyneura.** *The Nautilus* 2000, 114: 84-92.
33. Colgan DJ, Ponder WF, Beacham E, Macaranas JM: **Molecular phylogenetics of Caenogastropoda (Gastropoda: Mollusca).** *Mol Phylogenet Evol* 2006, in press.
34. Colgan DJ, Ponder WF, Beacham E, Macaranas JM: **Gastropod phylogeny based on six segments from four genes representing coding or non-coding and mitochondrial or nuclear DNA.** *Molluscan Research* 2003, 23: 123-148.
35. Salvini-Plawen L, Haszprunar G: **The Vetigastropoda and the systematics of streptoneurous gastropods (Mollusca).** *J Zool London A* 1987, 211: 747-770.
36. Bieler R: **Gastropod phylogeny and systematics.** *Annu Rev Ecol Syst* 1992, 23: 311-338.
37. Grande C, Templado J, Cervera JL, Zardoya R: **Molecular phylogeny of Euthyneura (Mollusca: Gastropoda).** *Mol Biol Evol* 2004, 21: 303-313.
38. Köhler A: **Beiträge zur Anatomie der Gattung Siphonaria.** *Zoologische Jahrbücher, Abteilung Anatomie* 1893, 7: 1-92.
39. Fretter V, Graham A: **The structure and mode of life of the Pyramidellidae, parasitic opisthobranchs.** *J Mar Biol Ass UK* 1949, 28: 493-532.
40. Gosliner TM: **Origins and relationships of primitive members of the Opisthobranchia (Mollusca: Gastropoda).** *Biol J Linn Soc* 1981, 16: 197-225.
41. Haszprunar G, Huber G: **On the central nervous system of Smeagolidae and Rhodopidae, two families questionably allied with the Gymnomorpha (Gastropoda: Euthyneura).** *J Zool* 1990, 220: 185-199.
42. Graham A: **Evolution within the Gastropoda: Prosobranchia. The Mollusca.** In: *Evolution.* Edited by Trueman ER, Clarke MR, vol. 10. New York, Academic Press 1985, 151-186.
43. Selmi MG, Giusti F, Manganelli G: **The fine-structure of the mature spermatozoon of Onchidella celtica (Cuvier) (Mollusca, Gastropoda) and its phylogenetic implications.** *J Ultrastruct Mol Struct Research* 1988, 99: 107-123.
44. Tillier S: **Relationships of Gymnomorph Gastropods Mollusca Gastropoda.** *Zool J Linn Soc* 1984, 82: 345-362.
45. Remigio EA, Hebert PDN: **Testing the utility of partial COI sequences for phylogenetic estimates of gastropod relationships.** *Mol Phylogenet Evol* 2003, 29: 641-647.

46. Thollesson M: **Phylogenetic analysis of Euthyneura (Gastropoda) by means of the 16s rRNA gene: use of a fast gene for higher-level phylogenies.** *Proc R Soc Lond B* 1999, 266: 75-83.
47. Boore JL: **Animal mitochondrial genomes.** *Nucleic Acids Res* 1999, 27: 1767-1780.
48. Boore JL, Medina M, Rosenberg LA: **Complete sequences of the highly rearranged molluscan mitochondrial genomes of the Scaphopod *Graptacme eborea* and the bivalve *Mytilus edulis*.** *Mol Biol Evol* 2004, 21: 1492-1503.
49. Dreyer H, Steiner G: **The complete sequences and gene organisation of the mitochondrial genomes of the heterodont bivalves *Acanthocardia tuberculata* and *Hiatella arctica* - and the first record for a putative *Atpase subunit 8* gene in marine bivalves.** *Front Zool* 2006, 3: 13.
50. Kim SH, Je EY, Park DW: ***Crassostrea gigas* mitochondrial DNA.** *Genbank accession number AF177226* 1999.
51. Okazaki M, Ueshima R: **Gender-associated mtDNA of *Tapes philippinarum*.** *Genbank accession number NC_003354* 2002.
52. Yokobori S, Fukuda N, Nakamura M, Aoyama T, Oshima T: **Long-term conservation of six duplicated structural genes in cephalopod mitochondrial genomes.** *Mol Biol Evol* 2004, 21: 2034-2046.
53. Yokobori S, Pääbo S: **Transfer RNA editing in land snail mitochondria.** *Proc Natl Acad Sci USA* 1995, 92: 10432-10435.
54. Ponder WF, Lindberg DR: **Gastropod phylogeny challenges for the 90S.** In: *Origin and evolutionary radiation of the Mollusca*. Edited by Taylor J, The Malacological Society of London. Oxford University Press 1996, 135-154.
55. Shao R, Dowton M, Murrell A, Barker SC: **Rates of gene rearrangement and nucleotide substitution are correlated in the mitochondrial genomes of insects.** *Mol Biol Evol* 2003, 20: 1612-1619.
56. McArthur AG, Koop BF: **Partial 28S rDNA sequences and the antiquity of hydrothermal vent endemic gastropods.** *Mol Phylogenet Evol* 1999, 13: 255-274.
57. Robertson R: **Four characters and the higher systematics of gastropods.** *Am Malac Bull Spetial Ed* 1985, 1: 1-22.
58. Haszprunar G: **The Heterobranchia- a new concept of the phylogeny and evolution of the higher Gastropoda.** *Z Zool Syst Evolut-forsch* 1985, 23: 15-37.
59. Ponder WF: **Marine valvatoideans gastropods _ implications for early heterobranch phylogeny.** *J Moll Stud* 1991, 57: 21-32.
60. Ponder WF, Warén A: **Classification of the Caenogastropoda and Heterostropha - a list of the family group names and higher taxa.** In: *Prosobranch phylogeny Malacological Rev Supp 4*. Edited by Ponder WF 1988, 288-328.
61. Franc A: *Traité de Zoologie. Opisthobranches*, vol. V, Sect. III. Paris, Masson 1968.

62. Taylor DW, Sohl NF: **An outline of gastropod classification.** *Malacologia* 1962, 1: 7-32.
63. Thompson TE: *Biology of Opisthobranch molluscs*, vol. I. London, The Ray Society 1976.
64. Pelseener P: **Recherches morphologiques et phylogénétiques sur les Mollusques Archaiques.** *Memoires couronnés et Memoires des savants étrangers, l'Academie Royale des Sciences de Belgique* 1899, 57: 1-112, I-XXIV.
65. Hyman LH: **The Invertebrates.** In: *Mollusca I.* vol. 6. New York, McGraw-Hill 1967, 1-792.
66. Huber G: **Zum cerebralen Nervensystem mariner Heterobranchia (Gastropoda).** *Dissertation am Institut für Zoologie der Universität Wien, Wien, Austria* 1987.
67. van Mol JJ: **Étude morphologique et phylogénétique du ganglion cérébroïde des gastéropodes pulmonés (mollusques).** *Mém Acad r Belg Cl Sci* 1967, 37: 1-168.
68. van Mol JJ: **Evolution phylogénétique du ganglion cérébroïde chez les gastéropodes pulmonés.** *Haliotis* 1974, 4: 77-86.
69. Barker GM: *The biology of terrestrial molluscs.* Hamilton, New Zealand, CABI publishing 2001.
70. Salvini-Plawen L: **Zur systematischen Stellung von Soleolifera und Rhodope (Gastropoda, Euthyneura).** *Zoologisches Jahrbuch für Systematik* 1970, 97: 285-299.
71. Smith MJ, Banfield DK, Doteval K, Gorski S, Kowbel DJ: **Gene arrangement in sea star mitochondrial DNA demonstrates a major inversion event during echinoderm evolution.** *Gene* 1989, 76: 181-185.
72. Lunt DH, Hyman BC: **Animal mitochondrial DNA recombination.** *Nature* 1997, 387: 247.
73. Moritz C, Brown WM: **Tandem duplications in animal mitochondrial DNAs: variation in incidence and gene content among lizards.** *Proc Natl Acad Sci U S A* 1987, 84: 7183-7187.
74. Macey JR, Schulte JA, 2nd, Larson A, Papenfuss TJ: **Tandem duplication via light-strand synthesis may provide a precursor for mitochondrial genomic rearrangement.** *Mol Biol Evol* 1998, 15: 71-75.
75. Rawlings TA, Collins TM, Bieler R: **Changing identities: tRNA duplication and remodeling within animal mitochondrial genomes.** *Proc Natl Acad Sci USA* 2003, 100: 15700-15705.
76. Sambrook J, Fritsch EF, Maniatis T: **Molecular Cloning.** In: *A Laboratory Manual.* Cold Spring Harbor, NY, Cold Spring Harbor Laboratory Press 1989.
77. Folmer O, Black M, Hoeh W, Lutz R, Vrijenhoek R: **DNA primers for amplification of mitochondrial cytochrome c oxidase subunit I from diverse metazoan invertebrates.** *Mol Mar Biol Biotech* 1994, 3: 294-299.
78. Palumbi S, Martin A, Romano S, McMillan WO, Stice L, Grabowowski G: *The simple fool's guide to PCR. Version 2.0.* Honolulu, HI 96822, Department of Zoology and Kewalo Marine Laboratory, University of Hawaii 1991.

79. Kocher TD, Thomas WK, Meyer A, Edwards SV, Paabo S, Villablanca FX, Wilson AC: **Dynamics of mitochondrial DNA evolution in animals: amplification and sequencing with conserved primers.** *Proc Natl Acad Sci USA* 1989, 86(16): 6196-6200.
80. Altschul SF, Lipman DJ: **Protein Database Searches for Multiple Alignments.** *Proc Natl Acad Sci U S A* 1990, 87: 5509-5513.
81. Maddison WP, Maddison DR: *MacClade: Analysis of phylogeny and character evolution*, Sinauer Associates, Sunderland, MA 1992.
82. Swofford DL: *PAUP*: phylogenetic analysis using parsimony (*and other methods), version 4.0.* Sunderland, MA, Sinauer Associates 2002.
83. Sasuga J, Yokobori S, Kaifu M, Ueda T, Nishikawa K, Watanabe K: **Gene content and organization of a mitochondrial DNA segment of the squid *Loligo bleekeri*.** *J Mol Evol* 1999, 48: 692-702.
84. Boore JL, Brown WM: **Complete DNA sequence of the mitochondrial genome of the black chiton, *Katharina tunicata*.** *Genetics* 1994, 138: 423-443.
85. Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG: **The CLUSTAL X windows interface: Flexible strategies for multiple sequence alignment aided by quality analysis tools.** *Nucleic Acids Res* 1997, 25: 4876-4882.
86. Rzhetsky A, Nei M: **A simple method for estimating and testing minimum-evolution trees.** *Mol Biol Evol* 1992, 9: 945-967.
87. Felsenstein J: **Confidence limits on phylogenies: an approach using the bootstrap.** *Evolution* 1985, 39: 783-791.
88. Abascal F, Zardoya R, Posada D: **ProtTest: selection of best-fit models of protein evolution.** *Bioinformatics* 2005, 21: 2104-2105.
89. Whelan S, Goldman N: **A general empirical model of protein evolution derived from multiple protein families using a maximum-likelihood approach.** *Mol Biol Evol* 2001, 18: 691-699.
90. Guindon S, Gascuel O: **A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood.** *Syst Biol* 2003, 52: 696-704.
91. Huelsenbeck JP, Ronquist FR: **MrBayes: Bayesian inference of phylogeny.** *Bioinformatics* 2001, 17: 754-755.
92. Shimodaira H: **An approximately unbiased test of phylogenetic tree selection.** *Syst Biol* 2002, 51: 492-508.
93. Shimodaira H, Hasegawa M: **CONSEL: for assessing the confidence of phylogenetic tree selection.** *Bioinformatics* 2001, 17: 1246-1247.
94. Simon DL, Larget B: **BADGER, version 1.02b.** Department of Mathematics and Computer Science, Duquesne University. Available from: <<http://badger.duq.edu/>>. 2004.

Figures

Figure 1. Phylogenetic hypothesis of gastropod relationships based on morphological data as proposed by Ponder and Lindberg [30]. For each lineage, available complete and incomplete (marked with an asterisk) mt genomes are listed together with their corresponding GenBank accession numbers.

Figure 2. Striking cloverleaf secondary structures of several tRNAs and tRNA-like structures deduced from the complete sequence of the five mitochondrial genomes sequenced in this study. A, B, and C show three proposed cloverleaf secondary structures found in the mt genome of *P. dolabrata*. Boxes in A, B, and C indicate the overlapping nucleotides of these tRNAs with their downstream genes (*cox1*, *trnF*, and *trnC*, respectively). Note that the nucleotide sequences of the *trnY* (A) and *trnK* of *P. dolabrata* (B) are almost identical. A truncated T Ψ C stem is shown in the *trnQ* of *P. dolabrata* (C). D and E indicate potential secondary structures found in the longest non-coding region of the mt genome of *O. celtica*. F and G show several proposed cloverleaf secondary structures found in the mt genome of *M. myosotis* and *S. pectinata* respectively. Boxes in F and G show the overlapping nucleotides between *trnY* and *trnW* in *M. myosotis* (F) and between *trnG* and *trnC* in *S. pectinata* (G).

Figure 3. Phylogenetic relationships among Gastropoda. The 50%-majority rule consensus of post-burn-in sampled trees from the Bayesian inference analysis based on a data set including the deduced amino acid sequences of all the protein coding genes (except *atp8*) of all gastropod complete mt genomes sequenced so far is shown. Branch lengths are mean estimates. Three species of cephalopods were used as outgroup taxa. Numbers in the nodes are from top to bottom: BI BPP, and ML, MP, and ME BP. Dashes indicate BP below 70%.

Figure 4. Hypothesized mitochondrial gene rearrangements during Gastropoda evolution based on observed gene orders and the recovered BI phylogenetic hypothesis. Inversion (indicated by the arrow) and transpositions of protein coding, tRNAs and rRNA genes are depicted among the different taxa (except between *H. rubra*-*L. digitalis* and *H. rubra*-Heterobranchia due to the

high number of changes). Genes encoded by the minor strand are underlined. Genes located in apomorphic arrangements are colored.

	Mt genomes sequenced before this study	Complete mt genomes sequenced in this study	GenBank Acc. numbers
Euthyneura	Opisthobranchia	<i>P. strigosa</i> (Kurabayashi and Ueshima, 2000a)	AB028237
		<i>R. europaea</i> (Grande et al. 2002)	AY083457
		<i>A. californica</i> (Knudsen et al. 2006)	AY569552
Heterobranchia		<i>A. fragilis</i>	AY098929
	Pulmonata	<i>A. coerulea</i> (Hatzoglou et al. 1995)	X83390
		<i>C. nemoralis</i> (Terrett et al. 1996)	U23045
		<i>E. herklotsi</i> (Yamazaki et al. 1997) *	Z71693-Z71701
	<i>B. glabrata</i> (DeJong et al. 2004)	AY380531	
Apogastropoda		<i>S. pectinata</i>	AY345049
		<i>O. celtica</i>	AY345048
		<i>M. myosotis</i>	AY345053
Orthogastropoda	Omalogyroidea	<i>O. atomus</i> (Kurabayashi and Ueshima, 2000b) *	-----
	Pyramidelloidea	<i>P. dolabrata</i>	AY345054
Eogastropoda	Other Heterostropha		
	Caenogastropoda	<i>L. saxatilis</i> (Wilding et al. 1999) *	AJ132137
		<i>I. obsoleta</i> (Simison et al. 2006)	DQ238598
	<i>L. cerithiformis</i> (Bandyopadhyay et al. 2006)	DQ284754	
	Vetigastropoda	<i>H. rubra</i> (Maynard et al. 2005)	AY588938
	Neritopsina		
	Cocculiniformia		
	Patellogastropoda	<i>L. digitalis</i> (Simison et al. 2006)	DQ238599

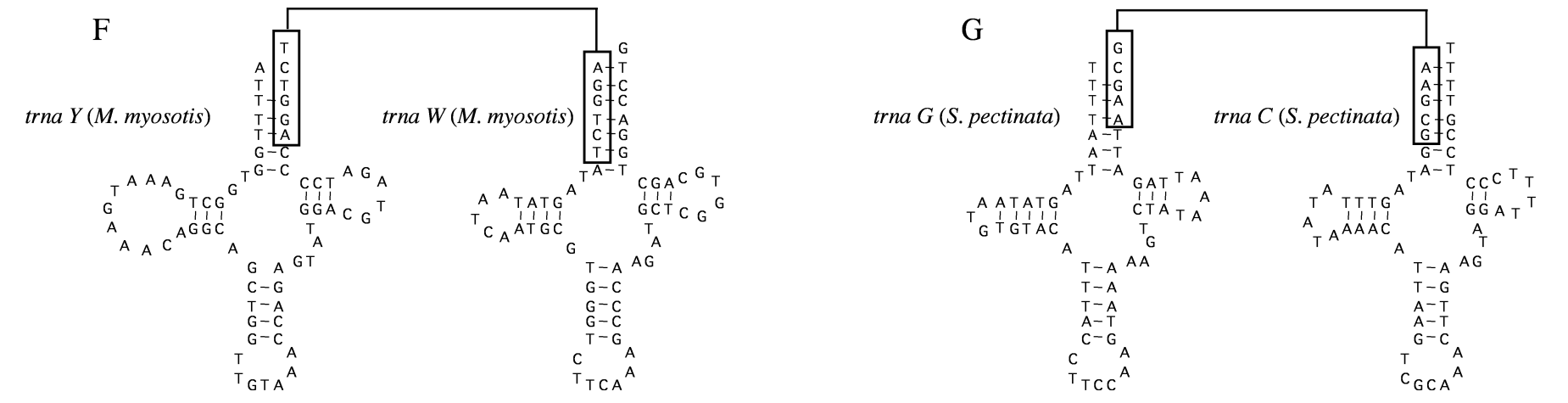
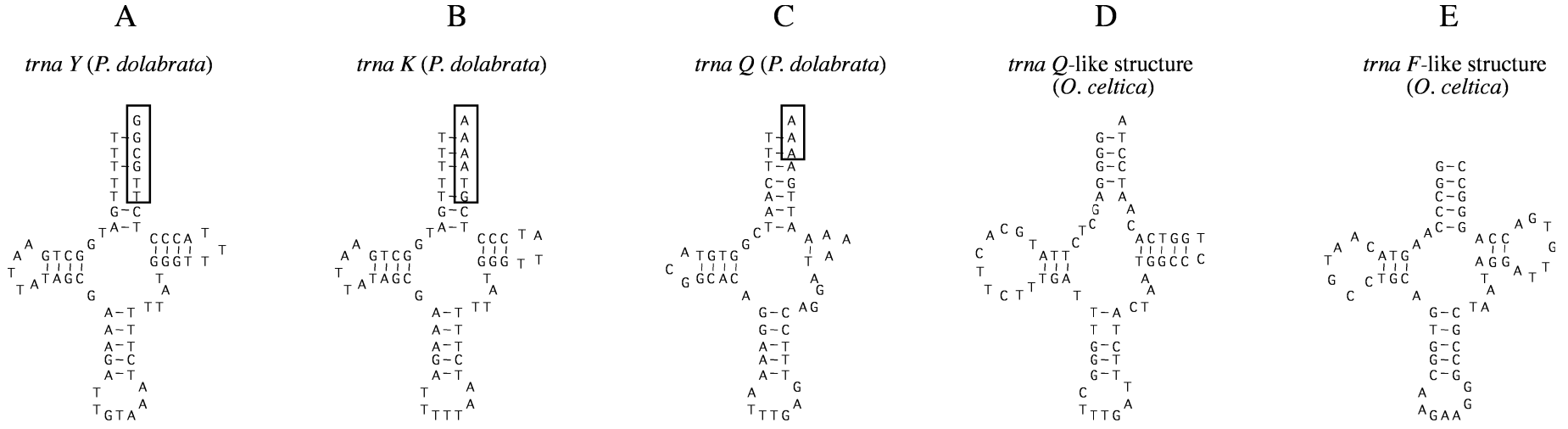
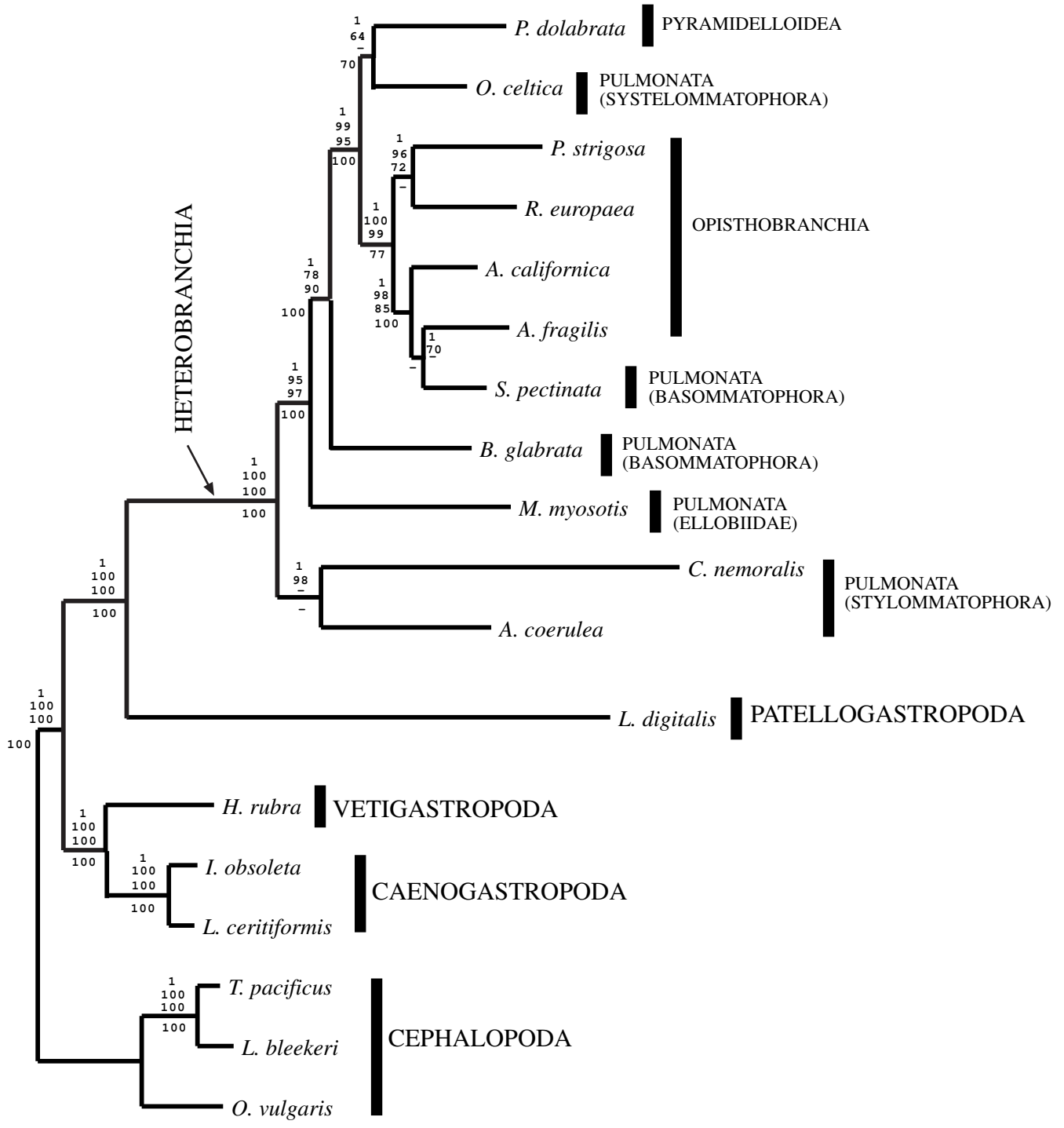


Figure 2



— 0.1 changes

