Molecular phylogenetic analysis of Pulmonata (Mollusca: Gastropoda) on the basis of Histone-3 gene

Vijaya Sai Ayyagari and Krupanidhi Sreerama*

Abstract

Background: In the present study, phylogenetic relationships within Heterobranchia in particular to Pulmonata were evaluated by means of Histone-3 (H3) gene sequence information. H3 gene is a slow evolving marker and is useful in resolving the deep level relationships. This is the first study to report the phylogeny of Pulmonata with more number of representatives from the group on the basis of H3 gene.

Results: The major groups within Heterobranchia viz. Lower Heterobranchia, Opisthobranchia, and Pulmonata were non-monophyletic. A few of the pulmonate groups’ viz. Planorbidae, Lymnaeidae, Siphonariidae, Veronicellidae, and Stylommatophora were recovered as monophyletic. The concepts of Eupulmonata and Geophila were not observed in the present study.

Conclusions: The present study was undertaken with an objective to study the phylogeny of Pulmonata reconstructed on the basis of H3 gene and its ability to resolve the deeper divergences in Pulmonata. However, the resolution at the deeper nodes is limited. There is a good resolution at the level of genera. In the future, inclusion of more number of taxa with increased sequence length of H3 marker may yield resolved topologies that may shed more insights into the phylogeny of Pulmonata.

Keywords: Gastropoda, Histone-3 gene, Phylogeny, Heterobranchia, Pulmonata

1 Background

Heterobranchia [1] is divided into Allogastropoda (=Lower Heterobranchia) and Euthyneura. The former is characterized by the presence of “Tri-ganglionate” state and the later (Euthyneura) by the “Penta-ganglionate” state by the acquisition of a pair of “Parietal ganglia” [1, 2]. The word “Euthyneura” was coined by Spengel [3] for uniting the taxa that reverted from the crossed state of visceral loop to uncross state during the process of detorsion. However, this is not considered as a synapomorphic character for “Euthyneura” due to the presence of this condition in few “Streptonerous” taxa [1]. Euthyneura is further divided into Opisthobranchia and Pulmonata. Each of these two groups is distinct on the basis of several synapomorphic characters of the nervous system [2]. Despite these synapomorphic characters, different molecular phylogenetic studies have yielded Lower Heterobranchia, Euthyneura, and Pulmonata as non-monophyletic [4–10].

Over the past few decades, with the advent of molecular techniques, there is growing information available with regard to the phylogenetic affinities among different groups of gastropods by means of widely used mitochondrial (16S, COI) and nuclear molecular markers (18S, 28S) in reconstructing their phylogenetic affinities. However, there is a need for new nuclear markers in studying gastropodan phylogenies especially with regard to deeper divergences. Histone-3 (H3) protein is one among the four highly conserved core nucleosome proteins that associate with DNA [11]. H3 gene, a nuclear marker is slow evolving when compared to the fast evolving mitochondrial markers like 16S and COI. Therefore, H3 gene is useful in studying the deep level relationships (for obtaining resolution at deeper nodes), while the widely used mitochondrial markers (16S, COI) are useful in resolving the relationships at the level of genera (Dinapoli et al. [12]). Colgan et al. [13] first used
H3 marker for studying the phylogenetic relationships among Arthropoda along with U2 snRNA gene.

With regard to the utility of H3 gene in the phylogenetics of Gastropoda, it was used either alone [12] or in combination with different molecular markers [14, 15]. In all these studies, taxa from Pulmonata were poorly represented. Further, there are no comprehensive phylogenetic studies available on Pulmonata on the basis of H3 gene or in combination with other molecular makers to date. Hence, the objective of the present study is to evaluate the phylogenetic affinities among different groups of Pulmonata by including more number of sequences available for major pulmonate groups from GenBank along with the sequences obtained from the present study by means of H3 sequence information. The present work is also carried out with a view to evaluate the utility of this marker to understand the phylogeny of Pulmonata, especially with regard to divergence of internal nodes.

2 Methods

2.1 Collection of specimens
Three different species of pulmonate gastropods viz., Laevicaulis alte, Gyraulus convexiusculus, and Lymnaea (Pseudosuccinea) acuminata form rufescens were collected for the present study. Laevicaulis alte was collected from the Selapadu village from the state of Andhra Pradesh, India. Gyraulus convexiusculus and Lymnaea (Pseudosuccinea) acuminata form rufescens were collected from the “Ganesha” pond located at Sri Sathyai Sai Prasanthi Nilayam Ashram, Andhra Pradesh, India. All the three specimens were sent to the Zoological Survey of India, Kolkata, for their identification.

2.2 Isolation of genomic DNA, PCR, and sequencing of H3 gene
The genomic DNA was isolated from the foot muscle tissue of snails using NucleoSpin® Tissue kit (Macherey-Nagel) by following the manufacturer’s instructions.

H3, a nuclear protein coding gene, was amplified using the primers H3F- ATGGCTCGTACCAAGCAGACVGC and H3R- ATATCCTTRGGCATRATRGTGAC [14]. The three species of gastropods were so far not sequenced for H3 gene and the analyzed sequences of this gene were submitted to DDBJ and NCBI GenBank and were assigned the following accession numbers LC198673, KX514447 - KX514448.

2.3 Multiple sequence alignment
Nucleotide sequences retrieved from GenBank were verified whether they were amplified by the same set of primers used in the present study or from the same regions of Histone 3 gene. After confirmation, nucleotide sequences obtained in the present study as well as those retrieved from NCBI GenBank (for few sequences) were trimmed at 5’ end in order to bring the sequences into reading frame. This ensures that stop codons are not present after the translation of nucleotide sequences to amino acids. Non-coding nucleotides were also removed from the 3’-end. Further, sequences obtained in the present study as well as those retrieved from NCBI GenBank were reverse complemented wherever the gene was in the complement (minus) strand.

Multiple sequence alignment was performed in MEGA v6.06 [17] using the inbuilt muscle program [18] with default parameters. Alignment of nucleotide sequences was guided by the translated amino acids. Regions of ambiguous alignments were excluded using online Gblocks server [19] under “less stringent” and “more stringent” options.

2.4 Model selection
Suitable substitution models for the datasets were determined using MrModeltest v2.3 [20] executed in PAUP* 4.0a152 [21] on the basis of Akaike Information Criterion (AIC). GTR+I+G model of nucleotide substitution was used for Dataset III in the present study.

2.5 Tests for substitution saturation
The following datasets were considered for the determination of substitution saturation using the entropy-
based test of Xia et al. [22] implemented in DAMBE v6.4.48 [23]. Apart from this, p-distances were plotted
against the distance obtained under GTR model in DAMBE for determining the saturation of a given loci.

1. Dataset I. Original unmasked sequence alignment
H3 gene was analyzed on the basis of its codons—codon 1, codon 2, codon 3, codons 1 & 2 combined, and codons 1, 2, & 3 combined. In Dataset 3, codon 1, codon 2, codons 1, 2, & 3 (combined), and codons 1 & 2 (combined) did not show saturation. For codon 3, only few taxa were saturated. Hence, the nucleotide sequences of codons 1, 2, & 3 were combined (dataset III) and subsequently used for phylogenetic analyses in the present study. The dataset III consisted of 231 nucleotides.

Plots of p-distances against the evolutionary distances obtained under GTR model for codons 1, 2, & 3 combined for all the three datasets showed saturation of transitions and showed similar pattern across all the three datasets. Similarly, for codons 1 & 2 combined for all the three datasets, transitions are saturated. Codon 3 of all the three datasets showed a high degree of substitution saturation for both the transitions and transversions.

2.6 Phylogenetic analyses
Bayesian inference (BI) of phylogeny (Fig. 1) and maximum likelihood (ML) analysis (Fig. 2) were carried out in MrBayes v3.2.5 [24] and raxmlGUI 1.5b2 [25] respectively under GTR+I+G model.

ML analysis was carried out for 1000 bootstrap replications. Bayesian analysis was carried out for 10 million generations with two independent runs. Each run consisted of four chains (3 heated and 1 cold) with default heating parameters. Sampling was done at the end of every 1000th generation with 25% of the initial trees discarded as burnin. The average standard deviation of the split frequencies at the end of 10 million generations was 0.006199. TRACER v1.6 [26] was used to determine the effective sample size (ESS) and it was found to be > 500.

The phylograms were visualized by importing the “.con” and “bipartitions.tre” files from MrBayes and RAxML respectively into iTOL v3 - http://itol.embl.de/ [27]. Apart from iTOL (Fig. 2), TreeGraph 2.14.0–771 beta was also used for annotating the phylogenetic tree (Fig. 1) [28]. In the subsequent section, bootstrap and posterior probability values were represented either as (1/100) which means a PP of 1 and a bootstrap value of 100 or in parenthesis next to the abbreviated form of the respective analysis or towards the end of a sentence discussing the same.

3 Results
3.1 Topology of the phylogenetic trees
ML tree (Fig. 2) was comparatively resolved than BI tree (Fig. 1). However, none of the deeper nodes in ML tree had bootstrap support compared to Bayesian tree. As a result, reliable inferences from ML tree could not be drawn regarding the overall relationships within Heterobranchia, except for a few bootstrap-supported external nodes. As a result, one should not consider the following conclusions inferred from the ML tree—paraphyly of Pulmonata, polyphyly of Opisthobranchia, and lower Heterobranchia.

From the topologies of the phylogenetic trees recovered in the present study, it was not possible to clearly assign the monophyly of lower Heterobranchia, Opisthobranchia, Pulmonata, and Euthyneura. The concept of Euthyneura seems to be non-existent as lower Heterobranchia was not basal in ML and BI trees.

3.2 Lower Heterobranchia
A total of five lower Heterobranch taxa were included in the present study.

*Pupa solidula*, *Turbonilla sp*, *Rictaxis punctocaelatus*, and *Micromelo undata* formed a monophyletic clade in BI (0.98) and ML (63) phylograms. This emerged as a sister clade to *Lymnaeidae* and *Planorbidae* with moderate support in BI (0.62). *Acteonoidae* recovered as non-monophyletic in BI and ML trees.

*Orbitestella sp* formed a monophyletic clade with an Opisthobranch *Haminoea hydatis* in both BI (0.99) and ML (56) analyses.

None of the lower Heterobranch taxa emerged basal in ML and BI phylograms.

3.3 Opisthobranchia
Of all the Opisthobranch taxa included in the present study, *Acanthodoris sps* formed a monophyletic clade. The relationships of the remaining Opisthobranch taxa were not clearly resolved.

*Alpinia: Akanthodoris brunnea* formed a basal clade in BI and ML trees. In BI tree, it formed a monophyletic clade with the outgroup taxa *Pila conica* (Caenogastropoda).

*Saccoglossa*: *Plakobranchus ocellatus* was the next basal taxon in BI (0.71) and ML trees.

*Cephalaspidea*: *Haminoea hydatis* formed a monophyletic clade with a lower Heterobranch taxon in BI (0.99) and ML (56) trees.

*Nudipleura*: *Anadoridoidea*: *Acanthodoris brunnea* and *Acanthodoris lutea* formed a monophyletic clade in BI (1) and ML (100) analyses and emerged as a sister clade to *Ellobiidae*.

*Bathydoridoidea*: *Bathydoris clavigera* formed a monophyletic clade with *Tylodina perversa* (Umbraculoidea - Tylodinidae) in ML (56) tree.

3.4 Pulmonata
*Siphonariidae*: *Siphonaria concinna* and *Siphonaria denticulata* formed a monophyletic clade in BI (0.94) and ML analyses.
**Hygrophila:**

*Lymnaeidae:* Two species viz., *Lymnaea acuminata* and *Lymnaea stagnalis* formed a monophyletic clade in BI and ML trees with strong node support values (0.99/89).

**Planorbidae:** Planorbidae represented by five species formed a monophyletic clade with strong support both in BI (1) and ML (89) trees.

*Lymnaeidae* and *Planorbidae* in turn formed a monophyletic clade in BI (0.84) and ML trees. This clade emerged as a sister clade to Lower Heterobranchia.

**Eupulmonata:**

*Ellobiidae:* Of the four Ellobiids *Pedipes jouani, Pedipes ovalis, Carychiunimum,* and *Ophicardelus ornatus* included in the present study, *Pedipes jouani* and *Pedipes ovalis* formed a monophyletic clade in BI (0.98) and ML (82) phylograms. They emerged as a sister clade to Opisthobranch taxa *Acanthodoris brunnea* and *Acanthodoris lutea* in BI (0.64) and ML trees.

*Carychiunimum, Ophicardelus ornatus,* and *Salinator solida* (Amphiboloidea) together formed a monophyletic clade in BI (0.55) and ML trees. These two clades were differently positioned in the phylogenetic trees.
Systellommatophora:

Onchidiidae: Of the six Onchidiids included in the present study, *Onchidium sp* emerged as a basal taxon in both BI (0.80) and ML analyses. Remaining five taxa formed a monophyletic clade in BI (1) and ML (91) trees.

Veronicellidae: *Laevicaulis alte* and *Veronicella cubensis* formed a monophyletic clade in both BI (1) and ML (76) analyses.

Most importantly, Onchidiidae and Veronicellidae did not group together in BI and ML phylograms to form a monophyletic Systellommatophora. Each of the two families was positioned differently in ML and BI trees.

Stylommatophora: Stylommatophora recovered monophyletic in both BI (0.52) and ML analyses.

4 Discussion

In the present study, H3 sequence information was used to assess the phylogenetic relationships within Heterobranchia, in particular to Pulmonata. In the present study, H3 gene sequences of taxa belonging to major pulmonate groups’ viz. Planorbidae, Ellobiidae, Onchidiidae, Veronicellidae, and Stylommatophora (represented by Helicidae and Orculidae) were retrieved from GenBank and included along with the sequences obtained in the present study for studying the phylogenetic relationships among pulmonates.

Relationships among Opisthobranchia, Lower Heterobranchia, and Pulmonata could not be determined from the topology of the trees recovered. ML tree is better resolved compared to Bayesian tree. However, none of the inner nodes were bootstrap supported. In ML and Bayesian trees, there was no step wise or progressive arrangement of taxa from Lower Heterobranchia to Pulmonata.

A few of the pulmonate groups’ viz. Planorbidae, Veronicellidae, and Orculidae recovered monophyletic with node support in both BI and ML trees. Stylommatophora formed a monophyletic clade in both ML and BI trees with high node support seen only for BI tree. Systellommatophora is non-monophyletic. It is interesting to observe that *Onchidium sp* emerged as a basal taxon from the rest of the Onchidiids in ML and BI trees. This is similar to the results obtained in the study of Dinapoli et al. [12] where *Onchidium sp* emerged separately from the remaining onchidiids.

The monophyly of Eupulmonata consisting of Systellommatophora and Stylommatophora and also the concept of Geophila [29] was not observed from the topologies of ML and BI trees as Stylommatophora and...
Systellommatophora were positioned differently in the phylograms. Taxa from Lower Heterobranchia emerged as a sister clade to Pulmonata and Opisthobranchia.

Further, the gene sequence information for H3 marker is not available for several taxa when compared to 16S, 18S, 28S, and COI. A simple search in the international nucleotide databases (DDBJ, NCBI, and EMBL) shall reveal this fact. As an example, Systellommatophora consists of three families viz. Onchidiidae under the superfamly Onchidioidae, Veronicellidae and Rathouisiidae under the superfamly Veronicelloidea. As at April 2019, there is only one sequence available for Veronicella cubensis belonging to the family Veronicellidae in NCBI GenBank. Hence, Veronicellidae is under-represented and the nucleotide sequence of H3 gene obtained from the present study for Lymnaea acuminata belonging to this family was reported for the first time. Similarly, there is only one Histone-3 gene sequence available for Lymnaea stagnalis under “Lymnaeidae/ Lymnaeidae”. In the present study, Lymnaea acuminata of this family was sequenced and this report is also the first of its kind.

Colgan et al. [14, 15] have utilized H3 gene in combination with other marker(s) for studying the phylogeny of Gastropoda. In their studies, pulmonates were poorly represented and consisted of only five pulmonates. Whereas, the present study included a large number of pulmonate taxa representing major groups in Pulmonata. In contrast to the results obtained in the present study, Euthyneura recovered monophyletic in these two studies. In the study of Colgan [14], within Euthyneura, opisthobranchs were paraphyletic with respect to pulmonates. When H3 gene alone was used, it resulted in an unresolved topology (polytomy). In the study of Colgan [15], Lower Heterobranchia (“Heterostropha”) and Heterobranchia are rarely monophyletic and monophyly of Opisthobranchia could not be determined due to tree topologies. This is more or less similar to the results obtained in the present study, where Lower Heterobranchia and Opisthobranchia did not emerge monophyletic.

Dinapoli et al. [12] had utilized the H3 gene alone in reconstructing the phylogeny of Heterobranchia. Unlike the present study, her study consisted of only nine pulmonates and most of the pulmonate groups were not represented. The phylograms lacked resolution as a result of which the monophyly of Pulmonata and Opisthobranchia were not obtained. Similar to the results obtained in the study, there is a good resolution at the level of genera.

5 Conclusion
The present study, for the first time reported the phylogeny of Pulmonata with increased number of taxa on the basis of H3 gene. This is in contrast to most of the phylogenetic studies, which focused on few of the widely used markers viz. 16S, 18S, 28S, and COI. Further, in order to obtain accurate results on the phylogenetic relationships among Heterobranchia and in particular to Pulmonata, there is a need for the utility of new molecular markers like H3 gene; particularly to study the divergences of deep nodes. However, phylogenies on the basis of H3 gene lacked resolution owing to high grade of conservation and high codon usage bias that results in partially obscured phylogenetic signal [12]. In the present study, a few of the major groups in Pulmonata recovered monophyletic. In the future, sequencing more number of taxa for H3 gene from different groups of Pulmonata, with increased sequence lengths may shed deeper insights into the phylogeny of Pulmonata and Heterobranchia.

Abbreviations
BI: Bayesian inference; ESS: Effective sample size; H3: Histone-3; ML: Maximum likelihood; NCBI: National Centre for Biotechnology Information

Acknowledgements
Authors acknowledge Drs. Tulika Biswas and Amit Kumar Mukhopadhyay from Zoological Survey of India, Kolkata, for identifying the gastropod specimens.

Authors’ contributions
VSA conceived the study, carried out experiments, performed phylogenetic analyses, and was a major contributor in drafting the manuscript. KS obtained funding, drafted, and corrected the manuscript. All authors read and approved the final manuscript.

Funding
This work was funded by the Science and Engineering Research Board (SERB), Department of Science and Technology (DST), Government of India (Sanction order SB/SO/AS-138/2012 dated 30 October 2013) for carrying out experimental work and also the “Junior Research Fellowship” provided to the first author for carrying out the research work.

Availability of data and materials
Nucleotide sequences were deposited to GenBank and DDBJ and are available online. The rest of the data, if any, will be provided by authors upon request.

Ethics approval and consent to participate
Not Applicable

Consent for publication
Not Applicable

Competing interests
The authors declare that they have no competing interests.

Received: 1 July 2019 Accepted: 30 September 2019
Published online: 18 November 2019

References
1. Haszprunar G (1985) The Heterobranchia – a new concept of the phylogeny of the higher Gastropoda. J Zool Syst Evol 23:15–37
2. Haszprunar G, Huber G (1990) On the central nervous system of Smeagolidae and Rhodopidae, two families questionably allied with the Gymnomorpha (Gastropoda: Euthyneura). J Zool 220:185–199
3. Spengel JW (1881) Die Geruchssorganen und das Nervensystem der Mollusken. Zeitschrift für Wissenschaftliche 35:333–383
4. Dayrat B, Conrad M, Balayan S, White TR, Albrecht C, Golding R et al (2011) Phylogenetic relationships and evolution of pulmonate gastropods
5. Dinapoli A, Klussmann-Kolb A (2010) The long way to diversity – phylogeny and evolution of the Heterobranchia (Mollusca: Gastropoda). Mol Phylogenet Evol 55:60–76
6. Dinapoli A, Zinssmeister C, Klussmann-Kolb A (2011) New insights into the phylogeny of the Pyramidellidae (Gastropoda). J Molluscan Stud 77:1–7
7. Grande C, Templado J, Cervera JL, Zardoya R (2004) Molecular phylogeny of Euthyneura (Mollusca: Gastropoda). Mol Biol Evol 21:303–313
8. Grande C, Templado J, Zardoya R (2008) Evolution of gastropod mitochondrial genome arrangements. BMC Evol Biol 8:1–15
9. Klussmann-Kolb A, Dinapoli A, Kuhn K, Streit B, Albrecht C (2008) From sea to land and beyond – new insights into the evolution of Euthyneuran Gastropoda (Mollusca). BMC Evol Biol 8:1–16
10. Thollesson M (1999) 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 Biol Sci 266:75–83
11. Marson R, Cohn R, Kedes L (1983) Expression and organisation of histone genes. Annu Rev Genet 17:239–277
12. Dinapoli A, Tamer C, Franssen S, Naduvilezhath L, Klussmann-Kolb A (2006) Utility of H3-gene sequences for phylogenetic reconstruction – a case study of heterobranch Gastropoda. Bonn Zool Beitr 55:191–202
13. Colgan DJ, McLauchlan A, Wilson GDF, Livingston SP, Edgecombe GD, Macaranas J et al (1998) Histone H3 and U2 snRNA DNA sequences and arthropod molecular evolution. Aust J Zool 46:419–437
14. Colgan DJ, Ponder WF, Eggler PE (2000) Gastropod evolutionary rates and phylogenetic relationships assessed using partial 28S rDNA and histone H3 sequences. Zool Sci 29:29–63
15. Colgan DJ, Ponder WF, Beacham E, Macaranas JM (2003) Gastropod phylogeny based on six segments from four genes representing coding or non-coding and mitochondrial or nuclear DNA. Molluscan Res 23:123–148
16. DNA Baser Sequence Assembler v4. 2013. Heracle BioSoft. www.DnaBaser.com/. Accessed June 2016
17. Tamura K, Stecher G, Peterson D, Filipski A, Kumar S (2013) MEGA6: molecular evolutionary genetics analysis version 6.0. Mol Biol Evol 30:2725–2729
18. Edgar RC (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res 32:1792–1797
19. Castresana J (2000) Selection of conserved blocks from multiple alignments for their use in phylogenetic analyses. Mol Biol Evol 17:540–552
20. Nylander JAA (2004) MrModeltest v2.3 program distributed by the author. Evolutionary Biology Centre, Uppsala University
21. Swoford DL (2002) PAUP*: Phylogenetic analysis using parsimony (*and other methods). Version 4. Sinauer associates, Sunderland, Massachusetts
22. Xia X, Xie Z, Salemi M, Chen L, Wang Y (2003) An index of substitution saturation and its application. Mol Phylogenet Evol 26:1–7
23. Xia X (2013) DAMBES: a comprehensive software package for data analysis in molecular biology and evolution. Mol Biol Evol 30:1720–1728
24. Ronquist F, Teslenko M, Van der Mark P, Ayres DL, Darling A, Höhna S et al (2012) MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. Syst Biol 61:539–542
25. Silvestro D, Michalak I (2012) raxmlGUI: a graphical front-end for RAxWL. Org Divers Evol 12:235–237
26. Rambaut, A., Suchard, M.A., Xie, W., Drummond, A.J. (2014) Tracer v1.6. Available from: http://beast.bio.ed.ac.uk/Tracer/. Accessed May 2017
27. Letunic I, Bork P (2016) Interactive tree of life (iTOl) v3: an online tool for the display and annotation of phylogenetic and other trees. Nucleic Acids Res 44(W2):W242–W245
28. Sñever B, Muller KF (2010) TreeGraph 2: combining and visualizing evidence from different phylogenetic analyses. BMC Bioinformatics 11:7
29. Fénusac JB (1819) Histoire Naturelle des Pulmonés sans Opercules. Arthus-Gertrand, Paris

Submit your manuscript to a SpringerOpen journal and benefit from:

- Convenient online submission
- Rigorous peer review
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

Submit your next manuscript at ➤ springeropen.com