

LBL-59179

Molluscan Evolutionary Genomics

W. Brian Simison and Jeffrey L. Boore

Molluscan Evolutionary Genomics

W. Brian Simison and Jeffrey L. Boore

Evolutionary Genomics Department, DOE Joint Genome Institute and Lawrence
Berkeley National Laboratory, 2800 Mitchell Drive, Walnut Creek, CA 94598

AND

Department of Integrative Biology, 3060 Valley Life Sciences Building, University of
California, Berkeley, CA 94720

I Introduction

In the last 20 years there have been dramatic advances in techniques of high-throughput DNA sequencing, most recently accelerated by the Human Genome Project, a program that has determined the three billion base pair code on which we are based. Now this tremendous capability is being directed at other genome targets that are being sampled across the broad range of life. This opens up opportunities as never before for evolutionary and organismal biologists to address questions of both processes and patterns of organismal change. We stand at the dawn of a new “modern synthesis” period, paralleling that of the early 20th century when the fledgling field of genetics first identified the underlying basis for Darwin’s theory. We must now unite the efforts of systematists, paleontologists, mathematicians, computer programmers, molecular biologists, developmental biologists, and others in the pursuit of discovering what genomics can teach us about the diversity of life.

Genome-level sampling for mollusks to date has mostly been limited to mitochondrial genomes and it is likely that these will continue to provide the best targets for broad phylogenetic sampling in the near future. However, we are just beginning to see an inroad into complete nuclear genome sequencing, with several mollusks and other eutrochozoans having been selected for work about to begin. Here, we provide an overview of the state of molluscan mitochondrial genomics, highlight a few of the discoveries from this research, outline the promise of broadening this dataset, describe upcoming projects to sequence whole mollusk nuclear genomes, and challenge the community to prepare for making the best use of these data.

II Mollusk Mitochondrial Genomes: Features, Discoveries, and Prospects

The typical animal mitochondrial genome is about 15,000 base pairs (bp) in size and contains 37 genes: 13 coding protein genes, 2 ribosomal RNA genes, and 22 tRNAs (Boore 1999). In addition to the genes, there is often one large non-coding region that is thought to contain the signals for controlling transcription and replication, although these signals, if they exist, are not well conserved. For some animals all genes are on one strand, but for others they are divided between both. Most animal mtDNAs are A+T-rich and there is often an asymmetry in nucleotide composition between the two strands, with one being more rich in G and T at the expense of C and A. In the few cases where it has been studied, all genes on each strand are expressed as a single polycistronic RNA which is then enzymatically cleaved to yield gene-specific messages (Ojala et al. 1980, 1981). Mitochondrial DNA (mtDNA), in general, is maternally inherited and therefore, does not experience allelic segregation.

There are several features that make it feasible and important to sample mitochondrial genomes across a broad phylogenetic range. Since they are closed circular, supercoiled, extrachromosomal DNAs, they can be physically isolated from the nuclear genome. They are known to play important roles in cellular metabolism, genetic disease, cellular processes such as apoptosis, and embryological development and their biochemistry is relatively well understood. Their diminutive size facilitates sampling, yet they contain many elements of a complete genomic system, including genes for all three primary transcript types (protein, tRNA, rRNA). They have been used extensively in

studies of population structure, conservation biology, and forensics. Many aspects of genome evolution are being modeled in these systems, including the role of biased nucleotide mutations on amino acid substitution patterns, codon usage changes, and mechanisms of gene order rearrangement.

Even the limited sampling of complete mollusk mtDNA sequences has revealed that they have many unusual characteristics. Some mollusks have very large mtDNAs, up to 42 kb. There is both loss and gain of genes. Although some have been stable in gene arrangement over long periods of evolutionary time, others have rearranged nearly every gene. At least some bivalves have a bizarre exception to maternal inheritance that has been dubbed “doubly-uniparental inheritance” whereby females have one type of mitochondrial genome and males have, in their gonads, another haplotype that differs by as much as NN% in sequence and can vary even in gene arrangement; male somatic tissues have a mixture of the two types, but only the male-specific haplotype is passed on to sons. This rich variation in mitochondrial features makes mollusks a great model system for understanding general principles of genome evolution and mitochondrial molecular biology.

Most molecular phylogenetic studies of mollusks to date have focused on the comparisons of a handful of partial nuclear and mitochondrial gene sequences, typically using from less than 1,000 to a few thousand bp per organism, and most commonly have used the sequences of small subunit (i.e., 18S) rRNA encoded by the nucleus. In

general, these studies have failed to demonstrate robust monophyly for any of the major classes except polyplacophorans (Winnepennickx et al, 1996) and to yield poorly resolved or contradictory phylogenetic results for the mollusks and for the broader Eutrochozoa. Studies using complete mitochondrial genomes have over 15 kb of sequence per organism with portions varying to different degrees providing different phylogenetic signals.

For most lineages, mtDNA has a higher substitution rate than nuclear DNA and thus is well suited to resolving relationships among closely related taxa, although has provided strong signal even for even very deep divergences.

Gene position can also be compared as a source of phylogenetic information (Boore and Brown, 1998). Several studies have documented that mollusks, unlike vertebrates, have a great deal of variation in gene order and content (Boore, 2004). This variation appears to offer useful phylogenetic data that may be advantageous for reconstructing molluscan evolutionary relationships. Complete mitogenomes offer complete ribosomal genes that can be folded into their natural secondary structures. Secondary structure has often been used in phylogenetic studies as a way to improve ribosomal sequence alignments, but it may also be used to generate additional phylogenetic characters by coding the secondary structure as a morphological entity (Lydeard et al. 2000).

III Mollusk Mitochondrial Genomes: Techniques

The most commonly used process of producing a complete mitochondrial genome sequence uses “shotgun sequencing” and involves six steps: 1) extraction of DNA from tissue; 2) PCR of long, overlapping fragments summing to the complete mtDNA; 3) creation of genomic libraries; 4) sequencing numerous, random clones from the libraries; 5) assembling the random sequencing reads into a complete mtDNA sequence; 6) gene annotation. The first step, DNA extraction, has historically presented numerous obstacles for malacologists due to the presence of mucopolysaccharides, which inhibit subsequent PCR, or the difficulty in obtaining sufficient high-quality mitochondrial DNA. Today, however, DNA extraction has become fairly straightforward procedure with several dependable kits available on the market for producing total cellular DNA (i.e., including both nuclear and mitochondrial genomes).

The next step is the most variable for success and often requires care and persistence. Primers must be designed to amplify overlapping regions of the circular DNA thereby ensuring complete coverage of the genome. While this sometimes works acceptably using primers designed to match conserved portions of the mtDNA, in other cases it is necessary to first amplify and sequence a small fragment in order to

design primers that are perfectly matched for amplifying long fragments (which, in general, requires more optimizing of reaction conditions).

Successful PCR products are broken into smaller fragments that are then ligated into plasmids to create a clone library. Because restriction enzyme digestion yields non-random fragments and sonication requires large amounts of input DNA, the method of choice is to shear the amplified DNA into small fragments (typically 1.5-3 kb) by passage through a narrow aperture using a Hydroshear device. These fragments are then inserted into plasmid vectors, which are then transformed into bacterial cells. These are incubated on a growth media with selection for plasmid integration where they multiply into colonies of large numbers of cells each making exact copies of the inserted DNA fragment. A random selection of clones is chosen and processed for sequence determination from each end of each inserted DNA fragment; see for protocols of how this is done at JGI.

Of course, there are other alternatives. Instead of using long-PCR amplification for the mtDNA, templates could be created by physical isolation of the mtDNA using ultracentrifugation or alkaline denaturation followed by differential reannealing. Instead of PCR amplification, the mtDNA could be cloned into phage vectors following by isolating DNA from this large insert clone. Products may be produced using rolling circle amplification (RCA) instead of PCR and work in progress may be able to clone mtDNA from a preparation of total DNA using transposon bombardment. As an alternative to sequencing random clones, some follow a strategy of primer-walking through long-PCR

amplified products. These and other alternative procedures are reviewed in Boore, Macey and Medina (2004).

The randomly determined sequences, each typically of ~700 bp, are assembled based on their overlapping regions. This is usually accomplished with specialized software like the Phred/Phrap/Consed suite. Although such software guides and streamlines the process, there is no substitute for human verification of proper assembly and correctness of sequence based on the extent of overlap, depth of coverage, agreement of reads, assigned quality scores, and appearance of chromatograms.

Annotation requires the identification of each gene based on similarity of sequence to those of other organisms and/or the potential for folding into specific secondary structures (for rRNA and tRNA genes). This is performed by using a variety of tools, including BLAST, specialized online annotators like DOGMA (Wyman, Jansen and Boore 2004) (<http://bugmaster.jgi-psf.org/dogma/>) and tRNAscan (<http://www.genetics.wustl.edu/eddy/tRNAscan-SE/>), and sequence processing tools such as MacVector (Accelrys).

IV Current Data Production in Mollusk Mitogenomics

Recent increases in the ability to sequence complete mitochondrial genomes has expanded their phylogenetic utility. Unfortunately, even for these diminutive genomes, malacology lags behind other disciplines such as mammalogy, herpetology, and entomology in data generation. For example, of the NNN complete animal mitochondrial genome sequences in Genbank as of June, 2005, only 13 (VERIFY) are mollusks, whereas NNN are vertebrates and NNN are insects. On the bright side, there

are currently at least four groups actively sequencing complete molluscan mitochondrial genomes:

(1) JLB leads the Evolutionary Genomics program at the DOE Joint Genome Institute (JGI), which includes mitogenomic researchers Monica Medina and J. Robert Macey (and others). Within this group, WBS is targeting deep level sampling of gastropods and has finished mtDNA sequencing for representatives of the caenogastropods, vetigastropods, and patellogastropods and is targeting a nerite and a heterobranch (all in collaboration with David Lindberg's group at Berkeley). Medina leads an effort that has completed 13 heterobranch mtDNAs. Previous publications from members of this group have included several mollusks, including a chiton (Boore and Brown, 1994), a bivalve, and a scaphopod. This is part of a broad sampling of mitochondrial genomes, including cnidarians, a brachiopod, a phoronid, bryozoans, a rotifer, chaetognaths, a tapeworm, nemertines, nematodes, a gastrotrich, crustaceans, myriapods, a horseshoe crab, apterygote hexapods, a pentastomid, annelids, an echiuran, a sipunculid, arachnids, echinoderms, a cephalochordate, sharks, salamanders, frogs, turtles, snakes, lizards, amphisbaenians, and primates. This research group is also developing software to streamline gene annotations and databases for mitogenomic comparisons.

(2) Rei Ueshima and colleagues at the University of Tokyo has sequenced X complete mitochondrial genomes of various molluscan groups.

(3) Timothy Collins at the Florida International University and his colleagues Timothy Rawlings and Rüdiger Bieler are using snail mitochondrial DNA as a model system in which to understand mechanisms and rates of mitochondrial gene order rearrangement and the consequences of gene order rearrangement for sequence-based phylogenetic analysis. They have focused in particular on the apple snails and vermetid worm snails in which gene rearrangements appear to be occurring at an accelerated pace (Rawlings et al., 2001; 2004). They are pursuing a strategy of sequencing portions of genomes that appear to be rearrangement hotspots, as well as select complete genomes.

(4) Gerhard Steiner at the University of Vienna is currently leading a 3-year mitogenomics project funded by the FWF (Austrian Science Fund) entitled "Mitochondrial Gene Arrangements of Selected Bivalves (Mollusca)." The group includes Dr. Hermann Dreyer (PostDoc), Martina Knapp and Miriam Satler (Diploma Students) and is primarily focused on bivalves, but is also sequencing other taxa from Scaphopoda, Aplacophora, and Polyplacophora. The group is investigating the high degree of gene rearrangements and doubly uniparental inheritance found in bivalves. They have finished five bivalves, one scaphapod and one chiton.

V Mollusk Nuclear Genomes

While mitochondrial genomes of mollusks are known to range in size from ~10 to ~42 kb, nuclear genomes are very much larger, ranging in size from ~400,000 kb for *Lottia gigantea* to ~5,900,000 kb for *Neobuccinum eatoni*. Although cytogenetic studies of chromosome number and structure have been used since the late nineteenth century for inferring molluscan phylogeny, the results have proven generally to be unreliable (Patterson, 1969; Nakamura, 1986; Thiriot-Quievreux, 2003). However, this classic literature records data from many hundreds of molluscan species.

We queried the BIOSIS¹ bibliographic database using a variety of keyword searches to show that malacology in general lags far behind the study of vertebrates or arthropods, and especially so in the fields of genomics and molecular evolution. There are approximately 130,000 extant mollusk species, 40,000 extant vertebrate species and probably over 9,000,000 arthropod species. Yet keyword searches on Mollusca, Vertebrata, and Arthropoda returns 124,968, 10,337,610 and 706,928 publications, respectively. That makes mollusk publications only 1.2% of vertebrate publications and 17.6% of arthropod publications. Searches that combine “Mollusca” or “Vertebrata” with specific molecular keywords² returns 7,403 and 976,258 publications, respectively, indicating that molluscan molecular publications are only 0.7% of vertebrate molecular publications. To see how malacological research is keeping up with the latest trends we performed a combined search of Mollusca with “Bayesian”, a word associated with a set of recent innovations in maximum likelihood estimations of phylogeny (Huelsenbeck and

¹ BIOSIS does not include all biological publication records and searches of its database do not reflect actual numbers of publications. Therefore all comparisons made are intended to illustrate relative trends in the fields of study.

² A combination of keywords that include DNA, molecular, mitochondria, nuclear, genome, ribosome, ribosomal, 16S, 12S, NAD, cytochrome oxidase.

Ronquist, 2001) and compared that to the same combined search for vertebrates. Mollusca returned only eight publications while vertebrates returned 350. Finally, we compared the keyword “genome” with “Mollusca” and “Vertebrata” which returned 290 and 45,544 publications, respectively. This BIOSIS analysis reveals that not only is malacology underrepresented in the biological sciences compared with the study of vertebrates and arthropods, it is falling further behind as new fields are developing.

Evolutionary genomics is a rapidly developing field and is rejuvenating the field of evolutionary biology. Malacology has an opportunity to exploit this burgeoning field and take the lead in studying evolution and natural selection in a new light using the new tools offered by evolutionary genomics. The primary source of opportunity here is the great diversity of mollusks to represent revolutions of animal bauplan.

Mollusks are abundant in marine, terrestrial, fresh water, and extreme environments, they are generally easy to observe and collect, and they are an extremely old and diverse group of organisms with an outstanding fossil record. Malacologists should consider integrating evolutionary genomics into their research programs whenever relevant to take advantage of this seminal field. If malacologists act

now and get funded for seminal research in evolutionary genomics, we will establish malacology as a critical component of the field and will lay the foundation for more funding and thus expanded recruiting of new students and eventually, an expanded presence of malacologists in biological sciences.

The most important goal for the future of malacology is for the research community to establish itself as an integral component of evolutionary genomics. The primary funding agencies such as the National Science Foundation, National Institute of Health, and the Department of Energy are looking for cooperation within research fields and encourage collaboration and broader participation. The cost of genome sequencing is still very high and prevents high throughput sequencing from all but a few well-funded projects. If individual researchers remain isolated from the rest of the malacological community and we do not form cooperative research objectives, malacology will fall further behind other fields and miss out on a remarkable opportunity to become a leader in a new and important field.

VI Conclusions

Determining complete mtDNA sequences has proceeded more slowly for mollusks than for some other phyla, notably vertebrates and arthropods. (However, there are other big and important groups that remain underrepresented, such as annelids, and numerous phyla even remain unsampled.) Even with this limited sampling, the field of molluscan mitogenomics is particularly promising. Gene rearrangements are common for some groups, bolstering confidence that this will provide a rich phylogenetic signal at some level and offering a dataset that may illuminate models of gene order change.

Mollusks themselves are highly diverse and the molecular mechanisms apparent in their mitochondrial systems seem to follow suit. Recent successes of several research groups suggest that this dataset is growing exponentially. The field of mitogenomics is at a formative stage and the time is now to establish mollusks as a model system for addressing many of the processes and patterns of genomic change in this model system.

Further, the imminent determination of the complete nuclear genome sequences of several mollusks and other eutrochozoans holds extraordinary promise. In addition to the enormous amount of phylogenetic information that will be made available, this will also reveal the genomic underpinnings of the developmental and morphological characters and that comprise the bulk of our traditional systematic studies.

Today's high-throughput genome sequencing centers lack the resources to comprehensively analyze anything but a small portion of these issues. It will take the devoted effort of a broad community with a wide spectrum of expertise to produce the most relevant and exciting biological insight. The community of malecologists must be ready.

ACKNOWLEDGEMENTS

This work was performed partly under the auspices of the US Department of Energy's Office of Science, Biological and Environmental Research Program, and by the University of California, Lawrence Berkeley National Laboratory under Contract No.

DE-AC02-05CH11231. This work was supported by NSF grant.

REFERENCES

- Boore, J. L. Animal mitochondrial genomes. *Nucleic Acids Res.* 27, 1767-1780 (1999).
- Boore, J. L. and Brown, W. M. Complete DNA sequence of the mitochondrial genome of the black chiton, *Katharina tunicata*. *Genetics* 138 (2), 423-443 (1994).
- Boore, J. L., J. R. Macey and M. Medina, 2005 Sequencing and comparing whole mitochondrial genomes of animals. In *Molecular Evolution: Producing the Biochemical Data, Part B* (E. A. Zimmer and E. Roalson, eds.). Volume 395 of the *Methods in Enzymology* series, Elsevier, Burlington, Massachusetts, pp. 311-348.
- Boore, J. L., Medina, M. and Rosenberg, L. A. Complete sequences of the highly rearranged molluscan mitochondrial genomes of the scaphopod *Graptacme eborea* and the bivalve *Mytilus edulis*. *Mol. Biol. Evol.* 21 (8), 1492-1503 (2004).
- DeJong, R. J., Emery, A. M. and Adema, C. M. The mitochondrial genome of *Biomphalaria glabrata* (Gastropoda: Basommatophora), intermediate host of *Schistosoma mansoni*. *J. Parasitol.* 90 (5), 991-997 (2004).
- Dreyer, H. and Steiner, G. The complete sequence and gene organization of the mitochondrial genome of the gadilid scaphopod *Siphonodontalium lobatum*. (Mollusca). *Mol. Phylogenet. Evol.* 31 (2), 605-617 (2004).
- Grande, C., Templado, J., Cervera, J. L. and Zardoya, R. The complete mitochondrial genome of the nudibranch *Roboastra europaea* (Mollusca: Gastropoda) supports the monophyly of opisthobranchs. *Mol. Biol. Evol.* 19 (10), 1672-1685 (2002).

- Hatzoglou, E., Rodakis, G. C. and Lecanidou, R. Complete sequence and gene organization of the mitochondrial genome of the land snail *Albinaria coerulea*. Genetics 140 (4), 1353-1366 (1995).
- Hoffmann, R. J., Boore, J. L. and Brown, W. M. A novel mitochondrial genome organization for the blue mussel, *Mytilus edulis* Genetics 131 (2), 397-412 (1992).
- Kim, S. -H., Je, E. -Y. and Park, D. -W. *Crassostrea gigas* mitochondrial DNA. Unpublished.
- Kurabayashi, A. and Ueshima, R. Complete sequence of the mitochondrial DNA of the primitive opisthobranch gastropod *Pupa strigosa*: systematic implication of the genome organization. Mol. Biol. Evol. 17 (2), 266-277 (2000).
- Lydeard, C, Holznagel, W. E., Schnare, M. N. and Gutell, R. R. Phylogenetic analysis of molluscan mitochondrial LSU rDNA sequences and secondary structures. Mol Phylogenet Evol. 15, 83-102 (2000).
- Maynard, B. T., Kerr, L. J., McKiernan, J. M., Jansen, E. S. and Hanna, P. J. Mitochondrial DNA sequence and gene organization of the Australian blacklip abalone, *Haliotis rubra* (Leach). Unpublished.
- Nahir, B., Kohn, A. B. and Moroz, L. L. Complete DNA sequence of the mitochondrial genome of the marine opisthobranch mollusc, *Aplysia californica*. Unpublished.
- Ojala, D., C. Merkel, R. Gelfand and G. Attardi, 1980 The tRNA genes punctuate the reading of genetic information in human mitochondrial DNA. Cell **22**: 393-403.
- Ojala, D., J. Montoya and G. Attardi, 1981 tRNA punctuation model of RNA processing in human mitochondria. Nature **290**: 470-474.

- Okazaki, M. and Ueshima, R. Gender-associated mtDNA of *Tapes philippinarum*.
Unpublished^a.
- Okazaki, M. and Ueshima, R. Evolutionary diversity between the gender-associate
mitochondrial DNA genomes of freshwater mussels. Unpublished^b.
- Patterson, C. M. Chromosomes of Molluscs. Proc. Symp. on Mollusks II Mar. Biol.
Assoc. India, pp. 635- 686 (1969).
- Perna NT, Kocher TD: Patterns of nucleotide composition at fourfold degenerate
sites of animal mitochondrial genomes. *J. Mol. Evol.* 1995, 41: 353-358.
- Sasuga, J., Yokobori, S. -I., Kaifu, M., Ueda, T., Nishikawa, K. and Watanabe, K. Gene
contents and organization of a mitochondrial DNA segment of the squid *Loligo*
bleekeri. *J. Mol. Evol.* 48 (6), 692-702 (1999).
- Serb, J. M. and Lydeard, C. Complete mtDNA sequence of the North American
freshwater mussel, *Lampsilis ornata* (Unionidae): An examination of the evolution
and phylogenetic utility of mitochondrial genome organization in Bivalvia
(Mollusca). *Mol. Biol. Evol.* 20 (11), 1854-1866 (2003).
- Terrett, J. A., Miles, S. and Thomas, R. H. Complete DNA sequence of the mitochondrial
genome of *Cepaea nemoralis* (Gastropoda: Pulmonata). *J. Mol. Evol.* 42 (2), 160-168
(1996).
- Thiriot-Quievreux, C. Review of the literature on bivalve cytogenetics in the last ten
years. *Cah. Biol. Mar.* 43, 17-26 (2003).
- Tomita, K., Ueda, T. and Watanabe, K. 7-Methylguanosine at the anticodon wobble
position of squid mitochondrial tRNA(Ser)GCU: molecular basis for assignment of

AGA/AGG codons as serine in invertebrate mitochondria. *Biochim. Biophys. Acta* 1399 (1), 78-82 (1998).

Tomita, K., Yokobori, S. -I, Oshima, T., Ueda, T. and Watanabe, K. The cephalopod *Loligo bleekeri* mitochondrial genome: Multiplied noncoding regions and transposition of tRNA Genes. *J. Mol. Evol.* 54 (4), 486-500 (2002).

Yamazaki, N., Ueshima, R., Terrett, J. A., Yokobori, S. -I., Kaifu, M., Segawa, R., Kobayashi, T., Numachi, K., Ueda, T., Nishikawa, K., Watanabe, K. and Thomas, R. H. Evolution of pulmonate gastropod mitochondrial genomes: comparisons of complete gene organization of *Euhadra*, *Cepaea* and *Albinaria* and implications of unusual tRNA secondary structure. *Genetics* 145(3):749-58 (1997).

Yokobori, S. -I., Fukuda, N., Nakamura, M., Aoyama, T. and Oshima, T. Long-term conservation of six duplicated structural genes in cephalopod mitochondrial genomes. *Mol. Biol. Evol.* 21 (11), 2034-2046 (2004).

Wyman, S., R. K. Jansen and J. L. Boore, 2004 Automatic annotation of organellar genomes with DOGMA. *Bioinformatics* **20(17)**: 3252-3255.

Table 1 – All available complete gene arrangements for mollusk mtDNA

All that are cited as “unpublished” are available in GenBank. tRNA genes are abbreviated by the one-letter code for the corresponding amino acid and anticodons are shown in parentheses to differentiate the two leucine and two serine tRNAs. A minus symbol indicates reverse transcriptional orientation.

Binomen and taxonomy	Gene arrangement	Reference
<i>Venerupis (Ruditapes) philippinarum</i> (Bivalvia; Heteroconchia; Veneridae)	cox1, L(nag), nad1, nad2, nad4L, I, cox2, P, cob, rrnL, nad4, H, E, S(nga), atp6, nad3, nad5, Y, M, M, D, V, nad6, K, V, F, W, R, L(yaa), G, Q, N, T, C, A, cox3, rrnS	Okazaki and Ueshima, unpublished ^a , NC_003354
<i>Inversidens japonensis</i> (female type) (Bivalvia; Palaeoheterodonta; Unionidae)	cox1, cox3, atp6, D, nad4L, nad4, -nad6, -G, -nad1, -L(yaa), V, -I, -C, -Q, nad5, -P, -F, -cob, -N, -L(nag), -rrnL, -Y, -T, -K, -rrnS, -R, -W, -E, -S(nga), -A, nad3, -M, -nad2, -S(rct), H, cox2	Okazaki and Ueshima, unpublished ^b , AB055625
<i>Inversidens japonensis</i> (male type) (Bivalvia; Palaeoheterodonta; Unionidae)	cox1, cox3, atp6, -D, nad4L, nad4, -nad6, -G, -nad1, -L(yaa), V, -I, -C, -Q, nad5, -F, -cob, -P, -N, -L(nag), -rrnL, -Y, -T, -K, -rrnS, R, -W, -M, -nad2, -E, -S, -S, -A, nad3, cox2, H	Okazaki and Ueshima, unpublished ^b , AB055624
<i>Lampsilis ornata</i> (Bivalvia; Palaeoheterodonta; Unionidae)	-cox1, -cox2, -nad3, -H, A, S(nga), E, nad2, M, W, R, rrnS, K, T, Y, rrnL, L(nag), N, P, cob, F, -nad5, Q, C, I, V, L(yaa), nad1, G, nad6, -nad4, -nad4L, -atp8, -D, -atp6, -cox3	Serb and Lydeard, 2003, NC_005335
<i>Mytilus edulis</i> (Bivalvia; Pteriomorphia;	cox1, atp6, T, nad4L, nad5, nad6, F, rrnS, G, N, E, C, I, Q, D, rrnL, Y, cob,	Boore, Medina and Rosenberg, 2004;

Mytilidae)	cox2, K, M, L(nag), L(yaa), nad1, V, nad4, cox3, S, M, nad2, R, W, A, S, H, P, nad3	Hoffmann, Boore and Brown, 1992, NC_006161
<i>Crassostrea gigas</i> (Bivalvia; Pteriomorphia; Ostreidae)	cox1, rrnL, cox3, I, T, E, cob, D, cox2, M, L(yaa), P, rrnS, K, C, N, rrnS, Y, atp6, G, V, nad2, R, H, nad4, nad5, nad6, Q, nad3, L(nag), nad1, nad4L, W	Kim, Je and Park, unpublished, NC_001276
<i>Loligo bleekeri</i> (Cephalopoda; Coleoidea; Loliginidae)	cox1, -C, -Y, E, N, cox2, -M, R, -F, -nad5, -nad4, -nad4L, T, -L(yaa), -G, A, D, atp8, atp6, -H, -L(nag), cox3, nad3, -S(nga), -cob, -nad6, -P, -nad1, -Q, I, -rrnL, -V, -rrnS, -W, K, S(nct), nad2	Tomita et al., 2002; Sasuga et al., 1999; Tomita, Ueda and Watanabe, 1998, NC_002507
<i>Todarodes pacificus</i> (Cephalopoda; Coleoidea; Ommastrephidae)	cox1, cox2, D, atp8, atp6, -F, -V, -rrnS, -C, -Q, cox3, K, R, S1, nad2, cox1, cox2, D, atp8, atp6, -nad5, -H, -nad4, -nad4L, T, -S2, -cob, -nad6, -P, -nad1, -L2, -L1, -rrnL, -M, -Y, -W, -G, -E, cox3, A, N, I, nad3	Yokobori et al., 2004, NC_006354
<i>Octopus vulgaris</i> (Cephalopoda; Coleoidea; Octopodidae)	cox1, cox2, D, atp8, atp6, -F, -nad5, -H, -nad4, -nad4L, T, -S(nga), -cob, -nad6, -P, -nad1, -L(yaa), -L(nag), -rrnL, -V, -rrnS, -M, -C, -Y, -W, -Q, -G, -E, cox3, K, A, R, N, I, nad3, S(nct), nad2	Yokobori et al., 2004, NC_006353
<i>Aplysia californica</i> (Gastropoda; Heterobranchia; Opisthobranchia; Aplysiidae)	cox1, V, rrnL, L(nag), A, P, nad6, nad5, nad1, Y, W, nad4L, cob, -D, F, cox2, G, H, -Q, L(yaa), -atp8, -N, C, -atp6, -R, -E, -rrnS, -M, -nad3, -S(nga), S(nct), nad4, -T, -cox3, I, nad2, K	Nahir, Kohn and Moroz, unpublished, NC_005827
<i>Pupa strigosa</i> (Gastropoda; Heterobranchia; Opisthobranchia; Acteonidae)	cox1, V, rrnL, L(nag), A, P, nad6, nad5, nad1, Y, W, nad4L, cob, D, F, cox2, G, H, -Q, -L(yaa), -atp8, -N, C, -atp6, -R, -E, -rrnS, -R, -nad3, -S(nga), S(nct), nad4, T, -cox3, I, nad2, K	Kurabayashi and Ueshima, 2000, NC_002176
<i>Roboastra europaea</i> (Gastropoda;	cox1, V, rrnL, L(nag), A, P, nad6, nad5, nad1, Y, W, nad4L, cob, D, F,	Grande et al., 2002, NC_004321

Heterobranchia; Opisthobranchia; Polyceridae)	cox2, G, H, C, -Q, -L(yaa), -atp8, -N, -atp6, -R, -E, -rrnS, -M, -nad3, -S(nga), S(nct), nad4, -T, -cox3, I, nad2, K	
<i>Halotis rubra</i> (Gastropoda; Vetigastropoda; Haliotidae)	cox1, cox2, atp8, atp6, -F, -nad5, -H, -nad4, -nad4L, T, -S(nga), -cob, -nad6, -P, -nad1, -L(yaa), -L(nag), -rrnL, -V, -rrnS, -M, -Y, -C, -W, -Q, -G, -E, cox3, D, K, A, R, I, nad3, N, S(nct), nad2	Maynard et al., unpublished, NC_005940
<i>Biomphalaria glabrata</i> (Gastropoda; Pulmonata; Planorbidae)	cox1, V, rrnL, L(nag), A, P, nad6, nad5, nad1, nad4L, cob, D, C, F, cox2, Y, W, G, H, Q, -L(yaa), -atp8, -N, -atp6, -R, -E, -rrnS, M, -nad3, -S(nga), S(nct), nad4, -T, -cox3, I, nad2, K	DeJong, Emery and Adema, 2004, NC_005439
<i>Albinaria coerulea</i> (Gastropoda; Pulmonata; Clausiliidae)	cox1, V, rrnL, L(nag), P, A, nad6, nad5, nad1, nad4L, cob, D, C, F, cox2, Y, W, G, H, -Q, -L(yaa), -atp8, -N, -atp6, -R, -E, -rrnS, -M, -nad3, -S(nga), S(rct), nad4, -T, -cox3, I, nad2, K	Hatzoglou, Rodakis and Lecanidou, 1995, NC_001761
<i>Euhadra herklotsi</i> (Sequence NOT complete) (Gastropoda; Pulmonata; Bradybaenidae)	cox1, V, rrnL, L(nag), P, A, nad6, nad5, nad1, nad4L, cob, D, C, F, cox2, G, H, Y, -W, -Q, -L(yaa), -atp8, -N, -atp6, -R, -E, -rrnS, -M, -nad3, -S(nga), S(rct), nad4, -T, -cox3, I, nad2, K	Yamazaki et al., 1997, Z71693-701
<i>Cepaea nemoralis</i> (Gastropoda; Pulmonata; Helicidae)	cox1, V, rrnL, L(nag), A, nad6, P, nad5, nad1, nad4L, cob, D, C, F, cox2, Y, W, G, H, -Q, -L(yaa), -atp8, -N, -atp6, -R, -E, -rrnS, -M, -nad3, -S(nga), -T, -cox3, S(nct), nad4, I, nad2, K	Yamazaki et al., 1997; Terrett, Miles and Thomas, 1996, NC_001816
<i>Katharina tunicata</i> (Polyplacophora; Chitonida; Mopaliidae)	cox1, D, cox2, atp8, atp6, -F, -nad5, -H, -nad4, -nad4L, T, -S(nga), -cob, -nad6, P, -nad1, -L(yaa), -L(nag), -rrnL, -V, -rrnS, -M, -C, -Y, -W, -Q, -G, -E, cox3, K, A, R, N, I, nad3, S(nct), nad2	Boore and Brown, 1994, NC_001636

<i>Graptacme eborea</i> (Scaphopoda; Dentaliidae)	cox1, S, N, nad2, cob, H, -cox2, -Q, G, -cox3, -Y, R, S, -nad6, -P, -nad1, -atp8, -I, -T, rrnS, -M, -rrnL, V, A, nad3, L(nag), L(yaa), E, W, -F, -K, -nad5, -D, -nad4, -nad4L, atp6, C	Boore, Medina and Rosenberg, 2004, NC_006162
<i>Siphonodentalium lobatum</i> (Scaphopoda; Siphonodentaliidae)	cox1, L(yaa), G, -T, R, nad2, nad4, I, nad1, nad5, -Y, -nad4L, -atp8, -H, -A, -W, -M, -V, -nad6, -Q, -K, rrnS, -P, -N, -S(nct), -cob, -cox2, -cox3, -C, -atp6, -S(nga), -nad3, -E, -D, -F, -rrnL, -L(nag)	Dreyer and Steiner, 2004, NC_005840