RESEARCH ARTICLE

The mitochondrial genome of the semi-slug *Omalonyx unguis* (Gastropoda: Succineidae) and the phylogenetic relationships within Stylommatophora

Leila Belén Guzmán*, Roberto Eugenio Vogler, Ariel Aníbal Beltramino*

Grupo de Investigación en Genética de Moluscos (GiGeMol), Instituto de Biología Subtropical (IBS), CONICET–UNaM, Posadas, Misiones, Argentina

* leilaguzman95@gmail.com (LBG); beltraminoariel@hotmail.com (AAB)

Abstract

Here we report the first complete mitochondrial genome of the semi-slug *Omalonyx unguis* (d’Orbigny, 1836) (Gastropoda: Succineidae). Sequencing was performed on a specimen from Argentina. Assembly was performed using Sanger data and Illumina next generation sequencing (NGS). The mitogenome was 13,984 bp in length and encoded the 37 typical Metazoan genes. A potential origin for mitochondrial DNA replication was found in a non-coding intergenic spacer (49 bp) located between *cox3* and *tRNA-Ile* genes, and its secondary structure was characterized. Secondary structure models of the tRNA genes of *O. unguis* largely agreed with those proposed for other mollusks. Secondary structure models for the two rRNA genes were also obtained. To our knowledge, the 12S-rRNA model derived here is the first complete one available for mollusks. Phylogenetic analyses based on the mitogenomes of *O. unguis* and 37 other species of Stylommatophora were performed using amino acid sequences from the 13 protein-coding genes. Our results located Succineoidea as a sister group of Helicoidea + Urocoptoidea, similar to previous studies based on mitochondrial genomes. The gene arrangement of *O. unguis* was identical to that reported for another species of Succineoidea. The unique rearrangements observed for this group within Stylommatophora, may constitute synapomorphies for the superfamily.

Introduction

Metazoa mitochondrial (mt) genome is a circular double stranded DNA that typically encodes two ribosomal RNAs (rRNAs), 22 transfer RNAs (tRNAs) and 13 protein-coding genes (i.e., cytochrome c oxidase subunits I-III; cytochrome b; ATP synthase subunits 6 and 8; NADH dehydrogenase subunits 1-6 and 4L), with a size ranging from 14 to 20 kb [1–4]. In recent years, the development of new technologies and the decrease in sequencing costs produced a significant increase in available animal mitogenomes, and currently more than 90,000 mitochondrial genomes are deposited in NCBI database [5–7]. Despite this and considering that mollusks are the second most specious animal phylum, they continue to be a poorly represented taxon with only 1,440 known mitogenomes (approximately 0.04% of species represented) [8].
 Unlike vertebrates with a highly conserved mitochondrial genome, the available mollusk mitochondrial genomes have shown exceptional features, including doubly uniparental inheritance, gene rearrangements, large sizes, gene duplications, different gene distribution between DNA strands, and a high degree of mtDNA variability [1, 9–11]. The mtDNA as a molecular marker has been widely used throughout animal groups for the study of phylogenetic relationships among taxa, resolution of taxonomic controversies, and population genetics [12–15]. In this sense, complete mitochondrial genomes as phylogenetic markers provide additional comparison features, including gene order and content, as well as structural and compositional features [1, 16]. In addition, several studies using gene order as a phylogenetic marker indicated a good level of resolution in phylogenetic relationships, because of its low probability of convergence [3, 17, 18].

Succineidae Beck, 1837 is a molluscan family included in the superfamily Succineoidea Beck, 1837 within the order Stylommatophora that comprise more than 30 genera. The members of the family are distributed worldwide and inhabit diverse environments [19, 20]. *Omalonyx* d’Orbigny, 1838 is a genus of succinids endemic to Central and South America and the Caribbean islands characterized by a significant reduction of the shell [19]. This group of semi-slugs is frequently found in backwater areas, on riparian vegetation, with some species considered to be a pest of some crops (e.g., *Nymphoides indica*, *Pennisetum purpureum*) [19, 21–23]. They are also natural and potential intermediate hosts for parasites of the genus *Leucochloridium* Carus, 1835 (Trematoda) and *Angiostrongylus* Kamensky, 1905 (Nematoda), respectively [24–27]. Based on morphological characteristics of the reproductive system, six species are currently recognized within *Omalonyx* [28, 29], including *Omalonyx unguis* (d’Orbigny, 1836) which inhabits the Paraná River basin [28, 30].

To date, phylogenetic studies in mollusks based on morphological and molecular data support Stylommatophora as a monophyletic group, and agree to distinguish two groups, “achatinoid” and “non-achatinoid” [31–33 and references therein]. Within the non-achatinoid group, the taxonomic position and phylogenetic affinities of succinids has been variable and still remains contentious [19, 31–35].

In this study we present the mitogenome of *Omalonyx unguis* from Argentina, the first mitogenome of a South American representative of Succineidae available for comparison, increasing the taxon sampling of succinids in the ongoing phylogenetic reconstructions of Stylommatophora based on complete mitogenomes. We provide structural and compositional features of the newly sequenced mitochondrial genome, as well as an update of Stylommatophora phylogenetic relationships inferred from complete mitochondrial genomes.

**Materials and methods**

**Sample and DNA extraction**

Five adult specimens of *Omalonyx unguis* were collected from Garupa Stream, Misiones Province, Argentina (-27.4786, -055.7933) at the Paraná River basin in 2017. Permission for collection was granted by Ministerio de Ecología y Recursos Naturales Renovables de la Provincia de Misiones (Disp. No. 027/2018). The individuals were relaxed in water with menthol crystals for 6 h, preserved in ethanol 96%, and deposited in the malacological collection at the Instituto de Biología Subtropical, CONICET–UNaM, Misiones Province, Argentina (IBS-Ma 073). The specimens were identified using morphological characters of the reproductive system [30, 36, 37]. Additionally, molecular data were used to confirm the taxonomic identity following Guzmán et al. [30]. For the mitochondrial genome amplification and sequencing, total genomic DNA was extracted from a portion of pedal muscle of a single specimen (IBS-Ma 073–3) using a cetyltrimethyl-ammonium bromide (CTAB) protocol [38].
Amplification and sequencing

Two strategies were combined to obtain the mitochondrial genome of *Omalonyx unguis*. Firstly, short PCR products (less than ~1.5 kb) were amplified and sequenced by Sanger sequencing, using pulmonate-specific primers pairs [39] and individual-specific primers pairs designed from previously sequenced regions (S1 Table). PCR reactions were performed in a total volume of 30 μl, containing 50–100 ng of template DNA, 0.4 M of each primer, 1× *Dream Taq* Green Buffer (ThermoScientific), 2 mM MgCl₂, 0.2 mM dNTPs, 0.32 mg/ml BSA, and 1.25 U *Dream Taq* DNA polymerase (ThermoScientific). PCR reactions were run on a T21 thermocycler (Ivema Desarrollos) with cycling conditions as follows: 5 cycles of initial denaturation at 94˚C for 2 min, denaturing at 94˚C for 40 s, annealing at 40˚C for 45 s, and extension at 72˚C for 1 min; followed by 30 cycles of denaturing at 94˚C for 40 s, annealing at 40–60˚C for 40 s (based on gradient PCR profile), and extension at 72˚C for a time set at 1 minute per kb of expected product, with a final extension at 72˚C for 3 min. PCR products were purified from solution by means of an AccuPrep PCR Purification Kit (Bioneer, Korea), and from 1.5% (w/v) agarose gels using an ADN PuriPrep-GP Kit (Inbio Highway, Argentina) in cases of co-amplification of nonspecific fragments. Finally, both DNA strands of the PCR products were directly cycle sequenced (Macrogen Inc., Seoul, Korea). Secondly, NGS was used for sequencing of the mitogenomic fragments. Paired-end sequencing (2x150 bp, 350 bp insert size) of total genomic DNA was performed by Novogene Corporation (Sacramento, CA, USA) using a HiSeq platform (Illumina).

Genome assembly and annotation

All fragments obtained by Sanger sequencing were edited and compared with reference sequences in GenBank using the BLASTn algorithm [40] to confirm the amplicons to be the target sequences. Then, sequences were assembled manually into two large contigs in a stepwise manner by concatenating the sequences and trimming off overlapping regions with the help of Bioedit 7.0.5 software [41]. Finally, the complete mitogenome was assembled from about 10 Gb raw data obtained by NGS with the software NOVOPlasty 3.8.2 (https://github.com/ndierckx/NOVOPlasty) [42] using the previously obtained contigs as starting seeds.

The mitogenome was annotated with MITOS Web Server [5] employing the invertebrate mitochondrial genetic code and corrected manually by comparison with the available mitochondrial genome for the congeneric species *Succinea putris* (Linnaeus, 1758) (JN627206). The tRNA genes were detected using MITOS, and manually checked. Additionally, their secondary structures were inferred with ARWEN 1.2 [43]. The protein-coding genes (PCGs) were validated using the NCBI ORF Finder resource (https://www.ncbi.nlm.nih.gov/orffinder/). Finally, the limits of rRNAs were extend to the boundaries of adjacent genes following Cameron [6]. Secondary structure of the 16S-rRNA and 12S-rRNA genes was predicted from previous models for mollusks [44] and arthropods [14, 45], respectively, with the help of RNAStructure 6.0.2 (https://rna.urmc.rochester.edu/RNAstructureWeb) [46]. Additionally, the potential origin of DNA replication (POR) folding was performed using RNAStructure. The annotated mitogenome was deposited in GenBank under accession number MT449229. The nucleotide composition and relative synonymous codon usage (RSCU) for PCGs were calculated by means of MEGA X [47]. The AT- and GC- skew values were calculated using the equations AT-skew = (A-T)/(A+T), and GC-skew = (G-C)/(G+C) [48]. The circular mitochondrial genome was generated using Mtviz online tool (http://pacosy.informatik.uni-leipzig.de/mtviz/).
Sequence alignments and phylogenetic analyses

Phylogenetic analyses were based on the 13 PCGs of *O. unguis* and 37 species of Stylophora available in GenBank (Table 1). *Galba pervia* (Martens, 1867) (Hygrophila), *Platevindex mortoni* Britton, 1984 (Systellommatophora) and *Carychium tridentatum* (Risso, 1826) (Ellobioidea) were used as outgroups (Table 1). All sequences were downloaded from NCBI with the R package AnnotationBustR [49]. Nucleotide sequences were translated into amino acid sequences in EMBOSS Transeq [50] with invertebrate mitochondrial code; subsequently, sequences of each protein-coding gene were aligned separately using Muscle as implemented in MEGA X [47]. Ambiguously aligned positions were removed using Gblocks 0.91b under relaxed settings [51]. Exceptionally, *atp8* and *nad4L* sequences were manually cleaned. Finally, the single alignments were concatenated into a final dataset consisting of 2,877 positions (71% of the original alignment).

Phylogenetic trees were inferred using Bayesian Inference (BI) and Maximum Likelihood (ML) following Uribe et al. [52] with some modifications. BI analysis was performed using MrBayes on XSEDE 3.2.6 [53] implemented in the CIPRES Science Gateway [54]. The substitution model was determined by PartitionFinder 1.1.1 [55] according to the Bayesian Information Criterion (BIC) using a greedy approach. The phylogenetic analysis was performed under the site-homogeneous mtREV+I+G model. Two independent runs with four Markov chains were set to run simultaneously for 10^6 generations, sampling every 1,000 generations, with a final burn-in of 25% [56]. Bayesian posterior probabilities (PP) were used as branch support values. ML analysis was conducted using IQ-TREE 1.6.12 [57]. The substitution models were estimated using ModelFinder [58] implemented in IQ-TREE. ML analysis was performed under the site-homogeneous mtZOA+F+R7 model. Additionally, nodal support values were evaluated with 10,000 replicates of ultrafast likelihood bootstrap (UFBoot) [59]. Finally, as particular Heterobranchia lineages are known to experience an acceleration of evolutionary rates that may lead to long-branch attraction artifacts (LBA) in phylogenies, an additional BI analysis using the site-heterogeneous CAT-GTR model was performed in PhyloBayes MPI 1.7b [60] implemented in the CIPRES Science Gateway following Uribe et al. [61]. The phylogenetic trees were visualized using FigTree 1.4.4 (http://tree.bio.ed.ac.uk/software/figtree/).

Results

Genome organization and features

The complete mitochondrial genome of *Omalonyx unguis* was 13,984 bp in length, and contained 13 PCGs, 22 tRNAs, and two RNA genes (Fig 1). Most genes (9 PCGs, 12 tRNAs, 1 rRNA) were found on the plus strand, while the remaining genes (4 PCGs, 10 tRNAs, 1 rRNA) were on the minus strand (Fig 1, Table 2). We identified 17 intergenic regions (113 bp in total) ranging from 1 to 49 bp. The largest one was located between *cox3* and *tRNA-Ile* genes with an AT content of 98% (Table 2). This sequence was identified as POR and its secondary structure is shown in S1 Fig. Additionally, there were 13 overlapping regions with 4–60 bp size, and six genes pairs were directly adjacent to one another. The base composition of this genome (33.38% A, 44.07% T, 10.59% C, 11.96% G) showed a high AT content (77.5%), with a negative AT-skew value (-0.14) and positive GC-skew value (0.06). The protein-coding genes accounted for 77.1% of the mitochondrial genome of *O. unguis*, with *nad5* and *atp8* being the longest and shortest genes, respectively (Table 2). The PCGs were initiated with ATG and the alternative initiation codons TTG, ATT or ATA, and terminated with stop codons TAA, TAG or incomplete stop codon T (Table 2). In total, PCGs consisted of 3,607 codons (not including start and stop codons), among which UUU (Phe, counted 393 times), UUA (Leu, counted 397 times)
Table 1. Information of the mitochondrial genomes of Stylommatophora analyzed in present paper.

<table>
<thead>
<tr>
<th>Species</th>
<th>Family</th>
<th>NCBI</th>
<th>Mitogenome size (bp)</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stylommatophora</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Omalonyx unguis</em></td>
<td>Succineidae</td>
<td>MT449229</td>
<td>13,984</td>
<td>This study</td>
</tr>
<tr>
<td><em>Succinea putris</em></td>
<td>Succineidae</td>
<td>NC016190</td>
<td>14,092</td>
<td>White et al. [39]</td>
</tr>
<tr>
<td><em>Achatinella fulgens</em></td>
<td>Achatinellida</td>
<td>MG925058</td>
<td>15,346</td>
<td>Price et al. [62]</td>
</tr>
<tr>
<td><em>Achatinella mustelina</em></td>
<td>Achatinellida</td>
<td>NC030190</td>
<td>16,323</td>
<td>Price et al. [91]</td>
</tr>
<tr>
<td><em>Achatinella sowerbyana</em></td>
<td>Achatinellida</td>
<td>KX356680</td>
<td>15,374</td>
<td>Price et al. [92]</td>
</tr>
<tr>
<td><em>Partula redfieldi</em></td>
<td>Achatinellida</td>
<td>MG925057</td>
<td>16,879</td>
<td>Price et al. [62]</td>
</tr>
<tr>
<td><em>Achatina fulica</em></td>
<td>Achatinellida</td>
<td>KM114610</td>
<td>15,057</td>
<td>He et al. [93]</td>
</tr>
<tr>
<td><em>Deroceras reticulatum</em></td>
<td>Agriolimacidae</td>
<td>NC035495</td>
<td>14,048</td>
<td>Ahn et al. [94]</td>
</tr>
<tr>
<td><em>Arion vulgaris</em></td>
<td>Arionidae</td>
<td>MN607980</td>
<td>14,548</td>
<td>Doğan et al. [7]</td>
</tr>
<tr>
<td><em>Arion rufus</em></td>
<td>Arionidae</td>
<td>KT626607</td>
<td>14,321</td>
<td>Romero et al. [95]</td>
</tr>
<tr>
<td><em>Aegista aubryana</em></td>
<td>Bradybaenidae</td>
<td>NC029419</td>
<td>14,238</td>
<td>Yang et al. [96]</td>
</tr>
<tr>
<td><em>Aegista diversifamilia</em></td>
<td>Bradybaenidae</td>
<td>NC027584</td>
<td>14,039</td>
<td>Huang et al. [97]</td>
</tr>
<tr>
<td><em>Dolichulelota formosensis</em></td>
<td>Bradybaenidae</td>
<td>NC027493</td>
<td>14,237</td>
<td>Huang et al. [97]</td>
</tr>
<tr>
<td><em>Mastigeulota kiansinensis</em></td>
<td>Bradybaenidae</td>
<td>NC029435</td>
<td>14,029</td>
<td>Deng et al. [68]</td>
</tr>
<tr>
<td><em>Camaena cicatricosa</em></td>
<td>Camaenidae</td>
<td>NC025511</td>
<td>13,843</td>
<td>Wang et al. [98]</td>
</tr>
<tr>
<td><em>Camaena poyuensis</em></td>
<td>Camaenidae</td>
<td>KT001074</td>
<td>13,798</td>
<td>Unpublished</td>
</tr>
<tr>
<td><em>Cerion incanum</em></td>
<td>Cerionidae</td>
<td>NC025645</td>
<td>15,177</td>
<td>González et al. [69]</td>
</tr>
<tr>
<td><em>Cerion tridentatum</em></td>
<td>Cerionidae</td>
<td>KT249249</td>
<td>15,409</td>
<td>Unpublished</td>
</tr>
<tr>
<td><em>Cerion uva</em></td>
<td>Cerionidae</td>
<td>KY124261</td>
<td>15,043</td>
<td>Harasewych et al. [99]</td>
</tr>
<tr>
<td><em>Ryssoa otaheitana</em></td>
<td>Chronidae</td>
<td>NC044784</td>
<td>13,888</td>
<td>Damatac and Fontanilla [88]</td>
</tr>
<tr>
<td><em>Albinaria caerulea</em></td>
<td>Clausiliidae</td>
<td>NC001761</td>
<td>14,130</td>
<td>Hatzoglou et al. [100]</td>
</tr>
<tr>
<td><em>Gastrocopta cristata</em></td>
<td>Gastrocoptidae</td>
<td>NC026043</td>
<td>14,060</td>
<td>Unpublished</td>
</tr>
<tr>
<td><em>Cerionella virgata</em></td>
<td>Geomitridae</td>
<td>NC030723</td>
<td>14,147</td>
<td>Lin et al. [70]</td>
</tr>
<tr>
<td><em>Heliella itala</em></td>
<td>Geomitridae</td>
<td>KT696546</td>
<td>13,967</td>
<td>Romero et al. [95]</td>
</tr>
<tr>
<td><em>Cepaea nemoralis</em></td>
<td>Helicidae</td>
<td>NC001816</td>
<td>14,100</td>
<td>Yamazaki et al. [77]</td>
</tr>
<tr>
<td><em>Cylindrus obtusus</em></td>
<td>Helicidae</td>
<td>NC017872</td>
<td>14,610</td>
<td>Groenenberg et al. [101]</td>
</tr>
<tr>
<td><em>Cornu aspersum</em></td>
<td>Helicidae</td>
<td>NC021747</td>
<td>14,050</td>
<td>Gaitán-Espitia et al. [4]</td>
</tr>
<tr>
<td><em>Helix pomatia</em></td>
<td>Helicidae</td>
<td>NC041247</td>
<td>14,072</td>
<td>Korábek et al. [102]</td>
</tr>
<tr>
<td><em>Orcula dolium</em></td>
<td>Orculidae</td>
<td>NC034782</td>
<td>14,063</td>
<td>Groenenberg et al. [103]</td>
</tr>
<tr>
<td><em>Oreohelix idahoensis</em></td>
<td>Oreohelicidae</td>
<td>NC043790</td>
<td>14,213</td>
<td>Linscott and Parent [89]</td>
</tr>
<tr>
<td><em>Naesiotus nux</em></td>
<td>Orthalicidae</td>
<td>NC028553</td>
<td>15,197</td>
<td>Hunter et al. [104]</td>
</tr>
<tr>
<td><em>Meghimatium bilineatum</em></td>
<td>Philomyicidae</td>
<td>NC035429</td>
<td>13,972</td>
<td>Xie et al. [86]</td>
</tr>
<tr>
<td><em>Philomyicus bilineatus</em></td>
<td>Philomyicidae</td>
<td>MG722906</td>
<td>14,347</td>
<td>Yang et al. [105]</td>
</tr>
<tr>
<td><em>Polygyra cereolus</em></td>
<td>Polygyridae</td>
<td>NC032036</td>
<td>14,008</td>
<td>Unpublished</td>
</tr>
<tr>
<td><em>Praticolleta mexicana</em></td>
<td>Polygyridae</td>
<td>KX240084</td>
<td>14,153</td>
<td>Minton et al. [106]</td>
</tr>
<tr>
<td><em>Pupilla muscorum</em></td>
<td>Pupillidae</td>
<td>NC026044</td>
<td>14,149</td>
<td>Unpublished</td>
</tr>
<tr>
<td><em>Microceramus pontificus</em></td>
<td>Urocoptidae</td>
<td>NC036381</td>
<td>14,275</td>
<td>Unpublished</td>
</tr>
<tr>
<td><em>Vertigo pusilla</em></td>
<td>Vertiginidae</td>
<td>NC026045</td>
<td>14,078</td>
<td>Unpublished</td>
</tr>
</tbody>
</table>

*Ellobiida*  
*Carychium tridentatum*  
*Hygrophiia*  
*Galba pervia*  
*Systellommaphora*  
*Platevindex mortoni*  

* denotes outgroups.

https://doi.org/10.1371/journal.pone.0253724.t001
and AUU (Ile, counted 360 times) were the most frequently used, while CCG (Pro, counted 1 time), GCG (Ala, counted 2 times) ACG, CGG, and AGC (Thr, Arg and Ser, respectively, counted 3 times) were the least used (S2 Table). The RSCU values are shown in Fig 2 and S2 Table. All 22 tRNAs were identified by both ARWEN and MITOS, and their length ranged from 57 to 75 bp. The tRNA genes in *O. unguis* showed a classical clover-leaf secondary structure with exception of tRNA-Leu<sup>2</sup>, tRNA-Lys, and tRNA-Ser<sub>2</sub>. The tRNA structures are shown in Fig 3. Twenty of the 22 tRNAs presented the standard anticodons, while tRNA-Trp and tRNA-Lys were identified with the anticodons UCA and UUU, respectively. In addition, seven tRNAs (tRNA-Ala, tRNA-Asp, tRNA-His, tRNA-Leu<sub>2</sub>, tRNA-Lys, tRNA-Pro and tRNA-Tyr) showed mismatched base pairs in the acceptor stem and three (tRNA-Val, tRNA-Trp and tRNA-Gly) in the anticodon stem, while three tRNAs (tRNA-Asp, tRNA-Gly and tRNA-Pro) showed G-U wobble base pairs in the acceptor stem, and another two (tRNA-Ile and tRNA-Ser<sub>2</sub>) in the anticodon stem. Finally, the lengths of the 12S-rRNA and 16S-rRNA genes were 771 and 1,013 bp, respectively, with the first located on the minus strand and the second on the plus strand. The inferred secondary structure models for 12S-rRNA and 16S-rRNA are provided in Figs 4 and 5, respectively.

**Phylogenetic analyses, gene order and rearrangements**

The results of the phylogenetic reconstructions under the ML and BI approaches from concatenated amino acid sequences of the 13 PCGs and carried out using site-homogeneous models are shown in Fig 6. The topology of both trees was not entirely identical and nodal
support values were generally higher in BI than in ML. Both approaches recovered all superfamilies as monophyletic groups with high support values and distinguished the “non-achatinoid” group from the “achatinoid” one represented by *Achatina fulica* Bowdich, 1822 (the only mitogenome available for achatinoids). In this study, the Succineoidea (represented by the succinids *Omalonyx unguis* and *Succinea putris*) showed a sister group relationship to

<table>
<thead>
<tr>
<th>Name</th>
<th>Position</th>
<th>Length (bp)</th>
<th>Strand</th>
<th>Start Codon</th>
<th>Stop Codon</th>
<th>Anticodon</th>
<th>ISR (bp)</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>cox1</em></td>
<td>1–1,530</td>
<td>1,530</td>
<td>+</td>
<td>TTG</td>
<td>TAA</td>
<td>+8</td>
<td></td>
</tr>
<tr>
<td><em>tRNA Val</em></td>
<td>1,539–1,604</td>
<td>66</td>
<td>+</td>
<td>TAC</td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>16S-rRNA</em></td>
<td>1,605–2,617</td>
<td>1,013</td>
<td>+</td>
<td>TGG</td>
<td>-4</td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>tRNA Pro</em></td>
<td>2,618–2,680</td>
<td>63</td>
<td>+</td>
<td>TAG</td>
<td>+2</td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>tRNA Leu1</em></td>
<td>2,677–2,740</td>
<td>64</td>
<td>+</td>
<td>TGC</td>
<td>-14</td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>tRNA Ala</em></td>
<td>2,743–2,805</td>
<td>63</td>
<td>+</td>
<td>AAT</td>
<td>TAA</td>
<td>-11</td>
<td></td>
</tr>
<tr>
<td><em>nad6</em></td>
<td>2,792–3,268</td>
<td>477</td>
<td>+</td>
<td>ATG</td>
<td>TAA</td>
<td>-60</td>
<td></td>
</tr>
<tr>
<td><em>nad4L</em></td>
<td>5,837–6,110</td>
<td>274</td>
<td>+</td>
<td>ATA</td>
<td>TAA</td>
<td>+9</td>
<td></td>
</tr>
<tr>
<td><em>cob</em></td>
<td>6,108–7,202</td>
<td>1,095</td>
<td>+</td>
<td>ATT</td>
<td>TAA</td>
<td>-10</td>
<td></td>
</tr>
<tr>
<td><em>tRNA Phe</em></td>
<td>7,193–7,257</td>
<td>65</td>
<td>+</td>
<td>GAA</td>
<td>+4</td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>tRNA Arg</em></td>
<td>7,262–7,324</td>
<td>63</td>
<td>+</td>
<td>ATG</td>
<td>TAA</td>
<td>-5</td>
<td></td>
</tr>
<tr>
<td><em>tRNA Gln</em></td>
<td>7,320–7,385</td>
<td>66</td>
<td>+</td>
<td>GCA</td>
<td>+1</td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>cox2</em></td>
<td>7,387–8,035</td>
<td>649</td>
<td>+</td>
<td>ATG</td>
<td>TAA</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td><em>tRNA Gly</em></td>
<td>8,036–8,110</td>
<td>75</td>
<td>+</td>
<td>TCC</td>
<td>-21</td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>tRNA His</em></td>
<td>8,090–8,155</td>
<td>66</td>
<td>+</td>
<td>GTG</td>
<td>+1</td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>tRNA Leu2</em></td>
<td>8,229–8,287</td>
<td>59</td>
<td>-</td>
<td>TTA</td>
<td>+7</td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>atp8</em></td>
<td>8,273–8,402</td>
<td>130</td>
<td>-</td>
<td>TTA</td>
<td>TAA</td>
<td>-15</td>
<td></td>
</tr>
<tr>
<td><em>tRNA Asn</em></td>
<td>8,404–8,466</td>
<td>63</td>
<td>-</td>
<td>GTT</td>
<td>+1</td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>atp6</em></td>
<td>8,457–9,122</td>
<td>666</td>
<td>-</td>
<td>ATG</td>
<td>TAG</td>
<td>-10</td>
<td></td>
</tr>
<tr>
<td><em>tRNA Arg</em></td>
<td>9,125–9,189</td>
<td>65</td>
<td>-</td>
<td>TCG</td>
<td>+8</td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>tRNA Glu</em></td>
<td>9,198–9,270</td>
<td>73</td>
<td>+</td>
<td>TTC</td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>12S-rRNA</em></td>
<td>9,271–10,041</td>
<td>771</td>
<td>+</td>
<td>CAT</td>
<td>+4</td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>tRNA Met</em></td>
<td>10,042–10,107</td>
<td>66</td>
<td>+</td>
<td>TGA</td>
<td>-6</td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>nad3</em></td>
<td>10,112–10,465</td>
<td>354</td>
<td>-</td>
<td>TTG</td>
<td>TAA</td>
<td>+1</td>
<td></td>
</tr>
<tr>
<td><em>tRNA Ser2</em></td>
<td>10,467–10,523</td>
<td>57</td>
<td>-</td>
<td>TGA</td>
<td>+1</td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>tRNA Tyr</em></td>
<td>10,518–10,584</td>
<td>67</td>
<td>+</td>
<td>GTG</td>
<td>+1</td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>tRNA Trp</em></td>
<td>10,590–10,655</td>
<td>66</td>
<td>+</td>
<td>TCA</td>
<td>+5</td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>tRNA Ser1</em></td>
<td>10,654–10,716</td>
<td>63</td>
<td>+</td>
<td>GCT</td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>nad4</em></td>
<td>10,718–12,028</td>
<td>1,311</td>
<td>+</td>
<td>TTG</td>
<td>TAA</td>
<td>-8</td>
<td></td>
</tr>
<tr>
<td><em>tRNA Thr</em></td>
<td>12,021–12,084</td>
<td>64</td>
<td>-</td>
<td>TGT</td>
<td>+1</td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>cox3</em></td>
<td>12,086–12,865</td>
<td>780</td>
<td>-</td>
<td>ATG</td>
<td>TAA</td>
<td>+49</td>
<td></td>
</tr>
<tr>
<td><em>tRNA Ile</em></td>
<td>12,915–12,984</td>
<td>70</td>
<td>+</td>
<td>GAT</td>
<td>TAA</td>
<td>+16</td>
<td></td>
</tr>
<tr>
<td><em>nad2</em></td>
<td>13,001–13,958</td>
<td>958</td>
<td>+</td>
<td>TTG</td>
<td>TAA</td>
<td>-33</td>
<td></td>
</tr>
<tr>
<td><em>tRNA Lys</em></td>
<td>13,926–13,983</td>
<td>58</td>
<td>+</td>
<td>TTT</td>
<td>+1</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* + and—denote plus and minus strands, respectively. ISR denotes the length of the intergenic spacer region, for which negative numbers indicate nucleotide overlapping between adjacent genes. The anticodons of tRNAs are reported in the 5' → 3' direction.

* *denotes incomplete stop codon.*

https://doi.org/10.1371/journal.pone.0253724.t002

https://doi.org/10.1371/journal.pone.0253724.t002
Helicoidea + Urocoptoidea with relatively high nodal support (UFBootS 78, PP 0.98).

Although Naesiotus nux (Broderip, 1832) and Oreohelix idahoensis (Newcomb, 1866) grouped together with high support values (Orthalicoidea + Punctoidea), the position of this group differed between the two approaches. On the other hand, there were differences in the relationships established among families within Helicoidea when comparing both trees. Helicidae grouped together with Geomitridae in both approaches, however, Polygyridae grouped together with Camaenidae in the BI tree (Camaenidae + Polygyridae) + (Helicidae + Geomitridae), but they grouped outside of Camaenidae + (Helicidae + Geomitridae) in the ML tree. The phylogenetic tree of the BI analysis performed under the site-heterogeneous model is shown in S2 Fig. The tree topology obtained using the CAT-GTR model, with branches showing high support values, was similar to that of the BI tree under the site-homogeneous MTREV +I+G model (Fig 6B).

Regarding the order and orientation of mitochondrial genes, some differences were observed among the species of the different stylommatophoran families (Fig 7). In particular, O. unguis presented a gene arrangement identical to that reported for S. putris. Both succinids differed from the other stylommatophoran by exhibiting the following rearrangements: tRNA-Pro–tRNA-Leu₁–tRNA-Ala, tRNA-Phe–tRNA-Asp–tRNA-Cys, and tRNA-Ser₂–tRNA–Tyr–tRNA-Trp–tRNA-Ser₁ (with tRNA-Tyr and tRNA-Trp on the minus strand).

**Discussion**

The mitogenome of the semi-slug Omalonyx unguis obtained in this study presented a size of 13,984 bp and constitutes the second complete mitochondrial genome available for
Fig 3. Putative secondary structure of the 22 tRNA genes identified in the mitogenome of *Omalonyx unguis*. The tRNAs are labeled with the abbreviations of their corresponding amino acids. Watson-Crick base pairings are indicated by dashes (−), and G·T(U) wobble base pairings are indicated by dots (•).

https://doi.org/10.1371/journal.pone.0253724.g003

Fig 4. Secondary structure of the 12S-rRNA gene of *Omalonyx unguis*. Domains are indicated with Roman numbers. Watson-Crick base pairings are indicated by dashes (−), and G·U wobble base pairings are indicated by dots (•). Adjacent genes are labelled at the 5′ and 3′ ends.

https://doi.org/10.1371/journal.pone.0253724.g004
Succineidae worldwide. The reduced size of this genome is consistent with that reported for Stylommatophora, which account for 38 mitogenomes to date with sizes ranging from 13,978 bp for *Camaena poyuensis* Zhou et al., 2016 to 16,879 bp for *Partulina redfieldi* (Newcomb, 1853) [2, 3, 62]. The 37 typical genes (13 PCGs, 22 tRNAs and 2 rRNAs) were present, and contained 17 intergenic spacer regions and 13 overlapping regions. Origins of mitochondrial replication are reported to be characterized by high AT content and stem-loop structures [63, 64]. These features were observed in the longest intergenic region (49 bp) of the *O. unguis* mitogenome located between *cox3* and tRNA-*lle* genes. The location of a potential origin of DNA replication (POR) in this region is in line with what has been observed in *S. putris* and other heterobranchs [3, 4, 39]. The compositional approach showed a high AT content.

Fig 5. Secondary structure of the 16S-rRNA gene of *Omalonyx unguis*. Domains are indicated with Roman numbers. Watson-Crick base pairings are indicated by dashes (−), and G-U wobble base pairings are indicated by dots (·). Adjacent genes are labelled at the 5′ and 3′ ends.

https://doi.org/10.1371/journal.pone.0253724.g005
Fig 6. Phylogenetic trees under site-homogeneous models based on the 13 PCGs dataset for stylommatophoran mollusks. A. Maximum Likelihood tree. B. Bayesian consensus tree. The trees were rooted with three outgroups (*Carychium tridentatum*, *Platevindex mortoni* and *Galba pervia*). The scale (0.3)
shows evolutionary distances. The ultrafast bootstrap (ML) and posterior probability (BI) support values are shown in the nodes. The GenBank accession numbers of the species are shown in Table 1. *Omalonyx unguis* was sequenced in this study.

https://doi.org/10.1371/journal.pone.0253724.g006

(77.5%) in *O. unguis*, with a value close to that described for the mitochondrial genome of *S. putris* (76.7%). This high content fits well with the expectations for Stylommatophora, which stand out for having the highest values of AT among the pulmonate gastropods [4]. Generally, the plus chain is characterized by a higher A and C content [65]. Within Mollusca, two groups have been characterized based on their compositional asymmetries. On one hand, cephalopods
and some gastropod species with positive AT and negative GC biases, and on the other hand, bivalves, and most gastropods with inverse values [65]. In agreement with most gastropods and other animal groups such as fish and arthropods, *O. unguis* showed a strong inversion of its asymmetries (AT-skew = -0.14, GC-skew = 0.06) [64, 66]. Some authors have suggested that one of the reasons for this finding could be related to a reversal of the origin of the replication [64, 66, 67].

The size of the protein-coding genes in *O. unguis* conformed well to the expected pattern for Eupulmonata, with *nad5* being the longest gene (1,683 bp) and *atp8* the shortest one (130 bp) [4]. The start and stop codons of the *cox2, cox3* and *nad5* genes corresponded to those identified for *S. putris* [39]. For the remaining genes, some start or stop codons matched with those of *S. putris*, while others corresponded to those informed for other pulmonate gastropods [3, 4, 39, 68–70]. Exceptionally, the *nad4L, cox2, nad2* and *atp8* genes showed a truncated stop codon (T—). This type of codon has been widely documented for mollusks and it was suggested that it is completed to TAA by post-transcriptional polyadenylation [71, 72]. In relation with mitogenomes that exhibited a high AT content, some Metazoa revealed a strong impact on the use of codons with this composition, even leading to the extinction of codons with GC content [73]. Accordingly, the analysis of codon uses in *O. unguis* revealed a higher use of those with high AT content, such as TTT (393), TTA (397), ATT (360), while those less represented had high GC content. Despite this, all possible codons were present at least once, as was the case with the CCG codon. The two most represented codon families were Ser2 and Leu2 with RSCU values of 4.77 and 4.65 respectively, while the least prevalent was Leu1 with a value of 1.35. The compositional bias in *O. unguis* was reflected in the RSCU values that showed no random use among families of codon synonyms, since a preferential use of codons with A or T in their third position was observed. It is known that codon bias influences the folding and differential regulation of proteins, as well as the efficiency of translation [74], however, to our knowledge the impact it has on gastropods mitogenomes and how it varies among its taxa is still unknown.

Recurrently, the variability in the sequence of tRNA genes generates difficulties for sequence annotation, so secondary structures are especially useful accessory tools [75]. tRNAs are generally characterized by a cloverleaf-shaped secondary structure that adopts an L-shaped tertiary structure to perform its function [76]. Most of the tRNAs of *O. unguis* showed the typical shape except for *tRNA-Leu2* and *tRNA-Lys* which showed the ΨC arm truncated, and *tRNA-Ser2* which showed the DHU arm truncated. However, loss of arms (ΨC or DHU) has been described as structural variation for Metazoa, including some species of mollusks, and it is suggested to be the result from selection pressure to reduce the size of mitochondrial genomes [3, 18, 77, 78]. Most of the characterized structures showed preserved sizes for the acceptor stem and the anticodon stem (7 bp and 5 bp, respectively), as expected for metazoans in general [70, 79]. However, the acceptor stem of the *tRNA-Ala, tRNA-Asp, tRNA-His, tRNA-Leu2*, *tRNA-Lys, tRNA-Pro* and *tRNA-Tyr*, and the anticodon stem of the *tRNA-Val, tRNA-Trp* and *tRNA-Gly* showed mismatched base pairs. Exceptionally, *tRNA-Ser2* showed a longer length than conventional in the anticodon stem (6 bp). These types of alterations have been frequently observed in other animal species involving diverse mitochondrial tRNAs [72, 77, 80]. Particularly for mollusks, Yamazaki et al. [77] concluded that tRNAs that overlap nucleotides with adjacent genes may have a lack of mating in the acceptor stem. This was observed in *O. unguis* (e.g., *tRNA-Ala, tRNA-Tyr* and *tRNA-Leu2*), although most tRNAs overlapping nucleotides were unmodified. To ensure the functionality of tRNAs that show modifications in their pairings, some authors proposed that post-transcriptional editing is responsible for rectifying these disappearances by Watson-Crick pairings [3, 77, 81, 82]. Despite the results provided here, more mitochondrial genomes of mollusks including secondary structures of these genes.
are required. These will allow comparative studies to help establish patterns of mitochondrial evolution, as well as to understand the functional importance of these structures.

The sizes of the large (16S-rRNA) and small (12S-rRNA) rRNA subunits in O. unguis were similar to those reported for other mollusks [39]. The 16S-rRNA structure model showed a relatively conserved secondary structure in relation to the available models for mollusks, and presented the six typical domains [44]. However, a stem-loop structure was not present in neither domain II or III, nor a bulge-stem-loop structure was present in domain V. These features are known to represent synapomorphies described for Heterobranchia and explain some reduction in mitochondrial size [44, 83]. Additionally, we generated the secondary structure model for the 12S-rRNA gene which, to our knowledge, represents the first complete model available for Mollusca. It was developed based on available models for arthropods and presented the four typical domains into which this gene is structurally divided [14, 72, 84]. This model is expected to contribute to further comparative studies aimed to investigate evolutionary questions among mollusks.

The BI and ML phylogenetic trees based on amino acid sequences of 13 PCGs recovered Stylommatophora as a monophyletic group. Both analyses showed a basal dichotomy within the order Stylommatophora that separated “non-achatinoid” (suborder Helicina) and “achatinoid” (suborder Achatinina), the latter represented only by Achatina fulica. This finding agrees well with previous studies in the literature based on mitochondrial genomes as well as individual genes, and represent widely accepted taxonomic clades [34, 85, 86]. Within Helicina, both trees reconstructed under site-homogeneous models recovered 11 superfamilies as monophyletic groups (UFBootS = 100, PP = 1.0), with similar topologies to studies carried out so far with mitogenome-based phylogenies. One of the previous mitogenomic studies grouped Succineoidae together with Arionoidea, which was interpreted as an artifact attributed to a scarce sampling of taxa [86]. Our results located Succineoidae as a sister group of Helicoidea + Urocoptoidea with Arionoidea grouping outside. These findings are consistent with the most recent results from Doğan et al. [7]. However, earlier phylogenies based on individual genes showed close relationships between Arionoidea and Limacoidea, while they failed to clarify the relationships of Elasmognatha (Succineidae + Athetaorocharophoridae) within “achatinoids” [32]. Previous phylogenetic studies in Stylommatophora, performed on individual genes, also showed a close relationship between Polygyridae and Camaenidae within the superfamily Helicoidea [32, 87]. However, some phylogenies based on mitochondrial genomes showed Polygyridae in the basal division of Helicoidea (Polygyridae + (Camaenidae (Helicidae + Geomitridae))) [7, 86, 88]. Our results are consistent with both Helicoidea topologies depending on the phylogenetic approach used, although with some low support values. Although Xie et al. [86] suggest that the topology that places Polygyridae in a basal position would make more sense from a biogeographic perspective, the topological location of this group and the relationships among helicoids remain unclear. In this study, we observed different positions of the group formed by N. nux (Orthalicoidea) and O. idahoensis (Punctoidea) within “non-achatinoids”, with low support values in both cases. The most recent phylogenetic studies with stylommatophoran mitogenomes were not concordant with respect to the relationships deduced between these groups. While the results of Xie et al. [86] grouped Orthalicoidea with Limacoidea, the results of Damatac and Fontanilla [88] grouped them with Succineoidae. The results of Linscott and Parent [89] grouped Orthalicoidea as follows: Punctoidea + (Orthalicoidea + (Succineoidae + Achatinelloidea)); however, the results of Doğan et al. [7] grouped Orthalicoidea + Arionoidea. Thus, this variation in the evolutionary affinities evidences an insufficient representativeness of the different taxonomic levels within Stylommatophora since the proposed relations are changing as new mitogenomes are made available. In addition, the BI analysis performed under the site-heterogeneous CAT-GTR model did not
improve the resolution of the trees inferred under the site-homogeneous models. Despite recovering all 11 superfamilies of Stylommatophora involved in this work as monophyletic groups, two nodes remained unresolved. Consequently, further research based on increased taxon sampling of Stylommatophora species is required to a better understanding of the evolutionary relationships within this group.

Among the additional phylogenetic markers provided by mitogenomes, several authors have mentioned the resolving potential of gene rearrangements [9, 90]. In this work we found the mitochondrial genome of *Omalonyx unguis* to present the same genetic arrangement as reported for *Succinea putris* [39]. To date and within Stylommatophora, Succineoidea is the only superfamily presenting the tRNA arrangements tRNA-Pro–tRNA-Leu1–tRNA-Ala and tRNA-Phe–tRNA-Asp–tRNA-Cys. In addition, an inversion and transposition of tRNA-Tyr–tRNA-Trp was observed, and they were located between tRNA-Ser2 and tRNA-Ser1 (tRNA--Ser2–tRNA-Tyr–tRNA-Trp–tRNA-Ser1), with both genes encoded on the minus strand instead of the plus strand. While only two mitogenomes are available for the group, the reported changes seem to represent conserved arrangements that may constitute synapomorphies for Succineoidea. Further research based on more succinids are required to test and validate this assumption.

**Conclusions**

This study provides the complete mitochondrial genome of the semi-slug *Omalonyx unguis* (Gastropoda: Succineidae), consisting of 13,984 bp and with a typical gene content. Genomic features were similar with those of other stylommatophoran mollusks. Additionally, we present, to our knowledge, the first structural model of the complete 12S-rRNA (small-subunit rRNA) within Mollusca. The gene rearrangement was identical to that reported for *Succinea putris*. Both BI and ML analyses supported Stylommatophora monophyly with *Achatina fulica* at the basal bifurcation. The reconstructions within Stylommatophora were similar to those reconstructed in earlier studies. The arrangements tRNA-Pro–tRNA-Leu1–tRNA-Ala, tRNA-Phe–tRNA-Asp–tRNA-Cys and tRNA-Ser2–tRNA-Tyr–tRNA-Trp–tRNA-Ser1 (with tRNA-Tyr and tRNA-Trp on the minus strand) were present only in Succineoidea, and these arrangements are suggested to represent synapomorphies for the superfamily. Further research based in increased taxon sampling of succinids is required to confirm this hypothesis.

**Supporting information**

S1 Fig. The putative stem-loop structure that is found in the longest non-coding region, located between *cox3* and tRNA-Ile genes.  
(TIF)

S2 Fig. Bayesian Inference tree under the site-heterogeneous CAT-GTR model based on the 13 PCGs dataset for stylommatophoran mollusks. The trees were rooted with three outgroups (*Carychium tridentatum*, *Platevindex mortoni* and *Galba pervia*). The scale (0.6) shows evolutionary distances. Posterior probability support values are shown in the nodes. The GenBank accession numbers of the species are shown in Table 1; *Omalonyx unguis* was sequenced in this study.  
(TIF)

S1 Table. Individual-specific primers of *Omalonyx unguis*. Primer sequences are in the 5’ to 3’ direction. The primers naming system follows that of White et al. [39]. The genes encoded in the minus strand are underlined.  
(DOCX)
S2 Table. Codon usage in *Omalonyx unguis*. RSCU: relative synonymous codon usage.

**Author Contributions**

**Conceptualization:** Leila Belén Guzmán, Ariel Aníbal Beltramino.

**Data curation:** Leila Belén Guzmán, Roberto Eugenio Vogler.

**Formal analysis:** Leila Belén Guzmán.

**Investigation:** Leila Belén Guzmán, Roberto Eugenio Vogler, Ariel Aníbal Beltramino.

**Methodology:** Leila Belén Guzmán.

**Project administration:** Leila Belén Guzmán, Ariel Aníbal Beltramino.

**Supervision:** Roberto Eugenio Vogler, Ariel Aníbal Beltramino.

**Validation:** Roberto Eugenio Vogler, Ariel Aníbal Beltramino.

**Visualization:** Leila Belén Guzmán.

**Writing – original draft:** Leila Belén Guzmán, Roberto Eugenio Vogler, Ariel Aníbal Beltramino.

**Writing – review & editing:** Leila Belén Guzmán, Roberto Eugenio Vogler, Ariel Aníbal Beltramino.

**References**


