



A foundation for comparative genomics and evolutionary studies in *Nucella lapillus* based on complete mitogenome assembly

Daniel García-Souto^{1,2,3} · Jonathan Fernández-Rodríguez⁴ · André Vidal-Capón⁴ · Neil Fuller⁵ · Juan Pizcueta⁴ · Juan J. Pasantes⁴ · Juan Galindo⁴

Received: 21 October 2023 / Accepted: 11 March 2024 / Published online: 4 April 2024
© The Author(s) 2024

Abstract

The Atlantic dog whelk, *Nucella lapillus*, is a marine snail that exhibits divergent evolution in response to habitat adaptation, resulting in distinct populations at the phenotypic, genotypic, and karyotypic levels. In this study, we utilized short- and long-read NGS data to perform a de novo assembly of the entire mitochondrial genome of *N. lapillus* and developed a multiplex PCR protocol to sequence most of its length using ONT sequencing. Our analysis revealed a typical circular configuration of 16,474 bp in length with 13 protein-coding genes, 22 different tRNA genes, 2 of them showing two copies, 2 rRNA genes, and a control region. Long-read sequencing enabled us to identify a 1826 bp perfect inverted repeat within the control region. Comparative analysis of the mitogenomes of related species in the Muricidae family revealed a conserved gene configuration for *N. lapillus*. We found a low genetic diversity, as well as a moderate genetic differentiation among the studied populations. Interestingly, there was no observed differentiation between the two chromosomal races, suggesting either introgression and permanent incorporation of the mitochondrial DNA haplotype from one of the chromosomal races into the other or a slower evolutionary rate of the mtDNAs with respect to that of the karyotypes. Our study serves as a foundation for comparative genomics and evolutionary investigations in this species.

Keywords Atlantic dog whelk · *Nucella lapillus* · Mitochondrial genome · Comparative phylogenomics · Divergent evolution · Genetic diversity

Introduction

Nucella lapillus (Linnaeus 1758) is a marine gastropod species belonging to the family Muricidae within the order Neogastropoda. It is widely distributed along the rocky shores of the North Atlantic, from Southern Portugal to Northwestern Iceland, and from Terra Nova (Canada) to Long Island (New York, USA) (Collins et al. 1996; Colson and Hughes 2007; Marko et al. 2014). This species exhibits remarkable intraspecific morphological and chromosomal polymorphisms that are linked to its habitat, primarily wave exposure, and to a lesser extent, predation (Staiger 1957; Bantock and Cockayne 1975; Pascoe 2006).

The two chromosomal races of *N. lapillus* are typically characterized by thick-shelled specimens with $2n > 26$ chromosomes found in sheltered sides and thin-shelled specimens with $2n = 26$ chromosomes occupying wave-exposed localities (Crothers 1985; Vazquez 2015). This variation arises from the fragmentation of five specific metacentric pairs in the $2n = 26$ karyotypes into ten acrocentric pairs,

Responsible Editor: M. Wegner.

✉ Daniel García-Souto
daniel.garcia.souto@usc.es

- ¹ Genomes and Disease, Centre for Research in Molecular Medicine and Chronic Diseases (CiMUS), Universidade de Santiago de Compostela, Santiago de Compostela, Spain
- ² Department of Zoology, Genetics and Physical Anthropology, Universidade de Santiago de Compostela, Santiago de Compostela, Spain
- ³ Department of Biological Sciences, College of Arts, Sciences and Education (CASE), School of Environment, Arts and Society, Florida International University, Miami, FL, USA
- ⁴ Centro de Investigación Mariña, Universidade de Vigo, Vigo, Spain
- ⁵ EA Engineering, Science, and Technology PBC, Hunt Valley, MD, USA

likely due to Robertsonian fusions/fissions (Pascoe 2006), and the different balanced combinations of intact and fragmented pairs in homo and heterozygosis result in karyotype intermediate between 26 and 36 chromosomes, which vary between locations and specimens (Pascoe 2006). As such, *N. lapillus* provides an excellent model for understanding the evolutionary process of species differentiation under divergent environmental conditions, making it ideal for studying chromosomal evolution, speciation, and biodiversity (Staiger 1957; Pascoe 2006). Several previous studies have shown that this phenotypic and chromosomal differentiation observed in *N. lapillus* is linked to a substantial genetic divergence between chromosomal races, as demonstrated by allozyme markers and microsatellites (Kirby et al. 1994; Guerra-Varela et al. 2009). However, the extent of this genetic differentiation remains a matter of debate, and it is yet to be determined whether these are isolated species in various degrees of hybridization or a single species in divergent evolution under different selective pressures (Kirby et al. 1994).

In this regard, the mitochondrial genome is a reliable molecular marker to infer the phylogenetic evolution between species and populations. Previous studies on *N. lapillus* have been limited to a fragment of the malate dehydrogenase I gene (Kirby 2000), and a comprehensive analysis of the mitochondrial genome is essential to determine distinct mitochondrial lineages between ecotypes or localities. Therefore, the present study aims to assemble and annotate the complete mitogenome of *N. lapillus* using two next-generation sequencing approaches, namely Illumina short-read and long-read nanopore multiplex PCR sequencing. The gene arrangement and sequence similarity will be compared to other Muricidae species, and differences between chromosomal races of this species will be assessed.

Materials and methods

Biological samples

A total of 40 adult specimens of *N. lapillus* were collected at each of the localities indicated in Table 1 and

transported to the laboratory of molecular cytogenetics of the University of Vigo (NW Spain). There, mitotic metaphases were obtained from gills and gonads according to García-Souto et al. (2018). In brief, after an *in vivo* overnight colchicine (0.005%) treatment in seawater, animals were euthanized and their soft tissues dissected and processed to obtain chromosome spreads and genomic DNA following previous protocols. For cytogenetics, gills and gonads were hypotonized in 50% and 25% seawater and fixed with methanol/acetic (3:1). Small pieces of fixed tissue were disaggregated with 60% acetic acid and dropped onto heated glass slides. These were stained with a mixture of DAPI (0.14 µg/mL) and PI (0.07 µg/mL), examined using a Nikon Eclipse-800 microscope equipped with an epifluorescence detection module and photographed with a DS-Qi1Mc CCD camera (Nikon) controlled by the NIS-Elements software (Nikon). Separated images for each metaphase and fluorochrome were obtained and merged with Adobe Photoshop CS6. In parallel, muscular tissue was fixed in ethanol and preserved at -20 °C, and total DNA was isolated using the EZNA Mollusc DNA Kit (Omega Bio-Tek). DNA integrity was evaluated using a Genomic DNA ScreenTape (Agilent) on a TapeStation 4200 (Agilent), and the DNA concentration was determined through the Qubit 4 fluorometer Broad Range Assay (Invitrogen).

Whole-genome short-read sequencing

We outsourced the sequencing of four *N. lapillus* DNA samples—two from Brest, France (2n = 26), and two from Udra, Spain (2n > 26)—to Macrogen in South Korea for whole-genome DNA Illumina paired-end sequencing. Whole-genome DNA libraries were prepared using an Illumina TruSeq DNA PCR-free kit with 350 bp insert size. Subsequently, the libraries were subjected to paired-end sequencing on an Illumina NovaSeq 6000 System, employing a 300-cycle NovaSeq 6000 SP Reagent Kit v 1.5 and resulting in a final yield of 8 (Brest samples) and 30 (Udra samples) Gb per sample. These accounted for circa 2000× of the complete mtDNA.

Table 1 Sampling locations of *N. lapillus*, indicating the chromosomal races

Country	Locality	Acronym	Latitude	Longitude	Chrom. Race
France	Roscoff	ROSC	48°43'35.71"N	3°59'21.62"W	2n > 26
France	Granville	GRAN	48°50'0.88"N	1°36'49.08"W	2n = 26
France	Brest	BRES	48°21'30.04"N	4°32'0.20"W	2n = 26
UK	Whitsand Bay	WITS	50°20'45.53"N	4°15'18.34"W	2n = 26
UK	Cawsand	CAWS	50°19'48.05"N	4°12'0.31"W	2n > 26
Spain	Senín	SENI	42°44'17.46"N	9°04'55.53"W	2n = 26 and 2n > 26
Spain	Udra	UDR	42°19'53.98"N	8°49'36.44"W	2n = 26 and 2n > 26

Whole-genome long-read sequencing

Whole-genome long-read libraries were obtained using a sequencing kit by Oxford Nanopore Technologies Ltd (ONT, SQK-LSK109) on gDNA from a single specimen from Udra (Spain, $2n > 26$) subjected to size selection using the Short Read Eliminator kit (Circulomics), end-prep repair (NEB-Next FFPE DNA Repair Mix, New England BioLabs) and adapter ligation (NEBNext Ultra II Ligation Module, New England BioLabs). These libraries were loaded in MinION R9.4 flowcells (Oxford Nanopore Technologies Ltd; FLO-MIN106 rev-D), controlled by MinKNOW v19.12.5. Sequencing raw data was converted to fastq with Guppy v6.5.7 in super accuracy mode. The final yield accounted for 1 Gb, representing approximately 70× of the mtDNA.

Mitochondrial DNA assembly

Primary assembly of the mitochondrial genome was conducted using MITObim v1.9.1 with default parameters (Hahn et al. 2013), leveraging short-read sequencing data and employing a fragment of the Cytochrome C Oxidase I gene from *N. lapillus* (EU391582.1) to prime the assembly. Subsequently, Illumina paired-end reads were realigned to the resulting draft mtDNAs through BWA-mem v0.7.17-r1188 (Li and Durbin 2010). The alignments were then sorted, quality filtered (applying a threshold of $-q60$) and indexed using samtools v1.10 (Li et al. 2009). The resulting alignment files were employed for two rounds of polishing the draft mtDNAs using Pilon v1.23 (Walker et al. 2014), after which we assessed the accuracy of SNPs by visual inspection of the mapped paired-end reads and structural artifacts by mapping nanopore reads onto the draft with minimap2 v2/2.24 (Li 2018) and variant calling using sniffles v2.2 (Sedlazeck et al. 2018). The unassembled portion of the control region was deduced from ONT long reads, as generated by flye v2.9 (Kolmogorov et al. 2019) and self-corrected with medaka 1.7.2 (<https://github.com/nanoporetech/medaka>). We employed the MITOS2 web server (Donath et al. 2019) to annotate all coding and RNA genes on the resulting polished mitochondrial genomes. The mitochondrial genetic code for invertebrates was applied during the annotation process and the RefSeq63 dataset for metazoans served as reference data. We manually curated the annotations to fit the ORF predictions by ORF-FINDER (Rombel et al. 2002) and other available Muricidae mitogenomes available in Genbank.

Mitochondrial multiplexed PCR

To generate a larger cohort for the mitochondrial DNA of *N. lapillus*, a set of partially overlapping PCRs were designed (Supplementary Table 1) using Primal Scheme

following previously established protocols (Quick et al. 2017). Genomic DNA was extracted from specimens of *N. lapillus* from different Spanish, French and British populations whose chromosomal number was previously established (Table 1). Their mtDNAs were then amplified in 26~650 bp fragments from two pooled multiplexed PCRs, using 1× Q5® High-Fidelity Master Mix (New England Biolabs). PCR products were purified, end-prep repaired, and ligated to ONT-specific barcoding adapters (Native Barcoding Expansion kits 1–12 EXP-NBD104 and 13–24, EXP-NBD114, Oxford Nanopore Technologies Ltd) using the NEB Blunt/TA Ligase Master Mix (New England Biolabs). After equimolar mixing and purification of the adapted fragments, library adapters (Adapter Mix II, Oxford Nanopore Technologies Ltd) were ligated to the libraries in a single reaction using Quick T4 DNA Ligase following the manufacturer's recommendations. The resulting libraries were then sequenced in MinION R9.4 flowcells (FLO-MIN106 rev-D), controlled by MinKNOW v19.12.5, and the data was base called in "high-accuracy" mode 8 (dna_r9.4.1_450bps_hac.cfg), de-multiplexed assuming barcodes in both ends of the inserts and filtered to ensure a Phred quality score > 7 with Guppy v6.5.7. All downstream processing, including trimming, alignment, variant identification and consensus recovery of mitochondrial DNA sequences, was performed using the ARTIC bioinformatics minion pipeline (Pater et al. 2021), prefiltering chimeric reads below 400 bp and above 700 bp and normalizing coverage to 400 reads per position.

Phylogenetic analyses

The mitochondrial sequences corresponding to the 13 protein-coding genes (PCGs) and the two ribosomal RNA (rRNA) genes from the 40 sequenced *N. lapillus* and the eight available species of the family Muricidae at Genbank (*Thais luteostoma* (TLU, NC_039165), *Chicoreus torrefactus* (CTO, NC_039165), *Boreotrophon candelabrum* (BCA, NC_046505, Tian et al. 2019), *Menathais tuberosa* (MTU, NC_031405, Sung et al. 2016), *Ocinebrellus falcatus* (OFA, NC_046052, Hao et al. 2019c), *Ocinebrellus inornatus* (OIN, NC_046577, Hao et al. 2019a), *Ceratostoma burnetti* (CBU, NC_046569, Hao et al. 2019b) and *Ceratostoma rorifluum* (CRO, NC_046526, Tian et al. 2019) were extracted, concatenated and aligned with MUSCLE v3.8.425 (Edgar 2004). We determined the best-fit nucleotide substitution model using JModelTest2 (Darriba et al. 2012). Subsequently, we constructed a maximum-likelihood phylogeny using PhyML (Guindon et al. 2010) with 100 bootstraps. To infer the evolutionary relationships among populations of *N. lapillus*, we employed the starBEAST multispecies coalescent model within BEAST v2.6.2 (Bouckaert et al. 2019). Our analysis utilized a Yule speciation prior and a strict molecular clock. The best-fit nucleotide substitution

model, determined by jModelTest2 (Darriba et al. 2012), was applied to the concatenated mitochondrial haplotypes, which included 13 PCGs and two rRNA genes. For this analysis, we defined the seven sampling points (Cawsand, Whitsand Bay, Brest, Granville, Roscoff, Udra and Senín) along with nine outgroup species as tips for the species tree. We ran a single MCMC simulation for ten million iterations, with samples collected every 1000 steps. A burn-in period of 10% was used to ensure that effective sample size values exceeded 200, as assessed with Tracer v1.7.1. The resulting posterior distributions of trees were visualized using DENSITREE v2.1 (Bouckaert 2010). Finally, a maximum clade credibility tree was generated with TreeAnnotator (Bouckaert et al. 2019) to summarize the topology information, with a 10% burn-in and common ancestors for node height calculations. We assessed the genetic distance within and between the sampling points based on the Kimura-2 parameter (Kimura 1980) in MEGA (Kumar et al. 2018). We conducted a PCA analysis based on the nucleotide diversity of the obtained mitogenomes with the R package *adegenet* (Jombart 2008) and an analysis of molecular variance (AMOVA) by means of the package *pegas* in R (Paradis 2010) to quantify how the genetic variability is partitioned among and within geographical populations and chromosomal races.

Results

Karyological analysis

Our cytogenetic analysis revealed specimens of exclusively $2n = 26$ chromosomes in the French populations from Brest and Granville and the British locality of Whitsand Bay, while all specimens examined from Roscoff (France) and Cawsand (UK) belonged to the $2n > 26$ chromosomal race. On the contrary, both chromosomal races were found to

coexist in the Spanish sampling sites of Udra and Senín (Table 1). Examples of metaphase plates with different chromosome numbers are shown in Fig. 1.

General features of *N. lapillus* mitochondrial genome

Preliminary drafts of the mitochondrial genome draft of four *N. lapillus* specimens based on short reads yielded single circular DNA molecules, 15,337 bp in length. Our analysis revealed that the gene arrangement of these specimens matched the content and ordination of the Muricidae family, which comprised 37 genes, including 13 PCGs, 22 transfer RNA (tRNAs) genes and 2 ribosomal RNA genes. Our analysis also revealed that the mitochondrial genome of *N. lapillus* specimens contained two copies of tRNA genes for serine (trnS1 and trnS2) and leucine (trnL1 and trnL2), as found in the vast majority of metazoans (Tomita et al. 1998). Notably, all of the features observed in our study were encoded on the D-strand, with the exception of eight tRNA genes, the seven tRNA gene array comprised between *atp6* and *rrnS*, and the *trnT*(tgt), which were found in antisense.

Upon closer inspection of the preliminary draft based on short reads only, it became apparent that the control region and its regulatory elements were mostly absent. Indeed, remapping of Illumina reads onto the preliminary draft assemblies revealed a small region devoid of coverage despite high coverage in the immediately adjacent regions (Fig. 2a). This gap suggested the presence of an artifact in the assembly, possibly due to a repeat in the control region that was longer than the Illumina library insert size and thus not fully resolved by neither Mitobim nor Pilon (Fig. 2a). Upon variant calling using sniffles on nanopore reads mapped on this first preliminary draft, a 1037 bp insertion became apparent. To address this issue, we utilized long-read sequencing technology to reconstruct this conflictive region, which included a 450 bp perfect inverted repeat

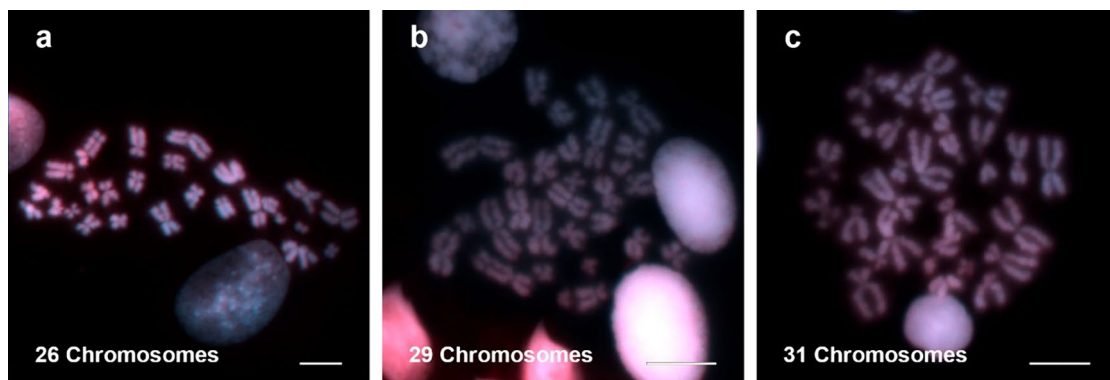


Fig. 1 Metaphase plates from *Nucella lapillus* specimens analyzed in this study holding 26 (a, Whitsand Bay—UK), 29 (b, Cawsand—UK) and 31 (c, Roscoff—France) chromosomes. Scale bars, 5 μ m

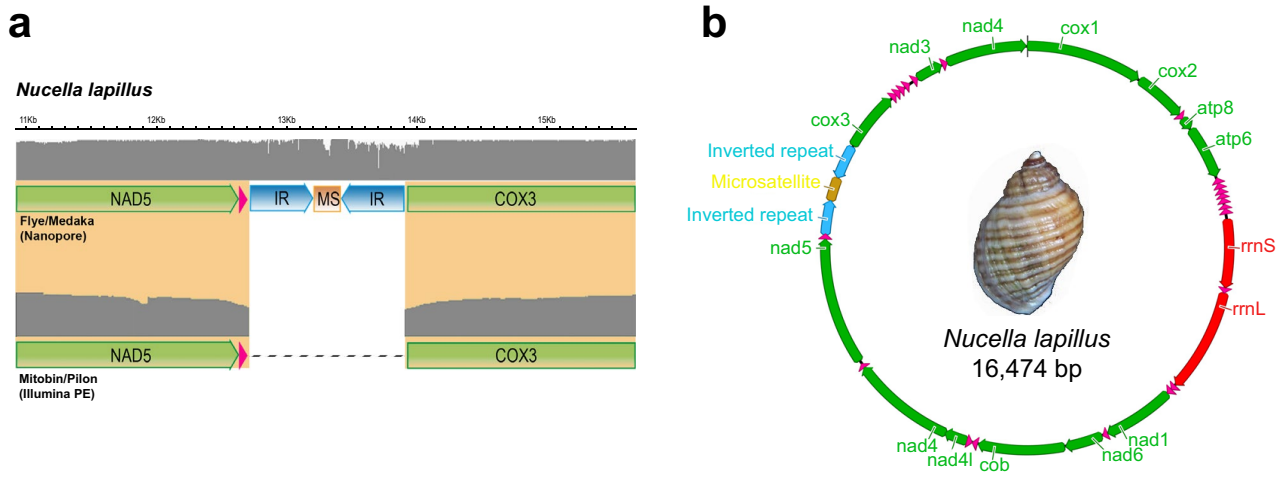


Fig. 2 Complete assembly of the mitochondrial genome of *N. lapillus*. **(a)** Schematic comparison of the control region of the mitochondrial genome drafts of *N. lapillus*, as recovered by nanopore long reads (Flye and Medaka) and Illumina paired-end short reads (Mitobin and Pilon), highlighting the inverted repeats (IR) and

AT microsatellite (MS) at the control region absent from the short-read only based assembly. In gray, mean read coverage (40×ONT, 500×Illumina). **(b)** Circular map of the complete mitochondrial genome of *N. lapillus*

flanking an AT microsatellite 299 bp in length (Fig. 2a). Incorporating this inverted repeat, the complete draft for the mitochondrial DNA of *N. lapillus* was found to be 16,474 bp in length (Fig. 2b).

As predicted by *mitos2*, some of the genes required manual curation. Such is the case for the ORF of the *nad4* gene that extends beyond its expected length and partially overlaps with the adjacent tRNAH sequence toward the 3' end, until reaching the nearest theoretical stop codon. This was due to an incomplete UAA codon, terminated upon polyadenylation of the resulting mRNA.

Multiplexed PCR genome sequencing

We used a targeted mitochondrial multiplexed PCR approach to sequence 40 *N. lapillus* specimens, including 19 with $2n=26$ and 21 with $2n>26$ chromosomes from different locations. The reads were aligned to the first *N. lapillus* mitochondrial draft genome, with a range of 2422 to 9714 reads recovered per genome (mean = 9364; SD = 3012), and the depth of coverage varied from 105.233X to 407.967X (mean = 393.34X; SD = 123.20X), according to analysis after normalization with the ARTIC bioinformatics pipeline (Supplementary Table 2). However, coverage decreased significantly in all samples due to the lack of primer efficiency at amplicon 10 and an almost complete dropout for amplicon 20, which spans the tandem inversion on the control region, and to the loss of both the 5' (32 nucleotides) and 3' (276 nucleotides) termini regions beyond the most outer primer binding sites (Fig. 3). We verified the accuracy of this multiplexed PCR nanopore sequencing approach for SNP calling

by comparing it with whole-genome Illumina sequencing for a single specimen from Udra, Spain (NLA_SP_U01). Following established strategies (Mariner-Llicer et al. 2021), we achieved nearly concordant consensus sequences between the two platforms, with 53 out of 54 true positive called SNPs and 3 out of 3 filtered true negatives, resulting in an F1-score of 0.99, high agreement (98.47%), accuracy (98.25%) and a perfect true positive rate (100%), as well as an observed error rate of 1.85%, thus demonstrating the reliability of this pipeline for sequence comparison and analysis. Despite incomplete genome coverage resulting from these coverage dropouts, this protocol allowed for sufficient recovery of almost the complete mtDNA, producing reliable consensus sequences for all tested samples to undergo phylogenetic analysis.

Phylogenetic analysis of mitochondrial sequences

Phylogenetic analysis was conducted on a sequence alignment encompassing 12.8 kb of most of the mitochondrial genomes (13 PCGs and 2 rRNA gene sequences) of 40 *N. lapillus* specimens and the other available Muricidae mitochondrial genomes in Genbank. The resulting phylogenetic tree shows the corresponding genetic relationships, with *N. lapillus* samples clustering predominantly by population rather than by their chromosomal races (Fig. 4a, Supplementary Fig. 1). *O. falcatus* and *O. inornatus* are sister taxa to *N. lapillus* and form a clade that is sister to a clade composed of *C. burnetti* and *C. rorifluum*. We also conducted a multilocus species tree analysis using a dataset comprising the 13 PCGs and the two rRNA genes from the mitochondrial genomes (Fig. 4b),

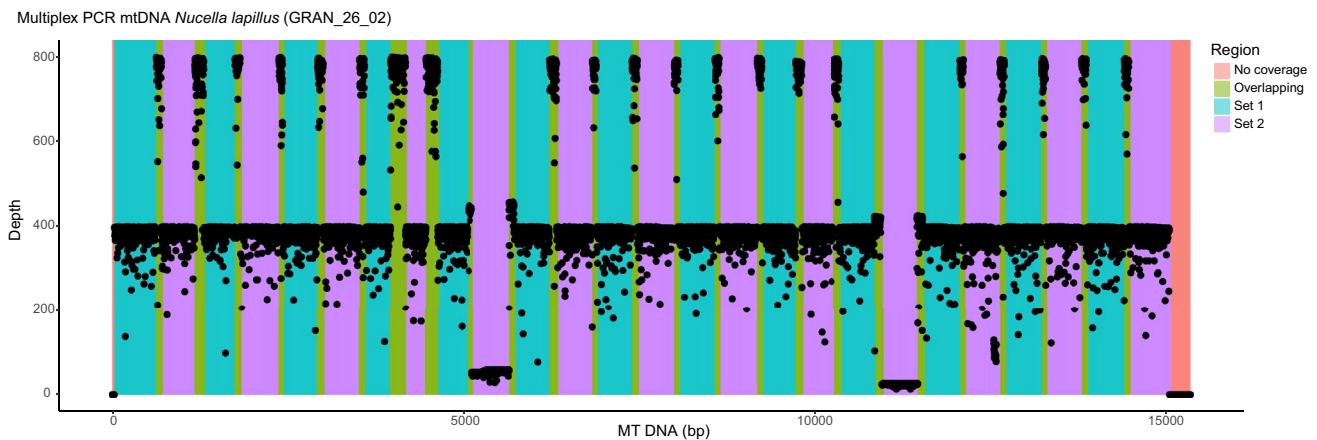


Fig. 3 Whole *N. lapillus* mitochondrial genome coverage plot. We performed nanopore amplicon sequencing of the entire *N. lapillus* mitochondrial genome, as PCR multiplexed with two sets of 700 bp partially overlapping amplicons (Set-1 and Set-2)

treating the collecting locations as separate entities for species classification. These results highlight that these populations exhibit a clearer clustering pattern based on geographical distinctions rather than by their chromosomal race. Most northern sampling points in the UK and France (Whitsand Bay, Cawsand, and Roscoff) form a distinct cluster separate from the Spanish locations, with the notable exception of Granville, which shows closer genetic affinity to the Spanish samples. Additionally, a PCA analysis based on pairwise genetic distances among all sequenced specimens further supports this population-based clustering (Fig. 4c).

Uncorrected K2P pairwise distances averaged 0.0018 across the population clades identified in Fig. 4b, underscoring a low genetic variation among the studied populations (Table 2). To further dissect genetic variation within and between populations of *N. lapillus* and assess the impact of chromosomal race on mitochondrial DNA variation, separate AMOVAs were conducted. Our results revealed a moderate genetic differentiation between populations, with 22.0% of the total variation attributable to differences between sampling points ($F_{st}=0.318$, $p < 0.05$), while the remaining 78.0% was due to differences within populations, suggesting a moderate genetic differentiation between populations. However, chromosomal race exerted no significant effect, explaining only 0.3% of the total variation ($F_{st}=0$, $p > 0.05$). These findings suggest that although there is moderate genetic differentiation between geographical populations, chromosomal race does not significantly influence the observed genetic variance in the mitochondrial DNA.

Discussion

This study presents the first assembly of mitochondrial DNA sequences for *N. lapillus*, laying the groundwork for future genomic comparative analyses in this species. The obtained sequences were used to investigate the mitochondrial genetic diversity and differentiation among seven populations spanning a broad geographic range and including both the $2n = 26$ and $2n > 26$ chromosomal races.

Assembling mitochondrial genomes from short-read data remains a formidable task, even with the use of high-throughput sequencing technology and rigorous data filtering (Formenti et al. 2021). The accuracy of the obtained sequences can be impacted by various factors, such as the coverage depth, read quality and the presence of repetitive elements that are longer than the average insert length and thus caution should be taken when interpreting the results, especially regarding the identification of rare or divergent haplotypes (Keraite et al. 2022). Our study also highlights the limitations of assemblies based uniquely on short-read technologies, as this approach was unable to generate a complete mitochondrial genome due to the presence of repetitive regions on it. In this regard, repeats in the control region of mitochondrial DNA, whether microsatellites or longer motifs, are in fact quite common in many taxa and are likely underrepresented in many of the available reference mitogenomes due to the intrinsic limitations of short-fragment assemblies (Formenti et al. 2021).

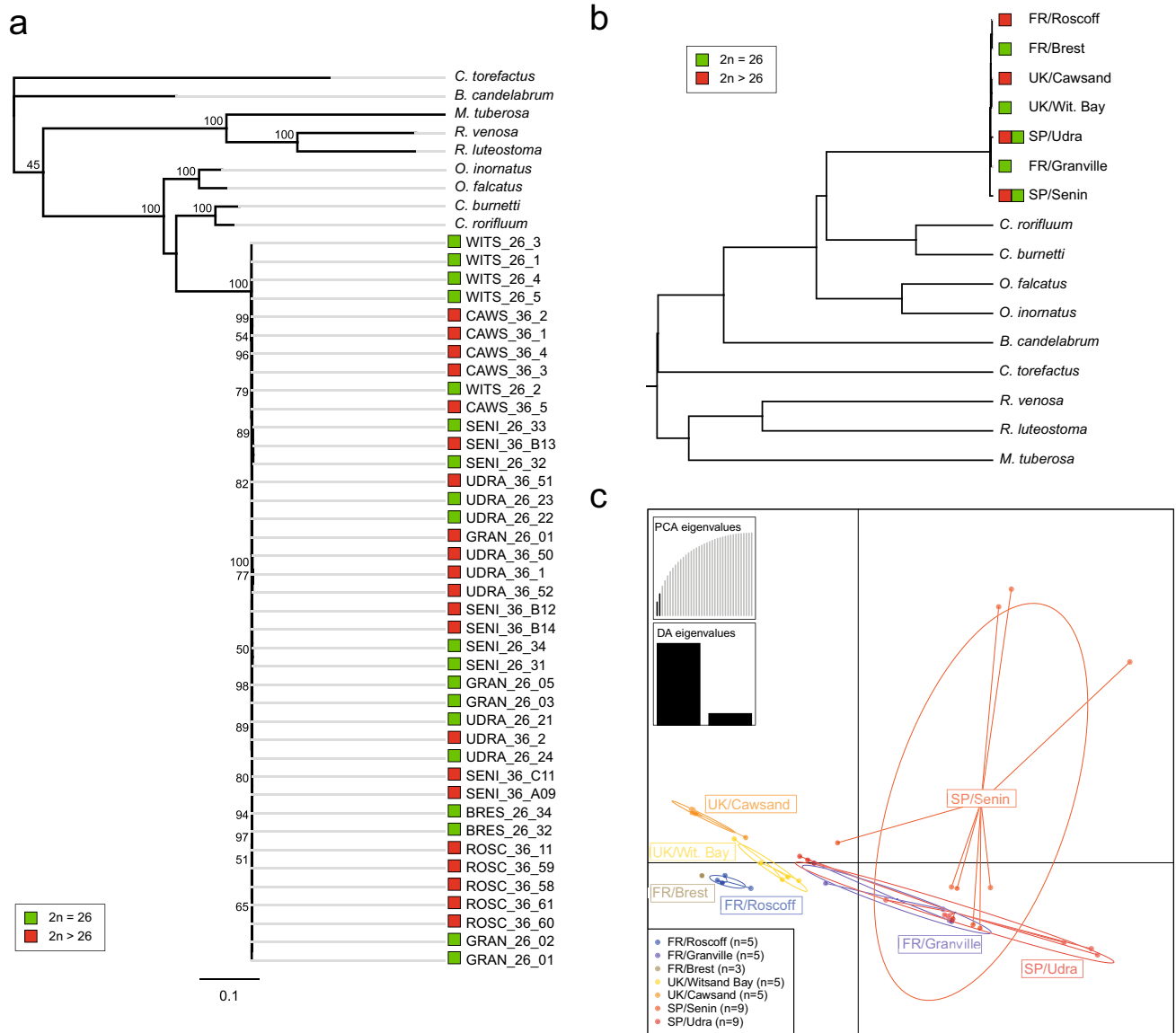


Fig. 4 Phylogeny, species tree and principal component analysis (PCA) of seven *N. lapillus* populations. **(a)** A maximum-likelihood phylogenetic tree based on the mitochondrial DNA sequences of seven populations of *N. lapillus* and the available Muricidae mitogenomes. ML bootstrap percentages > 50 are indicated at the nodes. Scale bar shows nucleotide substitutions per site. Chromosomal races

are indicated as green (2n=26 chromosomes) or red (2n>26 chromosomes) squares at the tips. A proportional branch transformed version of the same phylogenetic tree is provided in Supplementary Fig. 1. **(b)** Multispecies coalescent tree of the Muricidae family based on their mitochondrial DNAs. **(c)** Principal component analysis of *N. lapillus* populations

Table 2 Average evolutionary divergence estimates (K2P) based on the mitochondrial DNAs of *N. lapillus* within and between population sampling points

No	Population	Intrapopulation K2P distances	Interpopulation K2P distances							
			1	2	3	4	5	6	7	
1	Roscoff	0.00027								
2	Granville	0.00177	0.00151							
3	Brest	0.00000	0.00038	0.00163						
4	Whitsand Bay	0.00119	0.00127	0.00173	0.00120					
5	Cawsand	0.00070	0.00138	0.00180	0.00131	0.00130				
6	Senin	0.00320	0.00257	0.0028	0.00263	0.00269	0.00259			
7	Udra	0.00144	0.00160	0.0018	0.00154	0.00161	0.00169	0.00271		

This limitation may also apply to the other nine available Muricidae mitogenomes stored at Genbank, which were all obtained using short Illumina reads and present much-reduced control regions. The resolution of these conflicting regions often requires long-read sequencing for hybrid assembly to achieve complete resolution (De Coster et al. 2021). In fact, recent revisions of multiple vertebrate mitogenomes using long-read sequencing revealed widespread repeats and gene duplications that were previously undetected by short-read assemblies, particularly at the control region (Formenti et al. 2021).

Mitochondrial genes have been reported to exhibit incomplete stop codons ending in U or UA in some species, including humans. These are known to become complete termination codons (UAA) upon subsequent polyadenylation of the mRNA (Hou et al. 2006; Oh et al. 2007; Ki et al. 2010; Temperley et al. 2010). This feature has also been reported in the mitochondrial genomes of other Muricidae gastropods, including *Ocenebrellus falcatus* (NC_046052), *Ocenebrellus inornatus* (NC_046577), *Ceratostoma burnetti* (NC_046569) and *Ceratostoma rorifluum* (NC_046526), all with the notation “TAA stop codon is completed by the addition of 3' A residues to the mRNA” and “*transl_except*=(*pos*:10,711, *aa*:TERM).” The *nad4* gene in *N. lapillus* spans 1372 bp, equivalent to 457 triplets and a single terminal T, which would be completed to a TAA stop codon with the addition of two 3' A residues to the mRNA. This translational anomaly is worth noting, as it contributes to the diversity of mitochondrial genomes in the Muricidae family and may have implications for protein synthesis and function (Temperley et al. 2010).

We found no rearrangements of the mitochondrial genomes of *N. lapillus* with respect to the other available muricid mitogenomes, which is in contrast to the high level of variability that is typically observed in mitochondrial gene ordination in molluscs, even within the same genus (Varney et al. 2021). We speculate that this could be attributed to the recent radiation of the muricids, which may not have allowed enough time for their mitochondrial DNAs to undergo such structural variations. Nevertheless, it is important to note that this is but a minuscule representation of the species within this family, comprising more than 1600 recognized taxa (Barco et al. 2010). Therefore, a comprehensive analysis of the evolution of their mitogenomes would require a much larger effort encompassing a wider range of species.

The presence of inverted repeats in the control region of the mitochondrial genome of *N. lapillus* might be a common feature in the Muricidae family, since the available mitogenomes for this family at NCBI, based solely on short reads, apparently lack the necessary control region, likely due to these repeats collapsing during incomplete draft genome assembly. Otherwise, the control region provides valuable genetic information, accumulating more mutations as it is

less subjected to evolutionary constraints. In this regard, length polymorphisms at the DNA repeats on it can be used to discriminate between species and populations, and thus may have taxonomical applicability in discerning mitochondrial lineages and establishing intra- and interspecific relationships. Therefore, understanding and properly sequencing these control regions is crucial for comprehensive analysis of the mitochondrial genome and its evolutionary history in different taxa. For instance, DNA repeats have been found to be useful in distinguishing populations of other species, such as the green sea turtle *Chelonia mydas* (Encalada et al. 1996) and the Japanese eel *Anguilla japonica* (Sang et al. 1994), where they have been used to assess population structure and connectivity. The inverted repeats and an internal microsatellite at the control region of *N. lapillus* could vary in length between geographically or ecologically distant animals, providing a valid marker for population analysis. This possibility will be explored in future research.

Various PCR approaches have been successfully employed for the retrieval of complete mtDNAs across diverse species. These include, among others, PCR-tiling techniques, as demonstrated by Xin et al. (2017), or primer walking strategies, exemplified by Serb and Lydeard (2003). These targeted approaches facilitate the retrieval of complete mitochondrial genomes with a high level of accuracy, offering a cost-effective alternative to whole-genome sequencing (Quick et al. 2017). Beyond cost-efficiency and accuracy, these PCR-based methods boast scalability, enabling simultaneous analysis of multiple samples within a single sequencing run. This scalability not only further reduces overall costs, but also significantly enhances analysis throughput. In addition, the multiplex PCR approach is highly efficient and only requires small amounts of starting DNA, which is especially useful for the study of rare or endangered species, where sample availability can be limited. The application of multiplex PCR coupled with ONT sequencing for the recovery of mitochondrial DNAs presents a valid and cost-effective alternative to whole-genome NGS sequencing for phylogenetic analysis (Walsh et al. 2022; Karin et al. 2023). Overall, this approach represents a valuable and cost-effective approach to the recovery of mitochondrial DNA for phylogenetic analysis.

In this study, we observed a low genetic diversity and moderate geographical differentiation in mitochondrial DNA among *N. lapillus* populations, with no apparent differentiation between chromosomal races. This contrasts with previous findings from nuclear markers such as allozymes (Day et al. 1993; Kirby et al. 1994), microsatellites (Guerra-Varela et al. 2009) and mitochondrial genes (Kirby 2000), which implied partial genetic isolation between both races, leading us to anticipate at least some genetic differentiation on their mitochondrial DNAs. One possible explanation for this lack of mtDNA differentiation

could be an ancient introgression and replacement of either mitochondrial haplotype. Both chromosomal races likely underwent differentiation during a period of allopatric isolation. Subsequently, a secondary overlap and introgression would have mitigated the genetic differentiation on neutral markers, such as the mitochondrial DNA (Panova et al. 2011). After this, isolation by distance would have played a role in the genetic differentiation between populations at present, leading to the distinct evolution of the current mitochondrial haplotypes. Similar results have been reported in other marine gastropods, such as *Littorina* spp. A study by Johannesson et al. (1993) found evidence of mitochondrial introgression between subspecies of *L. saxatilis*, while other authors reported a similar process between *L. fabalis*, *L. obtusata* and *L. arcana* (Panova et al. 2011; Sotelo et al. 2020). Mitochondrial DNA introgression has also been documented in other sessile marine animals, such as corals (Fukami and Knowlton 2005) or sponges (Turon and López-Legentil 2004). These studies highlight the potential role of mitochondrial DNA introgression in driving genetic differentiation and speciation in marine organisms (Sotelo et al. 2020).

Nonetheless, it is also conceivable that, driven by specific selective pressures, the karyological evolution outpaced that of the mitochondrial DNAs, otherwise subjected to neutral selection. Indeed, a correlation between the chromosomal polymorphism in *N. lapillus* and Darwinian fitness has been suggested before (Pascoe 2006), although the extent to which this structural variation exerts, or has ever exerted, such an effect remains uncertain. The frequent occurrence of spontaneous somatic Robertsonian translocations in $2n > 26$ specimens at present (Dixon et al. 1994; Pascoe 2006) implies a genetic predisposition to such rearrangements, also potentially triggered or intensified by environmental factors, providing additional support to this hypothesis. Interestingly, this scenario would also open up the possibility of parallel karyological evolution, for which either chromosomal race might have evolved independently in multiple locations on similar ecological niches. Thus, the absence of significant genetic differences between the mitochondrial DNAs of both chromosomal races may be interpreted as a consequence of disparate evolutionary rates between the karyotypes and mitochondrial DNAs. Further research is deemed to explore the possible drivers of genetic differentiation and the potential fitness implications of the karyological differences between both chromosomal races of *N. lapillus*.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s00227-024-04424-3>.

Author contributions All authors contributed to the study conception and design. Material preparation, data collection and analysis were

performed by Daniel García-Souto, Jonathan Fernández-Rodríguez, André Vidal-Capón, Neil Fuller and Juan Galindo. The first draft of the manuscript was written by Daniel García-Souto and all authors commented on previous versions of the manuscript. All authors read and approved the final manuscript.

Funding Open Access funding provided thanks to the CRUE-CSIC agreement with Springer Nature. DG-S acknowledges support from the Consellería de Cultura, Educación e Ordenación 361 Universitaria, Xunta de Galicia, under the grant ED481D-2022-001. Funding was provided by the Ministerio de Ciencia e Innovación (PID2021-124930NB-I00) to JG, the ASSEMBLE PLUS (5th 363 call) programme to DG-S (NLA Chrom Evol), Xunta de Galicia (Centro singular de investigación de Galicia accreditation 2019-2022), the European Union (European Regional Development Fund - ERDF) and Xunta de Galicia (ED431C 2020-05).

Data availability All raw sequencing data generated for this project is stored at the NCBI SRA database under the Bioproject PRJNA1060882. All assembled mitogenomes are available at GenBank under the accession numbers PP147811-PP147850.

Declarations

Conflict of interest The authors declare no conflict of interest.

Ethical approval The Research Ethics Committee of the University of Vigo has confirmed that no ethical approval is required.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Bantock CR, Cockayne WC (1975) Chromosomal polymorphism in *Nucella lapillus*. *Heredity* 34:231–245
- Barco A, Claremont M, Reid DG, Houart R, Bouchet P, Williams ST, Cruaud C, Couloux A, Oliverio M (2010) A molecular phylogenetic framework for the Muricidae, a diverse family of carnivorous gastropods. *Mol Phylogenet Evol* 56(3):1025–1039. <https://doi.org/10.1016/j.ympev.2010.03.008>
- Bouckaert RR (2010) DensiTree: making sense of sets of phylogenetic trees. *Bioinformatics* 26(10):1372–1373. <https://doi.org/10.1093/bioinformatics/btq110>
- Bouckaert RR, Vaughan TG, Barido-Sottani J, Duchêne S, Fourment M, Gavryushkina A, Heled J, Jones G, Kühnert D, De Maio N, Matschiner M, Mendes FK, Müller NF, Ogilvie HA, du Plessis L, Poppinga A, Rambaut A, Rasmussen D, Siveroni I, Suchard MA, Wu CH, Xie D, Zhang C, Stadler T, Drummond AJ (2019) BEAST 2.5: an advanced software platform for Bayesian evolutionary analysis. *PLoS Comput Biol* 15(4):e1006650. <https://doi.org/10.1371/journal.pcbi.1006650>

- Collins TM, Frazer K, Palmer AR, Vermeij GJ, Brown WM (1996) Evolutionary history of northern hemisphere *Nucella* (Gastropoda, Muricidae): molecules, ecology, and fossils. *Evolution* 50:2287–2304. <https://doi.org/10.1111/j.1558-5646.1996.tb03617.x>
- Colson I, Hughes RN (2007) Contrasted patterns of genetic variation in the dogwhelk *Nucella lapillus* along two putative post-glacial expansion routes. *Mar Ecol Prog Ser* 343:183–191. <https://doi.org/10.3354/meps06879>
- Crothers JH (1985) Dog-whelks: an introduction to the biology of *Nucella lapillus* (L.). *Field Stud* 6:291–360
- Darriba D, Taboada G, Doallo R, Posada D (2012) jModelTest 2: more models, new heuristics and parallel computing. *Nat Methods* 9:772. <https://doi.org/10.1038/nmeth.2109>
- Day AJ, Leinaas HP, Anstensrud M (1993) Allozyme differentiation of populations of the dogwhelk *Nucella lapillus*, (L.): the relative effects of geographic distance and variation in chromosome number. *Biol J Linn Soc* 51:257–277. <https://doi.org/10.1111/j.1095-8312.1994.tb00961.x>
- De Coster W, Weissensteiner MH, Sedlazeck FJ (2021) Towards population-scale long-read sequencing. *Nat Rev Genet* 22:572–587. <https://doi.org/10.1038/s41576-021-00367-3>
- Dixon DR, Pascoe PL, Gibbs PE, Pasantes J (1994) The nature of Robertsonian chromosomal polymorphism in *Nucella lapillus*: a re-examination. In: Beaumont AR (ed) *Genetics and evolution of aquatic organisms*. Chapman & Hall, London, pp 389–399
- Donath A, Jühling F, Al-Arab M, Bernhart SH, Reinhardt F, Stadler PF, Middendorf M, Bernt M (2019) Improved annotation of protein-coding gene boundaries in metazoan mitochondrial genomes. *Nucleic Acids Res* 47(20):10543–10552. <https://doi.org/10.1093/nar/gkz833>
- Edgar RC (2004) MUSCLE: a multiple sequence alignment method with reduced time and space complexity. *BMC Bioinform* 5:113. <https://doi.org/10.1186/1471-2105-5-113>
- Encalada SE, Lahanas PN, Bjørndal KA, Bolten AB, Miyamoto MM, Bowen MW (1996) Phylogeography and population structure of the Atlantic and Mediterranean green turtle *Chelonia mydas*: a mitochondrial DNA control region sequence assessment. *Mol Ecol* 5(4):473–483. <https://doi.org/10.1111/j.1365-294X.1996.tb00340.x>
- Formenti G, Rhie A, Balacco J, Haase B, Mountcastle J, Fedrigo O, Brown S, Capodiferro MR, Al-Ajli FO, Ambrosini R, Houde P, Koren S, Oliver K, Smith M, Skelton J, Betteridge E, Dolucan J, Corton C, Bista I, Torrance J, Tracey A, Wood J, Uliano-Silva M, Howe K, McCarthy S, Winkler S, Kwak W, Korlach J, Functamman A, Fordham D, Costa V, Mayes S, Chiara M, Horner DSH, Myers E, Durbin R, Achilli A, Braun EL, Phillippy AM, Jarvis ED, Vertebrate Genomes Project Consortium (2021) Complete vertebrate mitogenomes reveal widespread repeats and gene duplications. *Genome Biol* 22:120. <https://doi.org/10.1186/s13059-021-02336-9>
- Fukami H, Knowlton N (2005) Analysis of complete mitochondrial DNA sequences of three members of the *Montastraea annularis* coral species complex (Cnidaria, Anthozoa, Scleractinia). *Coral Reefs* 24:410–417. <https://doi.org/10.1007/s00338-005-0023-3>
- García-Souto D, Alonso-Rubido S, Costa D, Eirín-López JM, Rolán-Álvarez E, Faria R, Galindo J, Pasantes JJ (2018) Karyotype characterization of nine Periwinkle species (Gastropoda, Littorinidae). *Genes* 9(11):517. <https://doi.org/10.3390/genes9110517>
- Guerra-Varela J, Colson I, Backeljau T, Breugelmanns K, Hughes RN, Rolán-Álvarez E (2009) The evolutionary mechanism maintaining shell shape and molecular differentiation between two ecotypes of the dogwhelk *Nucella lapillus*. *Evol Ecol* 23:261–280. <https://doi.org/10.1007/s10682-007-9221-5>
- Guindon S, Dufayard JF, Lefort V, Anisimova M, Hordijk W, Gascuel O (2010) New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Syst Biol* 59(3):307–321. <https://doi.org/10.1093/sysbio/syq010>
- Hahn C, Bachmann L, Chevreaux B (2013) Reconstructing mitochondrial genomes directly from genomic next-generation sequencing reads—a baiting and iterative mapping approach. *Nucleic Acids Res* 41(13):e129. <https://doi.org/10.1093/nar/gkt371>
- Hao Z, Tian Y, Yang J, Zhu J, Shi L, Cheng C, Chang Y (2019a) The complete mitochondrial genome of *Ceratostoma inornata* (Récluz, 1851). *Mitochondrial DNA Pt B* 4(1):1152–1153. <https://doi.org/10.1080/23802359.2019.1591196>
- Hao Z, Tian Y, Yang J, Zhu J, Shi L, Li Z, Chang Y (2019b) The complete mitochondrial genome of *Ceratostoma burnetti* (Adams et Reeve, 1849). *Mitochondrial DNA Pt B* 4(1):1159–1160. <https://doi.org/10.1080/23802359.2019.1591205>
- Hao Z, Tian Y, Yang J, Zhu J, Shi L, Li Z, Chang Y (2019c) The complete mitochondrial genome of *Pteropurpura falcatius* (Sowerby, 1834). *Mitochondrial DNA Pt B* 4(1):1179–1180. <https://doi.org/10.1080/23802359.2019.1591184>
- Hou WR, Chen Y, Wu X, Hu JC, Peng ZS, Yang J, Tang ZX, Zhou CQ, Li YM, Yang SK, Du YJ, Kong LL, Ren ZL, Zhang HY, Shuai SR (2006) A complete mitochondrial genome sequence of Asian black bear Sichuan subspecies (*Ursus thibetanus mupinensis*). *Int J Biol Sci* 3(2):85–90. <https://doi.org/10.7150/ijbs.3.85>
- Johannesson K, Panova M, Kempainen P, André C, Rolán-Álvarez E, Butlin R (1993) Repeated evolution of reproductive isolation in a marine snail: unveiling mechanisms of speciation. *Philos Trans R Soc b: Biol Sci* 340(1292):203–219. <https://doi.org/10.1098/rstb.2009.0256>
- Jombart T (2008) adegenet: a R package for the multivariate analysis of genetic markers. *Bioinformatics* 24(11):1403–1405. <https://doi.org/10.1093/bioinformatics/btn129>
- Karin BR, Arellano S, Wang L, Walzer K, Pomerantz A, Vasquez JM, Chatla K, Sudmant PH, Bach BH, Smith LL, McGuire JA (2023) Highly-multiplexed and efficient long-amplicon PacBio and nanopore sequencing of hundreds of full mitochondrial genomes. *BMC Genomics* 24:229. <https://doi.org/10.1186/s12864-023-09277-6>
- Keraite I, Becker P, Canevazzi D, Frias-López C, Dabad M, Tonda-Hernandez R, Paramonov I, Ingham MJ, Brun-Heath I, Leno J, Abulí A, García-Arumí E, Heath SC, Gut M, Gut IG (2022) A method for multiplexed full-length single-molecule sequencing of the human mitochondrial genome. *Nat Commun* 13:5902. <https://doi.org/10.1038/s41467-022-33530-3>
- Ki JS, Hwang DS, Park TJ, Han SH, Lee JS (2010) A comparative analysis of the complete mitochondrial genome of the Eurasian otter *Lutra lutra* (Carnivora; Mustelidae). *Mol Biol Rep* 37:1943–1955. <https://doi.org/10.1007/s11033-009-9641-0>
- Kimura M (1980) A simple method for estimating evolutionary rate of base substitutions through comparative studies of nucleotide sequences. *J Mol Evol* 16:111–120. <https://doi.org/10.1007/BF01731581>
- Kirby RR (2000) An ancient trans-specific polymorphism shows extreme divergence in a multitrait cline in an intertidal snail (*Nucella lapillus* (L.)). *Mol Biol Evol* 17:1816–1825. <https://doi.org/10.1093/oxfordjournals.molbev.a026282>
- Kirby RR, Bayne BL, Berry RJ (1994) Phenotypic variation along a cline in allozyme and karyotype frequencies, and its relationship with habitat, in the dog-whelk *Nucella lapillus*, L. *Biol J Linn Soc* 53:255–275. <https://doi.org/10.1006/bjil.1994.1071>
- Kolmogorov M, Yuan J, Lin Y, Pevzner PA (2019) Assembly of long, error-prone reads using repeat graphs. *Nat Biotechnol* 37(5):540. <https://doi.org/10.1038/s41587-019-0072-8>
- Kumar S, Stecher G, Li M, Knyaz C, Tamura K (2018) MEGA X: Molecular evolutionary genetics analysis across computing platforms. *Mol Biol Evol* 35(6):1547–1549. <https://doi.org/10.1093/molbev/msy096>

- Li H (2018) Minimap2: pairwise alignment for nucleotide sequences. *Bioinformatics* 34(18):3094–3100. <https://doi.org/10.1093/bioinformatics/bty191>
- Li H, Durbin R (2010) Fast and accurate long-read alignment with Burrows-Wheeler transform. *Bioinformatics* 26(5):589–595. <https://doi.org/10.1093/bioinformatics/btp698>
- Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R, 1000 Genome Project Data Processing Subgroup (2009) The sequence alignment/map format and SAM-tools. *Bioinformatics* 25(16):2078–2079. <https://doi.org/10.1093/bioinformatics/btp352>
- Mariner-Llicer C, Goig GA, Zaragoza-Infante L, Torres-Puente M, Villamayor L, Navarro D, Borrás R, Chiner-Oms Á, Comas I (2021) Accuracy of an amplicon-sequencing nanopore approach to identify variants in tuberculosis drug-resistance-associated genes. *Microb Genom* 7(12):000740. <https://doi.org/10.1099/mgen.0.000740>
- Marko PB, Moran AL, Kolotuchina NK, Zaslavskaya N (2014) Phylogenetics of the gastropod genus *Nucella* (Neogastropoda: Muricidae): species identities, timing of diversification and correlated patterns of life-history evolution. *J Mollusc Stud* 80:341–353. <https://doi.org/10.1093/mollusc/eyu024>
- Oh DJ, Kim JY, Lee JA, Yoon WJ, Park SY, Jung YH (2007) Complete mitochondrial genome of the rabbitfish *Siganus fuscescens* (Perciformes, Siganidae). *DNA Seq* 18(4):295–301. <https://doi.org/10.1080/10425170701248525>
- Panova M, Blakeslee AMH, Miller AW, Mäkinen T, Ruiz GM, Johannesson K, André C (2011) Glacial history of the North Atlantic marine snail, *Littorina saxatilis*, inferred from distribution of mitochondrial DNA lineages. *PLoS ONE* 6(3):e17511. <https://doi.org/10.1371/journal.pone.0017511>
- Paradis E (2010) pegas: an R package for population genetics with an integrated-modular approach. *Bioinformatics* 26(3):419–420. <https://doi.org/10.1093/bioinformatics/btp696>
- Pascoe PL (2006) Chromosomal polymorphism in the Atlantic dogwhelk, *Nucella lapillus* (Gastropoda: Muricidae): nomenclature, variation and biogeography. *Biol J Linn Soc* 87:195–210. <https://doi.org/10.1111/j.1095-8312.2006.00567.x>
- Pater AA, Bosmeny MS, White AA, Sylvain RJ, Eddington SB, Parasrampur M, Ovington KN, Metz PE, Yinusa AO, Barkau CL, Chilamkurthy R, Benzinger SW, Hebert MM, Gagnon KT (2021) High throughput nanopore sequencing of SARS-CoV-2 viral genomes from patient samples. *J Biol Methods* 8(COVID 19 Spec Iss):e155. <https://doi.org/10.14440/jbm.2021.360>
- Quick J, Grubaugh ND, Pullan ST, Claro IM, Smith AD, Gangavarapu K, Oliveira G, Robles-Sikisaka R, Rogers TF, Beutler NA, Burton D, Lewis-Ximenez LL, Goes de Jesus J, Giovanetti M, Hill SC, Black A, Bedford T, Carroll MW, Nunes M, Alcantara LC Jr, Sabino EC, Baylis SA, Faria NR, Loose M, Simpson JT, Pybus OG, Andersen KG, Loman NJ (2017) Multiplex PCR method for MinION and Illumina sequencing of Zika and other virus genomes directly from clinical samples. *Nat Protoc* 12(6):1261–1276. <https://doi.org/10.1038/nprot.2017.066>
- Rombel IT, Sykes KF, Rayner S, Johnston SA (2002) ORF-FINDER: a vector for high-throughput gene identification. *Gene* 282(1–2):33–41. [https://doi.org/10.1016/S0378-1119\(01\)00819-8](https://doi.org/10.1016/S0378-1119(01)00819-8)
- Sang TK, Chang HY, Chen CT, Hui CF (1994) Population structure of the Japanese eel, *Anguilla japonica*. *Mol Biol Evol* 11(2):250–260. <https://doi.org/10.1093/oxfordjournals.molbev.a040107>
- Sedlazeck FJ, Rescheneder P, Smolka M, Fang H, Nattestad M, von Haeseler A, Schatz MC (2018) Accurate detection of complex structural variations using single-molecule sequencing. *Nat Methods* 15(6):461–468. <https://doi.org/10.1038/s41592-018-0001-7>
- Serb JM, Lydeard C (2003) 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. <https://doi.org/10.1093/molbev/msg218>
- Sotelo G, Duvetorp M, Costa D, Panova M, Johannesson K, Faria R (2020) Phylogeographic history of flat periwinkles, *Littorina fabalis* and *L. obtusata*. *BMC Evol Biol* 20:23. <https://doi.org/10.1186/s12862-019-1561-6>
- Staiger H (1957) Genetical and morphological variation in *Purpura lapillus* with respect to local and regional differentiation of population groups. *Année Biol* 61:251–258
- Sung JM, Karagozlu MZ, Lee J, Kwak W, Kim CB (2016) The complete mitochondrial genome of *Menathais tuberosa* (Gastropoda, Neogastropoda, Muricidae) collected from Chuuk Lagoon. *Mitochondrial DNA B* 1(1):468–469. <https://doi.org/10.1080/23802359.2016.1186516>
- Temperley RJ, Wydro M, Lightowlers RN, Chrzanowska-Lightowlers ZM (2010) Human mitochondrial mRNAs-like members of all families, similar but different. *Biochim Biophys Acta Bioenerg* 1797(6–7):1081–1085. <https://doi.org/10.1016/j.bbabi.2010.02.036>
- Tian Y, Hao Z, Zhu J, Yang J, Shi L, Li Z, Chang Y (2019) The complete mitochondrial genome of *Boreotrophon candelabrum* (Reeve, 1848). *Mitochondrial DNA Pt B* 4(1):1142–1143. <https://doi.org/10.1080/23802359.2019.1591188>
- Tomita K, Ueda T, Watanabe K (1998) 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. [https://doi.org/10.1016/S0167-4781\(98\)00099-2](https://doi.org/10.1016/S0167-4781(98)00099-2)
- Turon X, López-Legentil S (2004) Ascidian molecular phylogeny inferred from mtDNA data with emphasis on the Aplousobranchiata. *Mol Phylogenet Evol* 33(2):309–320. <https://doi.org/10.1016/j.ympev.2004.06.011>
- Varney RM, Brenzinger B, Malaquias MAE, Meyer CP, Schrödl M, Kocot KM (2021) Assessment of mitochondrial genomes for heterobranch gastropod phylogenetics. *BMC Ecol Evol* 21:6. <https://doi.org/10.1186/s12862-020-01728-y>
- Vazquez KE (2015) Phenotypic variation in the dogwhelk *Nucella lapillus*: an integration of ecology, karyotype, and phenotypic plasticity. PhD Thesis, University of Pennsylvania
- Walker BJ, Abeel T, Shea T, Priest M, Abouelliel A, Sakthikumar S, Cuomo CA, Zeng Q, Wortman J, Young SK, Earl AM (2014) Pilon: an integrated tool for comprehensive microbial variant detection and genome assembly improvement. *PLoS ONE* 9(11):e112963. <https://doi.org/10.1371/journal.pone.0112963>
- Walsh DJ, Bernard DJ, Pangilinan F, Esposito M, Harold D, Parle-McDermott A, Brody LC (2022) Mito-SiPE is a sequence-independent and PCR-free mtDNA enrichment method for accurate ultra-deep mitochondrial sequencing. *Commun Biol* 5:1269. <https://doi.org/10.1038/s42003-022-04182-2>
- Xin ZZ, Liu Y, Zhu XY, Wang Y, Zhang HB, Zhang DZ, Zhou CL, Tang BP, Liu QN (2017) Mitochondrial genomes of two Bombycoidea insects and implications for their phylogeny. *Sci Rep* 7:6544. <https://doi.org/10.1038/s41598-017-06930-5>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.