frontiers | Frontiers in Marine Science

TYPE Original Research
PUBLISHED 20 June 2025
DOI 10.3389/fmars.2025.1600021

OPEN ACCESS

EDITED BY
Kerstin Johannesson,
University of Gothenburg, Sweden

REVIEWED BY
David Osca,
University of Las Palmas de Gran Canaria, Spain
Pierre De Wit,
University of Gothenburg, Sweden

*CORRESPONDENCE
Enrico Negrisolo
enrico.negrisolo@unipd.it

†Deceased

RECEIVED 25 March 2025
ACCEPTED 19 May 2025
PUBLISHED 20 June 2025

CITATION
Corrain D, Franch R, Babbucci M, Tagliapietra D, Guarneri I, Sigovini M, Bonfatti V, Patarnello T and Negrisolo E (2025) Three new sequences of *Ostrea stentina* and the evolution of the mitogenome of the Ostreinae clams (Ostreidae, Bivalvia).
Front. Mar. Sci. 12:1600021.
doi: 10.3389/fmars.2025.1600021

Three new sequences of *Ostrea stentina* and the evolution of the mitogenome of the Ostreinae clams (Ostreidae, Bivalvia)

Daniele Corrain[1,2], Rafaella Franch[1,2], Massimiliano Babbucci[1,2], Davide Tagliapietra[3†], Irene Guarneri[3], Marco Sigovini[3], Valentina Bonfatti[1], Tomaso Patarnello[1,2] and Enrico Negrisolo[1,2,4*]

[1]Department of Comparative Biomedicine and Food Science (BCA), University of Padova, Legnaro, Italy, [2]National Biodiversity Future Center (NBFC), Palermo, Italy, [3]Institute of Marine Sciences (ISMAR), National Research Council (CNR), Venice, Italy, [4]Department of Agronomy, Food, Natural Resources, Animals and Environment (DAFNAE), University of Padova, Legnaro, Italy

Oysters are a group of bivalves forming the family Ostreidae. The identification of oysters at species level is sometimes difficult. The use of molecular data has drastically improved the reliability of species identification and our understanding of their phylogenetic relationships. Markers obtained from mitochondrial genome have played and continue to play a key role in this process. Complete mitogenomes are still unavailable for many oyster species. We sequenced three complete mitogenomes of the dwarf oyster *Ostrea stentina*. We performed a comparative and evolutionary mitogenomic study of the new sequences combined with all available ones for the Ostreinae. The mitogenome of *O. stentina* exhibited the standard gene order of Ostreinae, which is different from those observed in other subfamilies of Ostreidae. The study of these mitogenomic arrangements identified gene blocks that were present in the mitogenome of the last common ancestor of the Ostreidae. The comparative analysis allowed identifying peculiar features of the mitogenomes of Ostreinae as well as of their protein coding genes, tRNAs genes, rRNA genes, and control regions. The genus *Ostrea* resulted polyphyletic in the mito-phylogenomic analysis. The stems and loops of several tRNAs contained short DNA motifs useful to identify single species/groups of species. Short sequences, playing the role of molecular signatures characterizing a single taxon or a group of species, were identified also in the intergenic spacers. The identification of these taxonomic and phylogenetic markers reinforces the crucial role of mitogenomes in elucidating the evolutionary history of oysters.

KEYWORDS

***Ostrea stentina*, Ostreinae, mitogenome, phylogenetics, mitochondrial genomics, molecular signatures**

1 Introduction

Oysters are a group of bivalve molluscs forming the family Ostreidae. This family is part of the order Ostreida, which is included in the subclass Autobranchia (Bivalvia, Mollusca) (WoRMS Editorial Board, 2025). The family Ostreidae is split into four subfamilies: Crassostreinae, Ostreinae, Saccostreinae, and Striostreinae (Salvi et al., 2014; Salvi and Mariottini, 2017; Li et al., 2021; Salvi and Mariottini, 2021; Spencer et al., 2022).

The identification of oysters at species level is sometimes a difficult task (Harry, 1985). Morphological traits used to define the species boundaries are primarily features of the shell (Harry, 1985). Oysters form thick reefs made of single/multiple species, where individuals of the same taxon may exhibit different morphologies, or conversely, specimens belonging to different species may have the same appearance (Lunetta et al., 2023). This intra/interspecific variability is the result of the plasticity of shell morphology, a feature that makes species identification challenging and contributes to taxonomic inflation (Harry, 1985). Habitat and environment factors affect shell shape (Lam and Morton, 2006). In addition, the high dispersal ability of individuals during the larval stages complicates species identification based on geographic collection site, as location does not necessarily reflect a distinct, species-specific distribution (Lapègue et al., 2002). This is even more true considering that human activities have altered the distribution of several species outside their original home range (e.g. Troost, 2010).

The availability of molecular data has drastically improved not only the reliability of species identification but also the understanding of the relationships among oyster species and their classification at higher taxonomic ranks (Salvi et al., 2014; Salvi and Mariottini, 2017; Li et al., 2021; Salvi and Mariottini, 2021; Spencer et al., 2022; Lunetta et al., 2023). Markers obtained from mitochondrial genome (hereafter, mitogenome) have played a key role in this molecular phylogenetic and taxonomic revolution.

The mitogenome of Mollusca is a double-helix circular molecule with a highly diverse size (Plazzi et al., 2016; Ghiselli et al., 2021). This is particularly evident among the members of the class Bivalvia, where it can reach a size exceeding 56 kbp in the ark clam *Anadara kagoshimensis* (Kong et al., 2020). In the family Ostreidae the size of the mitogenome varies from 16 to 20 kbp (e.g. Xiao et al., 2015; Li et al., 2021). The mollusc mitogenome contains 37 genes: 13 protein-coding genes (PCGs), 22 tRNA (one for each amino acid and 2 for Serine and Leucine that are duplicated), and two rRNA subunits. Initially, *atp8* was deemed to be absent in the bivalve mitogenome because its high level of divergence prevented detection (e.g. Milbury and Gaffney, 2005), but it was identified later (Breton et al., 2010). In Ostreidae the large subunit of ribosomal RNA is split into two halves (*rrnL 3'end* and *rrnL 5'end*), with the ribosome remaining functional (Milbury et al., 2010). In all oyster mitogenomes sequenced to date, a second copy of *trnM* exists (Milbury and Gaffney, 2005; Ren et al., 2009, 2010; Wu et al., 2010; Danic-Tchaleu et al., 2011; Yu and Li, 2011; Wu et al., 2012; Xiao et al., 2015; Cavaleiro et al., 2016; Ren et al., 2016). Furthermore, extra-copies of other tRNAs are present in the mitogenome of species of *Magallana* and *Talonostostrea* (e.g. Ren et al., 2010). In *Magallana* there is also a second copy of the small ribosomal subunit (e.g. Ren et al., 2010). All the genes are located on the same strand (Milbury and Gaffney, 2005; Ren et al., 2009, 2010; Wu et al., 2010; Danic-Tchaleu et al., 2011; Yu and Li, 2011; Wu et al., 2012; Xiao et al., 2015; Cavaleiro et al., 2016; Ren et al., 2016). They can be adjacent, overlapped, or separated by a variable number of nucleotides that form intergenic spacers (ISPs). The ISPs can be generated through a process of slippage during the replication of mitogenome, or formed during mitogenomic re-arrangements (Basso et al., 2017). In this latter case, they can retain remnants of the genes involved in the rearrangement process and provide valuable insights into how the event occurred (Basso et al., 2017). In oysters, the putative Control Region (CoRe), a non-coding sequence involved in the regulation of replication and transcription, is usually the longest intergenic spacer. The CoRe is variable for position and for base composition. Usually it contains AT-rich motifs, and stem-and-loop and cloverleaf secondary structures (Ghiselli et al., 2021). In oysters the mitochondrial inheritance follows the standard animal pathway and is strictly maternal, unlike in other bivalves (e.g. mussels) which exhibit doubly uniparental inheritance (Ghiselli et al., 2021).

The gene order (GO) is not conserved among the mitogenomes of oysters sequenced to date (Milbury and Gaffney, 2005; Ren et al., 2009, 2010; Wu et al., 2010; Danic-Tchaleu et al., 2011; Yu and Li, 2011; Wu et al., 2012; Xiao et al., 2015; Cavaleiro et al., 2016; Ren et al., 2016). The Ostreinae and Saccostreinae subfamilies exhibit two distinct gene orders (OstGO vs SacGO) while multiple GOs occur within the subfamily Crassostreinae, each characterizing different genera (CraGO, *Crassostrea*; MagGO, *Magallana;* TalGO, *Talonostrea*). Further GOs exist and are limited to single species of *Magallana* (data not shown). In the oysters, the different GOs are the result of the transposition of one or more genes, coupled in some cases with duplications/multiplications of additional genes. A tandem duplication random loss mechanism/event can partly explain these complicated rearrangements (Basso et al., 2017). The sequencing and analysis of the mitogenomes of Ostreidae allow not only to understand the phylogenetic relationships within this family but also to perform comparative and evolutionary genomic studies. However, our knowledge is still very fragmented and restricted to a limited number of species.

To expand the Ostreidae mitogenome data set we sequenced three complete mitogenomes for *Ostrea stentina* Payraudeau, 1826 (Ostreidae, Ostreinae). Payraudeau (1826) identified *O. stentina*, known as dwarf oyster or Provence oyster, from specimens collected in Corsica coasts. *O. stentina* is considered a complex of species (Hu et al., 2019). However, the taxonomic status and the phylogenetic relationships of the forms contained in this complex are not yet fully resolved. *O. stentina* has a broad distribution, as it has been collected in the Mediterranean Sea, in southern Argentina, western and eastern Atlantic coasts, Gulf of California and Asian

Pacific Ocean (Hu et al., 2019). In the past, *O. stentina* was confused with the juvenile form of other *Ostrea* species such as *Ostrea edulis* (Hamaguchi et al., 2017; Lunetta et al., 2023). In this work, we compared the new mitochondrial sequences of *O. stentina* with available mitogenomes of other oysters of the subfamily Ostreinae. The results of our comparative and evolutionary mitogenomic study are presented in the next sections.

Specifically, we focused on: (a) determining at least partially the gene order of the last common ancestor of Ostreidae; (b) exploring the key molecular features of the different type of markers encoded in the mitogenomes of the oysters; (c) reconstructing the phylogeny of Ostreinae; (d) identifying new markers capable of unambiguously distinguishing single species/group of species.

2 Materials and methods

2.1 *Ostrea stentina* sampling

The specimens of *O. stentina* used for this study were sampled in the Venice Lagoon (Italy), one of the widest coastal transitional ecosystems in the Mediterranean, and a complex mosaic of habitats. In particular, they were collected in the intertidal zone at the following locations: *O. stentina* NEWOS06 (Torson di sotto island: 45°20'55.2" N, 12°13'46.4" E); *O. stentina* NEWOS18 (Darsena dell'Arsenale: 45°26'14.5" N, 12°21'18.1" E); *O. stentina* NEWOS81 (Faro Rocchetta: 45°20'20.7" N, 12°18'39.6" E). Muscle tissue was stored in 100% ethanol and kept at -20°C until DNA extraction.

2.2 Mitogenomes sequencing and assembly

Genomic DNA was extracted using the commercial kit Invisorb Spin TissueMini Kit (Invitek, STRATEC Biomedical, 242 Germany), quantified using Qubit dsDNA BR Assay Kit (Invitrogen–ThermoFisher Scientific), and checked for quality on agarose gel electrophoresis.

Genomic libraries were constructed using the commercial kit Illumina DNA Prep (Illumina, Inc.), quantified using Qubit dsDNA HS (High Sensitivity) Assay Kit (Invitrogen–ThermoFisher Scientific), and checked for quality on an Agilent 2100 Bioanalyzer (Agilent Technologies, Santa Clara, California, USA), before sequencing. Libraries were equally pooled and sequenced on an Illumina HiSeq4000 platform with 150 bp pair-end read module at the UCDavis DNA Technologies & Expression Analysis Core (Davis, CA) in order to obtain 14 M of raw read-pairs/library.

Raw paired-reads obtained from Illumina sequencing were assessed for quality using FastQC v0.12 (Andrews, 2010), and consequently trimmed of any adaptors and low quality sequences using Trimmomatic v0.32 (Bolger et al., 2014); high quality reads of length ranging from 70 bp to 150 bp were retained. The whole mitogenome was assembled using the software Get Organelle version 1.7.7.1 with kmer values of 21, 45, 65, 85, 105 and "animal_mt" as seed database (Jin et al., 2020).

2.3 Annotation of mitogenomes and data set construction

The new mitogenomes of *O. stentina* were annotated following the strategy described in previous works from our laboratory (Babbucci et al., 2014; Montelli et al., 2016; Basso et al., 2017). Gene nomenclature followed standard naming for animal mitogenomes (Montelli et al., 2016; Basso et al., 2017; Ghiselli et al., 2021). A comparison was done with an ESTs library of *Magallana gigas* (*Crassostrea gigas* digestive gland subtracted library, multiple accession numbers, unpublished) available in GenBank (Clark et al., 2015) to ensure a correct identification of the 5' start of the protein coding genes that are particularly difficult to be identified in molluscs (Ghiselli et al., 2021). For the identification of the two halves of *rrnL* we followed Milbury et al. (2010). Finally, comparisons with other published and unpublished oyster mitogenomes were done to refine the annotations (Milbury and Gaffney, 2005; Ren et al., 2009, 2010; Wu et al., 2010; Danic-Tchaleu et al., 2011; Yu and Li, 2011; Wu et al., 2012; Xiao et al., 2015; Cavaleiro et al., 2016; Ren et al., 2016). The mitogenomes used in the analyses, and listed in Table 1 were re-annotated following the approach described above to ensure a consistent annotation across all the sequences. Different authors have assigned various names to the mitogenome strands (Basso et al., 2017). In the present paper, we refer to the strand encoding all the genes as the plus strand and the opposite strand as the minus strand. Because our study was focused on the subfamily Ostreinae we analyzed all available mitogenomes for this taxon, while for the subfamilies Crassostreinae and Saccostreinae, we restricted our analyses to selected species (Table 1). However, two mitogenomes of Ostreinae (*Ostrea denselamellosa* ON964460 and *O. edulis* CM063324) became available in GenBank (Clark et al., 2015) after the main phase of our analyses had been completed. Consequently, they were not fully integrated into the study but were included in selected downstream analyses (see below).

2.4 Inference of the ancestral gene order of the Ostreidae

Conserved gene blocks shared across two or more gene orders were identified through visual inspection of the gene arrangements of CraGO, MagGO, OstGO, SacGO, and TalGO. The analysis focused on gene blocks shared among taxa from different subfamilies. The ancestral GO for the Ostreidae was inferred manually, according to a principle of parsimony, by mapping GOs evolution along the reference phylogeny of the family (Salvi and Mariottini, 2017; Li et al., 2021; Salvi and Mariottini, 2021).

2.5 Multiple alignments of orthologous genes

Multiple alignments of the protein-coding genes (PCGs) were done in two steps. Initially, for each PCG an alignment of the amino

TABLE 1 List of the taxa and mitogenomes analyzed in the present paper.

Family	Subfamily	Species	GenBank	Size	Reference
Ostreidae	Crassostreinae	*Crassostrea tulipa* (Lamarck, 1819)	KR856227	17,685 bp	Cavaleiro et al., 2016
		Crassostrea virginica (Gmelin, 1791)	AY905542.2	17,244 bp	Milbury and Gaffney, 2005 (*)
		Magallana angulata (Lamarck, 1819)	KJ855248.1	18,225 bp	Ren et al., 2016 (*)
		Magallana ariakensis (Fujita, 1929)	KJ855251.1	18,420 bp	Ren et al., 2016 (*)
		Magallana hongkongensis (Lam and Morton, 2003)	FJ841963.1	18,620 bp	Wu et al., 2010 (*)
		Magallana gigas (Thunberg, 1793)	AF177226	18,224 bp	Unpublished
		Magallana sikamea (Amemiya, 1928)	KJ855258.1	18,243 bp	Ren et al., 2016 (*)
		Talonostrea talonata Li and Qi, 1994	KT353107	20,552 bp	Unpublished
	Saccostreinae	*Saccostrea cucullata* (Born Von, 1778)	KP967577.1	16,396 bp	Volatiana et al., 2016
		Saccostrea echinata (Quoy and Gaimard, 1835)	KU310913.1	16,281 bp	Unpublished (*)
		Saccostrea kegaki (Torigoe and Inaba, 1981)	KT936587.1	16,280 bp	Unpublished (*)
		Saccostrea malabonensis (Faustino, 1932)	ON649706.1	16,204 bp	Unpublished (*)
		Saccostrea glomerata (Gould, 1850)	KU310916	16,282 bp	Unpublished
		Saccostrea scyphophilla (Péron & Freycinet, 1807)	FJ841968	16,532 bp	Unpublished
	Ostreinae	*Ostrea stentina* Payraudeau, 1826. Isolate NEWOS06	**PV339533**	16,313 bp	**This paper**
		Ostrea stentina Payraudeau, 1826. Isolate NEWOS18	**PV345786**	16,304 bp	**This paper**
		Ostrea stentina Payraudeau, 1826. Isolate NEWOS81	**PV345787**	16,308 bp	**This paper**
		Ostrea edulis Linnaeus, 1758	JF274008	16,320 bp	Danic-Tchaleu et al., 2011
		Ostrea edulis Linnaeus, 1758	MT663266	16,356 bp	Hayer et al., 2021
		Ostrea edulis Linnaeus, 1758	OX387714	16,350 bp	Unpublished
		Ostrea edulis Linnaeus, 1758	CM063324	16,349 bp	Li et al., 2023 (**)
		Ostrea lurida Carpenter, 1864	KC768038	16,344 bp	Xiao et al., 2015
		Ostrea denselamellosa Lischke, 1869	HM015199	16,277 bp	Yu and Li, 2011
		Ostrea denselamellosa Lischke, 1869	ON964460	16,275 bp	Unpublished (**)
		Nanostrea pinnicola (Pagenstecher, 1877)	MT822277	16,315 bp	Unpublished
		Planostrea pestigris (Hanley, 1846)	MT822278	16,236 bp	Unpublished
		Dendostrea sandvichensis (Sowerby, 1871)	MT635133	16,338 bp	Unpublished

Size, genome size. (*), mitogenome considered only in gene order analysis; (**), mitogenome considered only in some analyses.
Accession numbers of newly sequenced mitogenomes are in bold.

acid sequences was performed with the online version of the MAFFT program (https://mafft.cbrc.jp/alignment/software/) (Katoh et al., 2002). Successively, through the TranslatorX server (http://161.111.160.230/index_v5.html), the codons of each orthologous set of PCGs were aligned using as template the corresponding amino acid multiple alignment (Abascal et al., 2010). Multiple alignments of the orthologous tRNAs were produced in two steps. Firstly, orthologous sequences were quickly aligned with ClustalW (Thompson et al., 1994). Successively, these alignments were improved manually (Montelli et al., 2016), considering the secondary structures predicted for each tRNA with the tRNA-scan software (Chan and Lowe, 2019). To provide a figure of the conservation level of each tRNA multiple alignment, a logo vignette was created with the WebLogo software (Crooks et al., 2004). The multiple alignments of rRNAs were produced with the MAFFT program (https://mafft.cbrc.jp/

alignment/software/) (Katoh et al., 2002). The secondary structure of *rrnL* of *O. stentina* was manually modeled using the published structure of *Magallana gigas* (Milbury et al., 2010) as template.

2.6 Characterization of intergenic spacers

The occurrence of the intergenic spacers was studied in all the mitogenomes of Ostreinae and their distribution was mapped on the reference mitogenome (see below). Alignments were produced for each group of ISPs. The shortest ISPs were aligned manually, while the longest were aligned with the online version of the MAFFT software (Katoh et al., 2002). ISPs are very fast evolving sequences, therefore in most of the cases the alignments were restricted to a single species or to closely related species (see below). We searched for the occurrence of one or more ISPs representing molecular signatures for the analyzed oysters. To qualify as a molecular signature, an ISP must exhibit a unique sequence characterizing the mitogenomes of one species or multiple taxa forming a monophyletic group in a phylogenetic tree.

2.7 Identification and characterization of the control region

In the mitogenomes of *O. stentina* the Control Region (CoRe) was identified as the longest intergenic spacer containing AT-rich motifs, stem-and-loop, and cloverleaf secondary structures (Ghiselli et al., 2021). The capability to produce secondary structures was tested with the software RNAstructure (Reuter and Mathews, 2010). Multiple alignment of the CoRes of Ostreinae was done with the T-Coffee. The web server version, using the M-Coffee option, aligns DNA sequences by combining the output of popular aligners (Di Tommaso et al., 2011).

2.8 Statistical analyses on DNA

The total number of codons and the relative abundance of each codon family used by the 13 PCGs were computed with the MEGA X program for all Ostreinae (Kumar et al., 2018). The codon distribution was expressed as number of codons per thousand codons (CDSpT). The Relative Synonymous Codon Usage (RSCU) values were calculated with the MEGA X program (Kumar et al., 2018). First codons, as well as stop codons, complete and incomplete, were excluded from the analysis to avoid biases due to unusual putative start codons and incomplete stop codons.

The A+T/G+C content and GC-skew = (G-C)/(G+C) and AT-skew = (A-T)/(A+T) (Perna and Kocher, 1995) were used to measure the compositional biases among analyzed sequences. The base compositions were computed with MEGA X (Kumar et al., 2018). The calculations of skews were performed with Excel program (Microsoft TM). The skews were computed for the whole mitogenomes, and for PCGs, tRNAs, rRNAs and CoRes. Differences in AT-skew vs A+T content and GC-skew vs the G+C content were plotted as scatterplots in Microsoft Excel.

We also tested whether the AT- and GC-skews of various PCGs, tRNAs, rRNAs and CoRes were statistically significantly different from those computed for the whole mitogenomes. Levene's test (Levene, 1960) revealed that the variances of AT- and GC-skews differed significantly across gene/regulatory regions (PCG, tRNA, rRNA, or CoRes), violating the assumption of homogeneity of variance required for linear models (AT-skew: $F_{(23,\ 240)} = 2.550$, $p < 0.001$; GC-skew: $F_{(23,\ 240)} = 3.169$, $p < 0.001$). For this reason, differences in AT-skew and GC-skew between the whole mitogenomes and the genes/regulatory regions were tested by bootstrapping. For each combination of mitogenome and gene/regulatory region, 50,000 bootstrapped data sets were generated. For each of these data sets, we calculated and stored the difference in AT-skew and GC-skew between the mitogenome and gene/regulatory region. To calculate the p-value of the difference in AT-skew and GC-skew between the mitogenome and genes/regulatory regions, we generated a distribution of the difference under the null hypothesis as follows: 1) for each combination of mitogenome and gene/regulatory region, we first calculated the difference in AT-skew and GC-skew in the original data set; 2) this value was added to the AT-skew or GC-skew of the gene/regulatory region, forcing the difference in AT-skew or GC-skew between the mitogenome and gene/regulatory region to be equal to zero; 3) from this new set of data, we created 50,000 bootstrapped data sets and, for each of these, calculated the difference in AT-skew or GC-skew between the mitogenome and gene/regulatory region, allowing us to obtain the distribution of the difference in AT-skew or GC-skew under the null-hypothesis. The p-value of the difference in AT-skew and GC-skew between the mitogenome and gene/regulatory region was calculated as the probability of obtaining a result equal to or more extreme than what was observed in the first bootstrapping, assuming the null hypothesis (no difference) was true.

2.9 Identification of hemi- and fully compensatory base changes in tRNAs

The nucleotide substitution pattern was tracked in the stems of the secondary structures of orthologous tRNAs (Montelli et al., 2016). We looked for the occurrence of (a) hemi-conservatory base changes, (b) type I fully compensatory bases changes, (c) type II fully compensatory base changes, and (d) mismatches (Montelli et al., 2016). These patterns were identified by visual inspection of multiple sequence alignments and taking into account the predicted secondary structure of tRNAs (Montelli et al., 2016).

Given the same pair of bases in a stem, a change with respect to the background condition for the multiple alignment can involve only one of the two bases, either at the 5′ or 3′ end, without altering the secondary structure of the stem (e.g., T•G vs. T–A) (Coleman, 2003; Montelli et al., 2016). This variation is referred to as a hemi-compensatory base change (Coleman, 2003; Montelli et al., 2016). The change can also involve both bases of the pair but the secondary structure remains intact (e.g., G–C vs. A–T). This type of change is

known as a fully compensatory base change because the substitution of both bases does not compromise the integrity of the secondary structure (Coleman, 2003; Montelli et al., 2016). There are two types of fully compensatory base changes: type I, which involves the substitution of a purine-pyrimidine pair with another purine-pyrimidine couple and vice versa (Montelli et al., 2016); and type II, which is characterized by a purine-pyrimidine vs. pyrimidine-purine substitution (Montelli et al., 2016). Type I occurs more easily than type II because its intermediate step is represented by a hemi-compensatory base change (Montelli et al., 2016). In contrast, type II is disfavored because its intermediate step involves a mismatch in the pair that jeopardizes the secondary structure of the stem (Montelli et al., 2016). Lastly, the change can involve a substitution pattern leading to a disruption of the secondary structure of the stem for the analyzed pair. Mismatches that do not prevent the formation of the cloverleaf structure or the tertiary structure are not uncommon in tRNAs. Various mechanisms of editing can correct mismatches in the stems, or alternatively, these mismatches can persist as unusual pairings (Cannone et al., 2002).

2.10 Homogeneity vs heterogeneity of the substitution process in the alignments

The level of homogeneity/heterogeneity in the substitution process in PCGs and their corresponding protein products, as well as in the two rRNAs multiple alignments, was tested with the AliGROOVE software (Kück et al., 2014). For the PCGs the AliGROOVE matrices were computed for: complete codons, the first plus the second position of each codon, single positions (first, second and third), and the translated amino acid sequences. AliGROOVE tests were performed also on the multigene concatenated data sets used in the final phylogenetic analyses (see below). In a matrix, obtained from AliGROOVE, a square ranging from brown to pink identifies a heterogeneous substitution process between the two compared sequences, while a square ranging from light to dark blue marks a homogenous substitution process (Kück et al., 2014).

2.11 Detection of the phylogenetic signal

The phylogenetic signal present in the different genes/multiple alignments was evaluated through two different methods: (a) the quartet puzzling analysis (Strimmer and von Haeseler, 1996) and (b) the boxplot graphics, which analyses the distribution of the pairwise distances computed according to the best-fit evolutionary model (e.g. Negrisolo et al., 2004). The quartet puzzling analysis was performed as implemented in IQ-TREE2 (Minh et al., 2020). The best fitting evolutionary models for DNA/proteins were identified with the ModelFinder program (Kalyaanamoorthy et al., 2017) implemented in IQ-TREE2 (Minh et al., 2020). IQ-TREE2 software was used also to compute the distances based on the best-fit models. The boxplots were created with the Excel software. The occurrence of phylogenetic signal was studied on the *rrnS* and *rrnL* and on each PCG. In this latter case, the analysis was done on the single positions of the codons (p1, first; p2, second; p3, third), on positions one and two (p12), overall codons (p123) and on the translated polypeptides. Finally, this analysis was extended to the concatenated data sets (see below).

2.12 Phylogenetic analyses on multiple markers data sets

For phylogenetic analyses, we created 10 concatenated data sets that are listed in Table 2. The concatenations were done with the MEGA X software (Kumar et al., 2018).

The phylogenetic analyses were performed according to the maximum likelihood (ML) method (Felsenstein, 2004). The ML trees were computed with the program IQ-TREE2 (Minh et al., 2020). In the tree search analysis, 50 independent runs were performed for the rRNA data set, 20 runs for 13PCGpro, 13PCGpro+rRNA, 13PCGnuc, 13PCGnuc+rRNA and 10 runs for 13PCGp1, 13PCGp2, 13PCGp3, 13PCGp12 and 13PCGp12+rRNA. The optimal partitioning scheme (Chernomor et al., 2016) and the best fitting evolutionary models (Kalyaanamoorthy et al., 2017) were selected with IQ-TREE2. The data sets 13PCGnuc, 13PCGnuc +rRNA and 13PCGp12+rRNA were analyzed also with a partition scheme that considered the single positions of the codons separately.

The Ultrafast Bootstrap Test (UBT) (Minh et al., 2013) (10,000 replicates) and the approximate Likelihood-Ratio Test for branches (aLRT) (Anisimova and Gascuel, 2006) (1,000 replicates) were used for evaluating the robustness of the tree topologies obtained in the various searches. A Robinson-Foulds distance data matrix (Llabrés et al., 2021) was computed for each set of the trees generated in every phylogenetic analysis to ensure that the top-ranked topologies had a null distance and the convergence had been reached in the tree searching.

To evaluate alternative phylogenetic hypotheses, topology tests were done according to the Approximately Unbiased (AU) test (Shimodaira, 2002), the Weighted Shimodaira and Hasegawa (WSH) test (Shimodaira and Hasegawa, 1999) and the Expected

TABLE 2 List of the data sets used in phylogenetic analyses.

Data set	Content	Type: aa/DNA	Length
13PCGpro	13 PCGs	aa	3,816
13PCGpro+rRNAs	13 PCGs + (*rrnS*+*rrnL*)	aa+DNA	6,276
13PCGnuc	13 PCGs, complete codons	DNA	11,448
13PCGnuc+rRNAs	13 PCGs + (*rrnS*+*rrnL*)	DNA	13,908
13PCGp1	p1–13 PCGs	DNA	3,816
13PCGp2	p2–13 PCGs	DNA	3,816
13PCGp3	p3–13 PCGs	DNA	3,816
13PCGp12	p12–13 PCGs	DNA	7,632
13PCGp12+rRNAs	p12–13 PCGs + (*rrnS*+*rrnL*)	DNA	10,092
rRNAs	rRNA (*rrnS*+*rrnL*)	DNA	2,460

Aa, amino-acid; length, length of the alignment.

Likelihood Weights (ELW) (Strimmer and Rambaut, 2002) method. Computations were performed with IQ-TREE2 (Minh et al., 2020).

3 Results

3.1 Structure of the mitogenome of *Ostrea stentina* and comparison with other Ostreidae

Three complete mitogenomes of *O. stentina* were sequenced for this work (Figure 1; Supplementary Table S1). The Illumina reads used to assemble these genomes spanned from a minimum of 12,236,243 to a maximum of 17,149,758 (Supplementary Table S1). The size of the *O. stentina* mitogenomes varied from 16,305 to 16,314 bp (Supplementary Table S1). This range was very similar to values obtained for the mitogenomes of Ostreinae and Saccostreinae, while values were higher in Crassostreinae also for the occurrence of extra genes (*Magallana* and *Talonostrea*) (Figure 2A; Table 1). The mitogenome of *O. stentina* contained a set of 38 genes: 13 PCGs, 23 tRNAs and 2 ribosomal RNAs (Figure 1; Supplementary Table S1). The gene order corresponded to the typical arrangement observed in the Ostreinae subfamily (OstGO), with all genes encoded on the plus strand (Figure 1). Genes were contiguous or separated by intergenic spacers (ISP) (Figure 1; Supplementary Table S1).

3.2 The ancestral gene order of Ostreidae

The mitogenomes of Ostreinae sequenced so far exhibited the same GO (Figure 2; Table 1). However, they varied at the microstructural level in the distribution of the ISPs (see below). OstGO exhibited the maximum level of synteny with SacGO of Saccostreinae as proved by the sharing of two large conserved gene blocks (Figure 2A). OstGO shared gene blocks also with CraGO, MaGO and TalGO (Figure 2A). Furthermore, blocks containing two or more genes were shared among all analyzed GOs (Figure 2A). Thus, by considering the distribution of the conserved blocks among different GOs and the reference phylogeny for the Ostreidae, we identified gene blocks that were present in the gene order of the last common ancestor of the Ostreidae (lcaO, Figure 2B).

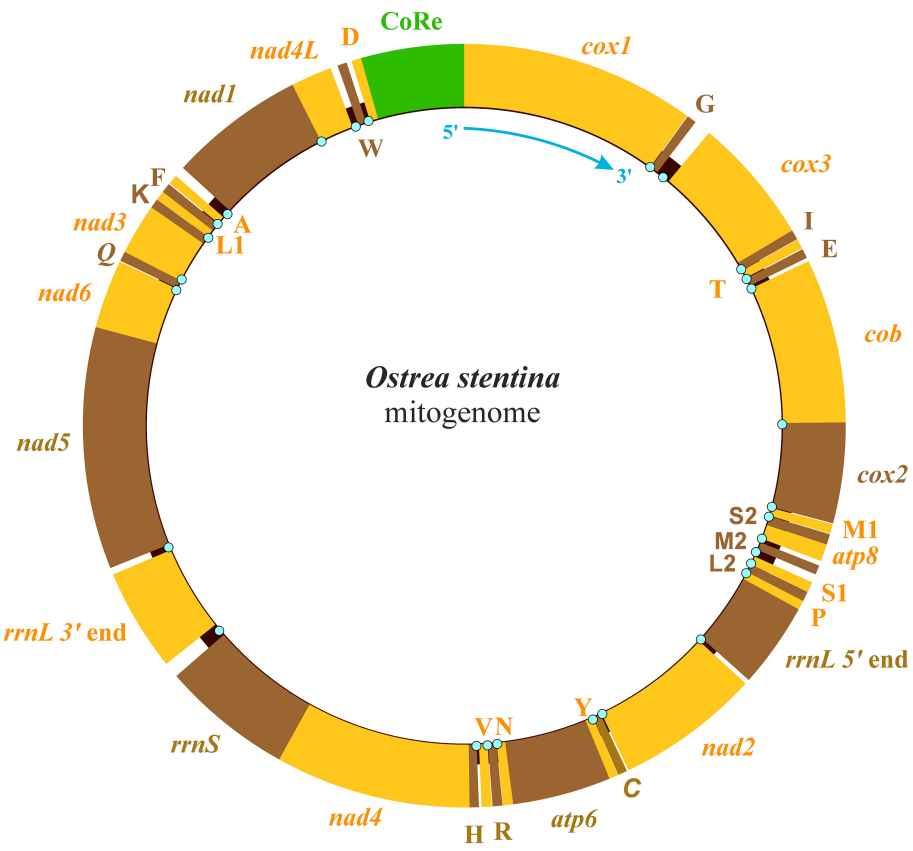

FIGURE 1
The structure of the mitogenome of *Ostrea stentina*. All genes are located on the plus strand. *atp6* and *atp8*, ATP synthase subunits 6 and 8. *cob*, cytochrome b. *cox1-3*, cytochrome c oxidase subunits 1–3. *nad1–6* and *nad4L*, NADH dehydrogenase subunits 1–6 and 4L. *rrnS* and *rrnL*, small and large subunit ribosomal RNA (rRNA) genes. X, transfer RNA (tRNA) genes, where X is the one-letter abbreviation of the corresponding amino acid, in particular L1 (CTN codon family) L2 (TTR codon family), S1 (AGN codon family) S2 (TCN codon family). CoRe, Control Region. The presence of a cyan dot indicates an intergenic spacer (ISP).

10.3389/fmars.2025.1600021

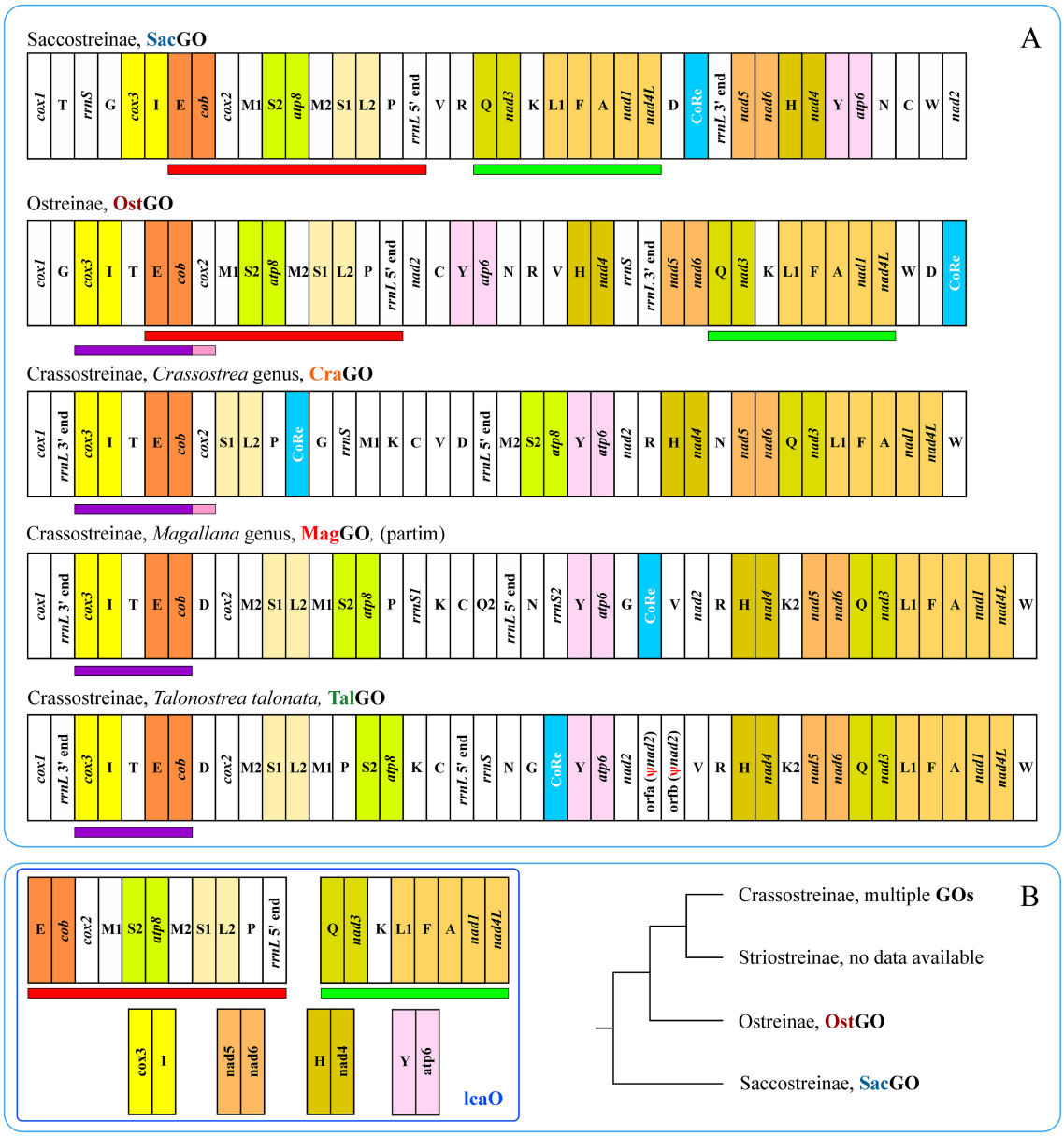

FIGURE 2
(A) Principal Gene Orders (GOs) occurring in Ostreidae. All genes are located on the plus strand. *atp6* and *atp8*, ATP synthase subunits 6 and 8. *cob*, cytochrome b. *cox1-3*, cytochrome c oxidase subunits 1–3. *nad1–6* and *nad4L*, NADH dehydrogenase subunits 1–6 and 4L. *rrnS* and *rrnL*, small and large subunit ribosomal RNA (rRNA) genes. X, transfer RNA (tRNA) genes, where X is the one-letter abbreviation of the corresponding amino acid, in particular L1 (CTN codon family) L2 (TTR codon family), S1 (AGN codon family) S2 (TCN codon family). CoRe, Control Region. Orf, Open reading frame. Ψ*nad2*, pseudogene *nad2*. Blocks of conserved genes are colored with different backgrounds. Red and green bars underline the two major conserved gene-blocks shared by SacGO and OstGO. The mitogenomes of all species listed in Table 1 were analyzed for identifying the different GOs. **(B)** Blocks of conserved genes inferred to occur in the GO (lcaGO) of the last common ancestor of Ostreidae.

3.3 Compositional biases and AT/GC-skews of the mitogenomes of the Ostreinae

The mitogenomes of *O. stentina* were A+T rich, negatively AT-skewed, and positively GC-skewed (Figure 3). This feature was shared by all oysters sequenced to date, limiting our comparison to Ostreinae (Table 1). Among Ostreinae, *Ostrea denselamellosa* presented the most diverging values for both A+T/G+C content (60.71%) and AT-skew (-0.153) (Figure 3). The range of variation of AT-skew was broader than that of GC-skew. Mitogenomes of Ostreinae exhibited a limited range of variation in A+T (60.71% - 65.41%)/G+C (34.58% - 39.29%) content and AT-/GC–skews (-0.153 – -0.128) (0.149 – 0.201) (Figure 3). However, the taxon coverage was very low.

3.3.1 Compositional biases and AT/GC-skews of PCGs

The comparisons among the compositional biases and AT/GC-skews of the PCGs vs those of the mitogenomes are presented in

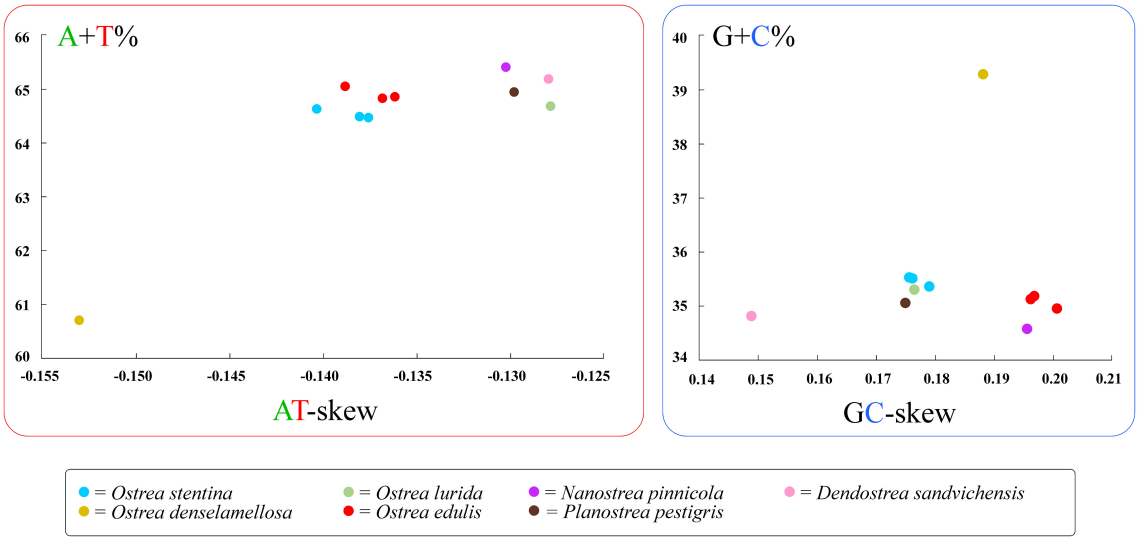

FIGURE 3
Genomic compositions and AT-/GC-skews in Ostreinae. AT-skew = (A-T)/(A+T). GC-skew = (G-C)/(G+C).

Supplementary Figures S1–S5. The A+T content of PCGs was slightly lower (*cox1-cox3, nad4L*), similar to (*atp6, cob, nad1, nad4, nad5, nad6*), or higher (*atp8*, *nad2*, *nad3*) than that of the whole mitogenomes. The G+C content showed the opposite pattern. AT- and GC-skew behaved as those of mitogenomes, but many PCGs exhibited significantly more negative skews ($p < 0.001$). *Cox2* sequences closely mirrored the pattern of mitogenomes, whereas *atp8* showed a contrasting trend, exhibiting the lowest AT-skews. The GC-skews were more variable in their pattern, but always positive. Some PCGs showed values significantly higher (e.g. *nad2*, $p < 0.001$) or lower (e.g. *cox1*, $p < 0.001$) than those of mitogenomes. Other PCGs did not differ from the mitogenomes (e.g. *cox2*). In general, the PGCs of *O. stentina* exhibited average values for both AT-/GC-skews and A+T G+C contents. The very low GC-skew of *atp8* and the very high GC-skew *nad4L* were notable exceptions.

3.3.2 Compositional biases and AT/GC-skews of tRNAs

The tRNAs showed high variability in the analyzed parameters (Supplementary Figures S6–S13), particularly in A+T and G+C contents with individual tRNAs exhibiting similar, higher, or lower enrichment compared to the whole mitogenomes (e.g. for A+T: *trnM2*, *trnY*, *trnM1*; for G+C: *trnN*, *trnF*, *trnT*). It was worth noting the extremely high A+T content of *trnG* in *Planostrea pestigris* and *Dendostrea sandvichensis*, which exceeded 80% and 77%, respectively. In contrast, the A+T content was exactly 50% in *trnF* of *O. denselamellosa* and in *trnM1*, of *Nanostrea pinnicola*, *O. lurida*, and *P. pestigris*. The lowest value (47.62%), the only one below 50%, occurred in t*rnM1* of *O. denselamellosa*. Many tRNAs had a G+C content higher than that of the whole mitogenomes, and this could possibly be associated with the increased stability that the G-C/C-G pairings provide in the stems of their secondary structure. 14 tRNAs showed one or more (up to all) sequences with positive AT-skew values, displaying a pattern opposite to that of the whole mitogenomes. The GC-skews patterns of tRNAs were still variable but not as much as those observed for AT-skews. Notably, six tRNAs had GC-skews values that were not significantly different from those of the full mitogenomes ($p > 0.05$). Furthermore, only in *trnG* (*D. sandvichensis*, *O. denselamellosa*, *P. pestigris*; Supplementary Figure S7), *trnT* (*D. sandvichensis* and *O. stentina*; Supplementary Figure S12) and *trnW* (*O. edulis* JF274008 and *O. edulis* OX387714; Supplementary Figure S13) multiple sequences presented negative values of GC-skews instead of the standard positive ones.

3.3.3 Compositional biases and AT-/GC-skews of rRNAs

In the mitogenome of oysters, the *rrnL* gene is split into two parts (Figure 2). The *rrnL* 5' ends were richer in A+T than the complete mitogenomes, while the opposite was true for the *rrnL* 3' ends (Supplementary Figure S14). When the two segments were merged into the complete *rrnL* gene, these discrepancies disappeared. In contrast, both the 5'/3'ends and the entire *rrnL* clearly differed from the complete mitogenomes in terms of their AT-skews. This was particularly evident for the 3' *rrnL* ends, which exhibited only positive AT-skew values. Most of the 5' *rrnL* ends showed negative AT-skews, while only the complete *rrnLs* of *O. stentina* presented slightly negative values (≥0.008). The *rrnSs* had a clearly lower A+T content and positive AT-skew values than the

10.3389/fmars.2025.1600021

FIGURE 4
Codon distribution **(A)** and Relative Synonymous Codon Usage (RSCU) **(B)** in the mitogenomes of *Ostrea stentina* and in the subfamily Ostreinae. Numbers to the right refer to the total number of codons. CDspT, codons per thousand codons. Codon families are provided on the x-axis.

complete mitogenomes. The 3' *rrnL* ends and *rrnSs* had higher G+C contents than the complete mitogenomes, while the 5' *rrnL* ends had much lower values. The latter exhibited also the highest GC-skews values.

3.3.4 Compositional biases and AT/GC-skews of control regions

Control regions of Ostreinae were particularly rich in A+T (75.11% on average) with only *O. denselamellosa* deviating from this pattern (Supplementary Figure S15). The three control regions of *O. stentina* ranked among those with the highest values (76.81%-77.75%). As direct consequence of the high A+T content, the G+C content was particularly low when compared to that of entire mitogenomes. AT-skews and GC-skews were rather variable and different from those of mitogenomes. In particular, the AT-skews were positive in more than half of the analyzed sequences included those of *O. stentina*, a behavior contrasting with the negative values of complete mitogenomes (Supplementary Figure S15).

3.4 The protein-coding genes in the mitogenomes of Ostreinae

O. stentina exhibited the whole set of PCGs usually present in animal mitogenome (Figures 1, 2). All the PCGs started with standard codons (ATT, ATG, GTG, and TTG) and ended with the canonical TAA codon, except for *cox3, atp6, nad5* and *nad3*, which ended with incomplete stop codons T(aa) or TA(a) (Supplementary Table S1). None of the PCGs overlapped. When the comparison was extended to all Ostreinae sequenced to date (Supplementary Table S2) it appeared that ATG was the most widespread codon followed by GTG. The genes using the most

10.3389/fmars.2025.1600021

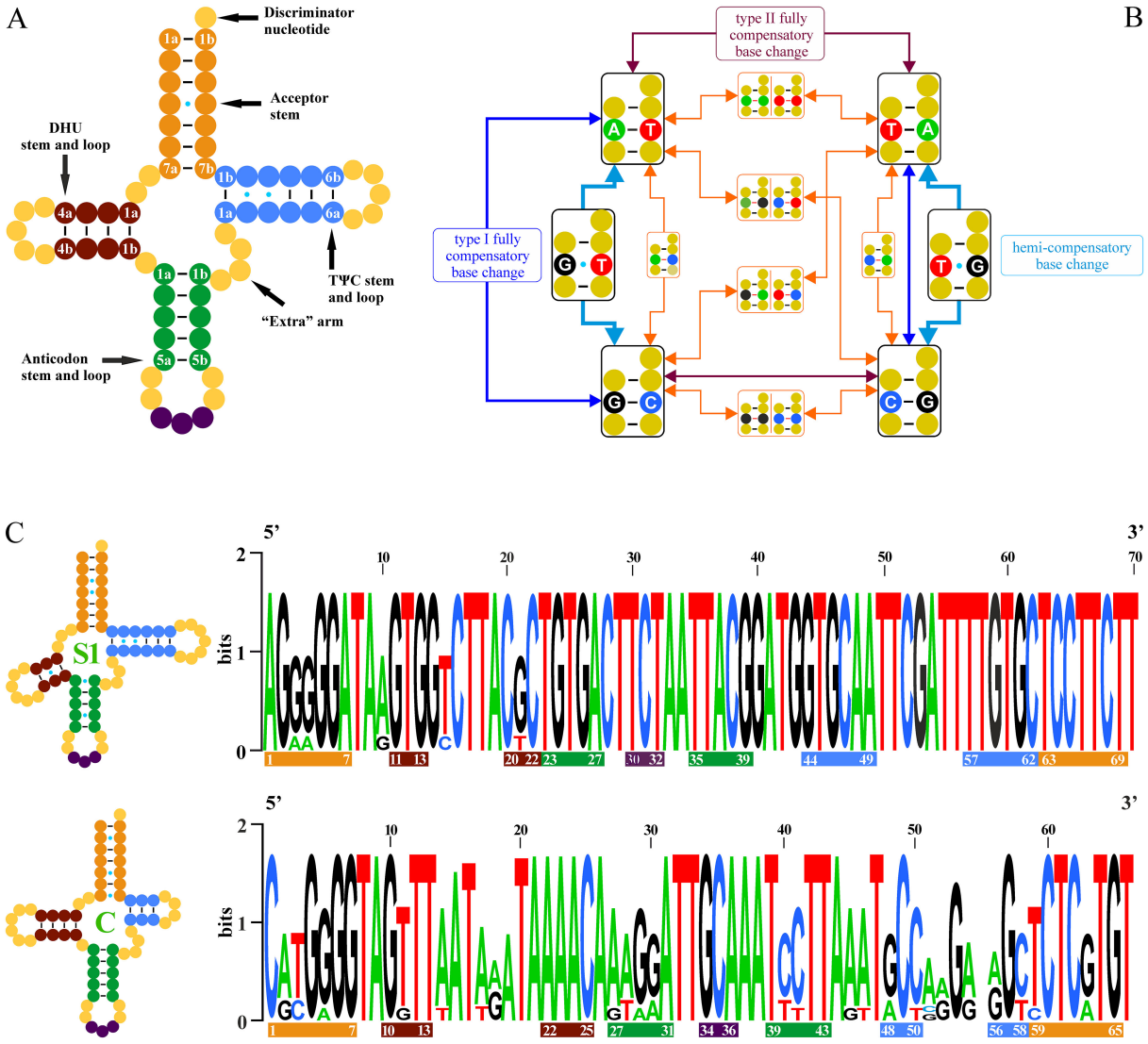

FIGURE 5
Comparative analyses of tRNAs. **(A)** Secondary structure, arms nomenclature and pairs numbering scheme. **(B)** Substitutional pathways leading to the different types of change of nucleotides in the pairs of the arms of a tRNA. **(C)** Logos of *trnS1* and *trnC*, the most conserved and the most variable tRNAs among the species of Ostreinae analyzed in this paper. The canonical Watson-Crick base pairings are figured with a black dash symbol. The wobble base pairings involving G and T are presented with a cyan dot symbol. The base pairings implying a mismatch are figured with a red dash symbol. (See Main text for further details).

variable sets of start codons were *cob* (5) and *nad2* (5), while the pair *cox1-cox2* invariably started with ATG.

The codon distribution and the Relative Synonymous Codon Usage (RSCU) in PCGs were analyzed for the different mitogenomes of oysters of the subfamily Ostreinae. The results are summarized in Figure 4 and Supplementary Figures S16, S17. The average number of codons for the subfamily Ostreinae was 3,710. The range of variation among the taxa was limited, spanning from 3,699 in *D. sandvichensis* to 3,717 codons in *O. lurida* and *O. edulis*. No intraspecific variation was observed for the multiple mitogenomes of *O. stentina* (Supplementary Figure S16A), while a very limited variability was detected in *O. edulis* (Supplementary Figure S16B). All Ostreinae exhibited a very consistent codon distribution and RSCU (Figure 4; Supplementary Figures S16, S17). The most abundant amino acids in mitochondrial proteins determined also the richest codon families. Ser, Leu, Val and Phe were present with more than 300 residues in all taxa, accounting for ≥ 42% of the whole set of amino acids across the 13 proteins. Gly, Ala, Ile, and Met occurred with more than 200 residues (25% of the whole set), while other amino acids were less abundant, with Gln being constantly the rarest. This distribution explained the pattern observed for the codon families in Figure 4, and Supplementary Figure S17. The Val codon family was the most abundant, as both Leu and Ser were split into two families, with Leu2 favored over Leu1, and Ser2 more represented than Ser1. The analysis of RSCU showed that the A+T rich codons were preferred over synonymous codons with a lower content in A+T (Figure 4; Supplementary Figure S17). This result was expected considering the compositional bias toward A+T exhibited by all PCGs (see above).

10.3389/fmars.2025.1600021

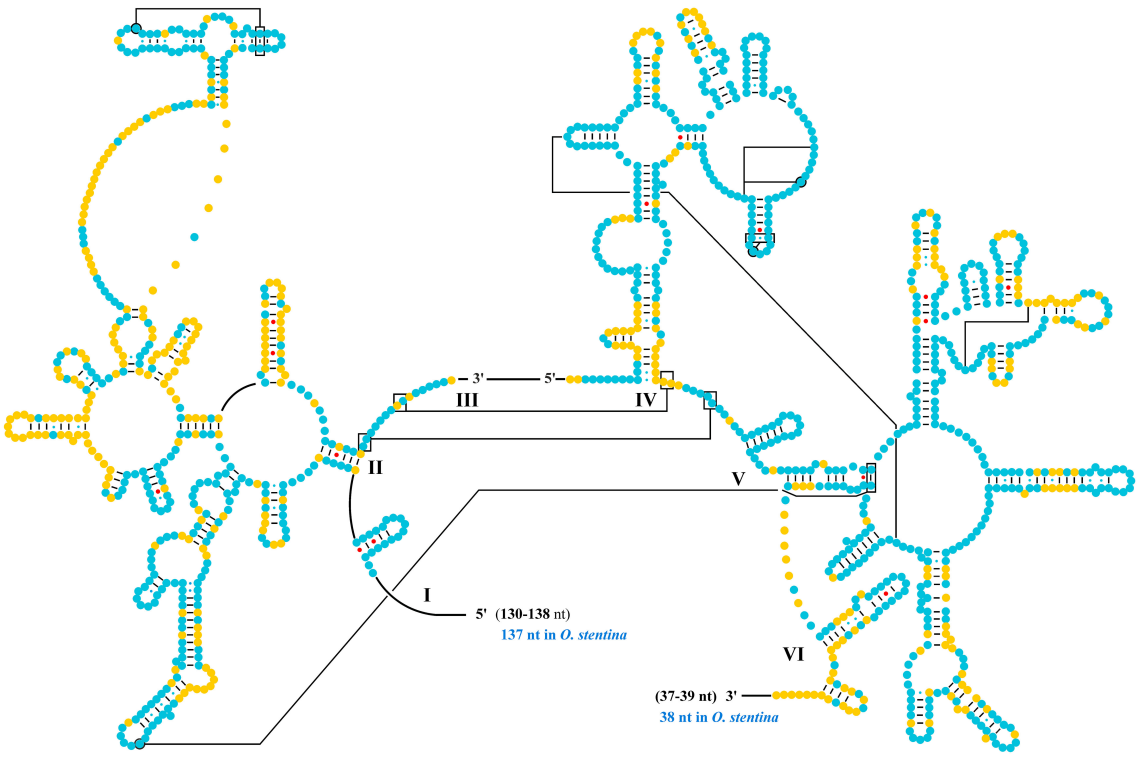

FIGURE 6
Secondary structure of the *rrnL* of *Ostrea stentina*. Roman numerals label the structural domains. Solid lines and boxes mark tertiary structures connections. Watson-Crick pairs are joined by dashes. GT wobble base pairs are joined by a cyan dot, while other non-canonical pairs are connected by a red dot. The fragmentation of the *rrnL* occurs between the 3' end of domain II and the 5' start of domain IV. The 5' nucleotides and 3' nucleotides un-modeled are listed. The 5' and 3' range of *rrnL* variability in Ostreinae is provided. A cyan background marks a position conserved in the multiple alignment of *rrnLs* of the Ostreinae, while a yellow background marks a position variable.

However, all codons were used at least once, as the compositional bias was not so extreme to determine the elimination of rare GC-rich codons. The combined effect of A+T richness, negative AT-skew and positive GC-skews of PCGs on the codon composition was well represented by the behavior of some fourfold-degenerated codon families where the abundance of the third codon base followed this order: T, A, G, C (e.g. Pro, Val). However, this pattern was not always consistent (e.g. Ser2) suggesting that other factors played a role in the final abundance of synonymous codons in the mitogenomes of oysters. The p-distances were computed for orthologous genes, including single codon positions, and proteins of Ostreinae (Supplementary Table S3). The most variable gene/protein was *atp8*, followed by *nad2* and *nad6* while the most conserved was *cox1*. At the intraspecific level, the variability was very low among the two genomes of *O. denselamellosa* (Supplementary Table S4), the four mitogenomes of *O. edulis* (Supplementary Table S5), and the three sequences of *O. stentina* (Supplementary Table S6).

3.5 The transfer RNA genes in the mitogenomes of Ostreinae

The mitogenome of *O. stentina* contained the full 22 tRNAs set of Metazoa plus a duplicated *trnM2*, a feature shared among all oysters sequenced to date (Figure 2). All tRNAs exhibited the clover-leaf secondary structure (Figure 5; Supplementary Figures S18–S20). The analysis of the multiple alignments of orthologous tRNAs (Figure 5; Supplementary Figures S18–S26) revealed different levels of conservation among Ostreinae. Most of the variable positions were located in the single helix portions of the tRNAs, i.e. DHU loop, "extra arm and TΨC loop, which were free to vary without hampering their structure (Figure 5; Supplementary Figures S18–S26). Some of these hyper-variable portions characterized single taxa (e.g. TΨC loop of *trnC* for *O. stentina* and *O. edulis;* Supplementary Figure S21). Base substitutions in the stems were prevalently hemi-compensatory and type I fully compensatory base changes (Figure 5), as they maintained the integrity of the stems, and the molecular pathways leading to them are favored (Montelli et al., 2016). Type II fully compensatory base changes, requiring intermediate mismatches, were much rare but occurred in *trnA* and *trnF* of *O. denselamellosa* (Supplementary Figures S21, S22), and in *trnH*, *trnL1*, *trnM1*, *trnN* and *trnW* of all Ostreinae (Supplementary Figures S21–S24, S26). Mismatches were also present in the stems and restricted to single species (e.g. *O. edulis*, *trnA*; *O. stentina*, *trnE*) (Supplementary Figures S21, S22) or common to all Ostreinae (e.g. *trnD*, *trnN*, *trnQ*, *trnR*, *trnV*) (Supplementary Figures S21, S24–S26).

The most conserved tRNAs was *trnS1* with only 5 variable positions over 70, whereas almost 50% of positions (30/66) changed in *trnC*, the most dynamic tRNA in Ostreinae (Figure 5; Supplementary Figures S21, S25). The tRNAs associated to the most abundant codon families were the least variable (Figure 4; Supplementary Figures S18–S20). The notable exception was represented by *trnQ*, which was associated to the least abundant amino acid (Figure 4) but was among the most conserved tRNAs (Supplementary Figures S19, S24). At the intraspecific level, only one base (T vs C) difference was found in the DHU loop of *trnR* of *O. stentina*. However, the three mitogenomes analyzed here were obtained from specimens of the same locality (Supplementary Figures S27–S30). In contrast, variable tRNAs were present in *O. denselamellosa*/*O. edulis* (Supplementary Figures S27–S30).

3.6 The ribosomal RNA genes in the mitogenomes of Ostreinae

The *rrnSs* were conserved among the analyzed Ostreinae (average p-distance = 0.167 ± 0.077). The G+C content was higher than both the 5' and 3' halves of *rrnL* (Supplementary Figure S14), which suggested a strong and important role of the G-C pairs in the 2D/3D structures of this molecule. The intraspecific variability was minimal for the three species with multiple sequences available (average p-distance = 0.002 ± 0.002 in *O. stentina*; average p-distance = 0.006 ± 0.007 in *O. edulis;* average p-distance = 0.001 in *O. denselamellosa*).

As secondary structure models existed for *rrnL* of the phylum Mollusca (Lydeard et al., 2000) and for the family Ostreidae (Milbury et al., 2010), we used these templates to infer the secondary structure of the *rrnLs* of Ostreinae.

The overall structure of *rrnL* of *O. stentina* is presented in Figure 6, while a detailed representation of the 2D structure is available in Supplementary Figure S31. The structure mirrored those available for other oyster species (Milbury et al., 2010) and more in general molluscs (Lydeard et al., 2000), with domain I and II located in the 5' half, domain III lacking, and domain IV-VI located in the 3' half. The *rrnL* structures inferred for other Ostreinae overlapped with that presented here for *O. stentina* (Figure 6). Among the Ostreinae, most of the variable positions in *rrnL* were located in the 5' half (average p-distance = 0.281 ± 0.112) while the 3' half was much more conserved (average p-distance = 0.148 ± 0.072). This higher level of conservation reflects the prominent structural role of the 3' half for the functioning of the whole *rrnL* molecule (Lydeard et al., 2000; Milbury et al., 2010). Furthermore, the 3' half was markedly GC-richer than the 5' half (Supplementary Figure S14), and the stability of the stems in its highly conserved domains IV-V was often guaranteed by the pairs G-C and C-G (Figure 6; Supplementary Figure S31). Intraspecific behavior mirrored that observed in the comparisons among different species of Ostreinae. The 5' half was more variable than the 3' segment in all tree species (Supplementary Figures S32–S34). However, the level of variation was very limited or non-existent as in case of the 3' halves of *O. denselamellosa* (Supplementary Figure S34B).

3.7 The control region of Ostreinae mitogenome

In the mitogenome of Ostreinae, the control region was located between *trnD* and *cox1* (Figure 2) and, for the first time, was characterized in detail in this paper for this subfamily of oysters. Its size varied from 688 bp (*O. denselamellosa* ON964460) to 742 bp (*N. pinnicola*). As stated above, the CoRe was extremely AT-rich (Supplementary Figure S15). The sequences of CoRe were highly diverging, as proved by the very low number of fully conserved positions in their multiple alignment (Supplementary Figure S35). The alignment showed large portions that were difficult to align with high accuracy (Chang et al., 2014), even when using a highly sophisticated software such as T-coffee (Di Tommaso et al., 2011). Despite the high variability, two segments appeared rather conserved in the alignment: one spanning positions 70 to 130, and the other ranging from positions 480 to 570 (Supplementary Figure S35). Two fully conserved motifs were present in the CoRes of Ostreinae. The first one (AAAGGGG) started at position 171 of the alignment (Supplementary Figure S35). This motif was present also in *rrnS*. It occurred in the mitogenome of other Osteidae (*Saccostrea*, *Magallana* and *Talonostrea*), but not in their CoRe. A second fully conserved motif of 11 nucleotides (CTATGTAAATA) extended from position 552 to position 562 (Supplementary Figure S35). This motif was exclusive of the CoRe of Ostreinae sequenced to date and did not occur in the mitogenomes of other oysters. The CoRe of *O. denselamellosa* presented a second copy of this motif (positions: 132-142, Supplementary Figure S35).

The CoRes of *O. stentina* ranged from 701 bp to 711 bp (Supplementary Table S1, Supplementary Figure S35) and their average p-distance was 0.010 ± 0.008. The four CoRes of *O. edulis* ranged from 695 bp to 700 bp, and two were identical (JF274008 and CM063324). Their average p-distance was 0.022 ± 0.023. Additionally, the two CoRes of *O. denselamellosa* differed in length by one nucleotide (688 vs. 689) (Supplementary Figure S35) and their p-distance was 0.017. A very peculiar case was observed for *D. sandvichensis* and *P. pestigris*, where the available mitogenomes exhibited identical CoRes.

Stretches of polyA, polyT and polyG, as well as polyAT, characterized the CoRes of Ostreinae (Supplementary Figure S35). These features are peculiar of, and specific to, the control regions of molluscs and, more in general, animals (Ghiselli et al., 2021). Finally, all CoRes of Ostreinae were able to form stem-and-loop secondary structures (Figure 7), and these structures were located in a highly variable portion of their multiple alignment. These structures are considered important for the replication and transcription of mitogenomes (Ghiselli et al., 2021).

3.8 Substitution patterns in the multiple alignments of orthologous sequences

The level of compositional heterogeneity occurring among orthologous sequences was evaluated for all 13 PCGs and rDNAs genes with the software AliGROOVE (Kück et al., 2014)

10.3389/fmars.2025.1600021

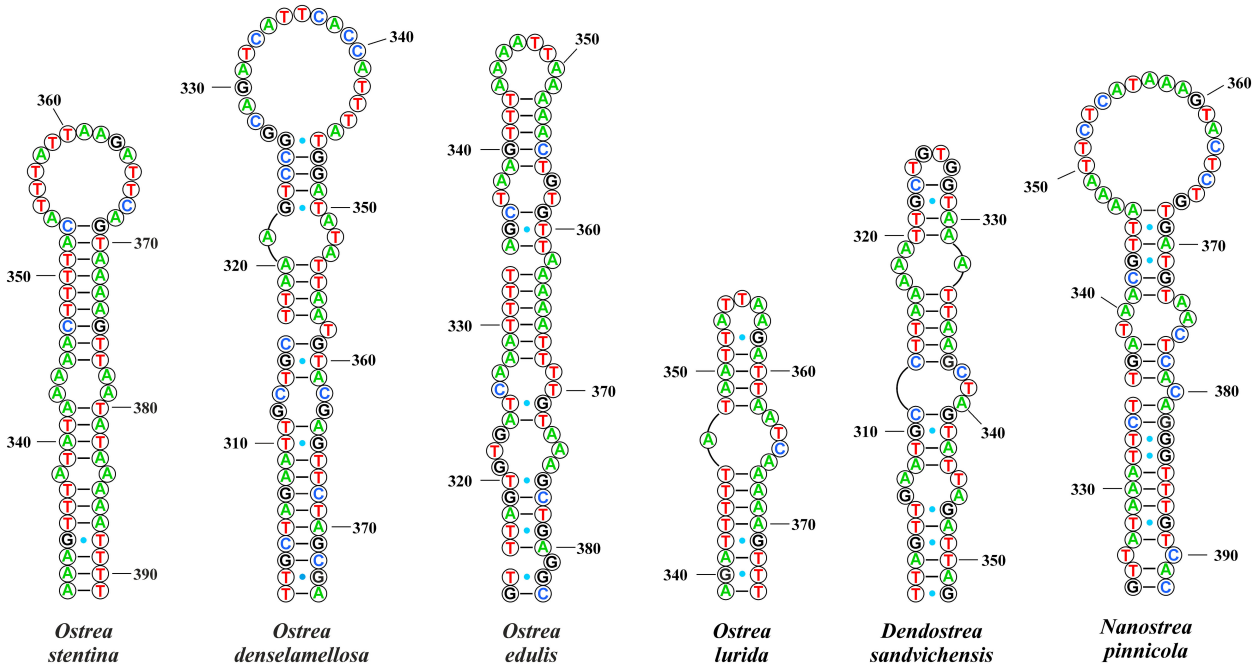

FIGURE 7
Secondary structures identified in the control regions (CoRes) of Ostreinae with the software RNAstructure. Numbers refer to the nucleotide positions in the CoRes sequences. Watson-Crick pairs are joined by dashes. GT wobble base pairs are joined by a cyan dot.

(Supplementary Figures S36–S43). The third codon positions of all PCGs exhibited high heterogeneous substitution patterns, while the most homogenous single positions were the second positions of several PCGs (i.e. *cob*, *cox1-cox2*, *nad1*, *nad3-nad5*). Amino acid sequences were very homogeneous in their substitution patterns with rare exceptions observed in *nad2* and *nad6* (Supplementary Figures S39, S42). Substitution patterns for *rrnL* and *rrnS* were homogeneous among Ostreinae species but were heterogeneous compared with the sequences of other subfamilies (Supplementary Figure S42). Additionally, when the 13PCG data sets were considered (Table 2), the substitution process was homogenous for amino acids, the first (mostly) and second positions of codons, as well as for first + second positions, plus a large part of whole codons. On the contrary, the substitution pattern was heterogeneous for third positions, except in intraspecific comparisons (Supplementary Figure S43).

3.9 Phylogenetic signal detection in the data sets

The phylogenetic signal for single PCG was the highest for amino acids and first + second codon positions, while third positions appeared highly saturated, as evidenced by the maximum likelihood distances, largely exceeding 1, and the lowest percentage of fully resolved quartets (Supplementary Figures S44–S46). Both *rrnL* and *rrnS* exhibited a good phylogenetic signal. Similarly, the best signal among the concatenated alignments was observed for 13PCGpro data set and 13PCGp12 data set (Table 2; Supplementary Figure S46), which contained respectively the amino acid sequences and the first + second position of the 13 PCGs.

3.10 Phylogenetic trees reconstruction

13PCGpro exhibited the best signal and the most homogeneous substitution pattern among the analyzed data sets (see above) (Table 2). The ML tree (hereafter Tree 1) obtained from this set is provided in Figure 8 (Supplementary Table S7). Most of the nodes and branches received very strong statistical corroboration. Within Ostreinae, the genus *Ostrea* appeared polyphyletic, with *O. stentina* sister species of *O. lurida* and *O. edulis* + *O. denselamellosa* forming a separated group nested within a second clade encompassing *N. pinnicola* + *P. plestigris*, their sister taxon, and *D. sandvichensis*. UBT/aLRT values strongly support this clade. 13PCGpro+rRNAs, 13PCGp2, 13PCGp12.a/b, and 13PCGp12+rRNAs.a/b produced also Tree 1 (Supplementary Figures S47–S52, and Supplementary Table S7).

Some data sets listed in Table 2 generated trees that differed from tree 1(Supplementary Figures S53–S59, and Supplementary Table S7). However, these alternative topologies lacked strong statistical support. We will analyze Tree 3, 4, and 6 (Supplementary Figures S54, S55, S57) in more detail here, as the *Ostrea* genus resulted monophyletic, although the most basal node, the critical one, did not receive strong statistical support. All these trees were the product of the analyses performed on data sets including third positions of codons and/or rRNAs (Table 2;

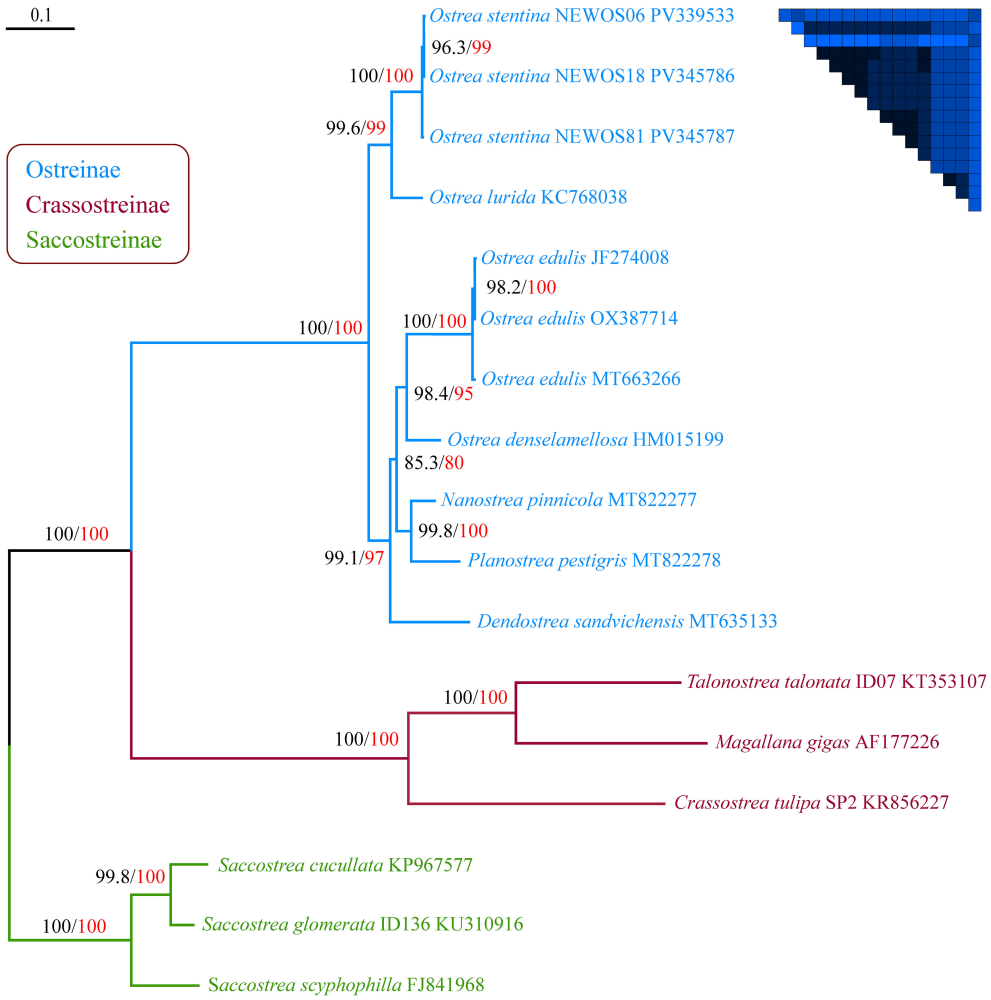

FIGURE 8
Best maximum likelihood tree (-ln = 38,261.6328) inferred from the 13PCGpro data set. The analysis was performed by applying the evolutionary models and the partitioning scheme listed in Supplementary Table S7.1. Black-colored numbers indicate UBT values whereas red-colored numbers indicate aLRT values, both expressed as percentage. The scale bar represents 0.09 substitutions/site. On the top right corner is figured the AliGROOVE matrix obtained from the 13PCGpro data set. This matrix contains only blue colored squares, thus denoting a high homogeneous substitution process among the sequences of the data set.

Supplementary Table S7). The third positions of codons exhibited a highly heterogeneous substitution pattern and their phylogenetic signal was mostly/completely lost, two factors that are highly detrimental for phylogenetic analyses (Negrisolo et al., 2004; Kück et al., 2014). Furthermore, they failed a test of stationarity or homogeneity ($p < 0.05$) computed with IQ-TREE2, raising serious concerns about the reliability of the trees derived from their analyses (Naser-Khdour et al., 2019). Tree 6 was the phylogenetic output of the rRNAs data set. The ribosomal markers exhibited good phylogenetic signals. However, the substitution pattern was not homogeneous between ingroup and outgroup sequences (Supplementary Figure S42), a factor that can influence the phylogenetic outputs (Kück et al., 2014).

The results of alternative topologies tests performed on the data sets analyzed in the present paper are summarized in Supplementary Table S8. Tree 1, our reference topology, was rejected only by 13PCGp3, the least reliable analyzed data set. Conversely, the most robust data set, i.e. 13PCGpro, rejected nearly all alternative topologies, except for Tree 4 in the AU test and Trees 4 and 5 in the highly conservative WSH test (Shimodaira, 2002).

3.11 Intergenic spacers in the mitogenomes of Ostreinae

The mitogenomes of *O. stentina* contained 28 ISPs, ranging from 1 (ISP *trnL2-trnP*) to 117-118 (ISP *trnG-cox3*) nucleotides (Figure 1; Supplementary Table S1). No genes overlapped. Similar patterns characterized the mitogenomes of Ostreinae sequenced to date (Figure 9). In several pair of consecutive genes, the behavior is fixed: they were either separated by an ISP (e.g. *cox1* and *trnG*, *trnK* and *trnL1*), or adjacent (e.g. *trnH* and *nad4*, *nad3* and *trnK*). In

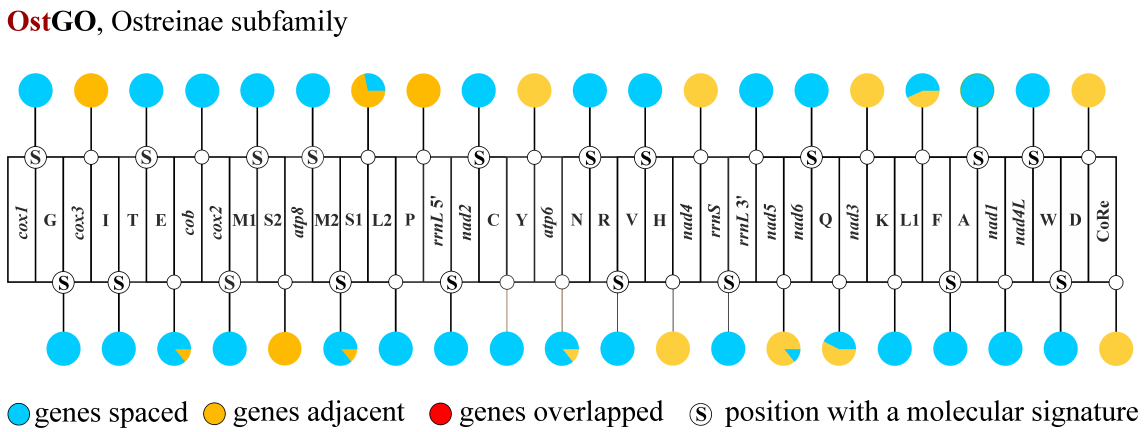

FIGURE 9
Distribution of intergenic spacers and occurrence of molecular signatures in the mitogenomes of the Ostreinae.

other cases, the pattern changed in different species. In particular, *trnL1* and *trnF* were spaced in *O. edulis*, *O. denselamellosa, N. pinnicola*, and *P. plestigris*, and adjacent in other oysters (Figure 9; Supplementary Figure S60G).

We analyzed the sequences of the ISPs by aligning them manually or with MAFFT. Notably, in 19 ISPs we identified sequences (Figure 8) that were exclusive to and characterized a single species or group of species located downstream to a well-supported node of the reference Tree 1 (e.g. ISP *trnG-cox3*; *O. lurida* + *O. stentina*) (Figure 8; Supplementary Figure S60).

4 Discussion

4.1 Gene order evolution in the mitogenomes of Ostreidae

Phylogenetic relationships among the four subfamilies of Ostreidae are well defined (Figure 2B) (Salvi and Mariottini, 2017; Li et al., 2021; Salvi and Mariottini, 2021). We identified gene blocks shared among the different subfamilies and, considering their phylogenetic relationships, we were able to partly infer the gene order arrangement of the mitogenome of the last common ancestor (lcaO) of all Ostreidae (Figure 2B). Saccostreinae and Ostreinae are not sister taxa, but share two large gene blocks, which represent a plesiomorphic condition for the family Ostreidae (Figure 2B). Similarly, the four blocks shared among the three subfamilies Crassostreinae, Ostreinae and Saccostreinae represent further plesiomorphies (Figure 2B). The complete transformational pathway that lead to the diversity of GOs observed today in oysters, particularly in Crassostreinae, remains to be fully understood. Sequencing the mitogenomes of Striostreinae, the sister group of Crassostreinae (Figure 2B), and the only subfamily without available mitochondrial sequences, is a priority to properly address this issue.

4.2 Mito-phylogenomics of the Ostreidae

Standard evolutionary models used in phylogenetic analyses assume that the substitution process among orthologous sequences in the multiple alignments is homogeneous and the violation of this assumption may generate misleading phylogenetic outputs (Kück et al., 2014). The lack of phylogenetic signal is another important source of distortive effects on phylogenetic results. We performed the quartet puzzling analysis (Strimmer and von Haeseler, 1996) and analyzed the distribution of the pairwise distances computed according to the best-fit evolutionary model to test this amount of phylogenetic signal (e.g. Negrisolo et al., 2004). It is well known that when the distribution of these distances is considerably greater than one, there is a substantial loss of phylogenetic signal in the analyzed dataset (e.g. Negrisolo et al., 2004). The best phylogenetic markers, when working at the taxonomic level of the family, proved to be the proteins. The 13PCGpro multiple alignment exhibited the best signal and the most homogeneous substitution pattern among the analyzed data sets. In contrast, third positions of codons showed very heterogeneous substitution patterns and substantial lack of phylogenetic signal.

The genus *Ostrea* was polyphyletic in our reference tree (Figure 8). We obtained also alternative topologies implying the monophyly of this taxon, but these were obtained from data sets that proved to be unreliable markers (third positions) or difficult to manage (ribosomal genes) due to non-homogeneous substitution patterns occurring between ingroup and outgroup (Kück et al., 2014). *Ostrea* resulted para/polyphyletic also in previous phylogenetic analyses based on both nuclear and mitochondrial genes (Salvi et al., 2014; Guo et al., 2018; Li et al., 2021; Salvi and Mariottini, 2021), suggesting that homoplasy characterizes the morphological evolution of the genus. Our results strongly support a polyphyletic nature for *Ostrea*. However, we worked with a limited taxon sampling. Therefore, a wider species coverage is necessary for corroborating this point.

4.3 Variability and molecular signatures in the mitogenomes of Ostreidae

The *atp8* was the most variable PCG, followed by *nad2* and *nad6*, while the most conserved was *cox1*. This result is very interesting and supports the hypothesis that the mitogenomes of oysters contain multiple PCGs that can be used for molecular identification of the species outside of *cox1* and further corroborates earlier findings (e.g. Xiao et al., 2015). At the intraspecific level, variability was limited in *O. stentina*. However, all mitogenomes were obtained from specimens collected in the Venice Lagoon, thus they do not represent the global diversity of the species.

The mitochondrial tRNAs harbor a considerable amount of taxonomic and evolutionary information that fully stands out when their secondary structure is considered (e.g. Simonato et al., 2013; Montelli et al., 2016). Unfortunately, these markers are often given only a cursory treatment. In this study, we analyzed in details the substitution process characterizing the multiple alignments of orthologous tRNAs. Particularly interesting are the base changes occurring in the stems of tRNAs (Coleman, 2003; Montelli et al., 2016). In our study, the tRNAs associated with the most abundant codon families were the least variable. This pattern of conservation supports the hypothesis that these tRNAs have a more constrained nucleotide substitution pattern, associated to their high frequency of usage in the protein synthesis.

Some tRNAs exhibited AT-/GC- skews values that differed greatly from those of the strand encoding them. This is not unique to oysters. A similar pattern was observed in the tRNAs of Cetacea (Montelli et al., 2016). It was not possible to identify a single cause (e.g. tRNAs associated to abundant amino acids) that explained this result. The short length of tRNAs likely played a role, as even a small number of substitutions can have a strong impact on their skew values.

The hyper-variable portions of DHU loop, "extra arm and TΨC loop of several mitochondrial tRNAs exhibited sequence motifs that characterized single/group of species of oysters. Fully compensatory base changes, as well as mismatches, were also present in the stems of tRNAs, either restricted to single oyster or, conversely, exclusive to the entire subfamily Ostreinae. Our taxon coverage is very sparse, but despite this limitation, these tRNAs features could serve as additional taxonomic tools for the family Ostreidae, where identification of species and taxa relationships are problematic (Harry, 1985), as observed in other groups of invertebrates (e.g. Simonato et al., 2013).

For *rrnS*, a secondary structure model did not exist for Ostreidae and we did not attempt to develop a new one. In contrast, we used the secondary structure models of *rrnL* available for the phylum Mollusca (Lydeard et al., 2000) and for the family Ostreidae (Milbury et al., 2010) to infer the secondary structure of the *rrnL* of O. *stentina*. The analyses of compositional biases and AT-/GC-skews of *rrnLs* and *rrnSs* suggests that structural constraints played a key role in shaping these features.

The CoRe of all analyzed mitogenomes contained the peculiar sequence motif CTATGTAAATA. If this motif is found to be exclusive to all Ostreinae, it might become a very useful marker to unambiguously identify this genomic portion, similar to other motifs identified in various animal groups (e.g. Lepidoptera; Salvato et al., 2008). A very peculiar case was observed for *D. sandvichensis* and *P. pestigris*, where the available mitogenomes exhibited identical CoRes. These sequences were produced by the same research group at different times. As shown above, CoRes are variable at the intraspecific level. Furthermore, *D. sandvichensis* and *P. pestigris* are not sister species (Figure 8). Therefore, the occurrence of an identical control region in their mitogenomes requires independent confirmation.

Many intergenic spacers located throughout the mitogenome (Figure 9) contain sequences characteristic of a single species or clade. These sequences are mito-signatures (Liu et al., 2022), i.e. molecular markers useful to define/identify taxa in a phylogenetic context, but cannot be considered true synapomorphyes. Uniqueness is the hallmark of a true apomorphy (Page and Holmes, 2009). However, it is very unlikely that an often short sequence of ISP could fulfill this stringent requirement. A mito-signature can be very useful to identify a species, a group of species, or even bigger taxa, within a well-established phylogenetic framework. This is particularly relevant in animals like oysters, as they are difficult to identify on a morphological basis (e.g. Harry, 1985). In the past, the value of ISP as intraspecific phylogenetic markers has been shown in the Crassostreinae (Ren et al., 2016). Our findings further corroborate this point and extend, at the interspecific level, the taxonomic/phylogenetic value of these short sequences for oysters, as already known in other groups of animals (e.g. Simonato et al., 2013; Basso et al., 2017; Liu et al., 2022).

5 Conclusions

For the first time, we provided at least a partial reconstruction of the gene arrangement in the mitogenome of the last common ancestor of the oysters. Our analysis revealed a complex molecular landscape of the different types of genes encoded in mitogenomes of these bivalves. Our phylogenomic analyses proved that multiple factors influence phylogenetic inference and supported previous findings indicating the polyphyly of the genus *Ostrea*. Finally, our study confirmed for the first time that, besides the widely used *cox1*, oyster mitogenomes contain several underutilized genetic markers with relevant phylogenetic/taxonomic information. These markers should be routinely used to identify species as well as to study their evolutionary relationships.

Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/Supplementary Material.

Ethics statement

The manuscript presents research on animals that do not require ethical approval for their study.

Author contributions

DC: Writing – original draft, Formal analysis, Visualization, Data curation. RF: Writing – original draft, Resources, Investigation. MB: Investigation, Resources, Writing – original draft, Data curation. DT: Conceptualization, Resources, Writing – original draft. IG: Investigation, Resources, Writing – original draft. MS: Resources, Writing – original draft. VB: Formal analysis, Writing – original draft. TP: Funding acquisition, Writing – original draft. EN: Data curation, Visualization, Supervision, Conceptualization, Investigation, Formal analysis, Resources, Funding acquisition, Writing – original draft.

Funding

The author(s) declare that financial support was received for the research and/or publication of this article. DC was supported by a PhD scholarship provided by Padua University. This project was funded by the grant BIRD191298/19 provided to EN by the BCA Department (University of Padova). Field activities were performed in the framework of the Venezia 2021 Research Program, coordinated by CORILA (Consortium for coordination of research activities concerning the Venice Lagoon system) and funded by the Ministero delle Infrastrutture e dei Trasporti (Provveditorato Interregionale per le Opere Pubbliche del Veneto - Trentino Alto Adige - Friuli Venezia Giulia), grant number 21/18/AC_AR02 (04/12/2018). The works was funded by the National Recovery and Resilience Plan (NRRP), Mission 4 Component 2 Investment 1.4 - funded by the European Union – NextGeneration EU; Award Number: Project code CN_00000033, Italian Ministry of University and Research, "National Biodiversity Future Centre-NBFC".

Acknowledgments

We thank L. Dametto for the technical support in field activities.

In memoriam

This paper is dedicated to the memory of Davide, whose untimely passing is a profound loss. His contributions and presence will be deeply missed.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Generative AI statement

The author(s) declare that no Generative AI was used in the creation of this manuscript.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmars.2025.1600021/full#supplementary-material

References

Abascal, F., Zardoya, R., and Telford, M. J. (2010). TranslatorX: multiple alignment of nucleotide sequences guided by amino acid translations. *Nucleic Acids Res.* 38, W7–W13. doi: 10.1093/nar/gkq291

Amemiya, I. (1928). Ecological studies of Japanese oysters, with special reference to the salinity of their habitats. *J. Coll. Agric. Univ.* 9, 333–382.

Andrews, S. (2010). *FastQC: a quality control tool for high throughput sequence data*. Available online at: https://www.bioinformatics.babraham.ac.uk/projects/fastqc/ (Accessed May 26, 2025).

Anisimova, M., and Gascuel, O. (2006). Approximate likelihood-ratio test for branches: a fast, accurate, and powerful alternative. *Systematic Biol.* 55, 539–552. doi: 10.1080/10635150600755453

Babbucci, M., Basso, A., Scupola, A., Patarnello, T., and Negrisolo, E. (2014). Is it an ant or a butterfly? Convergent evolution in the mitochondrial gene order of Hymenoptera and Lepidoptera. *Genome Biol. Evol.* 6, 3326–3343. doi: 10.1093/gbe/evu265

Basso, A., Babbucci, M., Pauletto, M., Riginella, E., Patarnello, T., and Negrisolo, E. (2017). The highly rearranged mitochondrial genomes of the crabs *Maja crispata* and *Maja squinado* (Majidae) and gene order evolution in Brachyura. *Sci. Rep.* 7, 4096. doi: 10.1038/s41598-017-04168-9

Bolger, A. M., Lohse, M., and Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30, 214–2120. doi: 10.1093/bioinformatics/btu170

Born Von, I. (1778). *Index rerum naturalium Musei Cæsarei Vindobonensis* (Vindobonae [Vienna]; (Kraus: Pars I.ma. Testacea. Verzeichniß der natürlichen Seltenheiten des k. k. Naturalien Cabinets zu Wien. Erster Theil. Schalthiere), 1–458.

Breton, S., Stewart, D. T., and Hoeh, W. R. (2010). Characterization of a mitochondrial ORF from the gender-associated mtDNAs of *Mytilus* spp. (Bivalvia: Mytilidae): identification of the 'missing' ATPase 8 gene. *Marine Genomics* 3, 11–18. doi: 10.1016/j.margen.2010.01.001

Cannone, J., Subramanian, S., Schnare, M. N., Collet, J. R., D'Souza, L. M., Du, Y., et al. (2002). The Comparative RNA Web (CRW) Site: an online database of comparative sequence and structure information for ribosomal, intron, and other RNAs. *BMC Bioinformatics;* 3, 2. doi: 10.1186/1471-2105-3-2

Carpenter, P. P. (1864). Diagnoses of new forms of Mollusca collected at Cape St. Lucas, Lower California. *Ann. Magazine Natural History*.

Cavaleiro, N. P., Sole-Cava, A. M., Melo, C. M. R., de Almeida, L. G., Lazoski, C., and Vasconcelos, A. T. R. (2016). The complete mitochondrial genome of *Crassostrea gasar* (Bivalvia: Ostreidae). *Mitochondrial DNA Part A* 27, 2939–2940. doi: 10.3109/19401736.2015.1060450

Chan, P. P., and Lowe, T. M. (2019). tRNAscan-SE: searching for tRNA genes in genomic sequences. *Methods Mol. Biol.* 1962, 1–14. doi: 10.1007/978-1-4939-9173-0_1

Chang, J.-M., Di Tommaso, P., and Notredame, C. (2014). TCS: a new multiple sequence alignment reliability measure to estimate alignment accuracy and improve phylogenetic tree reconstruction. *Mol. Biol. Evol.* 31, 1625–1637. doi: 10.1093/molbev/msu117

Chernomor, O., von Haeseler, A., and Quang Minh, B. (2016). Terrace aware data structure for phylogenomic inference from supermatrices. *Systematic Biol.* 65, 997–1008. doi: 10.1093/sysbio/syw037

Clark, K., Karsch-Mizrachi, I., Lipman, D. J., Ostell, J., and Sayers, E. W. (2015). GenBank. *Nucleic Acids Res.* 4, D67–D72. doi: 10.1093/2Fnar/2Fgkv1276

Coleman, A. W. (2003). ITS2 is a double-edged tool for eukaryote evolutionary comparisons. *Trends Genet.* 9, 370–375. doi: 10.1016/S0168-9525(03)00118-5

Crooks, G. E., Hon, G., Chandonia, J. M., and Brenner, S. E. (2004). WebLogo: a sequence logo generator. *Genome Res.* 14, 188–190. doi: 10.1101/gr.849004

Danic-Tchaleu, G., Heurtebise, S., Morga, B., and Lapègue, S. (2011). Complete mitochondrial DNA sequence of the European flat oyster *Ostrea edulis* confirms Ostreidae classification. *BMC Res. Notes* 4, 400. doi: 10.1186/1756-0500-4-400

Di Tommaso, P., Moretti, S., Xenarios, I., Orobitg, M., Montanyola, A., Chang, J.-M., et al. (2011). T-Coffee: a web server for the multiple sequence alignment of protein and RNA sequences using structural information and homology extension. *Nucleic Acids Res.* 39, W13–W17. doi: 10.1093/nar/gkr245

Faustino, L. A. (1932). Recent and fossil shells from the Phil. Islands. I. *Philippine J. Sci.*

Felsenstein, J. (2004). *Inferring phylogenies* (Sunderland (MA): Sinauer Associates).

Fujita, T. in Wakiya, Y. (1929). Japanese food oysters. *Japan J. Zoology* 2, 359–367.

Ghiselli, F., Gomes-dos-Santos, A., Adema, C. M., Lopes-Lima, M., Sharbrough, J., and Boore, J. L. (2021). Molluscan mitochondrial genomes break the rules. *Philos. Trans. R. Soc. B: Biol. Sci.* 376, 1825. doi: 10.1098/rstb.2020.0159

Gmelin, J. F. (1791). Vermes. *Caroli Linnaei Systema Naturae per Regna Tria Naturae* (Leipzig) 1, 3021–3910, 6.

Gould, A. A. (1850). [descriptions of new species of shells from the United States Exploring Expedition]. *Proc. Boston Soc. Natural History.* 3, 151-156, 169–172, 214-218, 252–256, 275–278, 292–296, 309–312, 343–348.

Guo, X., Li, C., Wang, H., and Xu, Z. (2018). Diversity and evolution of living oysters. *J. Shellfish Res.* 37, 755–771. doi: 10.2983/035.037.0407

Hamaguchi, M., Manabe, M., and Kajihara, N. (2017). DNA barcoding of flat oyster species reveals the presence of *Ostrea stentina* Payraudeau 1826 (Bivalvia: ostreidae) in Japan. *Marine Biodiversity Records* 10, 4. doi: 10.1186/s41200-016-0105-7

Hanley, S. (1846). A description of new species of Ostreae, in the collection of H. *Cuming Esq. Proc. Zoological Soc. London.* 13, 105–107.

Harry, H. (1985). Synopsis of the supraspecific classification of living oyster (Bivalvia: Gryphaeidae and Ostreidae). *Veliger* 28, 121–158.

Hayer, S., Brandis, D., Immel, A., Susat, J., Torres-Oliva, M., Ewers-Saucedo, C., et al. (2021). Phylogeography in an 'oyster' shell provides first insights into the genetic structure of an extinct *Ostrea edulis* population. *Sci. Rep.* 11, 2307. doi: 10.1038/s41598-021-82020-x

Hu, L., Wang, H., Zhang, Z., Li, C., and Guo, X. (2019). Classification of small flat oysters of *Ostrea stentina* species complex and a new species *Ostrea neostentina* sp. nov. (Bivalvia: Ostreidae). *J. Shellfish Res.* 38, 295. doi: 10.2983/035.038.0210

Jin, J.-J., Yu, W.-B., Yang, J.-B., Song, Y., dePamphilis, C. W., Yi, T.-S., et al. (2020). GetOrganelle: a fast and versatile toolkit for accurate *de novo* assembly of organelle genomes. *Genome Biol.* 21, 241. doi: 10.1186/s13059-020-02154-5

Kalyaanamoorthy, S., Quang Minh, B., Wong, T., von Haeseler, A., and Jermiin, L. S. (2017). ModelFinder: Fast model selection for accurate phylogenetic estimates. *Nat. Methods* 14, 587–589. doi: 10.1038/nmeth.4285

Katoh, K., Misawa, K., Kuma, K., and Miyata, T. (2002). MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Res.* 30, 3059–3066. doi: 10.1093/nar/gkf436

Kong, L., Li, Y., Kocot, K. M., Yang, Y., Qi, L., Li, Q., et al. (2020). Mitogenomics reveals phylogenetic relationships of Arcoida (Mollusca, Bivalvia) and multiple independent expansions and contractions in mitochondrial genome size. *Mol. Phylogenet. Evol.* 150, 106857. doi: 10.1016/j.ympev.2020.106857

Kück, P., Meid, S., Groß, C., Wägele, J. W., and Misof, B. (2014). AliGROOVE – visualization of heterogeneous sequence divergence within multiple sequence alignments and detection of inflated branch support. *BMC Bioinf.* 15, 294. doi: 10.1186/1471-2105-15-294

Kumar, S., Stecher, G., Li, M., Knyaz, C., and Tamura, K. (2018). MEGA X: molecular evolutionary genetics analysis across computing platforms. *Mol. Biol. Evol.* 35, 1547–1549. doi: 10.1093/molbev/msy096

Lam, K., and Morton, B. (2006). Morphological and mitochondrial-DNA analysis of the indo-west Pacific rock oysters (Ostreidae: *Saccostrea* species). *J. Molluscan Stud.* 72, 235–245. doi: 10.1093/mollus/eyl002

Lam, K., and Morton, B. (2003). Mitochondrial DNA and morphological identification of a new species of Crassostrea (Bivalvia: Ostreidae) cultured for centuries in the Pearl River Delta, Hong Kong, China. *Aquaculture* 228, 1–13. doi: 10.1016/S0044-8486(03)00215-1

Lamarck de, J.-B. M. (1819). *Histoire naturelle des animaux sans vertèbres* (Paris).

Lapègue, S., Boutet, I., Leitao, A., Heurtebise, S., Garcia, P., Thiriot-Quievreux, C., et al. (2002). Trans-Atlantic distribution of a mangrove oyster species revealed by 16S mtDNA and karyological analyses. *Biol. Bull.* 202, 232–242. doi: 10.2307/1543473

Levene, H. (1960). *In Contributions to Probability and Statistics: Essays in Honor of Harold Hotelling*. Ed. I. Olkin, S. G. Churye, W. Hoeffding, W. G. Madow and H.B. Mann (Stanford: Stanford University Press), 278–292.

Li, X., Bai, Y., Dong, Z., Xu, C., Liu, S., Yu, H., et al. (2023). Chromosome-level genome assembly of the European flat oyster (*Ostrea edulis*) provides insights into its evolution and adaptation. *Comp. Biochem. Physiol. - Part D: Genomics Proteomics* 45, 101045. doi: 10.1016/j.cbd.2022.101045

Li, C., Kou, Q. I., Zhang, Z., Hu, L., Huang, W., Cui, Z., et al. (2021). Reconstruction of the evolutionary biogeography reveal the origins and diversification of oysters (Bivalvia: Ostreidae). *Mol. Phylogenet. Evol.* 164, 107268. doi: 10.1016/j.ympev.2021.107268

Linnaeus, C. (1758). *Systema Naturae per regna tria naturae, secundum classes, ordines, genera, species, cum characteribus, differentiis, synonymis, locis. [The system of nature through the three kingdoms of nature, according to classes, orders, genera, species, with characters, differences, synonyms, places.]* Vol. 1 (Holmiae [Stockholm]: Impensis Direct. Laurentii Salvii), 824.

Li, X., and Qi, Z. (1994). Studies on the comparative anatomy, systematic classification and evolution of Chinese oysters (In Chinese). *Studia Marina Sinica.* 35, 143–173.

Lischke, C. E. (1869). Diagnosen neuer Meeres-Konchylien von Japan. *Malakozoologische Blätter.* 16, 105–109.

Liu, D., Basso, A., Babbucci, M., Patarnello, T., and Negrisolo, E. (2022). Macrostructural evolution of the mitogenome of butterflies (Lepidoptera, Papilionoidea). *Insects* 13, 358. doi: 10.3390/insects13040358

Llabrés, M., Rosselló, F., and Valiente, G. (2021). The generalized Robinson-Foulds distance for phylogenetic trees. *J. Comput. Biol.* 28, 1181–1195. doi: 10.1089/cmb.2021.0342

Lunetta, A., Albentosa, M., Nebot-Colomer, E., Pardo, B. G., Martínez, P., Villalba, A., et al. (2023). Assessment of *Ostrea stentina* recruitment and performance in the Mar Menor lagoon (SE Spain). *Regional Stud. Marine Sci.* 58, 102760. doi: 10.1016/j.rsma.2022.102760

Lydeard, C., Holznagel, W. E., Schnare, M. N., and Gutell, R. R. (2000). Phylogenetic analysis of molluscan mitochondrial LSU rDNA sequences and secondary structures. *Mol. Phylogenet. Evol.* 17, 83–102. doi: 10.1006/mpev.1999.0719

Milbury, C. A., and Gaffney, P. M. (2005). Complete mitochondrial DNA sequence of the eastern oyster *Crassostrea virginica*. *Marine Biotechnol.* 7, 697–712. doi: 10.1007/s10126-005-0004-0

Milbury, C. A., Lee, J. C., Cannone, J. J., Gaffney, P. M., and Gutell, R. R. (2010). Fragmentation of the large subunit ribosomal RNA gene in oyster mitochondrial genomes. *BMC Genomics* 11, 485. doi: 10.1186/1471-2164-11-485

Minh, B. Q., Nguyen, M. A., and von Haeseler, A. (2013). Ultrafast approximation for phylogenetic bootstrap. *Mol. Biol. Evol.* 30, 1188–1195. doi: 10.1093/molbev/mst024

Minh, B. Q., Schmidt, H. A., Chernomor, O., Schrempf, D., Woodhams, M. D., von Haeseler, A., et al. (2020). IQ-TREE 2: new models and efficient methods for phylogenetic inference in the genomic era. *Mol. Biol. Evol.* 37, 1530–1534. doi: 10.1093/molbev/msaa015

Montelli, S., Peruffo, A., Patarnello, T., Cozzi, B., and Negrisolo, E. (2016). Back to water: signature of adaptive evolution in cetacean mitochondrial tRNAs. *PloS One* 11, e0158129. doi: 10.1371/journal.pone.0158129

Naser-Khdour, S., Minh, B. Q., Zhang, W., Stone, E. A., and Lanfear, R. (2019). The prevalence and impact ofmodel violations in phylogenetic analysis. *Genome Biol. Evol.* 11, 3341–3352. doi: 10.1093/gbe/evz193

Negrisolo, E., Minelli, A., and Valle, G. (2004). The mitochondrial genome of the house centipede Scutigera and the monophyly versus paraphyly of Myriapods. *Mol. Biol. Evol.* 21, 770–780. doi: 10.1093/molbev/msh078

Page, R. D. M., and Holmes, E. C. (2009). *Molecular evolution: a phylogenetic approach*. (Oxford: John Wiley & Sons).

Pagenstecher, H. A. (1877). Mollusca. *Zoologische Ergebnisse einer im Aufträge der Königlichen Academie der Wissenschaften zu Berlin ausgeführten Reise die Küstengebiete Des. Rothen Meeres*. In R. Kossmann. 1 (2), 1–66, Leipzig.

Payraudeau, B. C. (1826). *Catalogue descriptif et méthodique des annelides et des mollusques de l'Ile de Corse; avec huit planches représentant quatre-vingt-huit espèces, dont soixante-huit nouvelles* Vol. vii + 218 (Paris: Imprimerie de J. Tastu), 1–8.

Perna, N. T., and Kocher, T. D. (1995). Patterns of nucleotide composition at fourfold degenerate sites of animal mitochondrial genomes. *J. Mol. Evol.* 41, 353–358. doi: 10.1007/BF00186547

Péron, F., and Freycinet, L. (1807). *Voyage de découvertes aux Terres Australes, exécuté par ordre de sa Majesté l'Empereur et Roi,. ... pendant les années 1800. 1801, 1802, 1803 et 1804* (Paris), 496.

Plazzi, F., Puccio, G., and Passamonti, M. (2016). Comparative large-scale mitogenomics evidences clade-specific evolutionary trends in mitochondrial DNAs of Bivalvia. *Genome Biol. Evol.* 8, 2544–2564. doi: 10.1093/gbe/evw187

Quoy, J. R. C., and Gaimard, J. P. (1835). Voyage de la corvette l'Astrolabe: exécuté par ordre du roi, pendant les années 1826-1827-1828-1829, sous le commandement de M. J. Dumont d'Urville. *Zoologie. 1: i-l 1-264; 2(1): 1-321 [1832]; 2(2): 321-686 [1833]; 3(1): 1-366 [1834]; 3(2): 367-954 [1835]; 4 [1833]; Atlas (Mollusques): pls 1-93 [1833] ...etc. In: Dumont d'Urville J.; 1834 Voyage Découvertes l'Astrolabe* (ParisJ. Tastu, Éditeur-Imprimeur).

Ren, J., Hou, Z., Wang, H., Sun, M.-A., Liu, X., Liu, B., et al. (2016). Intraspecific variation in mitogenomes of five *Crassostrea* species provides insight into oyster diversification and speciation. *Marine Biotechnol.* 8, 242–254. doi: 10.1007/s10126-016-9686-8

Ren, J., Liu, X., Jiang, F., Guo, X., and Liu, B. (2010). Unusual conservation of mitochondrial gene order in *Crassostrea* oysters: evidence for recent speciation in Asia. *BMC Evolutionary Biol.* 10, 394. doi: 10.1186/1471-2148-10-394

Ren, J., Liu, X., Zhang, G., Liu, B., and Guo, X. (2009). Tandem duplication-random loss" is not a real feature of oyster mitochondrial genomes. *BMC Genomics* 10, 84. doi: 10.1186/1471-2164-10-84

Reuter, J. S., and Mathews, D. H. (2010). RNAstructure: software for RNA secondary structure prediction and analysis. *BMC Bioinf.* 11, 129. doi: 10.1186/1471-2105-11-129

Salvato, P., Simonato, M., Battisti, A., and Negrisolo, E. (2008). The complete mitochondrial genome of the bag-shelter moth *Ochrogaster lunifer* (Lepidoptera, Notodontidae). *BMC Genomics* 9, 331. doi: 10.1186/1471-2164-9-331

Salvi, D., Macali, A., and Mariottini, P. (2014). Molecular phylogenetics and systematics of the bivalve family Ostreidae based on rRNA sequence-structure models and multilocus species tree. *PloS One* 9, e108696. doi: 10.1371/journal.pone.0108696

Salvi, D., and Mariottini, P. (2017). Molecular taxonomy in 2D: A novel ITS2 rRNA sequence-structure approach guides the description of the oysters' subfamily Saccostreinae and the genus *Magallana* (Bivalvia: Ostreidae). *Zoological J. Linn. Soc.* 179, 263–276. doi: 10.1111/zoj.12455

Salvi, D., and Mariottini, P. (2021). Revision shock in Pacific oysters taxonomy: the genus *Magallana* (formerly *Crassostrea* in part) is well-founded and necessary. *Zoological J. Linn. Soc.* 192, 43–58. doi: 10.1093/zoolinnean/zlaa112

Shimodaira, H. (2002). An approximately unbiased test of phylogenetic tree selection. *Systematic Biol.* 51, 492–508. doi: 10.1080/10635150290069913

Shimodaira, H., and Hasegawa, M. (1999). Multiple comparisons of log-likelihoods with applications to phylogenetic inference. *Mol. Biol. Evol.* 16, 1114–1116. doi: 10.1093/oxfordjournals.molbev.a026201

Simonato, M., Battisti, A., Kerdelhué, C., Burban, C., Lopez-Vaamonde, C., Pivotto, I., et al. (2013). Host and phenology shifts in the evolution of the social moth genus *Thaumetopoea*. *PloS One* 8, e57192. doi: 10.1371/journal.pone.0057192

Sowerby, G. B. II. (1871). Monograph of the genus Ostraea. *Conchologia Iconica illustrations shells molluscous Anim.* (L. Reeve & Co., London) 18, 1–33.

Spencer, H. G., Willan, R. C., Mariottini, P., and Salvi, D. (2022). Taxonomic consistency and nomenclatural rules within oysters: Comment on Li et al. (2021). *Mol. Phylogenet. Evol.* 170, 107437. doi: 10.1016/j.ympev.2022.107437

Strimmer, K., and Rambaut, A. (2002). Inferring confidence sets of possibly misspecified gene trees. *Proc. R. Soc. London. Ser. B: Biol. Sci.* 269, 137–142. doi: 10.1098/rspb.2001.1862

Strimmer, K., and von Haeseler, A. (1996). Quartet puzzling: a quartet maximum-likelihood method for reconstructing tree topologies. *Mol. Biol. Evol.* 13, 964–969. doi: 10.1093/oxfordjournals.molbev.a025664

Thompson, J. D., Higgins, D. G., and Gibson, T. J. (1994). CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* 22, 4673–4680. doi: 10.1093/nar/22.22.4673

Thunberg, C. P. (1793). Tekning och Beskrifning på en stor Ostronsort ifrån Japan. *Kongliga Vetenskaps Academiens Nya Handlingar* 14, 140–142.

Torigoe, K., and Inaba, A. (1981). On the scientific name of Japanese spiny oyster "Kegaki". *Venus* 40 (3), 126–134.

Troost, K. (2010). Causes and effects of a highly successful marine invasion: case-study of the introduced Pacific oyster *Crassostrea gigas* in continental NW European estuaries. *J. Sea Res.* 64, 145–165. doi: 10.1016/j.seares.2010.02.004

Volatiana, J. A., Fang, S., Kinaro, Z. O., and Liu, X. (2016). Complete mitochondrial DNA sequences of *Saccostrea mordax* and *Saccostrea cucullata*: genome organization and phylogeny analysis. *Mitochondrial DNA A* 27, 3024–3025. doi: 10.3109/19401736.2015.1063050

WoRMS Editorial Board (2025). *World Register of Marine Species* (Ostend: VLIZ). Available at: https://www.marinespecies.org (Accessed May 26, 2025).

Wu, X., Li, X., Li, L., Xu, X., Xia, J., and Yun, Z. (2012). New features of Asian *Crassostrea* oyster mitochondrial genomes: a novel alloacceptor tRNA gene recruitment and two novel ORFs. *Gene* 507, 112–118. doi: 10.1016/j.gene.2012.07.032

Wu, X., Xu, X., Yu, Z., Wei, Z., and Xia, J. (2010). Comparison of seven *Crassostrea* mitogenomes and phylogenetic analyses. *Mol. Phylogenet. Evol.* 57, 448–454. doi: 10.1016/j.ympev.2010.05.029

Xiao, S., Wu, X., Li, L., and Yu, Z. (2015). Complete mitochondrial genome of the Olympia oyster *Ostrea lurida* (Bivalvia, Ostreidae). *Mitochondrial DNA* 26, 471–472. doi: 10.3109/19401736.2013.834428

Yu, H., and Li, Q. (2011). Mutation and selection on the wobble nucleotide in tRNA anticodons in marine bivalve mitochondrial genomes. *PloS One* 6, e16147. doi: 10.1371/journal.pone.0016147