Title: Understanding the adaptive evolution of mitochondrial genomes in intertidal chitons

Dipanjana Dhar[1], Debayan Dey[2], Soumalee Basu[3*] and Helena Fortunato[4*]

[1] Graduate School of Science, Department of Natural History Sciences, Hokkaido University, Sapporo, Japan

[2] Graduate School of Life Science, Institute of Low Temperature Science, Hokkaido University, Sapporo, Japan

[3] Department of Microbiology, University of Calcutta, Kolkata, India

[4] Department of Natural History Sciences, Faculty of Science, Hokkaido University, Sapporo, Japan

* Corresponding authors: soumalee@gmail.com; helena@sci.hokudai.ac.jp

Abstract

Mitochondria are the centre of energy metabolism in eukaryotic cells and its genes are thus key to the evolution of molecular mechanisms that metabolize cellular energy. Intertidal zone is one of the most stressful environments with extreme shifts in temperature, salinity, pH and oxygen concentrations. Marine molluscs, particularly chitons belong to the ecologically dominant organisms in this extreme environment, symbolizing an ideal model to understand mitochondrial stress adaptation. Here, we used concatenated mitochondrial genetic components separately from seven chitons of the intertidal zone to reconstruct phylogenetic relationships among these species. We performed selection analyses considering sites and branches of individual protein-coding genes to identify potentially adaptive residues and localize them in the protein structures of mt subunits. Our results exhibited significant amino acid changes in sites under diversifying selection of all the protein-coding genes, indicative of the adaptive evolution of mitochondrial genome in chitons. Furthermore, we obtained sites in the transmembrane helices lining the proton translocation channel as well as in surrounding loop regions, providing implication towards functional modification of the OXPHOS proteins essential for survival in dynamic environment of the intertidal zone.

Keywords: Mitochondrial genome, Polyplacophora, selection studies, molecular phylogeny, intertidal, chitons, adaptive evolution.

Abbreviations:

mtDNA: mitochondrial DNA; OXPHOS: oxidative phosphorylation; PCGs: protein-coding genes; *r*RNA: ribosomal RNA; *t*RNA: transfer RNA; ATP: adenosine triphosphate; NAD: nicotinamide adenine dinucleotide; ND: NADH dehydrogenase; CytB: cytochrome *b*; COX: cytochrome *c* oxidase complex; ATPase: ATP synthase.

Introduction

Animal mitochondrial DNA is generally about 15,000-20,000 base pairs in size. However, much larger mitochondrial genomes are copies of duplicated genes rather than genetic variation in mtDNA **[1, 2]**. Mitochondrial DNA contains 13 protein-coding genes, 2 *r*RNAs of the mitochondrial ribosome and 22 *t*RNAs essential for translating OXPHOS proteins. Furthermore, presence of non-coding region(s) in the mtDNA is probably the site for controlling transcription and replication **[3]**. The 13 mitochondrial PCGs forming subunits of respiratory chain complexes: cytochrome *c* oxidase subunit 1-3 (*COX1*, *COX2*, *COX3*), cytochrome *b* (*CYTB*), NADH dehydrogenase subunit 1-6 (*ND1*, *ND2*, *ND3*, *ND4*, *ND5*, *ND6*), NADH dehydrogenase subunit 4L (*ND4L*), ATPase F0 subunit 6 (*ATP6*) and ATPase F0 subunit 8 (*ATP8*), are responsible for producing up to 95% of the energy of eukaryotic cells **[4]**. Complex I (*ND1*, *ND2*, *ND3*, *ND4*, *ND4L*, *ND5* and *ND6*), III (*CYTB*), IV (*COX1*, *COX2* and *COX3*) and V (*ATP6* and *ATP8*) are four of the five complexes required in the electron transport chain system and generation of ATP, combined with other subunits encoded by the nuclear DNA genome **[5]**.

Mitochondrial stress can have major consequences on the metabolic performance of an organism. Since mitochondrial proteins and mitochondria-derived stress signals regulate both oxidative phosphorylation and innate immune response, maintenance of the mitochondrial integrity and signaling therein are important for cellular homeostasis and survival **[6-11]**. All the events resulting to the impairment of mitochondrial functions have adverse effects on the amino

acid constituency of the OXPHOS proteins. Changes in these amino acid residues do not alter protein structure rather can have an impact on the functional domains, such as regions lining the proton translocation channel and subunit interacting sites, and thereby succour animals to adapt to challenging environments [12, 13]. Therefore, estimation of selection pressures acting on mt proteins could extend deep insight into the adaptive evolution of mitochondrial genome. Most of the researches have been conducted on the adaptive evolution of mitochondrial genome in vertebrates, specially mammals and fish [14, 15]. However, fundamental questions regarding mitogenome evolution of invertebrates, particularly marine invertebrates, remain largely unanswered.

Mollusca is the second largest animal phyla after arthropods consisting of eight classes: Bivalvia, Cephalopoda, Gastropoda, Monoplacophora, Polyplacophora, Scaphopoda, Caudofoveata and Solenogastres, as per the widely accepted "Aculifera hypothesis" [16-18]. Among all the classes, chitons (Polyplacophorans) are the most primitive marine molluscs with fossil record tracing back to the early Cambrian [19, 20]. In addition, this phylum constitutes a ubiquitous, heterogeneous and economically important group of invertebrates that function as ecosystem engineers [21]. Although the most conspicuous symptom of any ecosystem deterioration is the decline or disappearance of sensitive species, molluscs display an unusual resilience towards environmental changes [22]. Chitons, mostly inhabitants of the intertidal zones, are competent enough to survive and maintain mitochondrial homeostasis regardless of regular oscillations of immersion and emersion, extreme alternations of temperature, salinity, pH and hydrodynamic forces in their surrounding environment [23-25]. More detailed studies on mitochondrial genome evolution will allow conducting further analysis on metabolic adaptation of intertidal chitons.

In spite of possessing significant roles as a global food resource and marine bioindicator, only a small percentage of molluscs possess completely sequenced and annotated genomes, rendering the phylum with the lowest ratio between fraction of sequenced genomes and number of described species [26]. At the time of writing this article (December 2019), complete mt genome sequences of only 7 chiton species were available on GenBank [27-30]. We therefore considered seven mt genomes representing species of the order Chitonida from the class Polyplacophora to determine evolutionary selection acting on the 13 mitochondrial protein-coding genes. We

performed site and branch-site specific likelihood analyses to quantify the probability of positive selection on each site and branch of individual genes across all the seven chitons. Thereafter, we assessed changes in amino acids based on their physicochemical properties and mapped the positively selected sites in structure-based homology models to ascertain the effect of these changes in functionally important regions of the OXPHOS proteins. This work also focused on the reconstruction of phylogenetic relationships considering genetic components (whole genome, PCGs, *t*RNAs and *r*RNAs) of the mitochondrial DNA, separately. We also observed interspecific variations in the control region of the mitogenome of these polyplacophoran species, thus enhancing the existing knowledge on the mitogenomics of chitons.

Materials and Methods

Sequence retrieval

Complete mitochondrial genome sequences of seven Polyplacophoran species were downloaded from the National Center for Biotechnology Information (NCBI) database (https://www.ncbi.nlm.nih.gov/nucleotide/) **[31]**. The nucleotide and amino acid sequences of 13 mitochondrial PCGs for the same were also retrieved from the above mentioned database. Additionally, sequences pertaining to 2 *r*RNAs, 22 *t*RNAs and the non-coding control region were separately accumulated from the mitogenome of chitons. The details of the species considered for this study are provided in **Supp. Table 1**.

Sequence alignment & phylogenetic analysis

Four nucleotide sequence datasets were constructed with the mitogenomes obtained from 7 chiton species – (A) 13 mitochondrial protein-coding genes (*COX1*, *COX2*, *COX3*, *CYTB*, *ATP6*, *ATP8*, *ND1*, *ND2*, *ND3*, *ND4*, *ND4L*, *ND5* and *ND6*); (B) 22 *t*RNAs; (C) 2 *r*RNAs and (D) highly variable control region. A species from the class Gastropoda (*Conus textile*, accession ID: DQ862058.1) was chosen as outgroup. For each of the aforementioned datasets, codon-based (for PCGs) and multiple sequence alignments were performed using MUSCLE **[32]** integrated in the MEGA-X software **[33]**. Besides, the entire mitogenomes were aligned in the Geneious

Prime Version 2020.0.2 software (trial version), available at http://www.geneious.com. The aligned coding sequences for mitochondrial PCGs, *t*RNAs and *r*RNAs were then separately concatenated in MEGA-X, with default settings. The ambiguously aligned regions for individual alignments were removed using Gblocks server v0.91b **[34]**, prior to concatenation.

The appropriate nucleotide substitution models for each dataset were determined in jModelTest2 (v2.1.10), based on different information criterion calculations, starting with 11 substitution schemes and fixed BIONJ-JC option base tree for likelihood calculations **[35]**. We selected models according to Akaike Information Criterion (AIC) and corrected Akaike Information Criterion (cAIC). The best fit models obtained for whole mitogenome, PCGs, *t*RNAs, *r*RNAs and control region are TIM1+G, TVM+I+G, GTR+G, TVM+I+G and GTR respectively. Since most of the models obtained could not be implemented in MrBayes 3.2.6, they were substituted with General Time Reversible (GTR) model along with a proportion of invariable sites (I) and heterogeneity of substitution rates among sites modeled following a gamma distribution (G), wherever applicable **[36, 37]**. In addition, substitution saturation was calculated on each codon position of our dataset by Xia *et al.* statistical test **[38]**, executed in DAMBE software v7 **[39]**. Subsequently, the maximum log likelihood values for substitution matrix and transition/transversion rate ratio were calculated for the coding and non-coding regions of the mitogenome in seven chitons using MEGA-X.

Phylogenetic analyses, for each dataset, were performed using Bayesian inference with the aforementioned nucleotide substitution models. Phylogenetic trees were contrived using MrBayes v3.2.6 plug-in in the Geneious Prime Version 2020.0.2 software (trial version), via the CIPRES REST API **[40, 41]**. A Markov Chain Monte Carlo (MCMC) search with 5 million generations was performed, logging results every 500th generation. The first 25% of the trees were discarded as burn-in. A consensus tree and the Bayesian Posterior Probabilities (BPP) were estimated based on the remaining trees. All the derived phylogenies generated were visualized using Figtree v1.4.3 (http://tree.bio.ed.ac.uk/software/figtree/) and iTOL v5 **[42]**.

Selection Pressure Analyses

Estimation of ω *i.e.* the ratio of non-synonymous substitutions per non-synonymous sites (d_N) to synonymous substitutions per synonymous sites (d_S) indicate natural selection on protein coding genes. Under neutrality, the ratio is not expected to deviate significantly from 1 ($\omega = 1$), whereas notable changes in the ratio may be interpreted in case of positive selection ($\omega > 1$) or negative selection ($\omega < 1$). The ω value was determined using the codon-based maximum likelihood (CODEML) algorithm implemented in EasyCodeML v1.21 **[43, 44]**. Both site and branch-site models were used to identify the variation of selective pressures on individual mitochondrial PCGs of the chitons. Seven codon substitution models described as M0 (one-ratio), M1a (nearly neutral), M2a (positive selection), M3 (discrete), M7 (beta), M8 (beta and $\omega > 1$) and M8a (beta and $\omega = 1$) were investigated and four likelihood ratio tests (M0 vs. M3, M1a vs. M2a, M7 vs. M8 and M8a vs. M8) were performed to measure the likelihood of positive selection on each site for individual PCGs **[45-50]**. On the other hand, branch-site models enable ω to differ among sites as well as across branches (foreground-lineages) of the phylogenetic tree **[51]**. Thus, we tested for positive selection on each branch of the phylogeny, considering a single foreground branch at a time. Bayes Empirical Bayes analysis was used to compute posterior probabilities in order to recognize sites under positive selection on the selected branches if the LRT *p*-values are significantly correct.

Additionally, we used the HyPhy package v2.5 (available from www.hyphy.org) to assess codons under selection pressure **[52]**. MEME (Mixed Effects Model of Evolution) and FEL (Fixed Effects Likelihood) analyses were carried out to detect individual sites subjected to episodic diversifying selection and purifying selection, respectively **[53, 54]**. Both the analyses were inferred at *p*-value<0.05 significance level. Furthermore, the impact of potential positive selection was detected using TREESAAP software v3.2, which detects changes in local physicochemical properties due to amino acid replacements **[55, 56]**. A sliding window equal to 20 codons alongwith magnitude categories ≥ 6 and *p*-values < 0.001 (z-score > 3.09) significance level were considered for accurate assessment of amino acid property changes.

Protein modelling and structural analysis

In order to understand the influence of sites undergoing positive selection or exhibiting physicochemical property changes, we mapped them onto the three-dimensional structure of mitochondrial proteins. Owing to the unavailability of 3D structures of mt-subunits for molluscs in Protein Data Bank (PDB) **[57]**, we modeled the same for the Eastern beaded chiton *C.apiculata*. This species was selected among the seven chitons due to its ancestral mt genome architecture that presumably indicate the plesiomorphic condition for the phylum Mollusca. The structures of *COX* subunits and *CYTB* protein were modeled following significant similarities with templates sequences (PDB ID: 3ABM and 3H1I.1.C, respectively) using the SWISS-MODEL server **[58]**. The models for remaining mitochondrial proteins were generated in I-TASSER server, using default parameters **[59, 60]**. The overall structure of mitochondrial complex I and IV for *C.apiculata* were generated by structurally aligning the component subunits to corresponding complex I and IV templates from *Ovis aries* (PDB ID: 5LNK) and *Bos taurus* (PDB ID: 3ABM) respectively, using PyMOL software **[61]**. Additional information about the orientation of transmembrane domains of these proteins was obtained using the TMHMM server v2.0 **[62]**. PyMOL was further used for graphic modifications, visualization and depicting final illustrations.

Results & Discussions

Phylogenetic analyses & diversity

Based on the concatenated sequences of 13 mitochondrial PCGs, phylogenetic relationship among the polyplacophorans **(Fig. 1)** predominantly resembled the tree obtained by Guerra *et al*, 2018. The phylogenetic trees revealed family-wise clading of the chiton species with strong nodal support where families Mopallidae and Lepidochitonidae emerged sister to Chaetopleuridae and Chitonidae. However, the phylogeny derived from the whole mitogenomes of 7 species displayed contrasting topology **(Fig. 2)**, most probably owing to the alternative gene arrangement in *Sypharochiton sp*. This observation appeared highly congruent with the fact that family-specific rearrangements took place in Chitonidae due to inversion of $tRNA_{Val}$ and

tRNA$_{Trp}$ genes. Besides, other members exhibited alterations in their tRNA arrangement as follows: (1) inversion of tRNA cluster (tRNA$_{Met}$, tRNA$_{Cys}$, tRNA$_{Tyr}$, tRNA$_{Trp}$, tRNA$_{Gln}$, tRNA$_{Gly}$ and tRNA$_{Glu}$) in *C. caverna* and *N. californica*; (2) reversal of tRNA$_{His}$ in *C. stelleri* and (3) transposition of tRNA$_{Asp}$ particular to *K. tunicata*. It is worth mentioning that the clustering pattern of chitons in the phylogeny contrived using whole mitochondrial genome is reflective of the arrangement of their protein-coding genes and tRNAs **(Fig. 3)**.

Even though the entire mt-DNA remains linked and gets inherited as a single entity in the absence of recombination, coding and non-coding segments therein demonstrate differential substitution rates with higher selective forces acting on the functionally conserved protein-coding genes **[63]**. Therefore, we compared individual mitochondrial genetic components under varying selection pressure through phylogenetic clustering. Phylogeny inferred separately from concatenated nucleotide sequences of 22 tRNAs and 2 rRNAs **(Supp. Fig. 1a & 1b)** demonstrated identical clading pattern with the one contrived using PCGs. It can be stated that when all the tRNAs were taken into account, the phylogenetic relationship among the species was in accordance with other coding portions, regardless of their different arrangements in the mitogenome.

Similarly, we analyzed the highly polymorphic non-coding sequence (also called the control region) of the mitochondrial genome. This region is present between tRNA$_{Glu}$ and *COX3* in all the chitons studied here, excepting two members in the Lepidochitonidae family where the region is shifted between tRNA$_{Glu}$ and *12S rRNA*. Upon multiple sequence alignment **(Fig. 4)** we observed high sequence variability among these AT-rich regions, rendering it unsuitable for inferring spatial relationships from the control region alone. Thus, in spite of the utility of this mitogenomic part as a genetic marker for phylogenetic analysis in several metazoans **[64-66]**, the same is not applicable for polyplacophoran species considered here. Furthermore, we inferred incidences of nucleotide substitutions in the protein-coding genes and control region of chiton mitogenomes. Calculating transition/transversion bias (R) is important for deciphering genome evolution since it allows verifying the presence of bias in nucleotide conversions. In our datasets, transitions were more favoured over transversions **(Fig. 5)**. The maximum likelihood estimate of substitution matrix is depicted in **Supp. Table. 2**. Moreover, our results revealed the presence of

more transitions leading to synonymous substitutions in protein-coding genes so as to maintain functional constancy in the three-dimensional structures **[67]**. In case of the non-coding control region, though transversions seemed to be prevalent, high transitions were retained to ensure proper regulatory function.

Selection studies

Purifying selection is the prevalent force in the evolution of mitogenomes. Since mitochondria serves as the major site for energy metabolism and production in the cell, weak and/or episodic positive selection may act on its protein-coding genes allowing adaptation to changes in the surrounding environment **[68, 69]**. Marine molluscs constitute ecologically important organisms in the intertidal zone and have successfully acclimatized to a dynamic environment between the tides. Considering the dynamic habitat of intertidal zone which may influence the function of mitochondrial genes, we examined signatures of positive selection in the polyplacophoran species using the CODEML algorithm implemented in EasyCodeML.

The heterogeneous selective pressure among sites in each of the 13 mitochondrial PCGs was tested, using different site-specific models: M0 (one-ratio), M1a (nearly neutral), M2a (positive selection), M3 (discrete), M7 (beta), M8 (beta and $\omega > 1$) and M8a (beta and $\omega = 1$). According to the results obtained, purifying selection dominated the evolution of all mitogenes except for *ATP8* (**Supp. Table 3**). In other words, likelihood ratio test of M7 *vs*. M8 found an evidence of positive selection for only *ATP8* at codon position of 44 (LRT *p*-value = 0.000204865 and ω = 36.08464). The outcomes of comparing different site models (M0 *vs*.M3, M1a *vs*. M2a, M7 *vs*. M8, and M8a *vs*. M8) revealed the global molecular evolutionary rates of mitochondrial PCGs in the polyplacophorans under negative constraints. Furthermore, FEL analyses revealed that purifying selection was prevalent on all mitochondrial PCGs, with few sites evolving under neutrality ($p<0.05$). According to the FEL results, complex IV subunits (*COX1*, *COX2* and *COX3*) exhibited the highest percentage of codons under negative selection (**Fig. 6**). *CYTB* from complex III also displayed high levels of purifying selection whereas members of complex I and V showed comparatively more relaxed selection. Additionally, we identified nine positively selected sites in six genes (*COX1*, *CYTB*, *ND2*, *ND3*, *ND4* and *ND5*) using MEME ($p<0.05$)

which employs a mixed-effects maximum likelihood approach to identify individual sites subjected to episodic diversifying selection **(Supp. Table 4)**.

Since positive selection occurs at only few positions in a particular lineage for most genes, detection of natural selection becomes limited based on branch and site models separately. Therefore, we applied the more stringent branch-site model, which allows ω value to vary between branches and sites simultaneously, to measure divergent selective pressure on the specified foreground branch against the remaining background lineages. Using the Bayes empirical Bayes method, we detected 69 sites under potential positive selection in 12 PCGs with posterior probabilities ≥ 95% **(Fig. 7; Supp. Table 5)**. From the results, it is evident that complex I and V subunits possessed higher number of positively selected sites per branch compared to complex IV members.

Alterations of amino acid properties

The significance of ω>1 preferably indicates that positive selection would have been the driving force behind adaptive evolution of protein-coding DNA sequence. This presumption, ideally, has been found to be mostly conventional since PCGs will almost have more synonymous than nonsynonymous substitutions invariably, even if several sites have undergone positive selection pressure. Therefore, considerable changes in quantitative amino acid properties among protein residues in a phylogenetic tree were identified using the TreeSAAP software package, which compares random distributions of amino acid changes based on variable physiochemical properties with expected distributions of changes deviating from neutral conditions.

We identified several codon sites (at *p*<0.001 and category 6-8) exhibiting deviations in amino acid properties within 13 mitochondrial PCGs from seven chitons. Among these, most of the amino acid substitutions were restricted to changes in the ionization properties (*pK'*), except for *ATP8* and *ND6*. Changes in solvent accessible reduction ratio (R_a), buriedness (B_r), power to be at the C-terminal (α_c) and α-helical tendencies (P_a) were observed for *ATP6*, *COX1*, *ND1*, *ND2*, *ND4*, *ND5*; *ATP8*, *ND4*, *ND5*; *COX2*, *ND4*, *ND6* and *COX3*, *ND6*, respectively. Furthermore, *ND3* and *ND4* exhibited non-synonymous substitutions at few of their codons corresponding to

changes in power to be at the middle of α-helix ($α_m$) and surrounding hydrophobicity (H_p), respectively. At the biochemical level, most of the amino acid changes in these genes affected the ionization properties while few of the substitutions influenced the hydrophobicity or hydrophilicity of amino acids. Therefore, changes in these functional properties might disrupt protein's stability, solubility and structural organization. The alterations resulting from biochemical changes were analyzed by mapping these residues in the tertiary structures of mitochondrial protein complexes.

Mapping of positively selected sites in mitochondrial protein structures

The putative positively selected sites, identified by CODEML and MEME, were mapped on the corresponding mitochondrial subunits, to determine their positions and significance in protein structure and function. Additionally, sites displaying changes in physicochemical properties of amino acids inferred by TREESAAP were marked **(Fig. 8a-e)**. We delineated the functional domains of individual mitochondrial subunits using TMHMM and found that the majority of sites were localized within or in the vicinity of transmembrane region. It is worth mentioning that the subunits of complex I possess more evidence of positive selection in the proton translocation channel compared to other complexes **(Fig. 9)**. There is also a clear distinction in the distribution of sites undergoing positive selection and/or radical amino acid changes between loop regions and transmembrane helices for all mitochondrial subunits.

Complex I (proton-pumping NADH: ubiquinone oxidoreductase), the largest and most complicated enzyme set of the OXPHOS pathway, is responsible for NADH oxidation, ubiquinone reduction and proton pumping leading to ATP synthesis **[70]**. ND2, ND4 and ND5 are considered to be the pivotal members for maintaining proton pump due to their sequence homology with a class of Na^+/H^+ antiporters **[71]**. Non-synonymous substitutions in these subunits may interrupt efficacy of the proton pump. These could happen through biochemical changes that enhance/impair the proton translocation mechanism or improve/interrupt the redox linked conformational changes **[72]**. As described by de Fonseca *et al*. 2008, the high number of sites with radical amino acid changes observed in ND proteins can be attributed to their overall high mutation rate.

CYTB, a conserved protein constituting complex III of the electron transport chain, plays a fundamental role in energy production by catalyzing reversible electron transfer from ubiquinol to cytochrome *c* coupled to proton translocation **[73]**. We observed instances of positive selection in the functionally significant domains of *CYTB*, which may be correlated with the adaptation of chitons to the fluctuating environment of the intertidal zone. In concordance with previous literatures **[74, 75]**, we got substantial evidence of positive selection in the *COX* subunits (*COX1*, *COX2* and *COX3*) which is indicative of adaptive tolerance in intertidal organisms to intermittent hypoxia. Furthermore, sites under diversifying selection were identified in members of ATP synthase, the final enzyme complex of the respiratory chain that is coupled with the electrochemical gradient across the inner mitochondrial membrane to drive ATP production **[76-78]**. Therefore, the mutations in *ATPase* genes provide implications on adaptive evolution of the stress-tolerant intertidal chitons.

Intertidal molluscs display elevated sensitivity to environmental stressors like thermal stress, changing ion concentration, fluctuation in osmolarity and reduced oxygen availability. Besides, these marine invertebrates exhibit striking resilience to hydrogen sulphide (H_2S) toxicity, a potent mitochondrial poison that disrupts mitochondrial respiratory chain by inhibiting cytochrome *c* oxidase **[79, 80]**. At this juncture, survival of these organisms relies on a modified and adaptive energy metabolism. Our results revealed potential signatures of positive selection on mitochondrial PCGs of chitons. The study provides valuable insights into the molecular mechanisms underlying the adaptive strategies of these polyplacophorans to the intertidal habitat from a mitochondrial perspective.

Conclusion

Our study represents the first thorough investigation of evolutionary selection acting on the mitochondrial protein-coding genes of Polyplacophorans. Signatures of both positive and negative selection identified, may provide significant implications towards the adaptive evolution of mitogenomes in chitons inhabiting the intertidal zone. Additionally, mapping the positively selected sites in tertiary protein structures of OXPHOS subunits revealed their

presence in functional domains, predominantly in the proton translocation channel. In conclusion, our findings provide compelling evidence of positive selection in the mitochondrial PCGs, suggestive of their role in metabolic adaptations to contrasting environments.

References

1. Boyce TM, Zwick ME & Aquadro CF (1989). Mitochondrial DNA in the bark weevils: Size, structure and heteroplasmy. *Genetics*, 123: 825–836.

2. Fuller KM & Zouros E (1993). Dispersed length polymorphism of mitochondrial DNA in the scallop *Placopecten magellanicus* (Gmelin). *Current Genetics*, 23: 365–369.

3. Boore JL (1999). Animal mitochondrial genomes. *Nucleic Acids Research*, 27(8): 1767–1780.

4. Blier PU, Dufresne F & Burton RS (2001). Natural selection and the evolution of mtDNA-encoded peptides: evidence for intergenomic co-adaptation. *Trends in genetics: TIG*, 17(7): 400–6.

5. Saraste M (1999). Oxidative Phosphorylation at the fin de siècle. *Science*, 283(5407): 1488-1493.

6. Sokolova I (2018). Mitochondrial Adaptations to Variable Environments and Their Role in Animals' Stress Tolerance. *Integrative and Comparative Biology*, 58(3): 519–531.

7. Lartigue L & Faustin B (2013). Mitochondria: Metabolic regulators of innate immune responses to pathogens and cell stress. *International Journal of Biochemistry & Cell Biology*, 45(9): 2052-2056.

8. Monlun M, Hyernard C, Blanco P, Lartigue L & Faustin B (2017). Mitochondria as Molecular Platforms Integrating Multiple Innate Immune Signalings. *Journal of Molecular Biology*, 429(1): 1-13.

9. Bohovych I & Khalimonchuk O (2016). Sending Out an SOS: Mitochondria as a Signaling Hub. *Frontiers in Cell and Developmental Biology*, 4: 109.

10. Lane N (2005). Power, sex, suicide: Mitochondria and the meaning of life. New York, New York, USA: *Oxford University Press*. p. 368.

11. Vakifahmetoglu-Norberg H, Ouchida AT & Norberg E (2017). The role of mitochondria in metabolism and cell death. *Biochemical and Biophysical Research Communications*, 482(3): 426-431.

12. da Fonseca RR, Johnson WE, O'Brien SJ, Ramos MJ & Antunes A (2008). The adaptive evolution of the mammalian mitochondrial genome. *BMC genomics*, 9: 119.

13. Almeida D, Maldonado E, Vasconcelos V & Antunes A (2015). Adaptation of the Mitochondrial Genome in Cephalopods: Enhancing Proton Translocation Channels and the Subunit Interactions. *PLOS One*, 10(8): e0135405.

14. Foote AD, Morin PA, Durban JW, Pitman RL, Wade P & Willerslev E (2011). Positive selection on the killer whale mitogenome. *Biology letters*, 7(1): 116–8.

15. Garvin MR, Bielawski JP & Gharrett AJ (2011). Positive Darwinian Selection in the Piston That Powers Proton Pumps in Complex I of the Mitochondria of Pacific Salmon. *PLOS One*, 6(9).

16. Ponder WF & Lindberg DR (2008). Molluscan Evolution and Phylogeny: An Introduction. *In Phylogeny and Evolution of the Mollusca. University of California Press, Berkeley and Los Angeles, California*: 1-18.

17. Kocot KM, Cannon JT, Todt C, Citarella MR, Kohn AB, Meyer A, Santos SR, Schander C, Moroz LL, Lieb B & Halanych KM (2011). Phylogenomics reveals deep molluscan relationships. *Nature*, 477(7365): 452-6.

18. Vinther J, Jell P, Kampouris G, Carney R, Racicot RA & Briggs DEG (2012). The origin of multiplacophorans-convergent evolution in Aculiferan Molluscs. *Palaeontology*, 55(5): 1007-1019.

19. Vendrasco MJ, Wood TE & Runnegar BN (2004). Articulated Palaeozoic fossil with 17 plates greatly expands disparity of early chitons. *Nature*, 429: 288-91.

20. Sigwart JD & Sutton MD (2007). Deep molluscan phylogeny: synthesis of palaeontological and neontological data. *Proceedings. Biological Sciences*, 274(1624): 2413–2419.

21. Astorga MP (2014). Genetic considerations for mollusk production in aquaculture: current state of knowledge. *Frontiers in Genetics*, 5: 435.

22. Parker LM, Ross PM, O'Connor WA, Pörtner HO, Scanes E & Wright JM (2013). Predicting the response of molluscs to the impact of ocean acidification. *Biology*, 2(2): 651–692.

23. Helmuth B, Harley CDG, Halpin PM, O'Donnell M, Hofmann GE & Blanchette CA (2002). Climate Change and Latitudinal Patterns of Intertidal Thermal Stress. *Science*, 298(5595): 1015-1017.

24. Jensen MM & Denny MW (2016). Life in an extreme environment: Characterizing wave-imposed forces in the rocky intertidal zone using high temporal resolution hydrodynamic measurements. *Limnology and Oceanography*, 61(5): 1750-1761.

25. Richards JG (2011). Physiological, behavioral and biochemical adaptations of intertidal fishes to hypoxia. *Journal of Experimental Biology*, 214(2): 191-9.

26. Schell T, Feldmeyer B, Schmidt H, Greshake B, Tills O, Truebano M, Rundle SD, Paule J, Ebersberger I & Pfenninger M (2017). An annotated draft genome for *Radix auricularia* (Gastropoda, Mollusca). *Genome biology and evolution*, *9*(3): 585-592.

27. Boore JL & Brown WM (1994). Complete DNA sequence of the mitochondrial genome of the black chiton, *Katharina tunicata*. *Genetics*, 138: 423-443.

28. Irisarri I, Eernisse DJ & Zardoya R (2014). Molecular phylogeny of Acanthochitonina (Mollusca: Polyplacophora: Chitonida): three new mitochondrial genomes, rearranged gene orders and systematics. *Journal of Natural History*, 48: 2825-2853.

29. Veale AJ, Williams L, Tsai P, Thakur V & Lavery S (2016). The complete mitochondrial genomes of two chiton species (*Sypharochiton pelliserpentis* and *Sypharochiton sinclairi*) obtained using Illumina next generation sequencing. *Mitochondrial DNA*, 27: 537-538.

30. Guerra D, Bouvet K & Breton S (2018). Mitochondrial gene order evolution in Mollusca: Inference of the ancestral state from the mtDNA of *Chaetopleura apiculata* (Polyplacophora, Chaetopleuridae). *Molecular Phylogenetics and Evolution*, 120: 233-239.

31. National Center for Biotechnology Information (NCBI) [Internet]. Bethesda (MD): National Library of Medicine (US), National Center for Biotechnology Information; [1988] - [cited 2020 Feb 11]. Available from: https://www.ncbi.nlm.nih.gov/

32. Edgar RC (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic acids research*, 32(5): 1792-1797.

33. Kumar S, Stecher G, Li M, Knyaz C & Tamura K (2018). MEGA X: Molecular Evolutionary Genetics Analysis across Computing Platforms. *Molecular biology and evolution*, 35(6): 1547-1549.

34. Talavera G & Castresana J (2007). Improvement of Phylogenies after Removing Divergent and Ambiguously Aligned Blocks from Protein Sequence Alignments. *Systematic Biology*, 56(4): 564-577.

35. Darriba D, Taboada GL, Doallo R, & Posada D (2012). jModelTest 2: more models, new heuristics and parallel computing. *Nature methods*, 9(8): 772.

36. Lecocq T, Vereecken NJ, Michez D, Dellicour S, Lhomme P, Valterová I, Rasplus JY & Rasmont P (2013). Patterns of Genetic and Reproductive Traits Differentiation in Mainland vs. Corsican Populations of Bumblebees. *PLOS One*, 8: e65642.

37. Arenas M (2015). Trends in substitution models of molecular evolution. *Frontiers in Genetics*, 6: 319.

38. Xia X, Xie Z, Salemi M, Chen L & Wang Y (2003). An index of substitution saturation and its application. *Molecular phylogenetics and evolution*, 26(1): 1-7.

39. Xia X (2018). DAMBE7: New and Improved Tools for Data Analysis in Molecular Biology and Evolution. *Molecular Biology and Evolution*, 35(6): 1550-1552.

40. Ronquist F, Teslenko M, van der Mark P, Ayres DL, Darling A, Höhna S, Larget B, Liu L, Suchard MA & Huelsenbeck JP (2012) MrBayes 3.2: Efficient Bayesian Phylogenetic Inference and Model Choice Across a Large Model Space. *Systematic Biology*, 61: 539-542.

41. Miller MA, Schwartz T, Pickett BE, He S, Klem EB, Scheuermann RH, Passarotti M, Kaufman S, & O'Leary MA (2015). A RESTful API for Access to Phylogenetic Tools via the CIPRES Science Gateway. *Evolutionary bioinformatics online*, 11: 43-48.

42. Letunic I & Bork P (2019) Interactive Tree Of Life (iTOL) v4: recent updates and new developments. *Nucleic Acids Research*, 47(W1): W256-W259.

43. Yang Z (2007). PAML 4: Phylogenetic Analysis by Maximum Likelihood. *Molecular Biology and Evolution*, 24(8): 1586-1591.

44. Gao F, Chen C, Arab DA, Du Z, He Y & Ho SYW (2019). EasyCodeML: A visual tool for analysis of selection using CodeML. *Ecology and Evolution*, 9: 3891-3898.

45. Yang Z, Nielsen R, Goldman N & Pedersen AM (2000). Codon-substitution models for heterogeneous selection pressure at amino acid sites. *Genetics*, 155(1): 431-49.

46. Anisimova M, Bielawski JP & Yang Z (2001). Accuracy and power of the likelihood ratio test in detecting adaptive molecular evolution. *Molecular biology and evolution*, 18(8): 1585-92.

47. Suzuki Y & Nei M (2001). Reliabilities of parsimony-based and likelihood-based methods for detecting positive selection at single amino acid sites. *Molecular biology and evolution*, 18(12): 2179-85.

48. Wong WS, Yang Z, Goldman N & Nielsen R (2004). Accuracy and power of statistical methods for detecting adaptive evolution in protein coding sequences and for identifying positively selected sites. *Genetics*, 168(2): 1041-51.

49. Yang Z, Wong WS & Nielsen R (2005). Bayes empirical bayes inference of amino acid sites under positive selection. *Molecular biology and evolution*, 22(4): 1107-18.

50. Swanson WJ, Nielsen R & Yang Q (2003). Pervasive adaptive evolution in mammalian fertilization proteins. *Molecular biology and evolution*. 20(1): 18-20.

51. Zhang J, Nielsen R & Yang Z (2005). Evaluation of an improved branch-site likelihood method for detecting positive selection at the molecular level. *Molecular biology and evolution*, 22(12): 2472-9.

52. Kosakovsky Pond SL, Poon AFY, Velazquez R, Weaver S, Hepler NL, Murrell B, Shank SD, Magalis BR, Bouvier D, Nekrutenko A, Wisotsky S, Spielman SJ, Frost SDW & Muse SV (2020). HyPhy 2.5—A Customizable Platform for Evolutionary Hypothesis Testing Using Phylogenies. *Molecular Biology and Evolution*, 37(1): 295-299.

53. Murrell B, Wertheim JO, Moola S, Weighill T, Scheffler K & Kosakovsky Pond SL (2012). Detecting individual sites subject to episodic diversifying selection. *PLoS genetics*, 8(7): e1002764.

54. Pond SL & Frost SD (2005). Not so different after all: a comparison of methods for detecting amino acid sites under selection. *Molecular Biology and Evolution*, 22(5): 1208-22.

55. Woolley S, Johnson J, Smith MJ, Crandall KA & McClellan DA (2003). TreeSAAP: selection on amino acid properties using phylogenetic trees. *Bioinformatics*, 19(5): 671-2.

56. McClellan DA & Ellison DD (2010). Assessing and improving the accuracy of detecting protein adaptation with the TreeSAAP analytical software. *International Journal of Bioinformatics Research and Applications*, 6(2): 120-133.

57. Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, Shindyalov IN & Bourne PE (2000). The Protein Data Bank. *Nucleic acids research*, 28(1): 235-42.

58. Waterhouse A, Bertoni M, Bienert S, Studer G, Tauriello G, Gumienny R, Heer FT, de Beer TAP, Rempfer C, Bordoli L, Lepore R & Schwede T (2018). SWISS-MODEL: homology modelling of protein structures and complexes. *Nucleic acids research*, 46: W296-W303.

59. Yang J, Yan R, Roy A, Xu D, Poisson J & Zhang Y (2015). The I-TASSER Suite: Protein structure and function prediction. *Nature Methods*, 12: 7-8.

60. Zhang Y (2008). I-TASSER server for protein 3D structure prediction. *BMC Bioinformatics*, 9: 40.

61. The PyMOL Molecular Graphics System, Version 2.0 Schrödinger, LLC.

62. Krogh A, Larsson B, von Heijne G & Sonnhammer EL (2001). Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *Journal of molecular biology*, 305(3): 567-80.

63. Ballard JWO & Whitlock MC (2004). The incomplete natural history of mitochondria. *Molecular Ecology*, 13: 729-744.

64. Diniz FM, Maclean N, Ogawa M, Cintra IHA & Bentzen P (2005). The hypervariable domain of the mitochondrial control region in Atlantic spiny lobsters and its potential as a marker for investigating phylogeographic structuring. *Marine Biotechnology*, 7(5): 462-73.

65. Shao R, Barker SC, Mitani H, Aoki Y & Fukunaga M (2005). Evolution of duplicate control regions in the mitochondrial genomes of Metazoa: a case study with Australasian *Ixodes* ticks. *Molecular Biology and Evolution*, 22(3): 620-9.

66. Rutledge LY, Patterson BR & White BN (2010). Analysis of *Canis* mitochondrial DNA demonstrates high concordance between the control region and ATPase genes. *BMC Evolutionary Biology*, 10: 215.

67. Keller I, Bensasson D & Nichols RA (2007). Transition-Transversion Bias is not universal: a counter example from Grasshopper Pseudogenes. *PLoS Genetics*, 3(2): e22.

68. Shen YY, Liang L, Zhu ZH, Zhou WP, Irwin DM & Zhang YP (2010). Adaptive evolution of energy metabolism genes and the origin of flight in bats. *Proceedings of the National Academy of Sciences of the United States of America*, 107: 8666-8671.

69. Tomasco IH & Lessa EP (2011). The evolution of mitochondrial genomes in subterranean caviomorph rodents: adaptation against a background of purifying selection. *Molecular Phylogenetics and Evolution*, 61: 64-70.

70. Wirth C, Brandt U, Hunte C & Zickermann V (2016). Structure and function of mitochondrial complex I. *Biochimica et Biophysica Acta*, 1857: 902-914.

71. Mathiesen C & Hägerhäll C (2002). Transmembrane topology of the nuol, m and n subunits of nadh: quinone oxidoreductase and their homologues among membrane-bound hydrogenases and bona fide antiporters. *Biochimica et Biophysica Acta*, 1556: 121-132.

72. Brandt U (2006). Energy converting NADH: quinone oxidoreductase (complex I). *Annual Review of Biochemistry*, 75: 69-92.

73. Trumpower BL (1990). The protonmotive Q cycle: energy transduction by coupling of proton translocation to electron transfer by the cytochrome bc1 complex. *Journal of Biological Chemistry*, 265: 11409-11412.

74. Luo Y, Gao W, Gao Y, Tang S, Huang Q, Tan X, Chen J & Huang T (2008). Mitochondrial genome analysis of *Ochotona curzoniae* and implication of cytochrome *c* oxidase in hypoxic adaptation. *Mitochondrion*, 8: 352-357.

75. Sun S, Hui M, Wang M & Sha Z (2018). The complete mitochondrial genome of the alvinocaridid shrimp *Shinkaicaris leurokolos* (Decapoda, Caridea): Insight into the mitochondrial genetic basis of deep-sea hydrothermal vent adaptation in the shrimp. *Comparative Biochemistry and Physiology - Part D*, 25: 42-52.

76. Zhou T, Shen X, Irwin DM, Shen Y & Zhang Y (2014). Mitogenomic analyses propose positive selection in mitochondrial genes for high-altitude adaptation in galliform birds. *Mitochondrion*, 18: 70-75.

77. Hassanin A, Ropiquet A, Couloux A & Cruaud C (2009). Evolution of the mitochondrial genome in mammals living at high altitude: new insights from a study of the tribe Caprini (Bovidae, Antilopinae). *Journal of Molecular Evolution*, 68: 293-310.

78. Zhang B, Zhang YH, Wang X, Zhang HX & Lin Q (2017). The mitochondrial genome of a sea anemone *Bolocera sp.* exhibits novel genetic structures potentially involved in adaptation to the deep-sea environment. Ecology and Evolution, 7: 4951-4962.

79. Cooper CE & Brown GC (2008). The inhibition of mitochondrial cytochrome oxidase by the gases carbon monoxide, nitric oxide, hydrogen cyanide and hydrogen sulfide: chemical mechanism and physiological significance. *Journal of Bioenergetics and Biomembranes*, 40: 533-539.

80. Kraus DW & Doeller JE (2004). Sulfide consumption by mussel gill mitochondria is not strictly tied to oxygen reduction: measurements using a novel polarographic sulfide sensor. *Journal of Experimental Biology*, 207(21): 3667-3679.

Figures
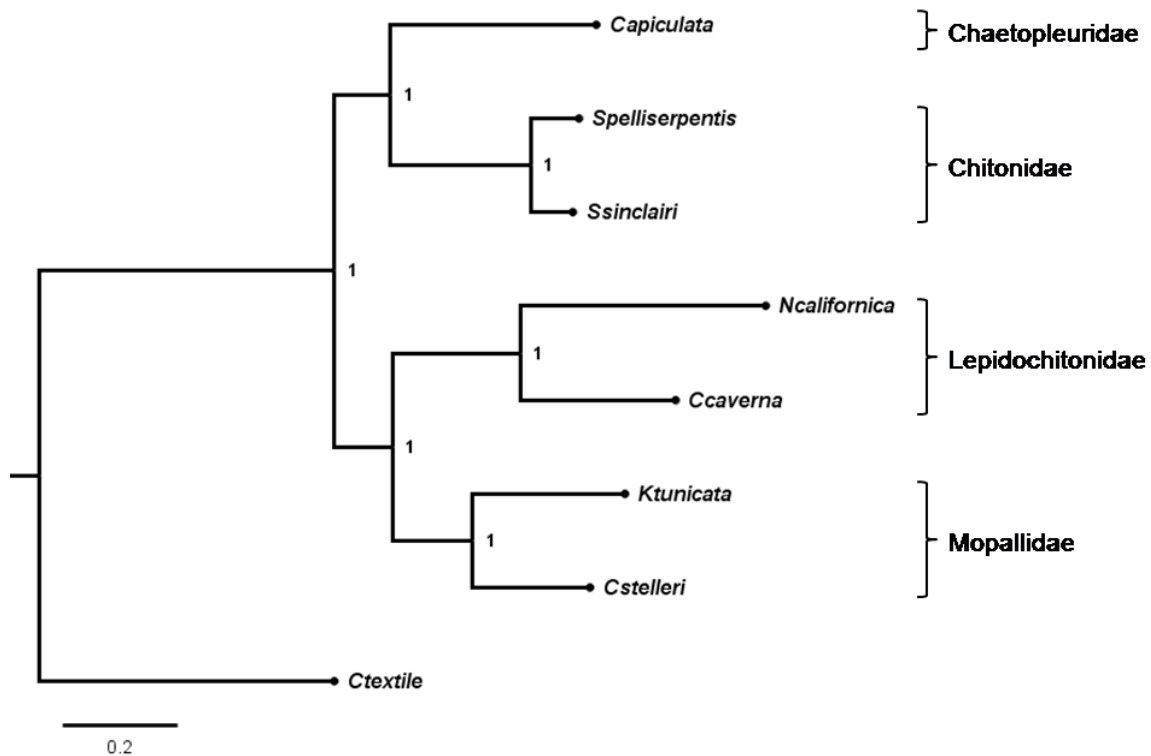
Figure 1: Consensus Bayesian phylogenetic tree based on the concatenated mitochondrial protein-coding genes from seven chitons. Posterior probabilities supporting each node are shown. *C. textile* (Gastropoda) was taken as outgroup. Families of the polyplacophorans corresponding to each clade are mentioned on the right side of the tree.

Figure 2: Bayesian phylogeny considering complete mitochondrial genome sequences of seven chitons. Posterior probabilities supporting each node are shown. *C. textile* (Gastropoda) was taken as outgroup.

Figure 3: Schematic representation depicting arrangement of genetic components (PCGs, *t*RNAs & *r*RNAs) of the mitogenome in the seven Polyplacophoran species. The accession IDs in the figure are as follows- NC_024173.1 (*Sypharochiton sinclairi*), NC_024174.1 (*Sypharochiton*

pelliserpentis), NC_026848.1 (*Cyanoplax caverna*), NC_026849.1 (*Nuttallina californica*), KY824658.1 (*Chaetopleura apiculata*), NC_001636.1 (*Katharina tunicata*) and NC_026850.1 (*Cryptochiton stelleri*).

```
Ccaverna         ---------CTTATTTATAAATCTTTTTGAAATAT---------CTTTCCCAGTTTTTTG
Ncalifornica     --------------ATTTATTTTTAATAGCATATAACGCTTGGCTTACCTGATAAATTA
Cstelleri        ------------------------GCTTATAAAT----------------ATAA-GTA
Capiculata       -----------ATTTTTAAGTTTAATTTATAT-----------------------TA
Ktunicata        -------------------GATAGCATTTATATAT--------------ATAT-ATA
Spelliserpentis  TGTTTGCTAACACTTTTTAATTGATATGTATATAT--------------ATAT-ATA
Ssinclairi       --------------TTTTTGATTGTATTAATATAT--------------ACATAATA
                                                *                        *

Ccaverna         CCTT------TATAAACCTGTTTATCATAAATTT----------------------TA
Ncalifornica     TTTAATACTCTAGTTATATCTTTATTATCTCCTCCCTTTACGTTACTAACACTTAAGTTA
Cstelleri        AATAAGCTCTTATCTATATATAT---------------------------------TA
Capiculata       TATTTTATTATATTAATAGATATA--------------------------------TA
Ktunicata        TATA------TATATATATATATA--------------------------------TA
Spelliserpentis  TATA------TAT-AATATATATA--------------------------------TA
Ssinclairi       TATT------TAT-AATATATATA--------------------------------TA
                  *          **   *    *  *                              **

Ccaverna         T-TGTTAATATTTAATTAGAAGTGTTTATAACTTTTATAATACAATTAGCACT-----TT
Ncalifornica     TACGTAAAGGGAGGAGATAATAAAGATATACTATGTTCATTAACTTAGCAAATAGTACTA
Cstelleri        TATAAAGTTATTATAATAAATAAAGATATA-------------------------TT
Capiculata       TATATTAATAC----ATATAAATAGATCCA-------------------------TC
Ktunicata        TATATATATAT----ATATATATATATATA-------------------------TA
Spelliserpentis  TATATATATAT----GTATATGTATATGTA-------------------------TG
Ssinclairi       TATATATATAT----ATATATATATATATA-------------------------TA
                  *                     *       *    *                    *

Ccaverna         TATATTTATAAAATACCTTATAATTTTCTTATAATTCATTCACTATATATATATATATAT
Ncalifornica     TATATAGATAAGGGAATATATAGGGTAATTATATTATATATA----ATATATATATATAC
Cstelleri        TATAT-----------TAATTAAAATGTTTACAATT--------TTATATATCTAAAAGT
Capiculata       CATAA-----------------------ACAACT-----------------TTAAAGT
Ktunicata        TATATATATATAATAGAATATGAC---CATATATTT----------ATATATCTTAAAGT
Spelliserpentis  TATATGTATACATAAATATGT-ATATGTAGATAAGT-ATATA-TAGATATAAATATAAAC
Ssinclairi       TATAT-----TATAAATATATCATAGATATATTATTAATGTA----TTATAAATAAAAAC
                  ***                             *                  *   *

Ccaverna         ATTAGCAC---------------------
Ncalifornica     ACAAAACAAAATAAGAAACAT---------
Cstelleri        GTTAGCAAAAAAAAAAAA-----------
Capiculata       GTTAGCAAAAAAAAAAAAAAAAAAGAAATG
Ktunicata        GCTAGCAAAAAAAAAAAAAAA-------
Spelliserpentis  ATAAACATAGATATAATGTTCAC-------
Ssinclairi       ATAAATATAAATATAAAA-----------
                  *
```
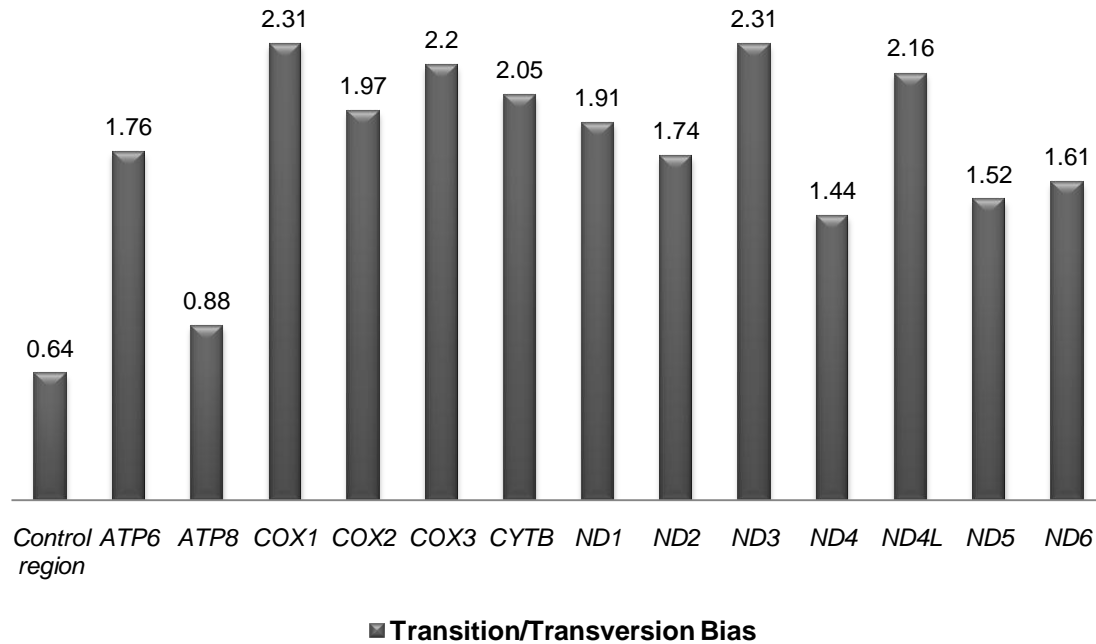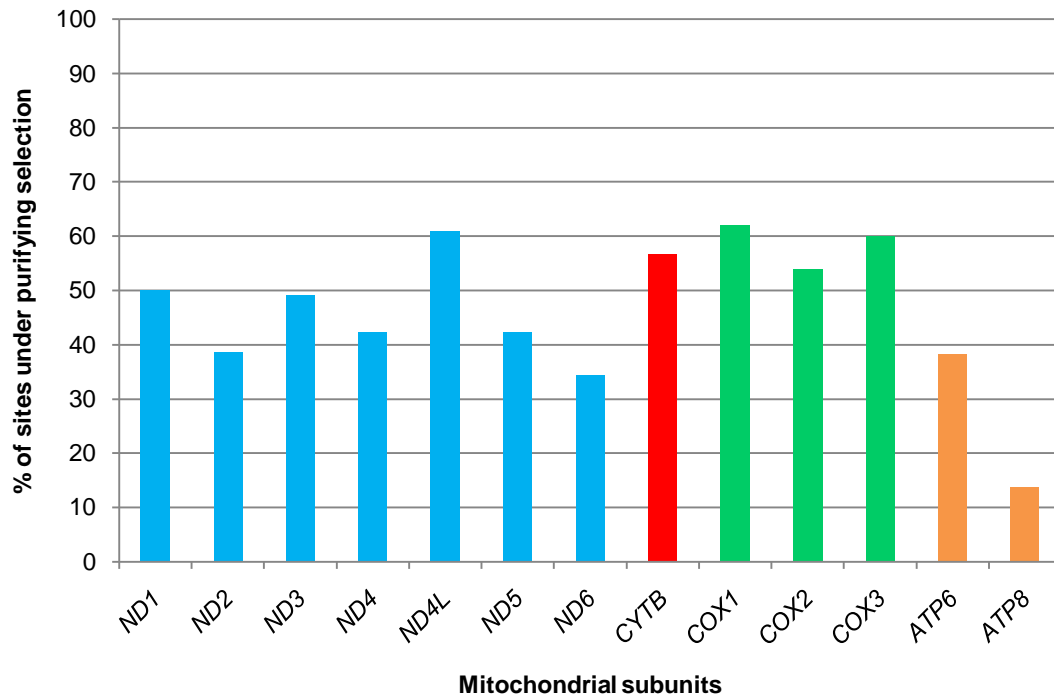
**Figure 4:** Multiple sequence alignment of the control region using MUSCLE.

Figure 5: Transition *vs.* transversion ratio for the 13 PCGs and control region of the chiton mitogenomes.

Figure 6: Percentage of sites under purifying selection in chiton mitochondrial subunits, detected by FEL test. The subunits are colored according to their corresponding complexes: complex I (blue), complex III (red), complex IV (green) and complex V (orange).
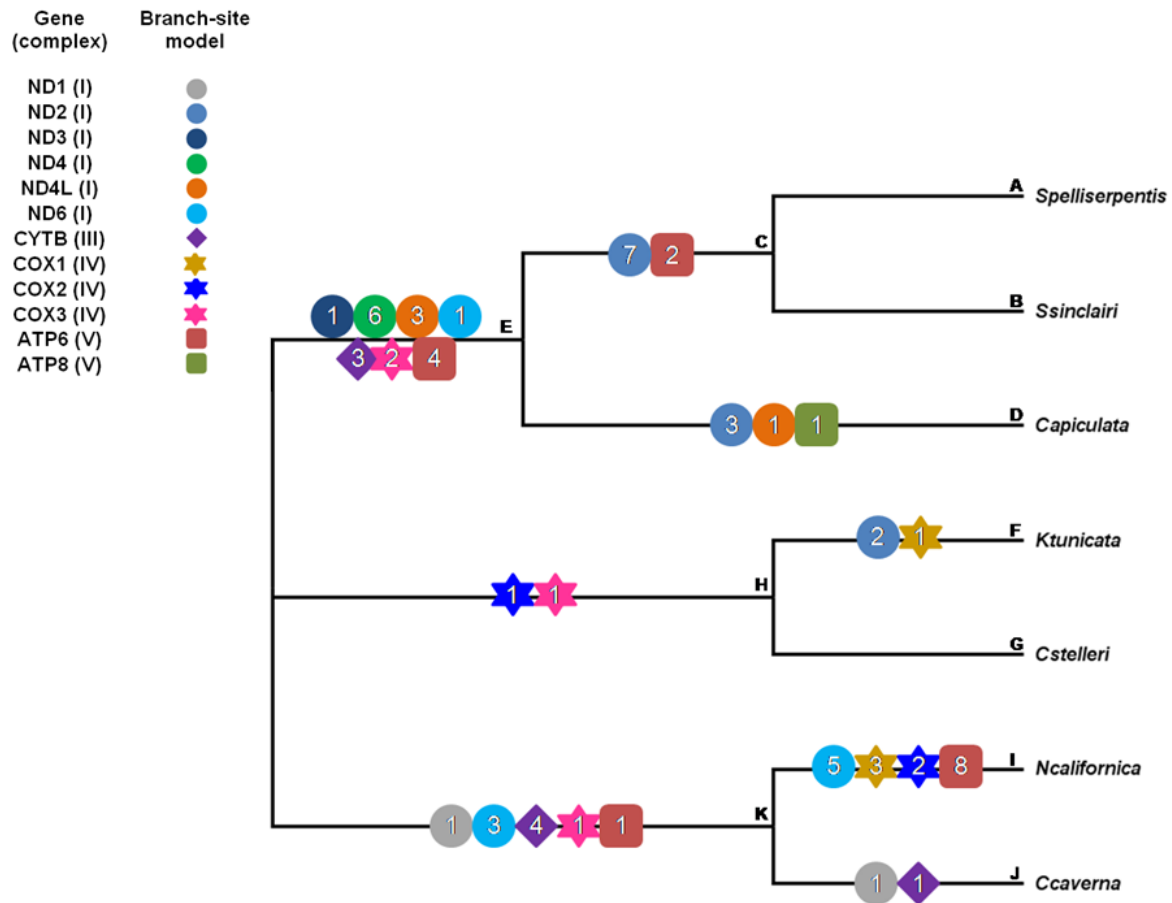
Figure 7: Working topology for analyzing selection pressures on each foreground branches, using the branch-site model in CODEML. The different branches of the phylogeny are labeled from A-K. Different colored shapes represent 13 mitochondrial PCGs. The number within each shape indicates the number of positively selected sites.
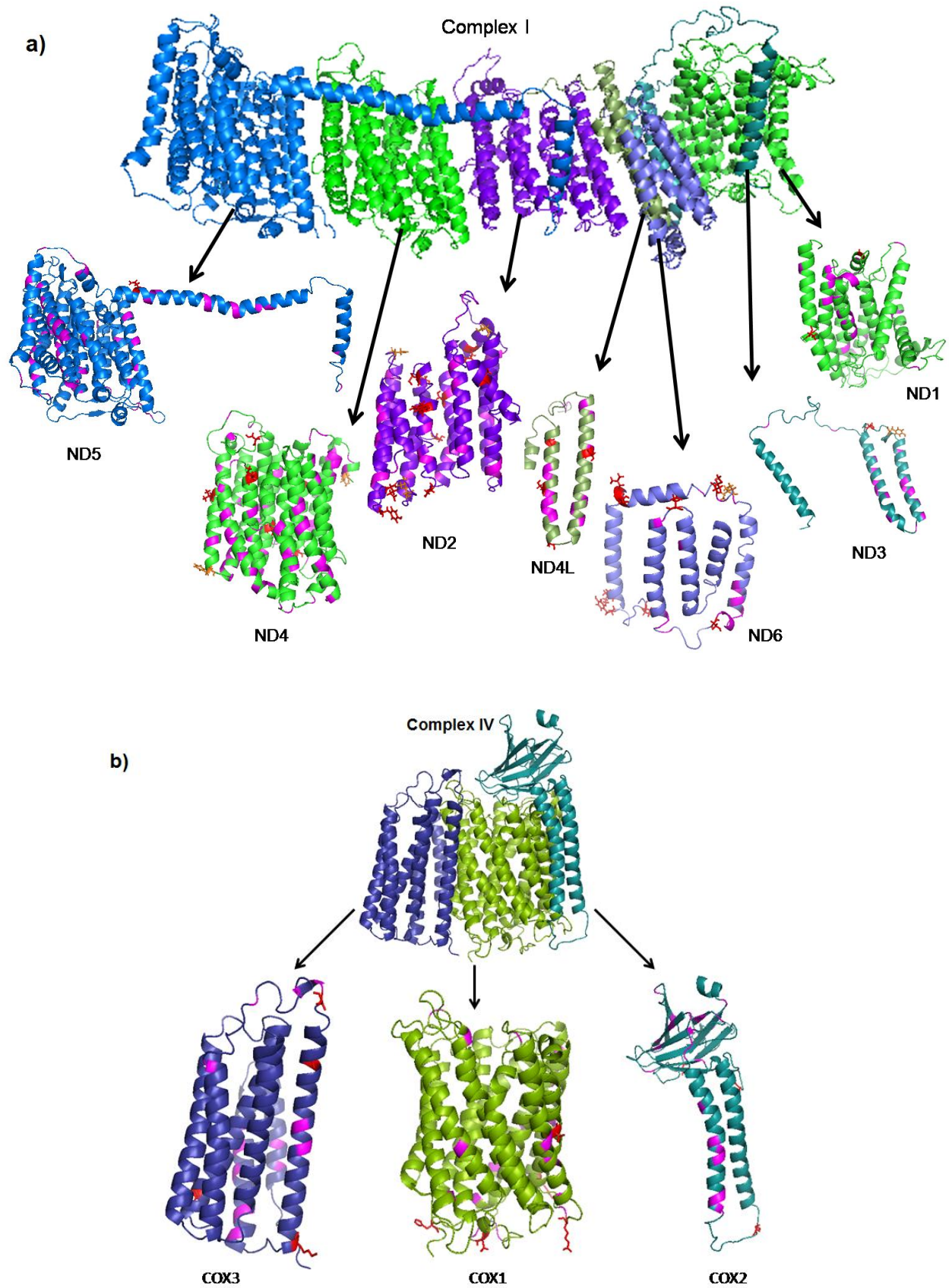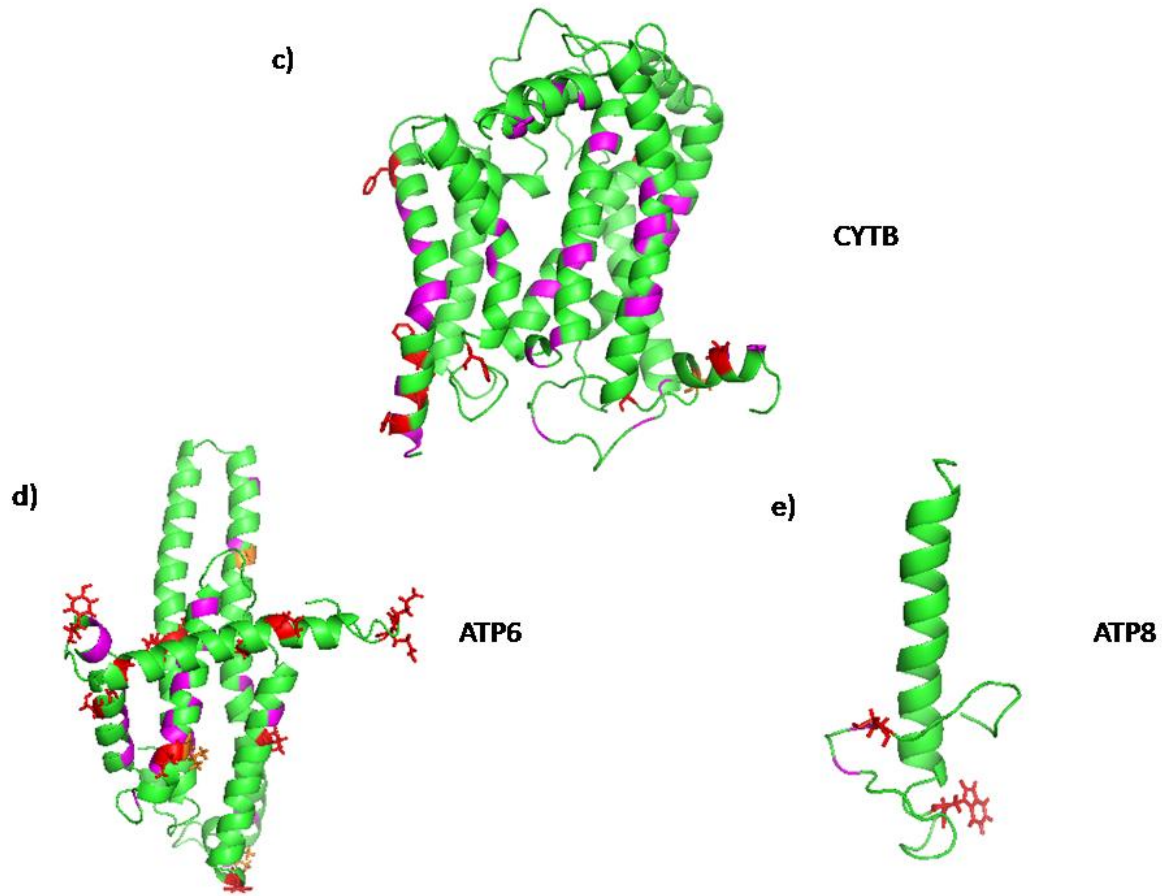
Figure 8: Atomic level structures of the computationally predicted mitochondrial (a) complex I, (b) complex IV, (c) CYTB subunit, (d) ATP6 subunit and (e) ATP8 subunit of *C. apiculata*. Positively selected sites inferred by CODEML and MEME are shown as red sticks whereas residues with amino acid substitutions revealed by TreeSAAP are marked in pink. Common sites identified by these methodologies are displayed as orange sticks.
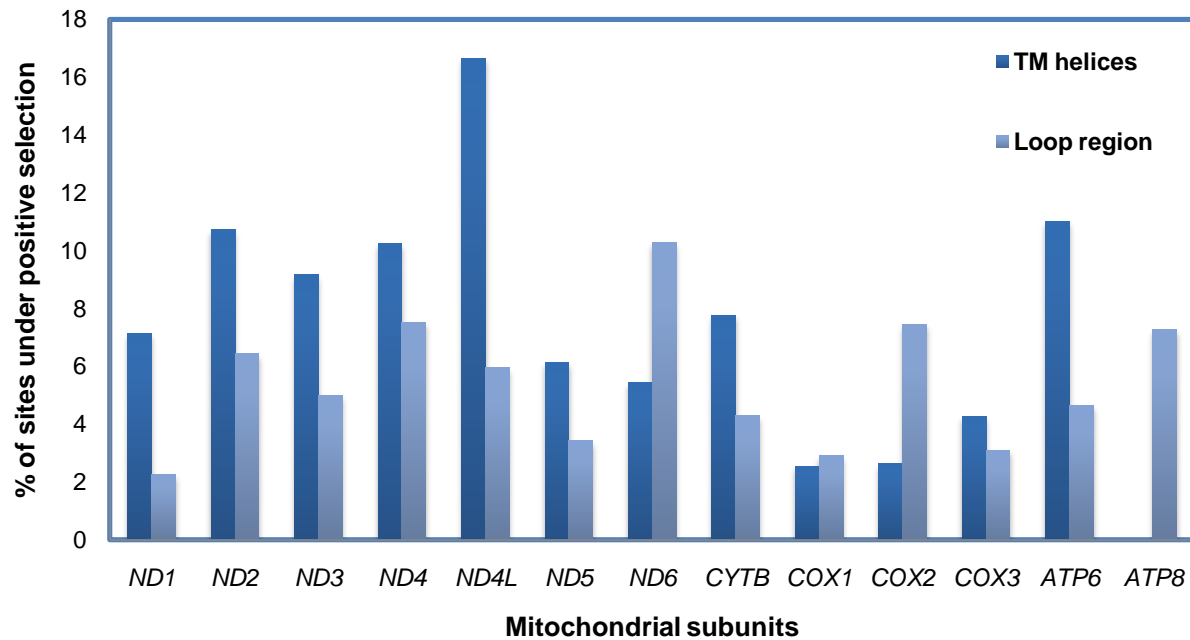
Figure 9: Distribution of sites, under positive selection and/or showing radical amino acid changes, between the transmembrane helices and the loop areas in individual mitochondrial subunits.