BMC Genomics

RESEARCH

Open Access

Complete mitogenomes reveal high diversity and recent population dynamics in Antarctic krill

Shuai Sun[1,2,3], Shuo Li[4,5], Inge Seim[6], Xiao Du[2,3,7], Xianwei Yang[2,8], Kaiqiang Liu[4,5], Zhanfei Wei[4,5], Changwei Shao[4,5*], Guangyi Fan[2,3,7*] and Xin Liu[1,7*]

Abstract

Background The Antarctic krill (*Euphausia superba*) is a keystone species in the Southern Ocean ecosystem, influencing food web dynamics and ecosystem functionality. Despite its ecological importance, further exploration is essential to understand their population dynamics.

Results In this study, we present the complete mitogenome of the Antarctic krill. The assembly is 18,926 bp, including a notably large 3,952 bp control region (CR). The CR features a satellite repeat spanning 2,289 bp, showcasing the effectiveness of long-read sequencing in resolving complex genomic regions. Additionally, we identified 900 nuclear-mitochondrial segments (NUMTs) totaling 2.79 Mb, shedding light on the dynamic integration of mitochondrial DNA (mtDNA) into the nuclear genomes. By establishing a dataset comprising 80 krill mitogenomes, we unveil substantial mitochondrial diversity, particularly within the *ND4* gene. While our analysis reveals no significant differentiation among four geographically distinct groups, we identify at least four maternal genetic clusters. Haplotype network analysis and demographic reconstructions suggest a recent population expansion, likely driven by favorable environmental conditions during the late Pleistocene. Furthermore, our investigation into selection pressures on mitochondrial genes reveals evidence of purifying selection across all 13 protein-coding genes, underscoring the pivotal role of mtDNA conservation in maintaining mitochondrial function under extreme environments.

Conclusions This study provides a repository of Antarctic krill mitogenomes and insights into the population genetics and evolutionary history of this ecologically important species from a mitogenomic perspective, with implications for krill conservation and management in the Southern Ocean.

Keywords Antarctic krill, Mitochondrial genome, Genetic diversity, Population dynamics, Natural selection

*Correspondence:
Changwei Shao
shaocw@ysfri.ac.cn
Guangyi Fan
fanguangyi@genomics.cn
Xin Liu
liuxin@genomics.cn
[1] College of Life Sciences, University of Chinese Academy of Sciences, Beijing 100049, China
[2] BGI Research, Qingdao 266555, China
[3] Shenzhen Key Laboratory of Marine Biology Genomics, BGI Research, Shenzhen 518083, China
[4] State Key Laboratory of Mariculture Biobreeding and Sustainable Goods, Yellow Sea Fisheries Research Institute, Chinese Academy of Fishery Sciences, Qingdao, Shandong 266071, China
[5] Laboratory for Marine Fisheries Science and Food Production Processes, Qingdao Marine Science and Technology Center, Qingdao, Shandong 266237, China
[6] Institute of Deep-Sea Science and Engineering, Chinese Academy of Sciences, Sanya 572000, China
[7] State Key Laboratory of Genome and Multi-Omics Technologies, BGI Research, Shenzhen 518083, China
[8] China National GeneBank, BGI Research, Shenzhen 518120, China

Introduction

The Antarctic krill (*Euphausia superba*), hereafter referred to as krill, is a polar stenothermic species found exclusively in the Southern Ocean that plays a critical role in the delicate balance of this ecosystem [1, 2]. With an estimated biomass of 379 million tons [3, 4], krill represents the dominant epipelagic zooplankton, serving as the cornerstone of the Antarctic food web. They connect primary producers, such as phytoplankton, to higher trophic levels and support marine life, including fish, seabirds, and marine mammals [5, 6]. Beyond their role as the food source, krill are vital drivers of biogeochemical cycles and influence the distribution of carbon, nitrogen, and iron [7]. The commercial fishery of krill in the last five decades has provided a rich source of products for animal feed and human health supplements [8, 9]. This economic importance further highlights the need to understand the genetic resources of krill for effective species protection and sustainable fisheries management.

Mitochondria are important for energy metabolism, stress tolerance, and cell apoptosis – processes vital for krill survival in the challenging polar environment [10, 11]. Although the massive krill nuclear genome, spanning approximately 48 gigabases (Gb), has recently been sequenced and assembled [12], a high-quality mitogenome is still required. Previous studies primarily relied on polymerase chain reaction (PCR) amplification of specific regions or short-read sequencing, yielding incomplete mitogenomes [13–16]. Advances in high-quality long-read sequencing technology now allow for the acquisition of complete and accurate mitogenomes, particularly for resolving long repeat regions [17, 18].

Genetic differentiation across temporal and spatial scales provides insights into the genetic heterogeneity and unique adaptations to local environments, which are essential for krill conservation and management [19, 20]. Studies based on nuclear DNA (nuDNA) markers, such as restriction site-associated DNA (RAD) markers and whole-genome single nucleotide polymorphisms (SNPs), have indicated limited geographic differentiation in krill populations [12, 21]. Mitochondrial DNA (mtDNA) remains a valuable population genomics resource because of its small size, maternal inheritance, and higher mutation rate than nuDNA, meaning that it can be often used to resolve population structure and dynamics [22–25]. Although previous studies employing mtDNA have also reported limited genetic differentiation between krill populations, they had technical and methodological limitations due to small sample numbers, restricted geographic sampling, or few cloned mitochondrial genes (i.e., *COX1*, *ND1*) [13–16, 19, 26, 27].

In this study, we present the assembly and analysis of a complete krill mitogenome. We then construct a mitogenome dataset comprising 80 high-quality mitogenomes, carefully curated to eliminate potential contamination from nuclear mitochondrial DNA segments (NUMTs). Using this dataset, we investigated population diversity, structure, and historical dynamics, providing new insights into the mitochondrial genomic landscape of krill.

Results

The architecture of Antarctic krill complete mitogenome

We assembled a complete, circularized Antarctic krill mitogenome reference using 128.4 Gb of PacBio circular consensus sequencing (CCS) reads (Table S1). This mitogenome spans 18,926 bp, with a GC content of 30.61%. Our assembly contains 37 genes encoding 13 proteins, two ribosomal RNAs (rRNAs), and 22 transfer RNAs (tRNAs) (Fig. 1a and Table S2). The major coding strand of the mitogenome encodes nine protein-coding genes (PCGs) and 14 tRNAs (Fig. 1a). Notably, we obtained a complete 3,952 bp control region (CR) sequence. Within this region, we identified a 2,289 bp satellite repeat composed of 13.8 repeat units of 166 bp each, accounting for 57.92% of the CR (Fig. 1a,b). In contrast, the previously published 16,591 bp assembly (NC_040987.1) [16] contains only 2.1 repeats in the CR (Fig. 1a,c), likely due to the limitations of PCR and Sanger sequencing. This highlights the advantage of long-read sequencing technologies for accurate and complete

(See figure on next page.)

Fig. 1 Landscape of the complete Antarctic krill mitogenome. **a** Circular map of the Antarctic krill mitogenome. The inner circle represents GC content (%), while the middle circle represents the percentage identity between our assembly and the published mitogenome (NC_040987.1) using non-overlapping 100 bp sliding windows. Boxes inside and outside the outer circle represent genes on the major coding and minor coding strands, respectively, with colors indicating gene functions. The directions of the arrows show the transcribed orientation from 5′ to 3′. **b** Self-alignment synteny of our assembled mitogenome. Green and red lines represent forward and reverse alignments, respectively. A tandem repeat of 13.8 repeats with 166 bp repeat units was observed in the CR (control region). **c** Synteny between our assembly and the published mitogenome. Green and red lines represent forward and reverse alignments, respectively. **d** Comparative analysis of mitogenomes among krill and the closely related species. The left panel shows the phylogenetic tree, and the right panel illustrates the organization of mitochondrial genes. *Saropogon luteus* is used as an outgroup, representing the ancestral mitochondrial gene order. Translocations are represented by green lines, and inverted translocations in red lines

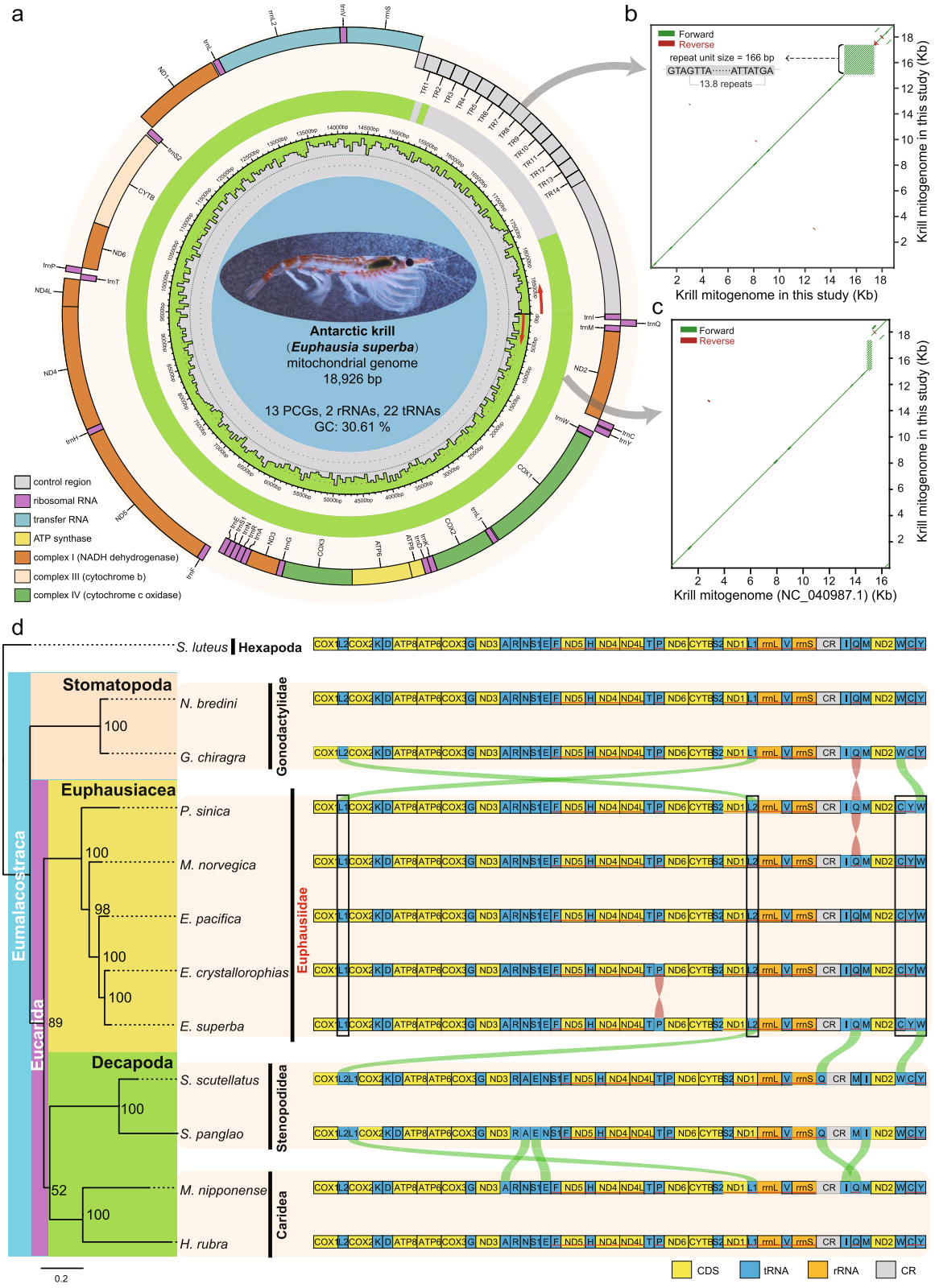

Fig. 1 (See legend on previous page.)

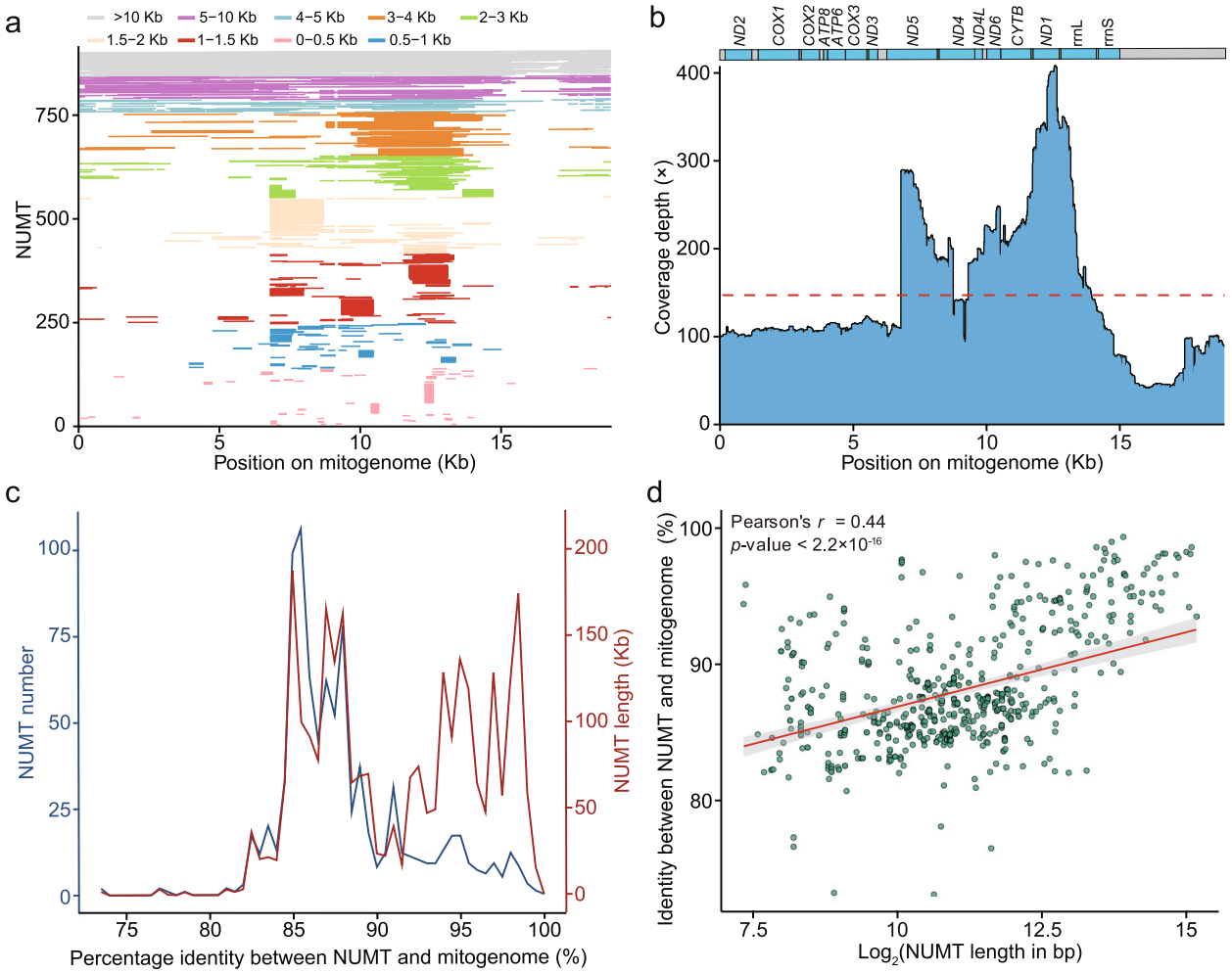

Fig. 2 Characteristics of NUMTs detected in the krill nuclear genome. **a** Locations of detected NUMTs plotted against their positions on the mitogenome. Each horizontal line represents one NUMT, with colors indicating the length of the NUMTs. **b** NUMT composition is illustrated by coverage depth on the mitogenome. The red line indicates the average coverage depth. The linear diagram above represents the mitochondrial genes, where 13 PCGs and 2 rRNAs are shown in blue and other regions in grey. **c** Distribution of percentage identity between NUMT and our mitogenome. The black line indicates the number of NUMTs, while the red line corresponds to the length of NUMTs. **d** Scatterplot showing the correlation of NUMT length and percentage identity. A simple linear regression line is shown as a red line with the grey shaded region showing the 95% confidence interval

assembly of complex genomic regions, particularly in non-coding regions like the CR.

Compared to the ancestral mitochondrial organizations of Pancrustacea (represented by *Saropogon luteus* in the subphylum Hexapoda) (Table S3), the krill mitogenome was almost identical except for two differences – the transposition of $tRNA^{Leu(CUN)}$ (L1) and $tRNA^{Leu(UUR)}$ (L2); and the translocation of $tRNA^{Trp}$ (W) to the downstream of $tRNA^{Cys}$ (C) and $tRNA^{Tyr}$ (Y) (Fig. 1d) – aligning with a previous study [14]. Additionally, comparing the mitochondrial gene order of closely related species revealed that these two gene arrangements were consistently observed in all five species belonging to the Euphausiidae family (Fig. 1d). This observation suggested that these gene order arrangements probably occurred in the early stages of Euphausiidae speciation.

Dynamic integration of mtDNA into the nuclear genome of krill

The integration of mitochondrial DNA (mtDNA) into the nuclear genome is a continuous and dynamic process that occurs both during and after speciation events [28–31]. In the krill nuclear genome [12], we identified 900 nuclear mitochondrial DNA segments (NUMTs) with a total length of 2.79 Mb, averaging 3.1 kb in length, and the longest segment reaching 36.9 Kb

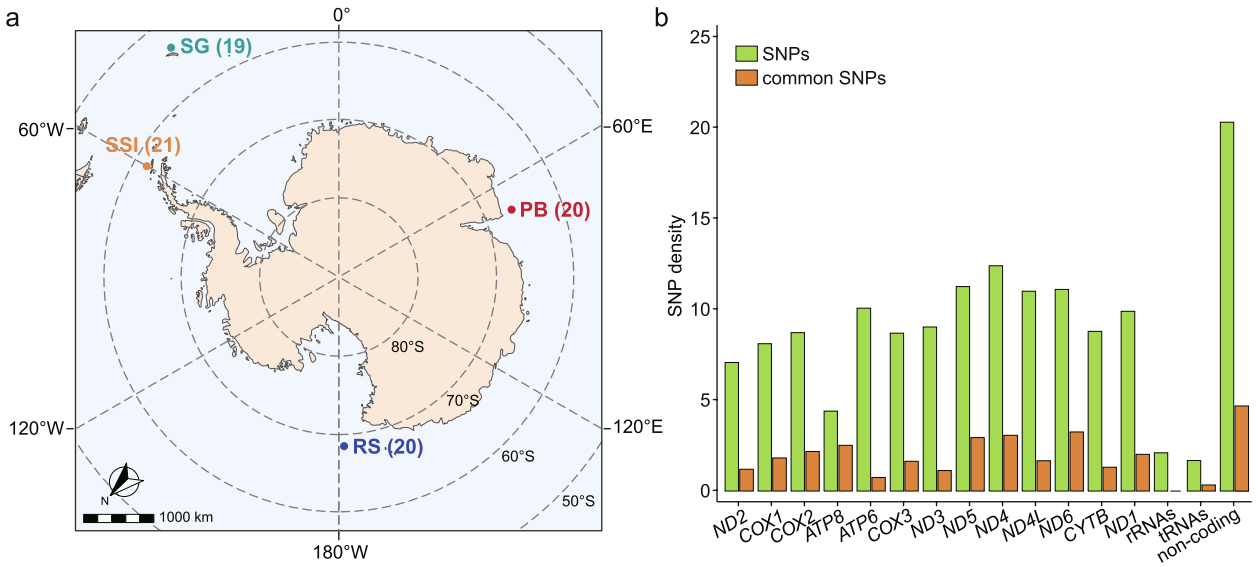

Fig. 3 Characteristics of 80 krill mitogenomes and intra-species variations. **a** Map of Antarctica illustrates the sampling locations for four distinct krill groups, denoted by colored points. Numbers in parentheses indicate the sample size for each group. **c** Distribution of SNP density (measured as the number of SNPs per 100 bp) is shown for the 13 PCGs, rRNA regions, tRNA regions, and non-coding regions. Blue bars represent the density of detected SNPs, while orange bars represent the density of common SNPs

(Fig. 2a and Table S4). Out of these NUMTs, 12 were found to be inserted within gene introns, while the majority were flanked by transposon elements (TEs) (Table S4). Interestingly, 75.33% of the NUMTs were flanked by long terminal repeat sequences (LTRs), despite LTRs accounting for only 13.14% of the entire nuclear genome (Table S5). This observation suggested that TEs, particularly LTRs, potentially affect the generation of NUMTs through mechanisms such as integration into unstable genomic regions or the subsequent post-insertion dynamics. The distribution of gene contents within the NUMTs was uneven, with *ND1* exhibiting the highest density (Fig. 2b), possibly indicating a higher susceptibility to insert into or spread in the nuclear genome.

The distribution of NUMT length revealed the distinct insertion pattern characterized by two prominent peaks (Fig. 2c). The first peak, centered at approximately 86% sequence identity, suggested that a substantial portion of NUMTs originated or close to krill speciation, given the average sequence identity of 87.14% between the mitogenomes of krill and its closely related species, the ice krill (*Euphausia crystallorophias*) (Table S6). The second peak, with sequence identities ranging from approximately 95 to 99%, represents more recent insertion events after speciation. Furthermore, we observed a significant positive correlation (Pearson's $r = 0.44$, p-value $< 2.2 \times 10^{-16}$) between NUMT length and percentage identity between NUMTs and our reference mitogenome (Fig. 2d). The longer NUMTs exhibited higher percentage identity with the mitogenome, suggesting their more recent insertion. In contrast, ancient NUMTs may have undergone fragmentation or elimination from the nuclear genome due to evolutionary pressures.

Krill genetic diversity revealed by mitogenomes

Utilizing whole-genome sequencing data from three newly added samples (Table S7) and previously sequenced samples [12], we assembled 78 additional krill mitogenomes. All mitogenomes assembled into a single contig, with an average length of 17,397 bp (Fig. S1a and Table S8). This size is slightly shorter than the complete krill mitogenome, primarily due to the expected difficulty in assembling the complete long tandem repeats within the CR from short-read data. Gene prediction in these 78 mitogenomes identified 37 genes, with the 13 PCGs exhibiting identical lengths to those in our complete mitogenome reference (Table S8). By integrating our assembled mitogenome reference and the previously published mitogenome (NC_040987.1) [16], we constructed a dataset of 80 high-quality krill mitogenomes, including 21 samples from South Shetland Island (SSI), 19 samples from South Georgia (SG), 20 samples from Prydz Bay (PB), and 20 samples from the Ross Sea (RS) (Fig. 3a).

Analysis of the 80 mitogenomes revealed no rearrangements among the 37 mitochondrial genes (Table S8). We examined various genetic variations, including single nucleotide polymorphisms (SNPs), short insertions and deletions (InDels), and structural variations (SVs). To

Table 1 The genetic diversity indices of mitogenome, 13 protein-coding genes (PCGs), and 2 ribosomal RNAs (rRNAs)

Region	Start (bp)	End (bp)	Length (bp)	Nucleotide diversity (π)	Nucleotide diversity (θ_w)	Number of haplotypes (H)	Haplotype diversity (H_d)
Mitogenome[a]	1	14,974	14,974	8.07×10^{-3}	1.59×10^{-2}	80	1.000
ND2	199	1,200	1,002	6.11×10^{-3}	1.43×10^{-2}	58	0.959
COX1	1,430	2,968	1,539	9.70×10^{-3}	1.64×10^{-2}	76	0.998
COX2	3,048	3,735	688	1.14×10^{-2}	1.76×10^{-2}	54	0.962
ATP8	3,873	4,031	159	1.15×10^{-2}	8.89×10^{-3}	14	0.705
ATP6	4,025	4,699	675	6.11×10^{-3}	2.03×10^{-2}	54	0.947
COX3	4,699	5,491	793	9.18×10^{-3}	1.76×10^{-2}	62	0.973
ND3	5,559	5,912	354	5.42×10^{-3}	1.83×10^{-2}	35	0.797
ND5	6,415	8,145	1,731	1.24×10^{-2}	2.27×10^{-2}	79	1.000
ND4	8,212	9,549	1,338	1.42×10^{-2}	2.51×10^{-2}	78	0.999
ND4L	9,543	9,842	300	8.93×10^{-3}	2.22×10^{-2}	37	0.877
ND6	9,982	10,503	522	1.37×10^{-2}	2.24×10^{-2}	55	0.978
CYTB	10,503	11,639	1,137	7.76×10^{-3}	1.78×10^{-2}	67	0.978
ND1	11,748	12,686	939	1.02×10^{-2}	2.00×10^{-2}	70	0.995
16S	12,769	14,094	1,326	6.20×10^{-4}	4.42×10^{-3}	27	0.596
12S	14,167	14,974	808	7.00×10^{-4}	4.00×10^{-3}	18	0.474

[a] The control region (CR) is excluded to prevent any potential errors that may arise from incomplete assembly of this complex region using WGS data

minimize potential errors arising from the complexity of the CR, variations located in this region were excluded from the analysis. A total of 1,186 SNPs and 17 InDels were identified, with no SVs detected (Table S9). Among the InDels, 12 were in non-coding regions, four in tRNAs, one in the rRNAs, and none in the 13 PCGs (Table S9). SNP density was highest in non-coding regions, followed by the 13 PCGs, and lowest in the rRNAs and tRNAs (Table S9). Among the 13 PCGs, *ND4* exhibited the highest nucleotide variability at 12.41% (calculated as the SNP number per 100 bp, also referred to as SNP density), while *ATP8* showed the lowest at 4.40% (Fig. 3b and Table S9). This contrasts with a previous study that reported that *ATP8* had the highest variability, likely due to the limited sample size of only five individuals [13]. Furthermore, most SNPs (57.25%) were specific to a single individual, while common SNPs (minor allele frequency >0.05) accounted for only 20.49% (Fig. S1b and Table S10), and *ND4* still displayed the highest density of common SNPs (Fig. 3b). These findings highlight the dynamic nature of krill mitogenomic variation, revealing a high level of genetic diversity within the species, particularly in the *ND4* gene.

To assess the adaptive evolution and ecosystem resilience of krill in the face of environmental changes [20], we analyzed their mitogenome diversity. Excluding the CR, we calculated a nucleotide diversity (π) of 8.07×10^{-3} and haplotype diversity (H_d) of 1.00 (Table 1). These values adhere to the most parsimonious model ($\pi = 0.0081\times H_d^2$), confirming the absence of potential methodological artifacts [32]. Notably, the observed genetic diversity is consistent with previous findings based on the *COX1* gene from 504 samples in the Scotia Sea, Southern Ocean [26]. Among the 13 PCGs and two rRNAs, *ND4* exhibited the highest nucleotide diversity ($\pi = 1.42\times10^{-2}$), whereas the *16S* displayed the lowest ($\pi = 6.20\times10^{-4}$) (Table 1). The H_d ranged from 0.705 (*ATP8*) to 1.000 (*ND5*) among the 13 PCGs, while it was lower for the *12S* (0.474) and *16S* (0.596) (Table 1). Overall, the krill population demonstrated a high level of mitochondrial genetic diversity, as indicated by the values defining the level of nucleotide diversity ($\pi = 0.005$) and haplotype diversity ($H_d = 0.5$) [33].

Krill population structure revealed by mitogenomes

Understanding genetic variation is essential for the effective management and conservation of threatened species, as it underlays the ability of populations to adapt, survive, and thrive in response to environmental challenges [34]. We observed low population pairwise differentiation (Φ_{ST}), with a maximum of 0.0185, none of which were statistically significant (Fig. 4a and Table S11). This finding aligned with earlier results based on the mitochondrial *ND1* gene [27]. Furthermore, the genetic diversity across the four studied groups was comparable, with π ranging from 7.93×10^{-3} to 8.37×10^{-3} (Fig. 4a), suggesting consistent levels of genetic diversity across the broad geographic distribution of the Southern Ocean. To further assess population structure, we employed discriminant analysis of principal components (DAPC) [35] and

constructed Bayesian phylogenetic trees using the 80 mitogenomes. Both analyses revealed almost complete genetic mixture among the individuals from the four geographic locations (Fig. 4b,c). These results indicated extensive genetic connectivity within the krill population distributed across the Southern Ocean, consistent with the findings from nuclear whole-genome SNPs [12].

Although no significant differentiation was detected among the geographic groups, four major maternal genetic clusters (M1-M4) were identified at the population level, as evident in both the DAPC and phylogenetic trees results. Furthermore, the pairwise Φ_{ST} values between the maternal genetic clusters ranged from 0.41 to 0.72 (Table S12). The Φ_{ST} values among the maternal genetic clusters were notably higher than those among the geographic groups. For instance, the highest differential loci ($\Phi_{ST} = 1.0$) displayed a distinct pattern of nucleotide composition (Fig. S2). Additionally, the nucleotide sites with high Φ_{ST} among the maternal genetic clusters were not concentrated in a specific region but rather dispersed throughout the mitogenome (Fig. 4d).

Krill population demography revealed by mitogenomes

Haplotype networks of the 13 PCGs exhibited complex reticulated networks observed in *COX1* (Fig. 5a), *ND4* (Fig. S3), and *ND5* (Fig. S3), or star-like networks in others, such as *ND3* and *ND4L* (Fig. S3). These networks exhibited no clear geographical partitioning among the geographic groups, and the presence of numerous rare haplotypes further indicated high genetic connectivity and a past population expansion. The unimodal or bimodal distributions of mismatches observed in the 13 PCGs, along with large differences in population sizes before (θ_0) and after expansion (θ_1) and non-significant values (p-value > 0.05) for the sum of squared deviations (SSD) and Harpending's raggedness index, support the sudden expansion model (Fig. S3 and Table S13). Significant negative values of Tajima's D and Fu's F_S neutrality tests were detected for almost all genes except *ATP8* (Fig. S4 and Table S13), reinforcing the signal of recent population expansion. Demographic reconstructions using extended Bayesian skyline plot (EBSP) analysis further revealed an increase in the effective population size (N_e) that occurred approximately over 100 thousand years ago (kya) (Fig. 5c). This expansion could be attributed to the colder climate during the late Pleistocene with larger areas of the sea ice serving as a habitat for larval krill, coinciding with previous findings [12, 27].

In summary, these findings suggested that krill in the Southern Ocean represented a single population with high levels of genetic connectivity. They have also undergone a recent population expansion, likely driven by favorable environmental conditions during the late Pleistocene. This genetic homogeneity underscores the importance of managing krill populations as a single, interconnected unit to ensure their long-term conservation.

Strong signals of purifying selection on krill mitochondrial genes

Mitochondrial genes are often subject to selection pressures that drive species adaptations to diverse environments [10, 11]. To further explore signals of natural selection, we employed the mixed effects model of evolution (MEME) and fast unconstrained Bayesian approximation (FUBAR) methods [36–38]. Only three amino acid sites exhibited evidence of either episodic or pervasive positive selection (Fig. 6a and Table S14). In contrast, we detected a significant number of 734 amino acid sites under purifying selection across the 13 PCGs (Fig. 6a and Table S14), indicating that krill mitogenome evolution is constrained at many codons to preserve essential functions. For comparison, we collected additional populations of four crustacea species and found similar patterns, with purifying selection being the dominant selective force across mitogenome (Fig. 6b,c and Table S15 - 16). This conservation of mtDNA is likely pivotal to maintaining mitochondrial function, as observed in a wide range of species, such as fishes and birds [39–42].

To further quantify selective pressure on mitochondrial genes, we conducted calculations of the ratio of nonsynonymous to synonymous polymorphisms (π_a/π_s) [43]. All 13 PCGs exhibited π_a/π_s values considerably less than 1 (Fig. 6d and Table S17), indicating the prevalence of purifying selection [44]. We observed relatively stronger selective pressure on *COX1*, *COX2*, *COX3*, *CYTB*, *ND1*,

(See figure on next page.)

Fig. 4 Genetic differentiation and population structure of four geographical krill groups. **a** The genetic diversity and differentiation among the four krill groups. Values within the circles represent the nucleotide diversity (π) at the group level, while values adjacent to the linked lines represent population differentiation (Φ_{ST}), and the corresponding p-values are provided in parentheses. **b** Scatterplot showing the discriminant analysis of principal components (DAPC) of four groups from different geographical locations. The axes represent the first two linear discriminants. Each dot represents an individual, and each circle represents one cluster. **c** Bayesian phylogenetic tree depicting the relationships among the 80 krill samples. Different groups are represented by different colors, and bootstrap values are shown at the nodes. **d** The Φ_{ST} values, calculated across the mitogenome, reflect the genetic differentiation among the four maternal genetic clusters and four geographical groups, respectively

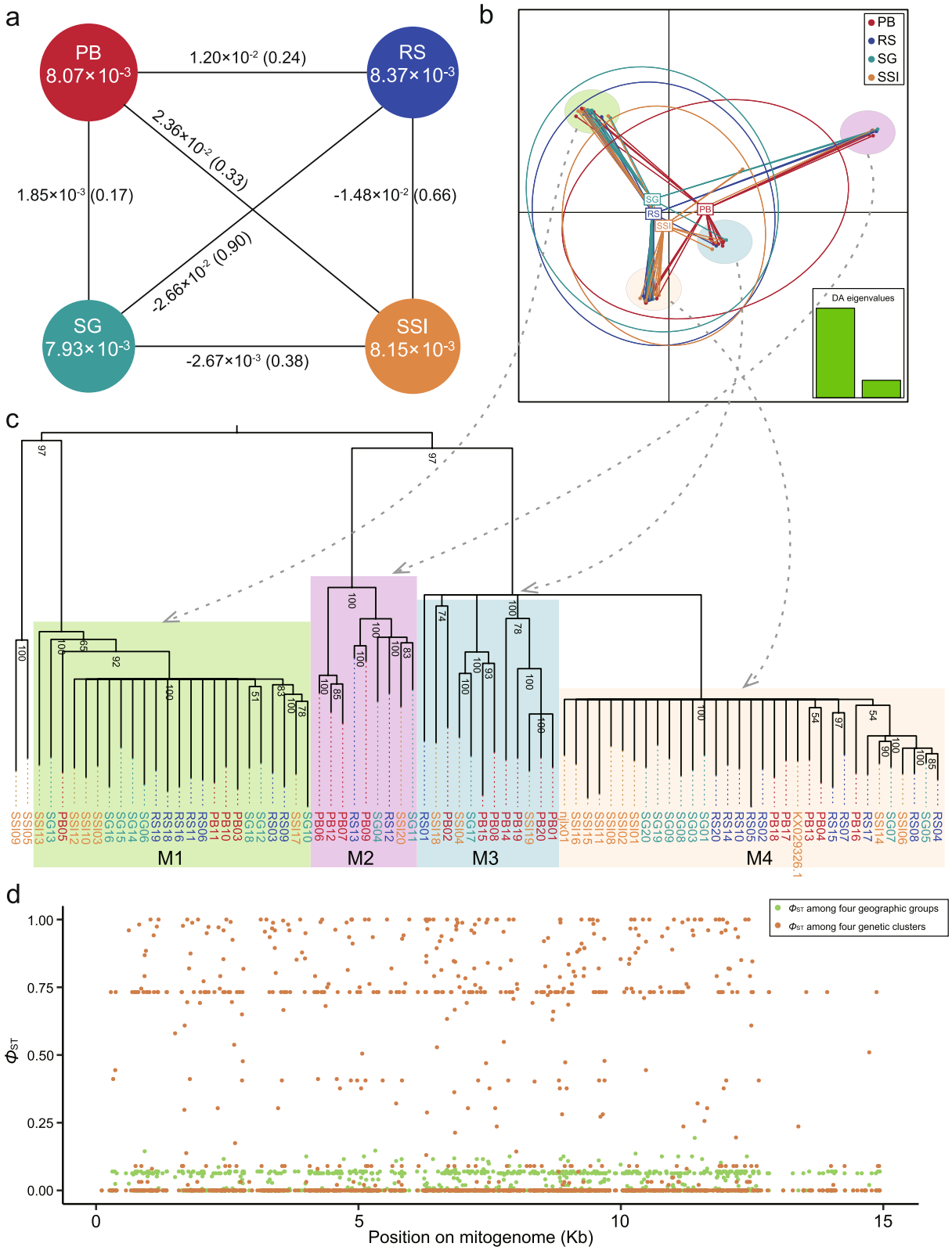

Fig. 4 (See legend on previous page.)

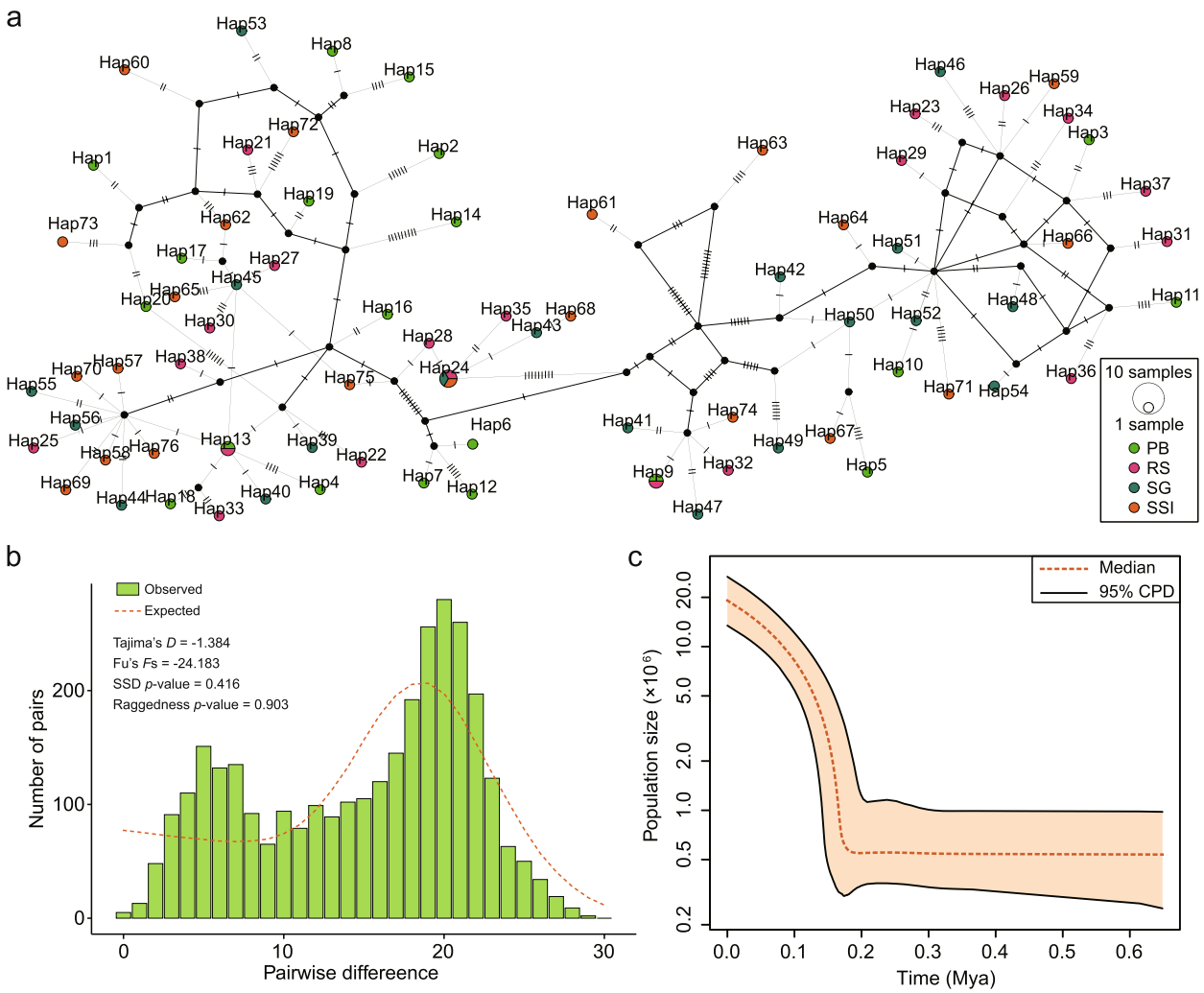

Fig. 5 Haplotype network, neutrality tests, and demography inference based on the mitochondrial PCGs. **a** The haplotype network of *COX1* is illustrated, where each cycle represents a haplotype, and the black dots represent inferred missing haplotypes. The size of the circles is proportional to the frequency of the haplotype, and different colors represent different geographical groups. The haplotype network of the other 12 PCGs were shown in Figure S3. **b** Distribution of pairwise mismatches under a sudden expansion model. Neutrality tests, including Tajima's *D* and Fu's F_S, are provided, along with the *p*-values of the sum of squared deviations (SSD) and Harpending's raggedness index in the goodness-of-fit tests. The neutrality tests and distribution of pairwise mismatches of the other 12 PCGs are shown in Figure S4. **b** Extended Bayesian skyline plot (EBSP) displays the historical changes in the effective population size (N_e). The dashed line represents the mean estimated population size, and the 95% central posterior density (CPD) intervals are indicated by the orange-shaded area

(See figure on next page.)

Fig. 6 Detection of natural selection across the 13 PCG. **a** Selection signatures across amino acid sites of the 13 PCGs were analyzed using the FUBAR and MEME methods. Red and yellow lines represent amino acid sites under pervasive positive and purifying selection, respectively, as detected by FUBAR method. Blue line represents amino acid sites under episodic positive selection, as detected by MEME method. **b** Map of the native distribution of species based on the occurrence data obtained from the GBIF website. Each species is represented by a different color. **c** Doughnut charts of the relative percentage of loci under different types of natural selection in krill and four other crustacean species. **d** Distribution of π_a/π_s values for the 13 PCGs in krill and the four crustacean species provides insights into the patterns of selective pressure on mitochondrial genes. Dashed horizontal lines indicate mean values for each species. Green-shaded areas highlight similar selection pressure patterns, while blue-shaded areas highlight opposing trends in selection pressure between krill and the other four crustacean species

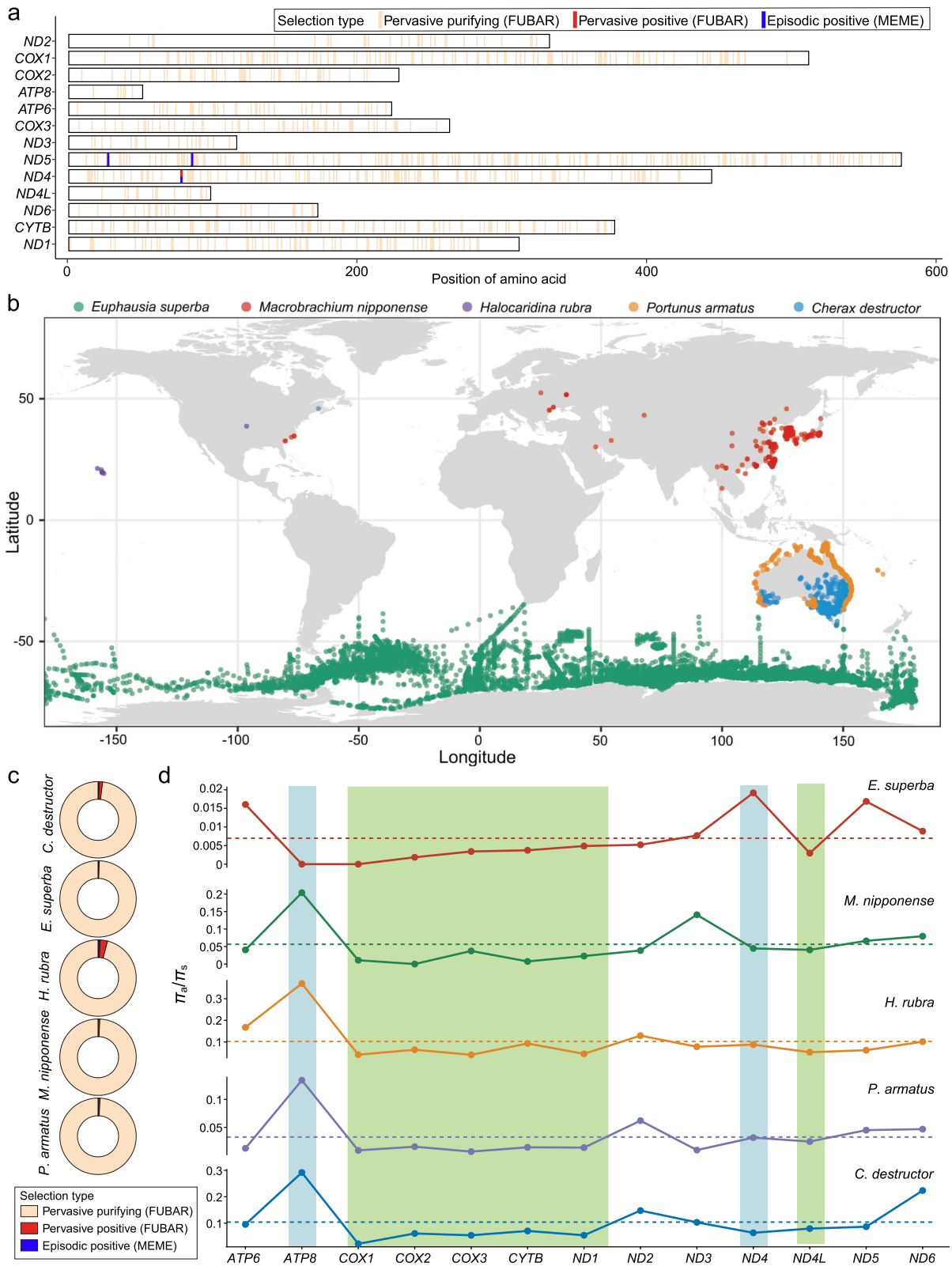

Fig. 6 (See legend on previous page.)

and *ND4L* across krill and four closely related species (Fig. 6d). Intriguingly, *ND4* exhibited relatively relaxed selective pressure in krill but stronger pressure in other species, while *ATP8* experienced stronger selective pressure in krill but more relaxed in other species (Fig. 6d). This difference may play a significant role in krill's adaptation to its extreme environments. Overall, our findings emphasize the importance of purifying selection in mtDNA, preventing the accumulation of deleterious mutations over generations, especially considering the high mutation rate of mtDNA [45].

Discussion

Antarctic krill play a critical role in the Southern Ocean ecosystem, driving food web dynamics and influencing overall ecosystem functionality due to their vast biomass [3–6]. In this study, we successfully generate a complete krill mitogenome using high-quality long sequencing reads, allowing an in-depth exploration of the CR, which included a long satellite of 2,289 bp. Our assembly surpasses previously published krill mitogenomes assembled from PCR or short sequencing reads [13–16]. The use of high-quality long reads shows great promise to reliably assemble the mitogenome, particularly for organisms with large and complex regions, such as mollusks, nematodes, and insects [17, 46–49]. Additionally, we generate a krill mitogenome dataset comprising 80 individuals widely distributed in the Southern Ocean. Although many of these mitogenomes lack the complete CR sequence due to assembly challenges from short whole-genome sequencing (WGS) reads, they remain valuable for examining krill diversity, population structure, and evolution.

Genetic diversity is an essential resource for biodiversity assessments and is increasingly important for guiding conservation and management programs [20, 34]. While the factors contributing to mitogenome diversity are complex and challenging to predict [50–52], our findings suggest that large effective population size, high levels of gene flow, and recent population expansion contribute to the high levels of genetic diversity observed in krill. These factors also result in the presence of numerous rare variants or haplotypes without fixation [19]. Furthermore, the pivotal role of mitochondrial genes in generating energy through oxidative phosphorylation (OXPHOS) leads to their high conservation under constrained evolutionary pressure [53, 54]. The intense purifying selection acting on these genes further reduces genetic variation.

Our analysis revealed a lack of population structure among the four studied locations distributed across the Southern Ocean, suggesting a genetically homogeneous population. This finding is supported by the analysis of 364.57 million nuclear genome SNPs [12]. The lack of population differentiation is likely attributed to the strong Antarctic Circumpolar Current (ACC), which prevents relative isolation in krill habitat areas [2, 55]. However, future research should investigate population structure across temporal periods, such as seasonal fluctuations or comparisons between historical and contemporary populations. Additionally, studying genes responsible for adaptation to temperature changes and potential differences in adaptation abilities among individuals or populations could provide insights into the evolutionary potential (EP) and species' responses to climate change [56, 57].

The demographic history revealed from our study, in conjunction with previous investigations [12, 26, 27], suggests that krill populations are influenced by global climate at the species level. Despite the lack of population differentiation and representative populations, effective population management, fisheries management, and real-time monitoring of species abundance remain necessary [8, 27, 58]. Moreover, the consistent results from population genetic analysis based on both the mitogenome and the nuclear genome [12] demonstrate that primary research programs relying solely on the mitogenomes are feasible, particularly for species with large or complex nuclear genome sizes, as is the case for many other crustaceans of the Southern Ocean and North Atlantic [59].

Conclusions

In this study, we report the complete Antarctic krill mitogenome and a comprehensive analysis of 80 mitogenomes, offering valuable insights into genetic diversity, population structure, and evolution. Our results highlight the necessity of managing krill populations as a single, interconnected unit to ensure their long-term conservation. Further research is needed to investigate the dynamics between krill mitogenomes and environmental or climate change to further elucidate the species' adaptive potential.

Materials and methods

Samples, sequencing, and data resources

We obtained a total of 80 krill individuals with whole-genome resequencing data, comprising 21 samples from South Shetland Island (SSI), 19 samples from South Georgia (SG), 20 samples from Prydz Bay (PB), and 20 samples from Ross Sea (RS) (Fig. 3a). The sequencing data comprised the following: (1) one South Shetland Island sample with PacBio CCS reads [12]; (2) 75 samples sequenced by short-read sequencing technology on DNBSEQ platform (obtained from the China National GeneBank DataBase (CNGB) under accession CNP0001930); (3) one sample retrieved from the

National Center for Biotechnology Information (NCBI) under accession NC_040987.1 [16]; and (4) three newly sequenced samples using short-read technology on DNBSEQ-T1 platform, following the protocol described in a previous study [12].

Assembly and annotation of mitogenome using PacBio CCS reads

The complete mitogenome was assembled using MitoHiFi (v3.2.1) [17], with the published krill mitogenome (NC_040987.1) as a reference sequence and an identity threshold set to 90% (parameters "-o 5 -a animal -p 90 -f NC_040987.1.fasta -g NC_040987.1.gb"). Following assembly, minimap2 (v2.24-r1122) [60] was then used to align the CCS reads to the assembled genome, and potential errors were manually checked using the Integrative Genomics Viewer (v2.16.0) [61]. The assembled mitogenome was rotated to initiate at tRNA-I using Rotate (accessed on 8 th May 2024) [62] with the parameter "-s AATGGAGTGCCTGATAAAGGAATGTCTTG". Gene prediction was performed using GeSeq [63] with the published krill mitogenome (NC_040987.1). Tandem repeats within the control region were annotated using TRF (v4.07b) with default parameters [64]. Finally, the mitogenome structure was visualized using Chloroplot [65].

Comparative mitogenomes of krill and closely related species

The self-synteny of our assembly, as well as the synteny between our assembly and the published mitogenome (NC_040987.1), was assessed using YASS (accessed on 17 th July 2024) [66] with an *E*-value threshold of 0.01.

To construct the phylogenetic tree of krill and the closely related species, the regions of 13 PCGs were extracted using "faidx" function in SAMtools (v1.6) [67]. Sequences for each gene were aligned using MUSCLE (v3.8.1551) [68] with default parameters. Gaps in the alignments were removed using trimAl (v1.4.rev22) [69] with the parameter "-nogaps". The resulting alignments were then concatenated. The phylogenetic tree was reconstructed using IQ-TREE (v2.3.5) [70] with the parameters "-m MFP -b 1000 -o Saropogon_luteus", with the outgroup *Saropogon luteus* for rooting the tree. The final phylogenetic tree, generated in NEWICK format, was visualized and modified using FigTree (v1.4.3).

Detection of NUMTs

We performed an alignment-based analysis combined with read-based validation to identify nuclear mitochondrial DNA segments (NUMTs) within the krill genome. First, we aligned the complete krill mitogenome against the published krill genome using NCBI BLAST + (v2.2.31 +) [71] with parameters optimized for NUMT detection ("-task blastn -gapopen 4 -gapextend 1 -penalty − 1 -reward 1"). To establish a reliable cutoff for NUMT identification, we generated a reversed krill mitogenome using SeqKit (v2.2.0) [72] with the parameter "seq –reverse"), and then aligned it to the reference genome, mimicking the scenario of a 'bogus' alignment. As the krill mitogenome is circular, we linearized the krill mitogenome at the 9,000 bp near the middle position and then followed the previous procedure. The results from the alignments were merged using BEDTools (v2.29.2) [73, 74].

To further validate NUMTs, we utilized read evidence from PacBio CCS reads. NUMTs within 300 Kb of each other were merged and extended by 100 Kb. We extracted the sequences of these NUMT regions using the "faidx" function in SAMTtools. Next, the PacBio CCS reads were aligned to the combined sequences of these NUMT regions and krill mitogenome using Minimap2 with the parameter "-x map-hifi -a -Y". We filtered out low-quality alignments with an identity of less than 98%. Reads spanning the junction of NUMT and no-NUMT regions were extracted, requiring an alignment length greater than 100 bp with NUMT regions and an alignment length greater than 1 kb with non-NUMT or non-mitogenome regions. The segments of reads that belong to the non-NUMT and non-mitogenome were then extracted and realigned to the mitogenome again using the same approach as in the NUMT detection step. Reads containing an unaligned segment of more than 1 kb and with an *E*-value cutoff of 0.05 were considered to support the NUMTs. This multi-step approach ensured the accurate identification and refinement of NUMTs in the krill nuclear genome.

To explore the location relationship between NUMTs and other genomic elements such as transposable elements (TEs), introns, and exons, the following approach was employed. Initially, the NUMTs were extended by 2 kb to encompass downstream and upstream regions. Subsequently, the overlaps between the extended NUMTs and the nuclear genome annotation GFF file, which included TEs, introns, and exons, were computed using BEDTools. The closest genomic elements to the NUMTs were extracted for further analysis.

To determine the identity and aligned positions between NUMTs and the krill mitogenome, a two-step BLAST approach was employed. First, the refined NUMTs were aligned to the krill mitogenome using NCBI BLAST + with the parameters "-task blastn -gapopen 4 -gapextend 1 -penalty -1 -reward 1". The best alignment result was retained, and the remaining portion was then extracted and realigned against the krill mitogenome using NCBI BLAST + with the same parameters.

The two alignment results were merged, and the average identity of NUMTs was calculated and weighted by the alignment length. To assess the relationship between NUMT length and identity, we $\log_2$-transformed the NUMT length, and the relationship between length and percentage identity was plotted in R using the "geom_point" function from "ggplot2" package [75], and a linear line was calculated using the "ggscatter" function from the "ggpubr" package (v0.6.0) [76].

Mitogenome assembly from WGS data

We assembled the mitogenomes of 78 samples from short-read whole genome sequencing data. Raw reads were filtered using fastp (v0.21.0) [77] with the parameter "–length_required 95" to obtain clean reads. These clean reads were then separately aligned to the krill mitogenome and a custom krill NUMTs database using BWA (0.7.17-r1188) [78] with the parameters "-A 2 -B 1 -O 4,4". The custom NUMT database was created by removing sequences with identity > 90% and alignment length > 100 bp against krill mitogenome. To avoid potential NUMT contamination reads likely originating from NUMTs were filtered out using Numt Parser (accessed on 9th May 2024) [79]. Subsequently, assembly was performed using GetOrganelle (v1.7.7.0) [80] with the parameters "-F animal_mt -k 41,61,81,91". For samples with fragmented assemblies, an additional round of assembly was run with increased sequencing data using the parameters "-F animal_mt -k 41,61,81,91 –reduce-reads-for-coverage inf –max-reads inf". Finally, all mitogenomes were rotated, starting at tRNA-I, using Rotate with the parameters "-s AATGGAGTGCCTGATAAAAGGAATGTCTTG".

Detection of mitochondrial sequence variations

To identify mitogenomic variations within the krill population, we aligned the assembled mitogenomes to our complete mitogenome reference, using minimap2 with the parameters "-a -x asm5 –cs". Primary and supplementary alignments were extracted and converted from SAM to PAF format using paftools.js (v2.24-r1122) [60] with the parameters "sam2paf -p -L". Sequence variations, including SNPs, InDelS, and SVs, were detected for each sample using paftools.js with the parameters "call -l 1000 -L 1000 -q 50". Variations of each sample were subsequently merged using BCFtools (v1.15) [81] with the parameters "merge – 0 -Ov".

Genetic diversity analysis of mitogenomes

Genetic diversity was assessed for the 13 protein-coding genes (PCGs) and two rRNAs (*12S* and *16S*) of the mitogenome. Sequences for each gene were aligned using MUSCLE (v3.8.1551) [68], and gaps were removed using trimAl (v1.4.rev22) [69] with the parameter "-nogaps". The resulting alignments in FASTA format were converted to NEXUS format using PGDSpider (v2.1.1.5) [82] with the parameters "-inputformat FASTA -outputformat NEXUS". Genetic diversity metrics, including the number of haplotypes (H), haplotype diversity (H_d), and nucleotide diversity (π and θ_w), were calculated using the "DNA Polymorphism" function in DnaSP (v5.1) [83]. These metrics were also calculated using the same approach for the entire mitogenome, excluding the control regions.

Population differentiation analysis based on mitogenomes

To investigate population differentiation, the whole mitogenome sequences, excluding control regions were aligned using MUSCLE with default parameters, and gaps were removed using trimAl with the parameter "-nogaps". The alignments in FASTA format were converted to Arlequin haplotype data format using the "Generate" function in DnaSP. Population pairwise comparisons of Φ_{ST} were calculated using Arlequin (v3.5.2.2) [84]. Statistically significant pairwise comparisons were tested with 10,000 permutations, and the false discovery rate (FDR) procedure was applied for multiple testing correction at a significance level of 0.05 [85, 86]. The Φ_{ST} value per site among the four genetic clusters and the four geographical groups was calculated using the function "diff_stats" in mmod packages (v1.3.3) [87] with the parameter "phi_st = TRUE".

Population structure analysis based on mitogenomes

Discriminant Analysis of Principal Components (DAPC) was performed to examine genetic population structure among geographic locations, a method that does not rely on assumptions of Hardy–Weinberg Equilibrium and panmixia [35]. FASTA alignments were converted to the "genind" class object using the "read.FASTA" and "as.matrix" functions in the "ape" R package (v5.7–1) [88] and the "DNAbin2 genind" function in the "adegenet" R package (v2.1.10) [89]. The optimal number of principal components (PCs) was determined based on alpha-score optimization using the "dapc" function in the "adegenet" R package. Two PCs and three discriminant functions were retained for analysis, and the DAPC results were visualized using the "scatter" function in the "ade4" R package (v1.7–22) [90].

The DNA sequences of 13 protein-coding genes (PCGs) and two rRNAs were extracted and aligned using MUSCLE with default parameters. To determine the optimal partitioning scheme and substitution models, PartitionFinder (v.2.1.1) [91] was employed using the Bayesian Information Criterion (BIC). The alignments of each gene were concatenated into a super-gene alignment. The phylogenetic tree was reconstructed using MrBayes (v3.2.7) [92] with the (TrN + I) model selected as the best-fit

model. Markov chains were run for 10,000,000 generations, sampling one tree every 1,000 generations, and allowing sufficient time for convergence using the parameters "mcmcp ngen = 10,000,000 samplefreq = 1000 diagnfreq = 1000". The initial 25% trees were discarded as part of a burn-in procedure (determined by the likelihood of stationarity), and the remaining 7500 sampling trees were used to construct a 50% majority rule consensus tree. The resulting NEWICK tree was visualized and modified using FigTree.

To construct haplotype networks for each PCG, sequences were aligned using MUSCLE with default parameters. Gaps were removed using trimAl with the parameter "-nogaps". The resulting alignments in FASTA format were converted to NEXUS haplotype data format using the "Generate" function in DnaSP. Haplotype networks were then calculated for each PCG using PopART (v1.7) [93], applying the Templeton, Crandall, and Sing (TCS) network algorithm [94] with 1×10^4 iterations.

Neutral tests, population demography inference, and natural selection detection

To investigate the evolutionary forces acting on the mitogenome, we conducted a series of analyses, including neutrality tests, population demography inference, and natural selection detection.

First, to distinguish between neutral evolution, natural selection, and demographic expansion or contraction, we employed two neutrality tests: Tajima's D [95] and Fu's F_S [96]. For each PCG, the DNA sequences were aligned using MUSCLE, and gaps were removed using trimAl with the parameter "-nogaps". The resulting alignments in FASTA format were converted to Arlequin haplotype data format using the "Generate" function in DnaSP. Tajima's D and Fu's F_S were calculated in ARLEQUIN with 10,000 permutations.

Next, we investigated changes in effective population size using the extended Bayesian skyline plot (EBSP) method [97] implemented in BEAST (v2.7.7) [98] based on the multi-locus datasets of 13 PCGs. We selected the general time reversible (GTR) nucleotide substitution model and strict clock model. Clock rates were set as follows: (1) *COX1* clock rate (per-linage-rate) was set as 0.01 by following the in the previous study of krill using *COX1* sequences with a pairwise rate of 2% per million [26]; (2) all other PCGs were set as 0.01 with estimation using a log normal prior. MCMC analyses were run for 2×10^8 generations to achieve an adequate effective sample size ($\geq$ 200), sampling every 1,000 generations and discarding the first 10% as burn-in. A linked phylogenetic tree and site model were used, and all other operator settings were left as default. BEAST outputs were visualized using the "plotEBSP.R" function with parameters "useHPD = FALSE, plotPopFunctions = FALSE, log = "y".

Species native distribution and natural selection detection

The occurrence data for the species were obtained from the GBIF website (*Macrobrachium nipponense,* https://doi.org/10.15468/dl.c6tgfk; *Cherax destructor,* https://doi.org/10.15468/dl.43ycc4*; Portunus armatus,* https://doi.org/10.15468/dl.7m8cwp*; Halocaridina rubra,* https://doi.org/10.15468/dl.atddr3*; Euphausia superba,* https://doi.org/10.15468/dl.fcuskn*). The map creation was performed using the "maps" R package (v3.4.1).*

To identify codons potentially under selection, two methods were employed to detect selection: 1) FUBAR: The Fast Unconstrained Bayesian AppRoximation (FUBAR) method [37] in HyPhy [38] assigns each codon a posterior probability (PP) of belonging to one of three dN/dS (ω) classes: $\omega < 1$, $\omega = 1$, and $\omega > 1$. Codons with PP > 0.9 and $\omega > 1$ or $\omega < 1$ were considered to have evolved under positive or purifying selection, respectively. 2) MEME: The Mixed Effects Model of Evolution (MEME) method [36] in HyPhy allows ω at each codon to vary across branches/lineages, facilitating the detection of episodic selection. Codons with $P < 0.05$ were considered to have experienced episodic positive selection. To further quantify selective pressure on mitochondrial genes, we calculated the ratio of non-synonymous to synonymous nucleotide diversity (π_a/π_s) of the 13 PCGs in the mitogenome using the function "Synonymous and NonSyn. Substitutions" in DnaSP.

Abbreviations

12S	12S ribosomal RNA
16S	16S ribosomal RNA
ACC	Antarctic Circumpolar Current
ATP6	ATP synthase 6
ATP8	ATP synthase 8
CCS	Circular consensus sequencing
COX1	Cytochrome *c* oxidase I
COX2	Cytochrome *c* oxidase II
COX3	Cytochrome *c* oxidase III
CR	Control region
CYTB	Cytochrome *b*
DAPC	Discriminant analysis of principal components
EBSP	Extended Bayesian skyline plot
EP	Evolutionary potential
FUBAR	Fast unconstrained Bayesian approximation
Gb	Gigabases
GC	Guanine cytosine
H_d	Haplotype diversity
InDels	Short insertions and deletions
kya	Thousand years ago
LTRs	Long terminal repeat sequences
MEME	Mixed effects model of evolution
mtDNA	Mitochondrial DNA
ND1	NADH dehydrogenase 1
ND2	NADH dehydrogenase 2
ND3	NADH dehydrogenase 3
ND4	NADH dehydrogenase 4

ND4L	NADH dehydrogenase 4L
ND5	NADH dehydrogenase 5
ND6	NADH dehydrogenase 6
N_e	Effective population size
nuDNA	Nuclear DNA
NUMTs	Nuclear mitochondrial segments
OXPHOS	Oxidative phosphorylation
PB	Prydz Bay
PCGs	Protein-coding genes
PCR	Polymerase chain reaction
RAD	Restriction site associated DNA
rRNAs	Ribosomal RNAs
RS	Ross Sea
SG	South Georgia
SNPs	Single nucleotide polymorphisms
SSI	South Shetland Island
SVs	Structural variations
TEs	Transposon elements
tRNAs	Transfer RNAs
WGS	Whole-genome sequencing
π	Nucleotide diversity
π_a/π_s	Ratio of nonsynonymous to synonymous polymorphisms
Φ_{ST}	Population pairwise differentiation

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12864-025-11579-w.

Additional file 1. Fig. S1 Assembly of krill mitogenomes and the distribution of minor frequency of SNPs. (a) The length and GC content of 78 krill mitogenome assemblies derived from WGS data. The black line denotes mitogenome length, while the red line represents GC content. Grey lines are used to represent the average values. (b) Distribution of minor allele frequency of SNPs. Fig. S2 Pattern of nucleotide composition for the highest differential 48 loci ($\Phi_{ST} = 1$). Boxes of different colors correspond to different nucleotides. Fig. S3 Haplotype network of the other 12 PCGs. Each cycle represents a haplotype, and the black dots represent inferred missing haplotypes. The size of the circles is proportional to the frequency of the haplotype, and different colors represent different geographical groups. Fig. S4 Distribution of pairwise mismatches of the other 12 PCGs under a sudden expansion model. Distribution of pairwise mismatches under a sudden expansion model. Neutrality tests, including Tajima's *D* and Fu's F_S, are provided, along with the *p*-values of the sum of squared deviations (SSD) and Harpending's raggedness index in the goodness-of-fit tests.

Additional file 2. Table S1. Summary of PacBio CCS reads used for mitogenome assembly. Table S2. Assembly and annotation of krill mitogenome. Table S3. Mitogenomes used in the comparation of mitochondrial organizations among krill and closely related species Table S4. Statistics of detected NUMTs. Table S5. Locations of NUMTs in krill nuclear genome. Table S6. BLASTN alignment results of mitogenomes of *Euphausia superba* and the closely related species *Euphausia crystallorophias*. Table S7. Statistics of newly sequenced three krill samples Table S8. Assembly and annotation of 78 mitogenomes assembly from WGS data. Table S9. The distribution of sequence variation on different regions of mitogenome. Table S10. Statistics of SNPs and InDels detected in krill mitogenome dataset. Table S11. Population differentiation (Φ_{ST}) and the significance test among four geographical groups. Table S12. Population differentiation (Φ_{ST}) and the significance test among four maternal genetic clusters. Table S13. Neutrality test and mismatch distribution for each protein coding gene (PCG) to infer population history. Table S14. Natural selection of each 13 PCGs detected. Table S15. Accessions of mitogenome populations of four Crustacea species. Table S16. Number of amino acid sites under natural selection for mitogenome populations of five Crustacea species. Table S17. The π_a/π_s values of mitogenome populations of five Crustacea species.

Acknowledgements
We would like to express our sincere gratitude to the colleagues of Changwei Shao at the Yellow Sea Fisheries Research Institute, Chinese Academy of Fishery Science, Qingdao, China, for providing samples for sequencing and granting us access to their krill genome and population data. We are grateful to China National GeneBank (CNGB) for their technical support in library construction and sequencing of the whole-genome short reads. Furthermore, we would like to extend our sincere gratitude to Andreas Wallberg at the Department of Medical Biochemistry and Microbiology, Uppsala University, Sweden, for providing the mitogenome of *Meganyctiphanes norvegica* (northern krill). The authors thank the anonymous reviewers for their peer review.

Authors' contributions
X.L., G.F., and C.S. initiated and designed the research. C.S. coordinated sampling and sequencing. S.S. carried out data analysis. S.S., S.L., X.D., and X.Y drafted the original manuscript, and prepared the figures and tables. X.L., G.F., C.S., S.S., I.S., K.L., and Z.W., edited and revised the manuscript. All authors contributed to and approved the final version of the manuscript.

Funding
This study was supported by the National Key Research and Development Program of China (2022YFC2807505); the Marine S&T Fund of Shandong Province for Pilot National Laboratory for Marine Science and Technology (Qingdao) (No.2022QNLM030004); the Central Public-interest Scientific institution Basal Research Fund, CAFS (No.2019QY01, No.2023 TD19); the Taishan Scholars Program (No.tstp20221149 to C.S.); and the Thousands Marine Species Genome Sequencing Project of Qingdao Free Trade Zone Management Committee.

Data availability
The whole genome sequencing reads of the three krill samples are deposited in the China National GeneBank (CNGB) with the project accession code CNP0006136. The newly assembled 79 mitogenomes have been deposited within the same project with assembly accession CNA0142445-CNA0142523. The complete mitogenome assembled using PacBio HiFi reads is submitted to the NCBI GenBank (PQ217826).

Declarations

Ethics approval and consent to participate
All experimental procedures were conducted in accordance with the guidelines approved by the institutional review board on bioethics and biosafety of BGI (IRB-BGI). The experiment was authorized by IRB-BGI (NO. BGI-IRB A20007), and the review procedures in IRB-BGI meet good clinical practice (GCP) principles.

Consent for publication
The authors declare that they have no conflicting agendas.

Competing interests
The authors declare no competing interests.

Received: 15 October 2024 Accepted: 7 April 2025
Published online: 29 April 2025

References

1. Atkinson A, Shreeve RS, Hirst AG, Rothery P, Tarling GA, Pond DW, Korb RE, Murphy EJ, Watkins JL. Natural growth rates in Antarctic krill (*Euphausia superba*): II. Predictive models based on food, temperature, body length, sex, and maturity stage. Limnology and Oceanography. 2006;51(2):973–987. https://doi.org/10.4319/lo.2006.51.2.0973
2. Atkinson A, Siegel V, Pakhomov EA, Rothery P, Loeb V, Ross RM, Quetin LB, Schmidt K, Fretwell P, Murphy EJ, et al. Oceanic circumpolar habitats of Antarctic krill. Mar Ecol Prog Ser. 2008;362:1–23. https://doi.org/10.3354/meps07498.

3. Bar-On YM, Phillips R, Milo R. The biomass distribution on Earth. Proc Natl Acad Sci. 2018;115(25):6506–11. https://doi.org/10.1073/pnas.1711842115.
4. Atkinson A, Siegel V, Pakhomov EA, Jessopp MJ, Loeb V. A re-appraisal of the total biomass and annual production of Antarctic krill. Deep Sea Res Part I. 2009;56(5):727–40. https://doi.org/10.1016/j.dsr.2008.12.007.
5. Hill SL, Murphy EJ, Reid K, Trathan PN, Constable AJ. Modelling Southern Ocean ecosystems: krill, the food-web, and the impacts of harvesting. Biol Rev. 2006;81(04):581–608. https://doi.org/10.1017/s1464793106007123.
6. Watanuki Y, Mori Y, Naito Y. *Euphausia superba* dominates in the diet of Adélie penguins feeding under fast sea-ice in the shelf areas of Enderby Land in summer. Polar Biol. 1994;14(6):429–32. https://doi.org/10.1007/bf00240264.
7. Cavan EL, Belcher A, Atkinson A, Hill SL, Kawaguchi S, McCormack S, Meyer B, Nicol S, Ratnarajah L, Schmidt K, et al. The importance of Antarctic krill in biogeochemical cycles. Nat Commun. 2019;10(1):4742. https://doi.org/10.1038/s41467-019-12668-7.
8. Nicol S, Foster J. The fishery for Antarctic krill: Its current status and management regime. Biology and Ecology of Antarctic Krill. 2016:387–421. https://doi.org/10.1007/978-3-319-29279-3_11
9. Krill fisheries [https://www.ccamlr.org/en/fisheries/krill-fisheries]
10. Lebenzon JE, Overgaard J, Jørgensen LB. Chilled, starved or frozen: insect mitochondrial adaptations to overcome the cold. Current Opinion in Insect Science. 2023;58: 101076. https://doi.org/10.1016/j.cois.2023.101076.
11. Sokolova I. Mitochondrial adaptations to variable environments and their role in animals' stress tolerance. Integr Comp Biol. 2018;58(3):519–31. https://doi.org/10.1093/icb/icy017.
12. Shao C, Sun S, Liu K, Wang J, Li S, Liu Q, Deagle BE, Seim I, Biscontin A, Wang Q, et al. The enormous repetitive Antarctic krill genome reveals environmental adaptations and population insights. Cell. 2023;186(6):1279-1294.e1219. https://doi.org/10.1016/j.cell.2023.02.005.
13. Johansson M, Duda E, Sremba A, Banks M, Peterson W. Assessing population-level variation in the mitochondrial genome of *Euphausia superba* using 454 next-generation sequencing. Mol Biol Rep. 2012;39(5):5755–60. https://doi.org/10.1007/s11033-011-1385-y.
14. Machida R, Miya M, Yamauchi M, Nishida M, Nishida S. Organization of the mitochondrial genome of Antarctic krill *Euphausia superba* (Crustacea: Malacostraca). Mar Biotechnol. 2004;6(3):238–50. https://doi.org/10.1007/s10126-003-0016-6.
15. Huang Y, Bian C, Liu Z, Wang L, Xue C, Huang H, Yi Y, You X, Song W, Mao X, et al. The first genome survey of the Antarctic krill (*Euphausia superba*) provides a valuable genetic resource for polar biomedical research. Mar Drugs. 2020;18(4):185. https://doi.org/10.3390/md18040185.
16. Zhao M, Zhao M, Ma C, Feng C, Wang L, Ma L. Complete mitochondrial genome of Antarctic krill *Euphausia superba* (Eucarida: Euphausiacea). Mitochondrial DNA Part B. 2017;2(2):928–9. https://doi.org/10.1080/23802359.2017.1413306.
17. Uliano-Silva M, Ferreira JGRN, Krasheninnikova K, Blaxter M, Mieszkowska N, Hall N, Holland P, Durbin R, Richards T, Kersey P, et al. MitoHiFi: a python pipeline for mitochondrial genome assembly from PacBio high fidelity reads. BMC Bioinformatics. 2023;24(1):288. https://doi.org/10.1186/s12859-023-05385-y.
18. Wenger AM, Peluso P, Rowell WJ, Chang P-C, Hall RJ, Concepcion GT, Ebler J, Fungtammasan A, Kolesnikov A, Olson ND, et al. Accurate circular consensus long-read sequencing improves variant detection and assembly of a human genome. Nat Biotechnol. 2019;37(10):1155–62. https://doi.org/10.1038/s41587-019-0217-9.
19. Jarman SN, Deagle BE. Genetics of Antarctic Krill. In: Siegel V editor. Biology and Ecology of Antarctic Krill. Cham: Springer. 2016;p. 247–277. https://doi.org/10.1007/978-3-319-29279-3_7
20. Hohenlohe PA, Funk WC, Rajora OP. Population genomics for wildlife conservation and management. Mol Ecol. 2020;30(1):62–82. https://doi.org/10.1111/mec.15720.
21. Deagle BE, Faux C, Kawaguchi S, Meyer B, Jarman SN. Antarctic krill population genomics: apparent panmixia, but genome complexity and large population size muddy the water. Mol Ecol. 2015;24(19):4943–59. https://doi.org/10.1111/mec.13370.
22. Bourgeois YXC, Warren BH. An overview of current population genomics methods for the analysis of whole-genome resequencing data in eukaryotes. Mol Ecol. 2021;30(23):6036–71. https://doi.org/10.1111/mec.15989.
23. Moritz C, Dowling TE, Brown WM. Evolution of animal mitochondrial DNA: Relevance for population biology and systematics. Annu Rev Ecol Syst. 1987;18(1):269–92. https://doi.org/10.1146/annurev.es.18.110187.001413.
24. Castellana S, Vicario S, Saccone C. Evolutionary patterns of the mitochondrial genome in Metazoa: Exploring the role of mutation and selection in mitochondrial protein–coding genes. Genome Biol Evol. 2011;3:1067–79. https://doi.org/10.1093/gbe/evr040.
25. Wallace DC. Mitochondrial DNA variation in human radiation and disease. Cell. 2015;163(1):33–8. https://doi.org/10.1016/j.cell.2015.08.067.
26. Goodall-Copestake WP, Pérez-Espona S, Clark MS, Murphy EJ, Seear PJ, Tarling GA. Swarms of diversity at the gene *cox1* in Antarctic krill. Heredity. 2010;104(5):513–8. https://doi.org/10.1038/hdy.2009.188.
27. Bortolotto E, Bucklin A, Mezzavilla M, Zane L, Patarnello T. Gone with the currents: lack of genetic differentiation at the circum-continental scale in the Antarctic krill *Euphausia superba*. BMC Genet. 2011;12(1):1–18. https://doi.org/10.1186/1471-2156-12-32.
28. Gaziev AI, Shaikhaev GO. Nuclear mitochondrial pseudogenes. Mol Biol. 2010;44(3):358–68. https://doi.org/10.1134/s0026893310030027.
29. Mourier T, Hansen AJ, Willerslev E, Arctander P. The Human Genome Project Reveals a continuous transfer of large mitochondrial fragments to the nucleus. Mol Biol Evol. 2001;18(9):1833–7. https://doi.org/10.1093/oxfordjournals.molbev.a003971.
30. Antunes A, Pontius J, Ramos MJ, O'Brien SJ, Johnson WE. Mitochondrial introgressions into the nuclear genome of the domestic cat. J Hered. 2007;98(5):414–20. https://doi.org/10.1093/jhered/esm062.
31. Grau ET, Charles M, Féménia M, Rebours E, Vaiman A, Rocha D. Survey of mitochondrial sequences integrated into the bovine nuclear genome. Sci Rep. 2020;10(1):2077. https://doi.org/10.1038/s41598-020-59155-4.
32. Goodall-Copestake WP, Tarling GA, Murphy EJ. On the comparison of population-level estimates of haplotype and nucleotide diversity: a case study using the gene *cox1* in animals. Heredity. 2012;109(1):50–6. https://doi.org/10.1038/hdy.2012.12.
33. Grant WS, Clark AM, Bowen BW. Why restriction fragment length polymorphism analysis of mitochondrial DNA failed to resolve sardine (*Sardinops*) biogeography: insights from mitochondrial DNA cytochrome *b* sequences. Can J Fish Aquat Sci. 1998;55(12):2539–47. https://doi.org/10.1139/f98-127.
34. Jost L, Archer F, Flanagan S, Gaggiotti O, Hoban S, Latch E. Differentiation measures for conservation genetics. Evol Appl. 2018;11(7):1139–48. https://doi.org/10.1111/eva.12590.
35. Jombart T, Devillard S, Balloux F. Discriminant analysis of principal components: a new method for the analysis of genetically structured populations. BMC Genet. 2010;11(1):1–15. https://doi.org/10.1186/1471-2156-11-94.
36. Murrell B, Wertheim JO, Moola S, Weighill T, Scheffler K, Kosakovsky Pond SL. Detecting individual sites subject to episodic diversifying selection. PLoS Genet. 2012;8(7): e1002764. https://doi.org/10.1371/journal.pgen.1002764.
37. Murrell B, Moola S, Mabona A, Weighill T, Sheward D, Kosakovsky Pond SL, Scheffler K. FUBAR: A fast, unconstrained Bayesian approximation for inferring selection. Mol Biol Evol. 2013;30(5):1196–205. https://doi.org/10.1093/molbev/mst030.
38. Kosakovsky Pond SL, Poon AFY, Velazquez R, Weaver S, Hepler NL, Murrell B, Shank SD, Magalis BR, Bouvier D, Nekrutenko A *et al*. HyPhy 2.5—A customizable platform for evolutionary hypothesis testing using phylogenies. Molecular Biology and Evolution. 2020;37(1):295–299. https://doi.org/10.1093/molbev/msz197
39. Ramos B, González-Acuña D, Loyola DE, Johnson WE, Parker PG, Massaro M, Dantas GPM, Miranda MD, Vianna JA. Landscape genomics: natural selection drives the evolution of mitogenome in penguins. BMC Genomics. 2018;19(1):1–17. https://doi.org/10.1186/s12864-017-4424-9.
40. Shen Y-Y, Shi P, Sun Y-B, Zhang Y-P. Relaxation of selective constraints on avian mitochondrial DNA following the degeneration of flight ability. Genome Res. 2009;19(10):1760–5. https://doi.org/10.1101/gr.093138.109.
41. Graham AM, Lavretsky P, Wilson RE, McCracken KG. High-altitude adaptation is accompanied by strong signatures of purifying selection in the mitochondrial genomes of three Andean waterfowl. PLoS ONE. 2024;19(1): e0294842. https://doi.org/10.1371/journal.pone.0294842.
42. Sebastian W, Sukumaran S, Zacharia PU, Muraleedharan KR, Dinesh Kumar PK, Gopalakrishnan A. Signals of selection in the mitogenome provide insights into adaptation mechanisms in heterogeneous habitats in a

widely distributed pelagic fish. Sci Rep. 2020;10(1):9081. https://doi.org/10.1038/s41598-020-65905-1.
43. Li L-L, Xiao Y, Wang X, He Z-H, Lv Y-W, Hu X-S. The *Ka* /*Ks* and *πa* /*πs* ratios under different models of gametophytic and sporophytic selection. Genome Biology and Evolution. 2023;15(8):evad151. https://doi.org/10.1093/gbe/evad151
44. Yang Z, Bielawski JP. Statistical methods for detecting molecular adaptation. Trends Ecol Evol. 2000;15(12):496–503. https://doi.org/10.1016/s0169-5347(00)01994-7.
45. Jeedigunta SP, Minenkova AV, Palozzi JM, Hurd TR. Avoiding extinction: Recent advances in understanding mechanisms of mitochondrial DNA purifying selection in the germline. Annu Rev Genomics Hum Genet. 2021;22(1):55–80. https://doi.org/10.1146/annurev-genom-121420-081805.
46. Zhang D-X, Hewitt GM. Insect mitochondrial control region: A review of its structure, evolution and usefulness in evolutionary studies. Biochem Syst Ecol. 1997;25(2):99–120. https://doi.org/10.1016/s0305-1978(96)00042-7.
47. Kern EMA, Kim T, Park J-K. The mitochondrial genome in Nematode phylogenetics. Front Ecol Evol. 2020;8:250. https://doi.org/10.3389/fevo.2020.00250.
48. Ghiselli F, Gomes-dos-Santos A, Adema CM, Lopes-Lima M, Sharbrough J, Boore JL. Molluscan mitochondrial genomes break the rules. Philosophical Transactions of the Royal Society B: Biological Sciences. 1825;2021(376):20200159. https://doi.org/10.1098/rstb.2020.0159.
49. Cai ZF, Hu JY, Yin TT, Wang D, Shen QK, Ma C, Ou DQ, Xu MM, Shi X, Li QL, et al. Long amplicon HiFi sequencing for mitochondrial DNA genomes. Mol Ecol Resour. 2023;23(5):1014–22. https://doi.org/10.1111/1755-0998.13765.
50. Nabholz B, Mauffrey J-Fo, Bazin E, Galtier N, Glemin S. Determination of mitochondrial genetic diversity in mammals. Genetics. 2008;178(1):351–361. https://doi.org/10.1534/genetics.107.073346
51. Clark A, Koc G, Eyre-Walker Y, Eyre-Walker A, Andres A. What determines levels of mitochondrial genetic diversity in birds? Genome Biology and Evolution. 2023;15(5):evad064. https://doi.org/10.1093/gbe/evad064
52. French CM, Bertola LD, Carnaval AC, Economo EP, Kass JM, Lohman DJ, Marske KA, Meier R, Overcast I, Rominger AJ, et al. Global determinants of insect mitochondrial genetic diversity. Nat Commun. 2023;14(1):5276. https://doi.org/10.1038/s41467-023-40936-0.
53. Das J. The role of mitochondrial respiration in physiological and evolutionary adaptation. BioEssays. 2006;28(9):890–901. https://doi.org/10.1002/bies.20463.
54. Bennett CF, Latorre-Muro P, Puigserver P. Mechanisms of mitochondrial respiratory adaptation. Nat Rev Mol Cell Biol. 2022;23(12):817–35. https://doi.org/10.1038/s41580-022-00506-6.
55. Thorpe SE, Murphy EJ, Watkins JL. Circumpolar connections between Antarctic krill (*Euphausia superba* Dana) populations: Investigating the roles of ocean and sea ice transport. Deep Sea Res Part I. 2007;54(5):792–810. https://doi.org/10.1016/j.dsr.2007.01.008.
56. Waldvogel A-M, Feldmeyer B, Rolshausen G, Exposito-Alonso M, Rellstab C, Kofler R, Mock T, Schmid K, Schmitt I, Bataillon T, et al. Evolutionary genomics can improve prediction of species' responses to climate change. Evolution Letters. 2020;4(1):4–18. https://doi.org/10.1002/evl3.154.
57. Franks SJ, Hoffmann AA. Genetics of climate change adaptation. Annu Rev Genet. 2012;46(1):185–208. https://doi.org/10.1146/annurev-genet-110711-155511.
58. Venverloo T, Duarte F. Towards real-time monitoring of insect species populations. Sci Rep. 2024;14(1):18727. https://doi.org/10.1038/s41598-024-68502-8.
59. Jeffery NW. The first genome size estimates for six species of krill (Malacostraca, Euphausiidae): large genomes at the north and south poles. Polar Biol. 2012;35(6):959–62. https://doi.org/10.1007/s00300-011-1137-4.
60. Li H. Minimap2: pairwise alignment for nucleotide sequences. Bioinformatics. 2018;34(18):3094–100. https://doi.org/10.1093/bioinformatics/bty191.
61. Robinson JT, Thorvaldsdóttir H, Winckler W, Guttman M, Lander ES, Getz G, Mesirov JP. Integrative genomics viewer. Nat Biotechnol. 2011;29(1):24–6. https://doi.org/10.1038/nbt.1754.
62. Durbin R, De Sanctis B, Blumer M. Rotate: A command-line program to rotate circular DNA sequences to start at a given position or string. Wellcome Open Research. 2023;8. https://doi.org/10.12688/wellcomeopenres.19568.1
63. Tillich M, Lehwark P, Pellizzer T, Ulbricht-Jones ES, Fischer A, Bock R, Greiner S. GeSeq – versatile and accurate annotation of organelle genomes. Nucleic Acids Res. 2017;45(W1):W6–11. https://doi.org/10.1093/nar/gkx391.
64. Benson G. Tandem repeats finder: a program to analyze DNA sequences. Nucleic Acids Res. 1999;27(2):573–80. https://doi.org/10.1093/nar/27.2.573.
65. Zheng S, Poczai P, Hyvönen J, Tang J, Amiryousefi A. Chloroplot: An online program for the versatile plotting of organelle genomes. Front Genet. 2020;11: 576124. https://doi.org/10.3389/fgene.2020.576124.
66. Noe L, Kucherov G. YASS: enhancing the sensitivity of DNA similarity search. Nucleic Acids Research. 2005;33(Web Server):W540-W543. https://doi.org/10.1093/nar/gki478
67. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R. The Sequence Alignment/Map format and SAMtools. Bioinformatics. 2009;25(16):2078–9. https://doi.org/10.1093/bioinformatics/btp352.
68. Edgar RC. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res. 2004;32(5):1792–7. https://doi.org/10.1093/nar/gkh340.
69. Capella-Gutiérrez S, Silla-Martínez JM, Gabaldón T. trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. Bioinformatics. 2009;25(15):1972–3. https://doi.org/10.1093/bioinformatics/btp348.
70. Minh BQ, Schmidt HA, Chernomor O, Schrempf D, Woodhams MD, von Haeseler A, Lanfear R. IQ-TREE 2: New models and efficient methods for phylogenetic inference in the genomic era. Mol Biol Evol. 2020;37(5):1530–4. https://doi.org/10.1093/molbev/msaa015.
71. Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, Madden TL. BLAST+: architecture and applications. BMC Bioinformatics. 2009;10(1):1–9. https://doi.org/10.1186/1471-2105-10-421.
72. Shen W, Le S, Li Y, Hu F. SeqKit: A cross-platform and ultrafast toolkit for FASTA/Q file manipulation. PLoS ONE. 2016;11(10): e0163962. https://doi.org/10.1371/journal.pone.0163962.
73. Quinlan AR. BEDTools: The swiss-army tool for genome feature analysis. Current Protocols in Bioinformatics. 2014;47(1):11.12. 11–11.12. 34. https://doi.org/10.1002/0471250953.bi1112s47
74. Quinlan AR, Hall IM. BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics. 2010;26(6):841–2. https://doi.org/10.1093/bioinformatics/btq033.
75. Wilkinson L. ggplot2: Elegant graphics for data analysis by WICKHAM. H Biometrics. 2011;67(2):678–9. https://doi.org/10.1111/j.1541-0420.2011.01616.x.
76. Kassambara A. ggpubr: 'ggplot2' based publication ready plots. Version 0.6.0 [R package]. 2023. https://doi.org/10.32614/CRAN.package.ggpubr
77. Chen S, Zhou Y, Chen Y, Gu J. fastp: an ultra-fast all-in-one FASTQ preprocessor. Bioinformatics. 2018;34(17):i884–90. https://doi.org/10.1093/bioinformatics/bty560.
78. Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics. 2009;25(14):1754–60. https://doi.org/10.1093/bioinformatics/btp324.
79. de Flamingh A, Rivera-Colón AG, Gnoske TP, Kerbis Peterhans JC, Catchen J, Malhi RS, Roca AL, Lopez J. Numt Parser: Automated identification and removal of nuclear mitochondrial pseudogenes (numts) for accurate mitochondrial genome reconstruction in Panthera. J Hered. 2023;114(2):120–30. https://doi.org/10.1093/jhered/esac065.
80. Jin J-J, Yu W-B, Yang J-B, Song Y, dePamphilis CW, Yi T-S, Li D-Z. GetOrganelle: a fast and versatile toolkit for accurate de novo assembly of organelle genomes. Genome Biol. 2020;21(1):1–31. https://doi.org/10.1186/s13059-020-02154-5.
81. Danecek P, Bonfield JK, Liddle J, Marshall J, Ohan V, Pollard MO, Whitwham A, Keane T, McCarthy SA, Davies RM *et al*. Twelve years of SAMtools and BCFtools. GigaScience. 2021;10(2):giab008. https://doi.org/10.1093/gigascience/giab008
82. Lischer HEL, Excoffier L. PGDSpider: an automated data conversion tool for connecting population genetics and genomics programs. Bioinformatics. 2012;28(2):298–9. https://doi.org/10.1093/bioinformatics/btr642.

83. Librado P, Rozas J. DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. Bioinformatics. 2009;25(11):1451–2. https://doi.org/10.1093/bioinformatics/btp187.
84. Excoffier L, Lischer HEL. Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. Molecular Ecology Resources. 2010;10(3):564–567. https://doi.org/10.1111/j.1755-0998.2010.02847.x
85. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. J Roy Stat Soc: Ser B (Methodol). 1995;57(1):289–300.
86. Narum SR. Beyond Bonferroni: Less conservative analyses for conservation genetics. Conserv Genet. 2006;7(5):783–7. https://doi.org/10.1007/s10592-005-9056-y.
87. Winter DJ. mmod: an R library for the calculation of population differentiation statistics. Mol Ecol Resour. 2012;12(6):1158–60. https://doi.org/10.1111/j.1755-0998.2012.03174.x.
88. Paradis E, Schliep K. ape 5.0: an environment for modern phylogenetics and evolutionary analyses in R. Bioinformatics. 2019;35(3):526–528. https://doi.org/10.1093/bioinformatics/bty633
89. Jombart T. adegenet: a R package for the multivariate analysis of genetic markers. Bioinformatics. 2008;24(11):1403–5. https://doi.org/10.1093/bioinformatics/btn129.
90. Dray S, Dufour A-B. The ade4 package: Implementing the duality diagram for ecologists. Journal of Statistical Software. 2007;22(4):1–20. https://doi.org/10.18637/jss.v022.i04
91. Lanfear R, Frandsen PB, Wright AM, Senfeld T, Calcott B. PartitionFinder 2: New methods for selecting partitioned models of evolution for molecular and morphological phylogenetic analyses. Mol Biol Evol. 2016;34(3):772–3. https://doi.org/10.1093/molbev/msw260.
92. Ronquist F, Teslenko M, van der Mark P, Ayres DL, Darling A, Höhna S, Larget B, Liu L, Suchard MA, Huelsenbeck JP. MrBayes 3.2: Efficient Bayesian phylogenetic inference and model choice across a large model space. Systematic Biology. 2012;61(3):539–542. https://doi.org/10.1093/sysbio/sys029
93. Leigh JW, Bryant D. popart: full-feature software for haplotype network construction. Methods Ecol Evol. 2015;6(9):1110–6. https://doi.org/10.1111/2041-210x.12410.
94. Templeton AR, Crandall KA, Sing CF. A cladistic analysis of phenotypic associations with haplotypes inferred from restriction endonuclease mapping and DNA sequence data. III Cladogram estimation Genetics. 1992;132(2):619–33. https://doi.org/10.1093/genetics/132.2.619.
95. Tajima F. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. Genetics. 1989;123(3):585–95. https://doi.org/10.1093/genetics/123.3.585.
96. Fu Y-X. Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. Genetics. 1997;147(2):915–25. https://doi.org/10.1093/genetics/147.2.915.
97. Drummond AJ, Rambaut A, Shapiro B, Pybus OG. Bayesian coalescent inference of past population dynamics from molecular sequences. Mol Biol Evol. 2005;22(5):1185–92. https://doi.org/10.1093/molbev/msi103.
98. Bouckaert R, Vaughan TG, Barido-Sottani J, Duchêne S, Fourment M, Gavryushkina A, Heled J, Jones G, Kühnert D, De Maio N *et al*. BEAST 2.5: An advanced software platform for Bayesian evolutionary analysis. PLOS Computational Biology. 2019;15(4):e1006650. https://doi.org/10.1371/journal.pcbi.1006650

Publisher's Note