

RESEARCH ARTICLE

Open Access



# Landscape genomics: natural selection drives the evolution of mitogenome in penguins

Barbara Ramos<sup>1,2</sup>, Daniel González-Acuña<sup>3</sup>, David E. Loyola<sup>4,5</sup>, Warren E. Johnson<sup>6</sup>, Patricia G. Parker<sup>7</sup>, Melanie Massaro<sup>8</sup>, Gisele P. M. Dantas<sup>9</sup>, Marcelo D. Miranda<sup>1</sup> and Juliana A. Vianna<sup>1,10\*</sup>

## Abstract

**Background:** Mitochondria play a key role in the balance of energy and heat production, and therefore the mitochondrial genome is under natural selection by environmental temperature and food availability, since starvation can generate more efficient coupling of energy production. However, selection over mitochondrial DNA (mtDNA) genes has usually been evaluated at the population level. We sequenced by NGS 12 mitogenomes and with four published genomes, assessed genetic variation in ten penguin species distributed from the equator to Antarctica. Signatures of selection of 13 mitochondrial protein-coding genes were evaluated by comparing among species within and among genera (*Spheniscus*, *Pygoscelis*, *Eudyptula*, *Eudyptes* and *Aptenodytes*). The genetic data were correlated with environmental data obtained through remote sensing (sea surface temperature [SST], chlorophyll levels [Chl] and a combination of SST and Chl [COM]) through the distribution of these species.

**Results:** We identified the complete mtDNA genomes of several penguin species, including *ND6* and 8 tRNAs on the light strand and 12 protein coding genes, 14 tRNAs and two rRNAs positioned on the heavy strand. The highest diversity was found in NADH dehydrogenase genes and the lowest in *COX* genes. The lowest evolutionary divergence among species was between Humboldt (*Spheniscus humboldti*) and Galapagos (*S. mendiculus*) penguins (0.004), while the highest was observed between little penguin (*Eudyptula minor*) and Adélie penguin (*Pygoscelis adeliae*) (0.097). We identified a signature of purifying selection ( $Ka/Ks < 1$ ) across the mitochondrial genome, which is consistent with the hypothesis that purifying selection is constraining mitogenome evolution to maintain Oxidative phosphorylation (OXPHOS) proteins and functionality. Pairwise species maximum-likelihood analyses of selection at codon sites suggest positive selection has occurred on *ATP8* (Fixed-Effects Likelihood, FEL) and *ND4* (Single Likelihood Ancestral Counting, SLAC) in all penguins. In contrast, *COX1* had a signature of strong negative selection. *ND4*  $Ka/Ks$  ratios were highly correlated with SST (Mantel,  $p$ -value: 0.0001; GLM,  $p$ -value: 0.00001) and thus may be related to climate adaptation throughout penguin speciation.

**Conclusions:** These results identify mtDNA candidate genes under selection which could be involved in broad-scale adaptations of penguins to their environment. Such knowledge may be particularly useful for developing predictive models of how these species may respond to severe climatic changes in the future.

**Keywords:** Comparative mitogenomics, Selection, Penguins, Adaptation

\* Correspondence: [jvianna@uc.cl](mailto:jvianna@uc.cl)

<sup>1</sup>Departamento de Ecosistemas y Medio Ambiente, Facultad de Agronomía e Ingeniería Forestal, Pontificia Universidad Católica de Chile, Av. Vicuña Mackenna, 4860 Santiago, Chile

<sup>10</sup>Centro de Cambio Global UC, Santiago, Chile

Full list of author information is available at the end of the article



## Background

The mitochondrial genome (mtDNA) is under continuous selection since the 13 protein-coding genes produce polypeptide products that work in association with nuclear-encoded subunits of protein complexes involved in oxidative phosphorylation (OXPHOS) [1]. Three types of hypotheses have been proposed to explain mitogenome evolution: 1) mitochondrial OXPHOS generates both energy and heat, and environmental temperature imposes a selective trade off between ATP production and heat [2–6]; 2) risk of starvation is associated with restricted food availability, since starvation can generate more efficient coupling of energy production by OXPHOS pathway [7–10]; 3) a more recent and debated hypothesis of immune response to pathogens [11–15]. Therefore, chlorophyll (Chl), as a measurement of primary productivity [16], and sea surface temperature (SST), are important environmental variables for marine organisms that are directly related to the efficiency of the OXPHOS and relevant to understanding mtDNA gene selection patterns. As OXPHOS is the terminal stage of cellular respiration, disruption of this process will be lethal for individual cells and if it occurs widely in an individual, it will adversely impact its fitness [3]. Consequently, the mitochondrial genome on one hand evolves under continuous purifying selection (or negative selection) that removes deleterious variation from populations [17], while on the other hand, beneficial variation in extreme temperature environments may be under positive selection, increasing in frequency in a population or species. It is possible that selection may operate similarly among species in similar environments or it may drive adaptations in different genes.

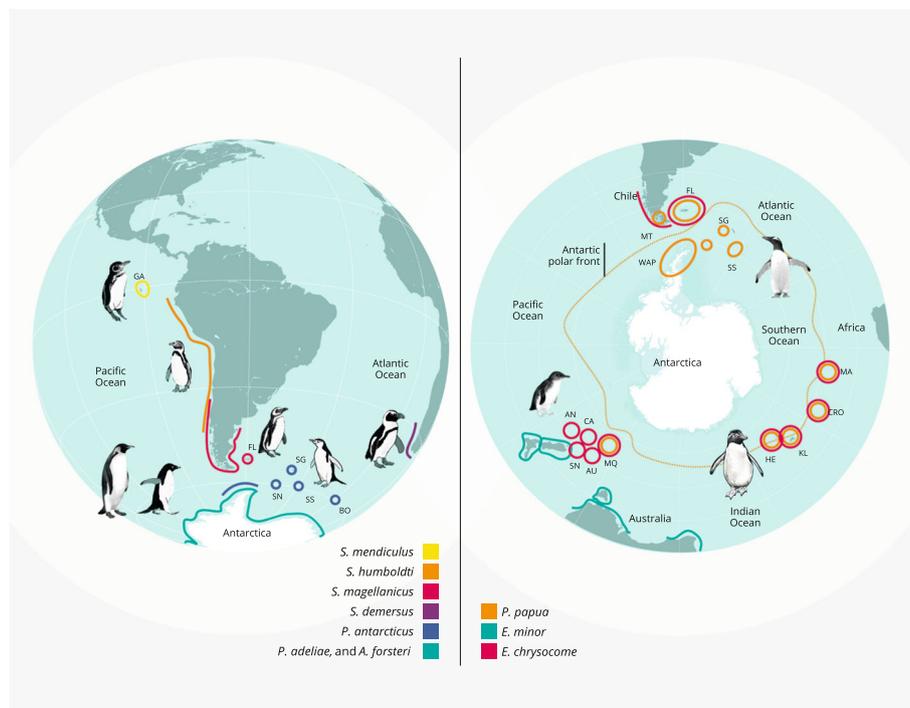
Patterns of adaptation in the mtDNA of organisms have been evaluated mostly at the population level by associating each population or lineage with local climatic conditions [8, 18–21]; however, only relatively few have incorporated environmental data in the genetic analyses and generally only using nuclear genes or whole-genome data [22, 23]. Environmental temperature is the main variable that has been linked to selection in mtDNA genes [8, 18–20], along with altitude [24–26] and oxygen availability [27]. Although most studies on the role of selection on mtDNA genome evolution in vertebrate species have focused on mammals at the intraspecific level [28–31], recently there has been an increased focus on birds [32–34]. Here, we provide one of the first species-level analyses in birds linking patterns of mtDNA selection with environmental variables (temperature, chlorophyll) which are associated with the evolution of this genome. Chlorophyll (Chl) is an excellent indirect measurement of food availability because primary productivity is associated with species distributions at trophic levels, including krill and penguins [16].

Penguins inhabit a broad latitudinal distribution in the Southern Hemisphere from the tropics to the extreme

cold of the Antarctic. While all penguin species can live and forage at sea in cold, nutrient-rich water (including those breeding on the Galapagos Islands), they breed on land under varying climatic conditions. The northernmost species is the Galapagos penguin (*Spheniscus mendiculus*), while the southernmost distributed species, the Adélie penguin (*P. adeliae*) and Emperor penguin (*Aptenodytes forsteri*), breed in the Antarctic, south of the Antarctic Convergence (Polar Front). Emperor penguins breed on stable fast ice during the Antarctic winter when temperatures drop regularly to  $-50^{\circ}\text{C}$ . As the evolution, speciation and extinction of penguins are closely related to historical climatic changes [35–38], the broad distribution of penguins, which includes different and extreme climatic conditions, provides a unique opportunity to identify molecular genetic patterns associated with these climatic conditions.

We used Next Generation Sequencing approaches to study patterns of selection in the mitochondrial genome of penguins from two genera (*Spheniscus* and *Pygoscelis*) and correlated the results with remote sensing spatial data for each species from throughout their distribution (sea surface temperature, chlorophyll levels, and both variables combined). The four *Spheniscus* penguin species are distributed along the coast of southern Africa (African penguin *S. demersus*) and along the South America coast from Galapagos Islands in the tropics (Galapagos penguin *S. mendiculus*) to southern Patagonia (Humboldt *S. humboldti* and Magellanic penguin *S. magellanicus*) (Fig. 1). Galapagos and Humboldt penguin populations have experienced high mortality rates [39, 40] and decreased breeding success [39, 41] from increased water temperatures during El Niño Southern Oscillations (ENSO) [42]. The three *Pygoscelis* penguins breed on sub-Antarctic islands and the coast of Antarctica and they have sympatric distributions in the Western Antarctic Peninsula (WAP). The Gentoo penguin (*P. papua*) has the northernmost range in the genus, with a distribution along sub-Antarctic Islands and the Antarctic Peninsula. The Chinstrap penguin (*P. antarcticus*) has a more southern distribution, almost exclusively around the Antarctic peninsula. The Adélie penguin is the most southern and pagophilic species, with a circumpolar distribution [43]. Over the last decades climate change has impacted the WAP, increasing sea surface temperatures (SST) and changing patterns of penguin food availability (e.g. krill) [44–46]. Concurrently, Adélie and Chinstrap penguin populations have been decreasing, while Gentoo penguin populations have increased in numbers and distributions [44–46].

Historical climate changes have been linked with penguin species diversification. Severe declines in Antarctic temperature that began approximately 12 mya coincide with the divergence of the four major penguin groups, around 11–16 mya [38]. This penguin phylogeny is well



**Fig. 1** Map of studied penguin distribution. Locality where samples were collected (indicated with a star): Galapagos (red), Humboldt (blue), Magellanic (green), African (orange), Gentoo (pink), Chinstrap (dashed line), Adélie penguin (yellow). Antarctic Polar Front (yellow line), locations are indicated as: Martillo Island (MT), Falkland/Malvinas Islands (FL), South Shetland Islands and Western Antarctic Peninsula (WAP), Scotia Arc (SN, South Orkney Islands; SS, South Sandwich Islands; SG, South Georgia), Bouvet Island (BO), Crozet Islands (CRO), Kerguelen Islands (KE), Marion and Prince Edward Islands (MA), Heard Islands (HE), Macquarie Islands (MQ), Snares Islands (SN), Auckland Island (AU), Antipodes (AN), Campbell (CA). By Juan Pablo Bravo

characterized, making it possible to assess evidence of selection and adaptive evolution of the mtDNA and to explore possible links with environmental changes, including genes involved in penguin thermal adaptation. Sequences of entire mitochondrial genomes are available for five of the 18 extant species of penguins [47–50], including the Adélie, Chinstrap, African, Little, and Southern Rockhopper (*E. chrysocome*) penguins and more recently the entire genome became available for the Adélie and Emperor penguin [51]. The mtDNA genome of *P. antarcticus* consists of 15,972 bp and has 94% identity with its sister species *P. adeliae* [50]. Subramanian et al. [49] compared the mitochondrial genomes of modern and ancient (up to 44,000 years old) *P. adeliae*, documenting higher evolutionary rates than described earlier by phylogenetic studies. However, those studies evaluated the genome without a broader comparative approach to understand the evolution of the group and assess the evidence of signatures of selection on the mitochondrial genome.

We hypothesize that selection operates differentially on the mtDNA coding genes of penguin species in response to environmental variables. We evaluated variation in mitochondrial genomes at different taxonomic

levels in seven species of penguins from South America and Africa (all four *Spheniscus* species) and Antarctica (all three *Pygoscelis* species). We then compared our results with two penguin species from temperate and sub-Antarctic regions (*E. minor* and *E. chrysocome*) and with *A. forsteri*, which survives and breeds during the most-extreme cold conditions of the Antarctic winter. We evaluated evolutionary rates and signatures of natural selection for all 13 mtDNA protein-coding genes and across the mtDNA genome phylogeny for the penguin species. Lastly, we assessed the genetic data for each gene against the spatial environmental data (sea surface temperature, chlorophyll and both variables combined) from remote sensing and GIS databases for each species along its distributional range.

## Methods

We sequenced the entire mtDNA genome of 12 individuals from 6 species of penguins: the Galapagos (*S. mendiculus*,  $n = 1$ ), Humboldt (*S. humboldti*,  $n = 2$ ), Magellanic (*S. magellanicus*,  $n = 3$ ), Adélie (*P. adeliae*,  $n = 2$ ), Gentoo (*P. papua*,  $n = 2$ ) and Chinstrap penguins (*P. antarcticus*,  $n = 2$ ). We sequenced individuals from different latitudes or the extreme limits of their distribution (Fig. 1, Tables 1, 2).

**Table 1** Latitudinal range, adult body mass, adult length and broad climatic preference for each penguin species [79, 110], [Vianna pers. comm.]

Genus	Species	Common Name	Geographic range		Body Mass (kg)		Length (cm)	Climate
			Latitude (S)	Longitude (W)	range	Average		
<i>Pygoscelis</i>	<i>antarcticus</i>	Chinstrap Penguin	54°	64°	3.5–5.5	4.1	61–76	Temperate/Polar
	<i>papua</i>	Gentoo Penguin	46°	65°	3.3–8.8	6.1	56–83	Temperate/Polar
	<i>adeliae</i>	Adélie Penguin	54°	75°	4.6–5.4	5	46–71	Polar
<i>Spheniscus</i>	<i>mendiculus</i>	Galápagos Penguin	0°	0°	1.4–2.6	2	46–55	Tropical
	<i>humboldti</i>	Humboldt Penguin	5°	42°	4.2–5.0	4.6	67–72	Tropical/Temperate
	<i>magellanicus</i>	Magellanic Penguin	40°	55°	3.2–5.2	4	48	Temperate
	<i>demersus</i>	African Penguin	25°	33°	2.7–3.4	3.1	60–70	Subtropical/Temperate
<i>Eudyptula</i>	<i>minor</i>	Little Penguin	30°	47°	0.7–1.3	1.2	33	Temperate
<i>Eudyptes</i>	<i>chrysocome</i>	Southern Rockhopper Penguin	46°	54°	2–3.8	3.4	45–55	Temperate
<i>Aptenodytes</i>	<i>forsteri</i>	Emperor Penguin	64°	77°	22–40	28	120	Polar

**Table 2** Details of the individuals sequenced

Genus	Species	Locality	Geographic position		Total genomic reads	MtDNA		Coverage	Genome length (bp)		Genbank
			Latitude (S)	Longitude (W)		Total Reads	Used Reads		all genes	+ D-loop	
<i>Pygoscelis</i>	<i>P. antarcticus</i>	Antarctica, South Shetland Islands, Narebski base	62° 13'	58° 45'	80,496,366	15,788	11,347	36	15,491	16,664	KU356673
		Antarctic Peninsula, Kopaitic	63° 19'	57° 54'	52,990,428	9298	6359	20	15,490	16,653	KU356674
	<i>P. papua</i>	Elephant Island, South Shetlands Islands	61° 06'	55° 09'	7,341,640	2630	1647	7 <sup>a</sup>	–	–	–
	<i>P. adeliae</i>	Antarctic Peninsula, Gabriel Gonzalez Videla base	64° 48'	62° 51'	158,565,602	34,092	25,947	82	15,500	16,704	KU356677
		Antarctica, South Shetland Islands, Arctowski base	62° 09'	58° 28'	52,997,551	8162	5049	16	15,493	16,764	KU356675
		Antarctica, Ross Sea, Cape Crozier	77° 29'	69° 20'	68,998,340	13,000	7956	26	15,493	16,765	KU356676
<i>Spheniscus</i>	<i>S. mendiculus</i>	Galapagos Islands, Isabela Island	00° 35'	91° 05'	69,908,826	13,006	7360	24	15,516	16,367	KU361807
	<i>S. humboldti</i>	South America, Peru, Punta San Juan	15° 21'	75° 11'	64,663,956	21,786	13,908	43	15,519	16,658	KM891593
		South America, Chile, Cachagua	32° 35'	71° 27'	55,087,204	14,606	9269	28	15,519	16,660	KU361805
	<i>S. magellanicus</i>	South America, Chile, Chiloé Island, Puñihuil	41° 55'	74° 02'	61,046,791	11,456	5984	19	15,520	16,427	KU361804
		South America, Chile, Magdalena Island	52° 51'	70° 36'	54,372,256	11,108	6490	20	15,514	16,659	KU361806
		South America, Argentina, Puerto Deseado	47° 45'	65° 43'	67,969,432	10,640	5483	18	15,520	16,549	KU361803

<sup>a</sup>Genome not included in all analyses due the low coverage

Adélie samples were obtained from the two divergent lineages identified by Ritchie et al. [52] using D-loop mtDNA.

### SOLiD sequencing

Genomic DNA was isolated from blood samples obtained from all penguins and preserved in ethanol using the salt method [53]. DNA from six individuals of the three species was quantified and quality checked with Invitrogen's PicoGreen<sup>®</sup> assay kit (fluorometry). Genomic sequencing in ABI-SOLiD<sup>™</sup> 5500 XL was performed at Omics Solution, a Next Generation Sequencing facility. DNA was desalted then concentrated using standard EtOH/sodium acetate precipitation at 20 °C for 2 h, followed by two 70% EtOH washes. Then the pooled DNA was re-dissolved in low TE as per standard protocol for ABI SOLiD sequencing of genomic DNA fragment libraries. DNA samples were sheared in Covaris<sup>™</sup>S220 System, which fragments the input DNA, by sonication, into small fragments with a mean size of 160 bp.

The fragmented DNA was then purified with SOLiD Library Column Purification Kit, and libraries were prepared according to standard SOLiD protocols. Fragment libraries for the twelve penguins were prepared separately, P1 and P2 adapters were ligated, and each sample was tagged with a different barcode (a known sequence of ten bp in adapters). Prepared libraries were quantified by real-time PCR on a Light Cycler<sup>®</sup>Nano (Roche) using Invitrogen's Quantification Kit for SOLiD. The double stranded library was added at a concentration of 0.2 pg/μL to the emulsion with 2400 million beads, according to manufacturers' instructions. Thirty percent of the beads were P2 positive (contained amplified library fragments) before enrichment and 90% of the beads were P2 positive after enrichment, yielding 790 million beads deposited on the Flow Chip. Library beads were sequenced on a SOLiD<sup>™</sup> 5500 XL using standard chemistry for paired-end fragment libraries and 75–35-bp read lengths.

### SOLiD sequence alignment

The color-space reads (di-base encoded) were aligned with LifeScope<sup>™</sup> (Applied Biosystems) using the genome assembly of *P. adeliae* [51] as a reference. The reference was translated into color-space with the aim of mapping the reads. The color-space reads helped to improve the quality of each base call, since each base was read twice during the sequencing step.

The consensus sequence was built from the binary alignment map (BAM) files obtained in the previous step. We used SAMtools [54] to obtain all bases mapped to each position, BCF tools to get the most probable genotype per position, and VCF utils to build the consensus sequence in FASTQ format. The FASTQ file was then converted to FASTA using SEQTK.

### Mitochondrial sequence assembly

The mitochondrial genomes were assembled with the Denovo2 pipeline developed by Life Technologies using Velvet assembler version 1.2.09 [55] in color-space mode. ASiD (Assembly Assistant for SOLiD) was used to fill gaps between contigs that formed scaffolds and to perform a color-space to base-space translation for the final assembly. Individual genes were selected from each of the penguins' contigs using Sequencher software 5.1 (Gene Codes Corp., MI) and annotated by homology with genes of mitogenomes of the penguins previously published [48, 51, 56, 57]. Due to the low number of reads from one individual of *P. papua* (Table 2), this individual was not included in the data analyses. All 11 mitogenomes were submitted to GenBank (accession numbers KM891593; KU361803-KU361806; KU356673-KU356677).

### Molecular evolution and selection

Sequences were aligned and polymorphic sites were confirmed by visual review of the chromatograms using Sequencher v. 5.1. (Gene Codes, Ann Arbor, MI, USA) and multiple alignments were done using ClustalX v. 2.1 [58]. For interspecific data analysis we compared the eleven mtDNA genomes with four other penguins genomes: *S. demersus* obtained from KwaZulu-Natal Province of South Africa (African penguin) [56], *E. minor* obtained from Nelson in New Zealand (Little blue penguin) [48], *E. chrysocome* (Rockhopper penguin) [57] and *A. forsteri* (Emperor penguin) [51].

The number of indels, polymorphic sites (*S*), nucleotide diversity ( $\pi$ ), synonymous mutation rates (*K*<sub>s</sub>), nonsynonymous mutation rates (*K*<sub>a</sub>) and the *K*<sub>a</sub>/*K*<sub>s</sub> (or  $\omega$ ) ratio using the DNAsp v. 5 [59] were estimated for all 13 protein-coding genes from the complete mtDNA. For analyses of the reading frame of the genes, such as *K*<sub>a</sub>/*K*<sub>s</sub>, only the *ND3* gene was considered a different reading frame as described by Mindell et al. [60]. All *K*<sub>a</sub> and *K*<sub>s</sub> estimates and corresponding ratios were obtained between pairwise species within *Spheniscus* and *Pygoscelis* genera, as well as between all five genera (Additional file 1 Table S1). Since we sequenced several species of *Spheniscus* and *Pygoscelis* genera, the values between genera were averages of the pairwise species *K*<sub>a</sub> and *K*<sub>s</sub> values for the genera comparison. We also calculated the pairwise distance between all 15 penguins studied here (10 species) for all 13 concatenated genes of the mtDNA genomes using the Maximum Likelihood model (Additional file 2 Table S2) with Mega v. 6.06 software [61, 62] and Arlequin software in R package [63, 64]. All positions containing gaps were eliminated, remaining a total of 10,874 positions in the final dataset.

Phylogenetic reconstruction was used to evaluate selection on each codon, gene and phylogenetic branch. We combined each of the 13 protein-coding genes from the mtDNA genome from the studied penguin species into a

concatenated data set of 11,409 bp. The phylogeny was also estimated excluding the *ND6* gene (a total of 10,886 bp), which is encoded by the light-strand [49] and has a significantly different base composition from the heavy-chain [65]. Phylogenetic relationships were inferred using Bayesian inference methods (BA) with Markov chain Monte Carlo (MCMC) sampling as implemented in MrBayes v. 3.1.2 [66, 67]. The evolutionary model selected GTR + I + G, using jModelTest v. 0.1.1 [68] and Akaike Information Criterion (AIC). The Mr. Bayes analyses were run for one million generations while sampling every 100 generations. At this point the standard deviation of split frequencies was <0.01, indicating that both runs had converged. Phylogenetic reconstructions were visualized using Figtree v. 1.2.2 [69]. Although positive selection is frequently inferred by genetic data analysis as  $Ka/Ks > 1$  (and purifying selection by  $Ka/Ks < 1$ ) [70], a gene with  $Ka/Ks$  lower than 1 may show signature of positive selection at codon level. Evidence for signatures of selection at codon sites was evaluated for each gene with the concatenated data set using the BA phylogeny and the selected model of evolution without outgroups with the three algorithms available in HyPhy software [71–73]. For comparison, we used the Single Likelihood Ancestral Counting (SLAC), Fixed-Effects Likelihood (FEL) and the Random Effects Likelihood (REL) methods using the best fit nucleotide model [71]. The single likelihood ancestor counting (SLAC) method is based on weighting across all possible ancestral reconstructions, or employs sampling from ancestral reconstructions. The fixed effects likelihood (FEL) method estimates nonsynonymous and synonymous substitution rates at each site. Random effects likelihood (REL) models variation in nonsynonymous and synonymous rates across sites based on a predefined distribution, where a Bayesian approach are used to infer selection pressure at an individual site [71]. We also performed the branch Site REL [72] analysis in HyPhy software to estimate branch-specific episodes of positive selection along the phylogeny.

#### Environmental data and species distribution

Seasonal Chlorophyll concentrations (Chl) and Sea Surface Temperatures (SST) were acquired from Sea-viewing Wide Field-of-view Sensor (SeaWiFS) and Moderate Resolution Imaging Spectroradiometer (MODIS-Aqua) global seasonal databases that are based on a spatial resolution of  $9 \times 9$  km [74]. The standard SeaWiFS chlorophyll product,  $Chl_{OC4}$ , is estimated based on an empirical maximal band-ratio algorithm that uses remote sensing reflectance measurements at 443, 490, 510 and 555 nm [75]. This algorithm relates the maximum of three remote sensing reflectance band-ratios (443/555, 490/555 and 510/555) to field measures of sea chlorophyll concentration [76] SST is based on MODIS-calibrated mid- and far-infrared radiances (IR),

estimated using an algorithm that exploits differences in atmospheric transmissivity in the different IR bands [77]. The reliability of SST data is supported by its high correlation with sea surface temperatures measured during both oceanographic cruises and at local stations, with a slight positive bias at higher temperatures [78]. Here we use Chl and SST as proxies for the environmental conditions along the distributional range of each penguin species.

Satellite data (Chl and SST) were processed and saved in image format using SAGA - System for Automated Geoscientific Analyses (SAGA Development Team 2008). Penguin at-sea distribution during breeding were obtained from Borboroglu and Boersma [79] and digitized in vector format using a Geographic Information System QGIS 2.6 [80]. Mean and standard deviation of seasonal Chl and SST products were obtained for each species by intersecting seasonal images with the vector files representing the geographical distribution of each species (Fig. 2). All data were re-scaled into a range from 0 to 1 to be comparable the Chl (Fig. 2a, c) and SST data (Fig. 2b, d).

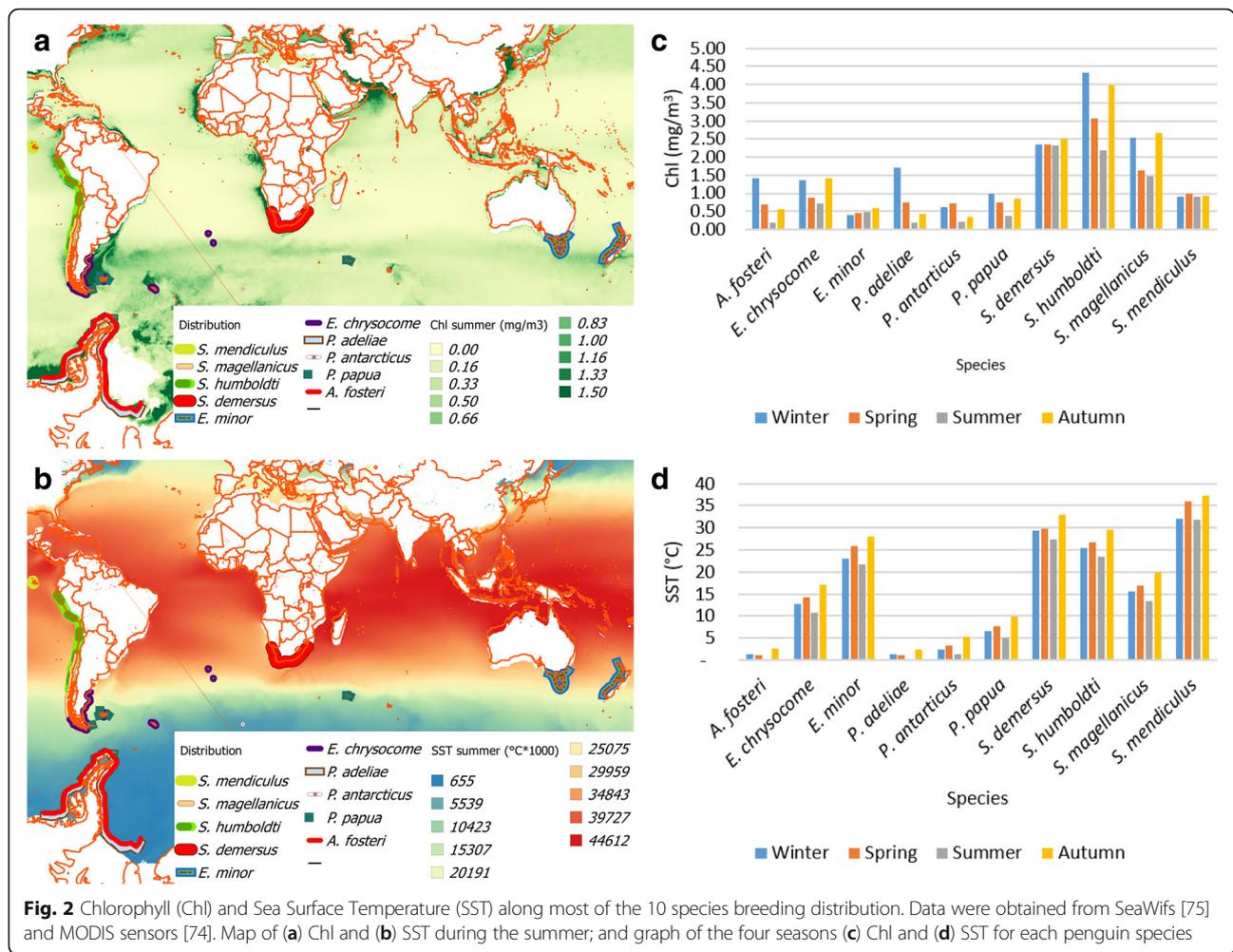
Using this information, between-species pairwise Euclidian distances for Chl, SST and Chl-SST (combined environmental distance COM) were calculated using the equation Eq. (1):

$$D_{xy} = \sqrt{(W_x - W_y)^2 + (A_x - A_y)^2 + (SP_x - SP_y)^2 + (S_x - S_y)^2} \quad (1)$$

where  $D_{xy}$  is the Chl, SST or COM distance between two penguin species  $x$  and  $y$  based on the seasonal variability of Chl and SST.  $W$ ,  $A$ ,  $SP$  and  $S$  are the values of Chl, SST or COM in the winter, autumn, spring and summer season respectively.

#### Mantel test and GLM models

Mantel tests and GLM models (Generalized Linear Model) were implemented to explore the relationships between distance matrices of genetic parameters ( $\pi$  and  $ka/ks$ ) and environmental parameters (Chl, SST and COM) as explanatory variables [81, 82]. To study the relationship between genetic patterns and the environmental parameters (the distances across Chl, SST and COM dimension), the Mantel matrix distances were divided into three sub-matrices (Example in Additional file 3 Table S3), each describing the patterns between pairs of species within a bounded interval of Chl, SST and COM distances [81]. These bounded intervals consisted of a) short pairwise distances corresponding with the minimum to 33% of the data distance, b) middle distances corresponding with pairwise distances between 33% and 66% of the data distance, c) large distances consisting of the largest 33% of the pairwise distances and d) all pairwise data. For Mantel tests,  $p$ -values were obtained based on 10,000 simulations. Normality of the data ( $Ka/$



Ks and π) was tested applying the Shapiro-Wilk test. The GLMs *p*-values and AICs were calculated considering a Gaussian error structure for all data under normal distribution and gamma error structure for data without normal distribution. All statistical analyses were implemented using R packages [64].

## Results

### MtDNA genomes of penguins

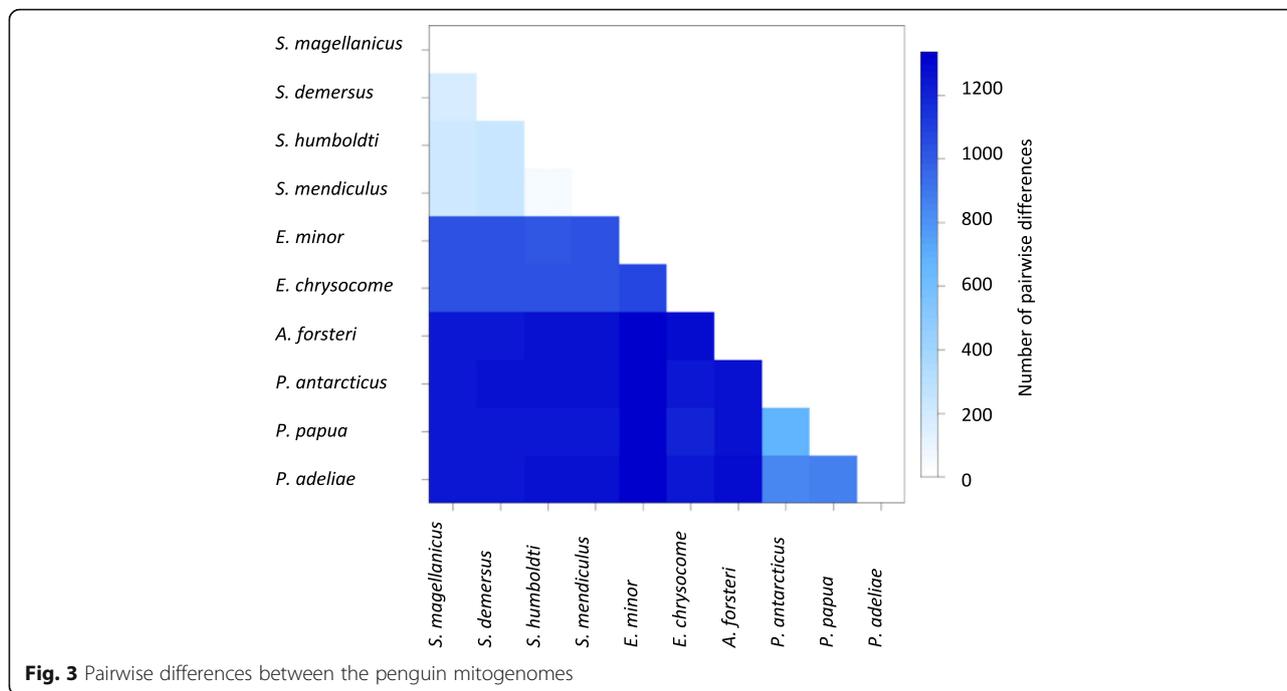
The number of reads for each of the 12 genomes ranged from 52,990,428 to 158,565,602. One individual of *P. papua* was excluded because of insufficient reads (7,341,640). The number of reads attributed to mtDNA genomes ranged from 5049 to 25,947 (1647 for *P. papua*) and the percent coverage varied from 16 to 82 (Table 2). We recovered the full sequence from all mtDNA regions except the repetitive region of the D-loop.

These are the first reported near-complete mtDNA genomes of *S. humboldti*, *S. magellanicus*, *S. mendiculus* and *P. papua*. The complete mitochondrial genome, without the *D*-loop, varied between 15,514 bp and 15,520 bp

among the three *Spheniscus* ssp. and between 15,490 bp and 15,500 bp among the three *Pygoscelis* species. Including the partial *D*-loop, lengths varied between 16,367 bp and 16,660 bp among the three *Spheniscus* species, and between 16,653 bp and 16,765 bp among the three *Pygoscelis* species. These mt-genome lengths are comparable to the mtDNA genomes of *P. antarcticus* (15,972 bp), *S. demersus* (17,346 bp), *E. minor* (17,611 bp), *E. chrysocome* (16,930 bp), and *P. adeliae* (17,486 bp).

The lowest evolutionary divergence among species was between *S. humboldti* and *S. mendiculus* (0.004), with six of the thirteen genes being 100% identical (*COX1*, *COX2*, *ATP8*, *ATP6*, *ND3*, *ND4L*), followed by *S. demersus* and *S. magellanicus* (0.010). The highest divergence was observed between *E. minor* and *P. adeliae* (0.097, Fig. 3, Additional file 2 Table S2).

The *ND6* and 8 tRNAs are transcribed from the light strand, while the other 12 protein coding genes and 14 tRNAs and two rRNAs are positioned on the heavy strand. This pattern of transcription has been described for *P. adeliae*, *S. demersus* and other bird species [49,



56, 83]. Indels, when comparing all ten species, were observed in five of the thirteen mtDNA genes (*ATP8*, *CYTB*, *ND4*, *ND5* and *ND6*). All indels comprised three consecutive bases, such as for *Pygoscelis* and *A. forsteri* at *ATP8*; for *E. chrysocome* at *CYTB*; for *A. forsteri* at *ND4* and *ND6*; and for *E. minor* at *ND5*.

At the interspecific level, higher diversity was found in NADH dehydrogenase genes (Table 3), among which *ND6* was the most diverse between all species ( $\pi = 0.11385$ ),

followed by *ATP8* ( $\pi = 0.11203$ ). The cytochrome c oxidase genes (*COX2*, *COX3* and *COX1*) were the least diverse among the mtDNA genes ( $\pi = 0.05816$ , 0.06 and 0.07417, respectively). Comparing the two genera, *ND6* was also the most diverse among *Pygoscelis* ( $\pi = 0.0763$ ), while *ND1* was the most diverse among the *Spheniscus* ( $\pi = 0.02152$ ) species (Table 3). *ND3* was the second most diverse among *Pygoscelis* ( $\pi = 0.0733$ ), however, it had the lowest nucleotide diversity among *Spheniscus* ( $\pi = 0.0622$ ). The numbers

**Table 3** Interspecific diversity across whole data set (all 15 penguins of 10 species) for all 13 protein-coding genes

MtDNA Gene Complex Length (bp)	Genera									
	<i>Pygoscelis</i>				<i>Spheniscus</i>				Among all species	
	S	$\pi$	S	$\pi$	S	$\pi$	S	$\pi$		
ND1	I	978	108	0.06074	49	0.02152	262	0.08941		
ND2	I	1039	115	0.06141	36	0.01457	305	0.09337		
COX1	IV	1551	161	0.05642	32	0.01038	344	0.07417		
COX2	IV	684	61	0.04766	13	0.00919	144	0.05816		
ATP8	V	165/168	17	0.05833	4	0.01212	52	0.11203		
ATP6	V	684	68	0.05468	23	0.01682	176	0.08619		
COX3	IV	784	56	0.03852	21	0.01215	157	0.06		
ND3	I	352	46	0.0733	6	0.00622	89	0.08423		
ND4L	I	297	27	0.04714	6	0.01058	71	0.08325		
ND4	I	1379/1382	163	0.06425	51	0.01796	369	0.08703		
ND5	I	1818/1821	180	0.05369	55	0.01435	502	0.08794		
CytB	III	1143/1146	123	0.05836	32	0.01267	285	0.08026		
ND6	I	519/522	72	0.0763	36	0.0211	177	0.11385		
Overall		11,408	1197		364		2933			

of polymorphic sites were higher within *Pygoscelis* ( $S = 1197$ ) compared to *Spheniscus* ( $S = 364$ ).

**Gene evolution and natural selection**

The general pattern of Ka/Ks ratios for the different genes was similar between pairwise species at inter-genera comparisons (*Spheniscus*, *Pygoscelis*, *Eudyptula*, *Eudyptes*, *Aptenodytes*; Fig. 4; see Additional file 1 Table S1). The Ka/Ks ratios were all <1 (values under 1 suggest negative selection and over 1 positive selection using this method). *ATP8* had the highest values and variability of Ka/Ks ratio among all species comparison between genera (Fig. 4) with values between 0.1 and 0.4. *ATP8* Ka/Ks ratios were lower between *Pygoscelis* and *Spheniscus* (0.09 to 0.15), and between *A. forsteri* and *Spheniscus* (0.14 to 0.16). High *ATP8* Ka/Ks ratios were observed between *E. minor* and most other genera (0.24 to 0.41).

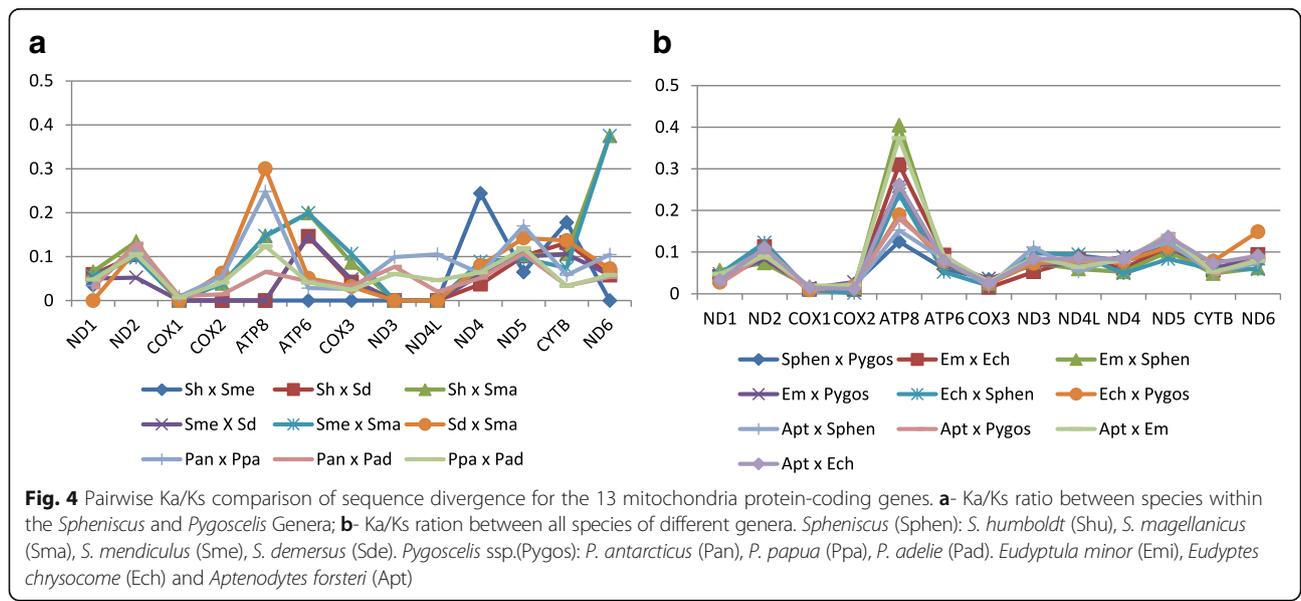
Markedly low values of Ka/Ks ratios were found for within and between pairwise genera comparisons and were observed for *COX1*, followed by *COX2*, *COX3* and *ND1*, suggesting negative selection (Fig. 4). In the intra-genus comparisons, the Ka/Ks ratios were zero for *COX1*, *ND3* and *ND4L* among all *Spheniscus* species and were also low for *COX2*, *ND1* and *COX3*. Similarly, within *Pygoscelis* comparisons, Ka/Ks ratios were lowest for *COX1*, followed by *COX3*, *COX2*, *ND1* and *ATP6*. Therefore, *COX1* had a signature of purifying selection for amino acid sequence conservation for all pairwise comparisons among penguin genera and species.

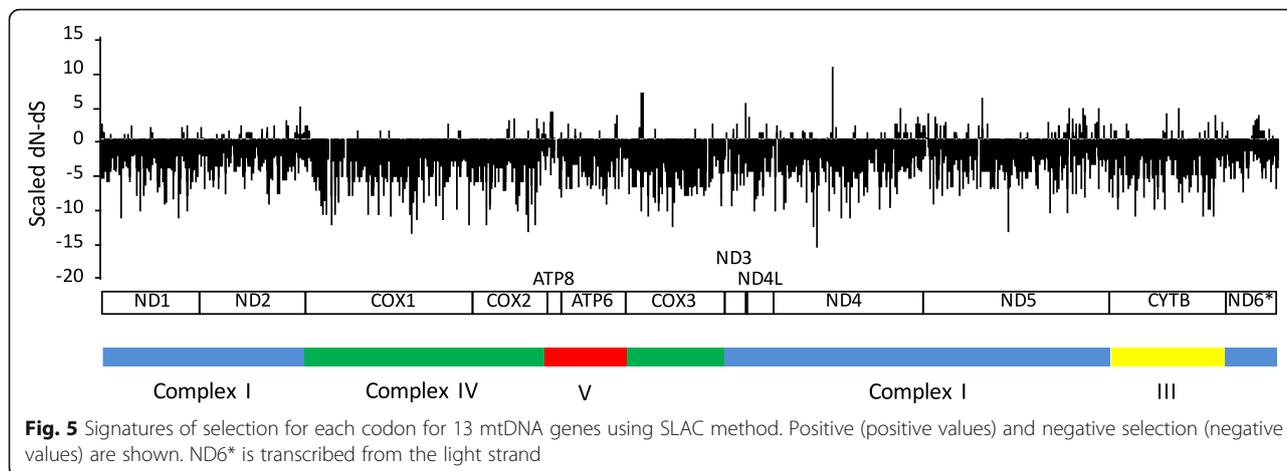
Similar results were detected using the SLAC, REL and FEL algorithms. Overall, significant levels of positive selection were detected for 69 different codon sites in four genes (*COX1*, *ATP8*, *ND4L*, *ND4*). Only one codon was under positive selection for *ND4* using SLAC (codon 187)

and one codon for *ATP8* (codon 36) using FEL. We detected a large number of codons under positive selection for *COX1* (a total of 39 codons) and 28 codons for *ND4L* using REL. Although REL methods have higher power and are more computationally intensive than SLAC methods of detecting selection, they have higher rates of false positives with small data sets [71, 84]. FEL approaches appear to be more robust with fewer sequences than REL and SLAC that may underestimate the substitution rate [71]. Indeed, no codon had positive and significant values by SLAC for *ATP8*. The largest number of codons with positive, but not significant signatures were identified for this gene using this method (18% of codons), in contrast with patterns observed in other genes (1.4% in *COX1* to 11% in *ND2*, *ND5* and *ND6*, Figs. 5, 6). Low values of negative selection were obtained for *ATP8* and *ND4L* using SLAC and FEL, and the most significant values were obtained for *COX1* (Fig. 6). We identified 375 codons with significant negative selection using SLAC algorithms and 1787 codons using FEL algorithms. All of the codons identified by SLAC overlapped with those identified using FEL algorithms. We did not detect lineages under positive selection using a Branch REL analysis based on the phylogeny.

**Environmental data**

Chlorophyll concentration (Fig. 2c) was high in the area occupied by *S. humboldti*, and varied highly among seasons (2.2 mg/m<sup>3</sup> and 4.3 mg/m<sup>3</sup> in winter and summer respectively). Chlorophyll concentrations in the distributional range of *S. demersus* were also high, but more stable (from 2.3 mg/m<sup>3</sup> to 2.5 mg/m<sup>3</sup>). Highest SSTs were recorded throughout the range of *S. mendiculus* (32 °C and 37 °C during winter and summer respectively), followed by lower SSTs throughout areas occupied by *S. demersus* and *S. humboldti* (27.4–

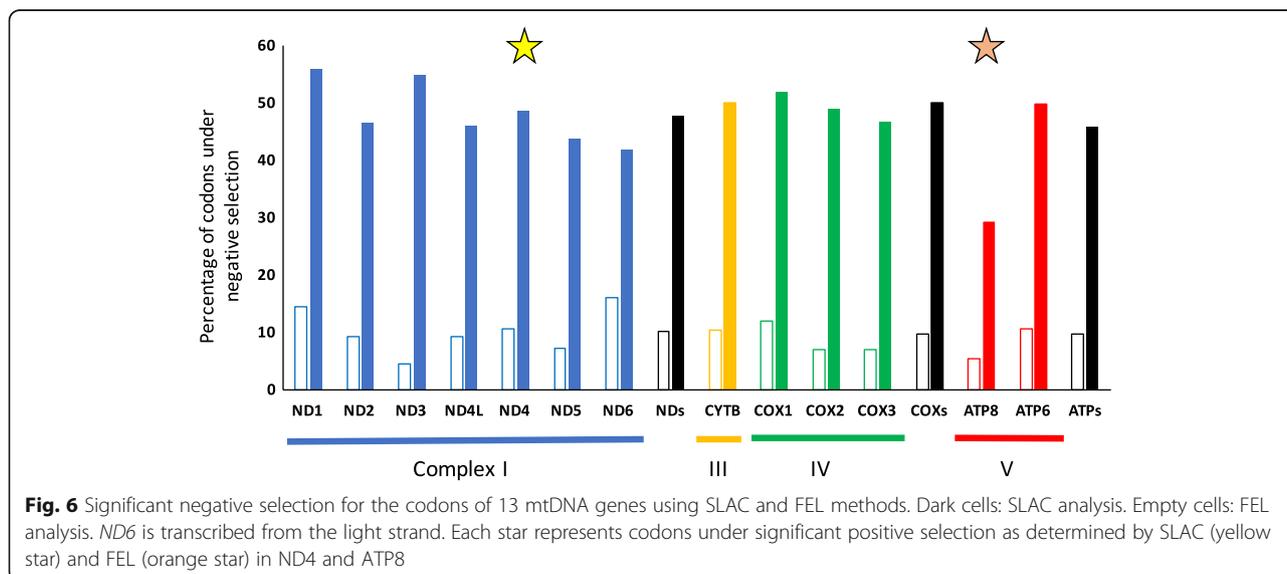




32.8 °C; and 23.6–29.6 °C during winter and summer respectively), with the lowest SST in areas where *P. adeliae*, *A. fosteri* and *P. antarcticus* are distributed (0.5 °C and 5.2 °C during winter and summer respectively) (Fig. 2d).

Nucleotide diversity ( $\pi$ ) was not significantly correlated with environmental data at different spatial distances in any of 13 mtDNA genes for Mantel test (Additional file 4 Table S4). The Ka/Ks ratio was related with environmental data at different spatial scales for four genes supported by the significant values of Mantel test and GLM (*ND4*, *ND3*, *ND5* and *Cytb*; Table 4) and three other genes only for GLM (*ND1*, *COX1* and *ATP8*). A significant positive correlation between an environmental variable and Ka/Ks ratio for a specific mtDNA gene may suggest that species inhabiting similar environments undergo adaptation processes for the same gene (similar protein reflected by Ka/Ks) and when inhabiting different environments undergo different gene adaptation processes (different protein reflected by Ka/Ks). The highest significant correlation was

between SST and Ka/Ks ratio for *ND4* (Mantel,  $p$ -value: 0.0001; GLM,  $p$ -value: 0.00001) for all distances. Although, chlorophyll concentration was significantly correlated with Ka/Ks ratio of *ND1*, *COX1*, *ND5* for GLM, only *COX1* showed signature of selection and low  $p$ -value for Mantel ( $p$ -value: 0.063) at short distances. The combined environmental data (COM) were correlated with the Ka/Ks ratio for three genes supported by significant Mantel test and GLM: *ND3* (Mantel,  $p$ -value: 0.0304; GLM  $p$ -value: 0.0188), *ND5* (Mantel,  $p$ -value: 0.0232; GLM  $p$ -value: 0.0095), and *Cytb* at short distance (Mantel,  $p$ -value: 0.0463; GLM  $p$ -value: 0.0103) and middle distance (Mantel,  $p$ -value: 0.0064; GLM  $p$ -value: 0.0112). All significant  $p$ -values for the Mantel test were also significant for GLM (Additional file 4 Table S4). Although we detected few significant correlations between Ka/Ks ratio and environmental variables, we cannot reject the possibility that these results are influenced by other variables not evaluated here.



**Table 4** Significant values of GLM and Mantel Test for Ka/Ks ratio for each gene and environmental data

Genes	Environmental variable	Class distance	Ka/Ks					
			Mantel			GLM		
			Expectation	Observation	P	Variance	AIC	P
COX1	CHL	0–33	0.0000	0.2216	0.0631	0.0162	-122.01	0.0298*
ND3	COM	33–66	-0.0014	0.2642	0.0304*	0.0125	-53.22	0.0188*
<b>ND4</b>	<b>SST</b>	<b>0–100</b>	<b>-0.0006</b>	<b>0.2978</b>	<b>0.0001*</b>	<b>0.0029</b>	<b>-248.99</b>	<b>0.00001*</b>
ND5	COM <sup>1</sup>	0–100	0.0000	0.1312	0.0232*	0.0038	-222.34	0.0095**
CYTB	COM	0–33	-0.0008	0.2476	0.0463*	0.0149	-63.94	0.0103*
	COM <sup>1</sup>	33–66	0.0013	0.4184	0.0064*	0.0137	-72.91	0.0112*

\*significant values; bold the gene with lowest *p*-value; <sup>1</sup>negative correlation

Nucleotide diversity among penguin species was not correlated with the environmental variables (no significant value for Mantel test), nevertheless Ka/Ks ratios were significantly correlated. This supports the premise that broadly, genetic diversity of mtDNA genes are conserved, but that specific signals of selection related to environmental variables could be identified by comparative estimation of Ka/Ks ratios among genes.

Assessing genes under selection, *ND4* and *COX1* had the highest Ka/Ks ratio between species from distinct environments and the lowest ratios between species from the same environment (Figs. 4, 7). In contrast, *ATP8* had the highest Ka/Ks ratio between similar environments and the lowest between different ones.

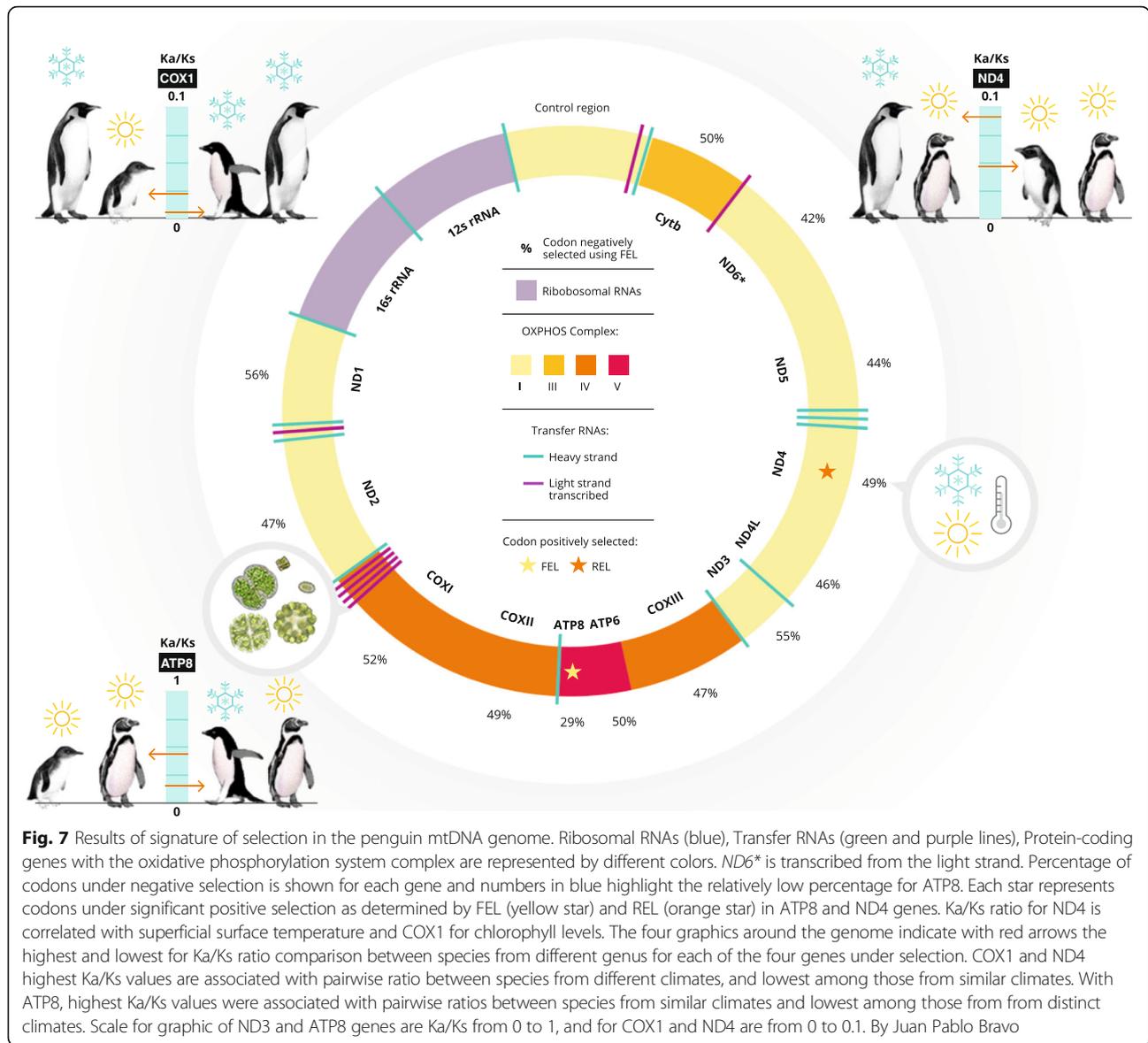
## Discussion

Our analyses revealed strong and contrasting signals and selection patterns in the coding mtDNA genes of closely related penguin species, some of which correlated with patterns of sea surface water temperatures and chlorophyll concentrations. We identified a widespread signature of purifying selection (Ka/Ks < 1) across the mitochondrial genome, consistent with expected patterns if purifying selection were constraining the mitogenome evolution to maintain the OXPHOS proteins and functionality [20, 70]. These patterns included a signature of purifying selection in *COX* genes, especially *COX1*, which had Ka/Ks ratios close to zero for all comparison among species and genera. In contrast, one codon under positive selection was evidenced for *ND4* and one for *ATP8* (SLAC and FEL analysis respectively). Sea surface temperature (SST) was strongly correlated with the Ka/Ks ratio of *ND4* and chlorophyll concentration was correlated with Ka/Ks ratios of *COX1*. The Ka/Ks ratio patterns for *ATP8* varied widely among genera and there was only weak or no evidence of a correlation between genetic diversity of *ATP8* and measured environmental factors. These results identify mtDNA candidate genes under selection which could be involved in broad-scale adaptations of penguins to their environment.

## Evidence linking selection and gene function *NADH dehydrogenase*

In metazoans, mitogenome complexes I through V are physically positioned relative to each other in accordance with their function, with the complex II genes residing in the nucleus [85]. A positively selected codon was detected by SLAC model on the mitogenome of penguins in the Complex I or *NADH* dehydrogenase gene (*ND4*). Complex I through IV generate the proton gradient in the intermembrane space of mitochondria. Complex I is the main entrance point of electrons and the largest complex of the respiratory chain. It catalyzes the transfer of two electrons from NADH to ubiquinone and translocates four protons across the inner mitochondrial membrane [7]. Complex I consists of different protein subunits, with fourteen homologous subunits present in all organisms, and which constitute the enzyme.

ND1 and ND2 are situated at the connection between the peripheral and the membrane arm, whereas ND4 and ND5 are at the distal end [86, 87]. If ND2, ND4 and ND5 are the actual proton pumping devices [86], mutations in these genes may alter the efficiency of the proton-pumping process. The concentration of codons under positive selection in specific domains of the mitogenome are very likely to be related to protein function [20, 88] and a meta-analysis of 237 vertebrates showed that most of the genes under positive selection were located in Complex I [89]. *NADH* dehydrogenase genes have also shown signatures of positive selection at a population level in birds (*Parus* spp.) in the *ND2* gene [90], in lineages of an Australian bird (*Eopsaltria australis*) [20] in *ND4* and *ND4L*. Similarly the Chinese snub-nosed monkey (*Rhinopithecus* ssp.) showed evidence of adaptive variation related with physiological hypoxia and cold stress in altitude in *ND2* and *ND6* [91]. Few studies have assessed if selection on mtDNA genes is correlated with environmental variables [92]. In the Atlantic herring (*Clupea harengus*), significant positive selection was observed in *ND2*, *ND4* and *ND5*. Correlations between genetic diversity and variation of environmental factors (temperature, salinity and latitude) were generally weak,



although the authors addressed some correlation between temperature and genetic distances with *ND3* and *ND4L* [92].

**ATP synthase**

In our study, the Ka/Ks ratio for *ATP8* was lower than 1, suggesting negative selection for the gene, although it was higher than any other value for mtDNA genes in penguins. However, one codon of *ATP8* (Complex V) showed evidence of positive selection (under model FEL). Complex V (or ATP synthase) is the final enzyme in the OXPHOS pathway and is ubiquitous throughout all forms of life and biological functions. The enzyme uses the energy stored in a proton gradient across a membrane to drive the synthesis of ATP from ADP and phosphate. Complex V uses the proton gradient resulting from the activity of the respiratory

chain enzymes complexes I, III and IV for ATP synthesis, generating 95% of cellular ATP. Complex V is a multi-subunit complex with two functional domains, F1 and F0, for which *ATP8* encodes a core subunit of the F0 component. Across mammalian evolution, *ATP6* is more conserved, in contrast to the *ATP8* subunit which has some highly variable sites with extreme amino acid changes of properties, suggesting that there are differences in its regulatory function among species [88].

**Cytochrome c oxidase**

*COXI*, complex IV, showed purifying selection in penguins, as has been observed across all Aves [32]. We estimated Ka/Ks ratios close to zero for all comparison among species and genera, and similar results were obtained using both the SLAC and FEL methods. These genes from complex IV

represent the terminal complex of the respiratory chain. It catalyzes electron transfer from cytochrome c to molecular oxygen, reducing the latter to water. It has been postulated that Complex IV of the OXPHOS, in which *COXI* is one of the two catalytic subunits, has the property of intrinsic uncoupling [93], through which it regulates and increases the coupling efficiency to produce ATP or heat. This mechanism is consistent with the negative selection observed for this gene, as adaptations can occur by changing the configuration of the protein, a mechanism that does not occur in complexes I and V.

Further studies are required to completely understand the factors constraining the evolutionary rate of mitogenomes in penguins. Evolution of protein sequences depends on functional constraints including differential gene expression, protein folding and connectivity, functional importance, and expression breadth among tissues [94]. Here, we discussed the implications of functionality of different mtDNA genes on the evolutionary rate of a protein. Alternatively, Zang and Yang [94], suggest that this rate may be predominantly influenced by its expression level rather than functional importance.

#### Mitogenomics and ecological-evolutionary hypotheses to explain selection

There are two hypotheses generally offered to explain patterns of mitochondrial genome evolution. The first is that temperature is a primary driver of selection and the evolution of mtDNA genes. For penguins, speciation and adaptation are seemingly strongly correlated with temperature differences of ocean water between the Antarctic and the tropics [35–38]. The mitochondrial OXPHOS generates both energy and heat, with the balance of each depending on the coupling efficiency of the electron transport chain to the *ATP* synthase activity [2, 3]. Less efficient energy production and more heat production would be beneficial for survival in cold environmental conditions [2]. A reduced proton leak would lead to incremental cold sensitivity that effectively constrains the biogeographic distribution of species with this pattern of mtDNA variation [4, 5]. Selection on mtDNA variation influencing mitochondrial bioenergetics would thus be plausible drivers of adaptation and speciation [6], and indeed we detected correlation between *Ka/Ks* from *ND4* and sea surface water temperature which should be further investigated.

A second hypothesis is that the risk of starvation associated with restricted food availability (less abundant or less reliable food source) drives mtDNA gene evolution, since starvation can generate more efficient coupling of energy production by the OXPHOS pathway [7]. Three mechanisms have been suggested for the evolution of starvation resistance and each may be influenced by mitochondrial diversity: (a) greater energy reserves (lipid stores); (b)

reduced rate at which reserves are utilized; (c) and lowering of the minimal resources required for survival [7, 9]. Reducing energy requirements may occur by reducing ROS and heat production and thus increasing OXPHOS efficiency [8, 10]. Chl can be considered to be a measure or correlate of primary productivity and therefore resource availability in the Antarctic marine food web. Penguins as top predators, are dependent on the productivity of lower trophic levels. Therefore, the high variation in productivity of the habitats throughout the *Spheniscus* species distributional ranges may have led to the evolution of more efficient coupling in those species, compared to the stable environments of Emperor and Adélie penguins. For example, high variability in levels of Chl is observed within the range of *S. humboldti*, and becomes further exacerbated during El Niño events, leading to a higher probability of mortality due to starvation. In this circumstance, selection may favor OXPHOS efficiency rather than heat production. During El Niño events, the upwelling of nutrients is interrupted, reducing food availability and increasing water temperature in Chilean-Peruvian waters [95]. This affects fish abundance, causing increased Galapagos and Humboldt penguin mortality, nest desertion and large-scale movements [40, 42, 96]. Although Chl could be an important variable for mtDNA evolution, only a weak correlation between *Ka/Ks* ratio of *COXI* and Chl across small geographic scales was detected here. Therefore, our results should be further investigated in penguins not only at a species level, but also in populations of each penguin species across a gradient of sea surface temperatures (e.g. across latitudes) and chlorophyll levels.

Although *ATP8* was not correlated to SST and Chl here, the *Ka/Ks* ratio values were higher than in any other gene compared. Moreover, it showed a remarkable pattern for all comparisons among species distributed in extremely different environments (tropical and hot temperate climate versus polar climate). We found contrasting results for *ATP8*. *A. forsteri* and *Spheniscus* had the lowest values for *ATP8* genes in pairwise comparisons among all genera. This may be indicative of different mechanisms of adaptation among genes and species depending on their individual habitat requirements. However, we cannot discount the possibility that this gene is associated with other selective variables, which in turn are correlated with climatic and productive variations observed in the distributional ranges of penguins. Indeed, it was recently suggested that evolution of mitochondrial genes might be linked with pathogen variability [11].

An increasing number of studies have implicated mitochondria with innate immune responses, including the control of apoptotic signaling, sterile inflammation, and antibacterial and antiviral response [12–15]. Although it is uncertain what direct role mitochondrial encoded genes have

in immune response, it bears further evaluation. Antarctica and the tropics differ considerably in virulence, ensemble, abundance, development and diversity of pathogens. Just recently, the influenza virus was detected in penguins in Antarctica [97–99]. Antarctic wildlife is particularly susceptible to infectious agents because of its geographic isolation and extreme environment, which increases the likelihood that the fauna has evolved under low “pathogen pressure” and are thus immunologically naïve [11].

### Implications for future adaptation to changing environments

Ongoing human-induced climatic changes are predicted to have large impacts on the world biota, including penguin species [44–46, 100]. *Pygoscelis* species appear to be changing their distributions. *P. papua* populations have been increasing along the Antarctic Peninsula in association with an abrupt increase of water temperatures, while populations of *P. adeliae* and *P. antarcticus* have been decreasing [44–46]. The decline and disappearance of an *A. forsteri* colony was associated with the rise in local mean annual air temperature and decline in seasonal sea ice duration [101]. *S. magellanicus* populations have experienced increasing reproductive failure, measured in chicks’ mortality associated with starvation, rain, heat, and predation, which has been associated with climate change and an increase in frequency and intensity of storms [102]. *S. humboldti* and *S. mendiculus* have been historically affected by strong El Niño events causing high mortality, nesting delay, nest flooding and breeding failure [40–42, 96, 103]. El Niño events are becoming more intense and frequent due to global climate change [104, 105]. The decreasing food availability for penguins in South America during El Niño events [95], and for penguins in Antarctica during the past 50 years [106] has caused not only direct mortality. Increased mortality has also been attributed to compromised immune systems, since adequate nutrition is fundamental to maintaining a functioning immune system [107]. The expanding human activities in Antarctica also increase the risk of introduction of infectious diseases into this naïve wildlife community [108]. Therefore, selective pressures on mtDNA genes by environmental factors such as temperature, food availability and novel pathogens could intensify significantly with climate change.

The effect of climate change on gene evolution does not have to be direct. As penguins are top marine predators, changes at all trophic levels will affect penguin population demography [109]. However, understanding the genes involved in adaptation to climate by penguin species may permit us to predict how the species are physiologically limited and how they may be able to respond to environmental changes such as climate warming.

### Conclusions

Here, we identify mtDNA candidate genes under selection which could be involved in broad-scale adaptations of penguins to their environment. Our results suggest a strong negative selection for *COXI*, and a positive selection at codon sites on *ATP8* and *ND4*. We also found a high correlation between *ND4* and sea surface temperature. Further studies on these genes at the intra-specific level are required to completely understand penguin adaptation across different environmental conditions. This is a novel study that illustrates the applicability of large-scale remote sensing as a useful approach to comprehend adaptation to the environment occurring at a molecular level. Integration of these different data provide insights into how penguins species have adapted to their environments and therefore how they may respond to future, human-induced changes to their environment.

### Additional files

**Additional file 1: Table S1.** Pairwise Ka/Ks comparison of sequence divergence for the 13 mitochondria protein-coding genes. (DOCX 44 kb)

**Additional file 2: Table S2.** Estimates of Evolutionary Divergence between Sequences. (DOCX 21 kb)

**Additional file 3: Table S3.** Example of matrix between the environmental distance (ENDIST), chlorophyll (CHL) in this example, and the the Ka/Ks. Then environmental data (CHL) were grouped in three class distances for the GLM and Mantel tests. (DOCX 35 kb)

**Additional file 4: Table S4.** Significant values of GLM and Mantel Test for Ka/Ks ratio and nucleotide diversity ( $\pi$ ) for each gene and environmental data (SST, CHL, COM) at different class distance. (DOCX 26 kb)

### Acknowledgements

Not applicable.

### Funding

The research was funded by the Sea World and Busch Garden Conservation Fund, Fondecyt N° 11110060, 1150517 INACH N° G06–11, T12–13 and 12–14, UNAB DI-410-13/I, and CNPq 482501/2013–8 and 490403/2008–5, FAPESP 2009/08624. Sampling of Adélie penguins at Cape Crozier was supported through funding from the National Science Foundation to David G. Ainley (ANT 0944411) and logistics supplied by the U.S. Antarctic Program. MM was funded by a New Zealand’s Ministry of Science and Innovation grant (C01X1001). PP was funded by the Saint Louis Zoo and the Des Lee Collaborative Vision.

### Availability of data and materials

All genome and sequences were submitted to GenBank (accession numbers KM891593; KU361803-KU361806; KU356673-KU356677).

### Authors’ contributions

BR and JAV conceived the entire study, DEL implemented the genomics bioinformatics. JAV, PGP, MM, DGA, GPMD contributed to the sample collection. BR, WEJ, GPMD, JAV analyzed and interpreted the genetic data. MDM conceived the acquisition of data, analysis and interpretation of spatial environmental data from remote sensing and GIS databases. All authors were involved in drafting or revising the content of the manuscript. All authors read and approved the final manuscript.

### Ethics approval and consent to participate

All samples were collected with permits in accordance to Annex II, Article 3 of the Protocol on Environmental Protection to the Antarctic Treaty, and the regulation from the Scientific Committee on Antarctic Research (SCAR) provided by the Chilean Antarctic Authorities (permits INACH 44/2012, 45/2016, 46/2016) and the Brazilian Antarctic Authorities through the PROANTAR and Environmental Ministry (N°21/2011 and N 20/2012), and other permits from Subsecretaría de Pesca (Chile, 04/2008, 110/2012, 2086/2014) from other locations. The permits incorporated authorization for sample collection for all locations including special authorization to the studied protected areas. A bioethics permit was provided by Universidad de Concepción and Pontificia Universidad Católica de Chile, which was required to obtain INACH permits and Fondecyt project.

### Consent for publication

Not applicable.

### Competing interests

The authors declare that they have no competing interests.

### Publisher's Note

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

### Author details

<sup>1</sup>Departamento de Ecosistemas y Medio Ambiente, Facultad de Agronomía e Ingeniería Forestal, Pontificia Universidad Católica de Chile, Av. Vicuña Mackenna, 4860 Santiago, Chile. <sup>2</sup>Facultad de Ecología y Recursos Naturales, Universidad Andrés Bello, República 252, Santiago, Chile. <sup>3</sup>Departamento de Ciencias Pecuarias, Facultad de Ciencias Veterinarias, Universidad de Concepción, Av. Vicente Méndez 595, 3780000 Chillán, CP, Chile. <sup>4</sup>Centro Nacional de Genómica y Bioinformática, Portugal 49, Santiago, Chile. <sup>5</sup>I+DEA Biotech, Av. Central 3413, Padre Hurtado, Santiago, Chile. <sup>6</sup>Smithsonian Conservation Biology Institute, National Zoological Park, 1500 Remount Road, Front Royal, VA 22630, USA. <sup>7</sup>University of Missouri St Louis and Saint Louis Zoo, One University Blvd., St. Louis, MO 63121-4400, USA. <sup>8</sup>School of Environmental Sciences and Institute for Land, Water and Society, Charles Sturt University, PO Box 789, Albury, NSW, Australia. <sup>9</sup>Pontificia Universidade Católica de Minas Gerais, Av. Dom José Gaspar 500, Coração Eucarístico, Belo Horizonte, MG, Brazil. <sup>10</sup>Centro de Cambio Global UC, Santiago, Chile.

Received: 5 April 2017 Accepted: 28 December 2017

Published online: 16 January 2018

### References

- Boore JL. Animal mitochondrial genomes. *Nucleic Acids Res.* 1999;27:1767–80.
- Pörtner HO. Climate variability and the energetic pathways of evolution: the origin of Endothermy in mammals and birds. *Physiol Biochem Zool.* 2004;77:959–81.
- Wallace DCA. Mitochondrial paradigm of metabolic and degenerative diseases, aging, and cancer: a Dawn for evolutionary medicine. *Annu Rev Genet.* 2005;39:359–407.
- Ballard JWO, James AC. Differential fitness of mitochondrial DNA in perturbation cage studies correlates with global abundance and population history in *Drosophila Simulans*. *Proc R Soc B Biol Sci.* 2004;271:1197–201.
- Katewa SD, Ballard JWO. Sympatric *Drosophila Simulans* flies with distinct mtDNA show difference in mitochondrial respiration and electron transport. *Insect Biochem Mol Biol.* 2007;37:213–22.
- Gershoni M, Templeton AR, Mishmar D. Mitochondrial bioenergetics as a major motive force of speciation. *BioEssays.* 2009;31:642–50.
- Ballard JWO, Melvin RG. Linking the mitochondrial genotype to the organismal phenotype. *Mol Ecol.* 2010;19:1523–39.
- Mishmar D, Ruiz-Pesini E, Golik P, Macaulay V, Clark AG, Hosseini S, et al. Natural selection shaped regional mtDNA variation in humans. *Proc Natl Acad Sci.* 2003;100:171–6.
- Rion S, Kawecki TJ. Evolutionary biology of starvation resistance: what we have learned from *Drosophila*: starvation resistance in *Drosophila*. *J Evol Biol.* 2007;20:1655–64.
- Ruiz-Pesini E, Mishmar D, Brandon M, Procaccio W, Wallace DC. Effects of purifying and adaptive selection on regional variation in human mtDNA. *Science.* 2004;303:223–6.
- Dobson A, Foutopoulos J. Emerging infectious pathogens of wildlife. *Philos Trans R Soc B Biol Sci.* 2001;356:1001–12.
- Arnoult D, Carneiro L, Tattoli I, Girardin SE. The role of mitochondria in cellular defense against microbial infection. *Semin Immunol.* 2009;21:223–32.
- Ohta A, Nishiyama Y. Mitochondria and viruses. *Mitochondrion.* 2011;11:1–12.
- West AP, Shadel GS, Ghosh S. Mitochondria in innate immune responses. *Nat Rev Immunol.* 2011;11:389–402.
- Zhou R, Yazdi AS, Menu P, Tschopp JA. Role for mitochondria in NLRP3 inflammasome activation. *Nature.* 2011;469:221–5.
- Friedlaender AS, Johnston DW, Fraser WR, Burns J, Patrick NH, Costa DP. Ecological niche modeling of sympatric krill predators around Marguerite Bay, western Antarctic peninsula. *Deep Sea Res Part II Top Stud Oceanogr.* 2011;58:1729–40.
- Bazin E. Population size does not influence mitochondrial genetic diversity in animals. *Science.* 2006;312:570–2.
- Balloux F, Handley L-JL, Jombart T, Liu H, Manica A. Climate shaped the worldwide distribution of human mitochondrial DNA sequence variation. *Proc R Soc B Biol Sci.* 2009;276:3447–55.
- Harrisson K, Pavlova A, Gan HM, Lee YP, Austin CM, Sunnucks P. Pleistocene divergence across a mountain range and the influence of selection on mitogenome evolution in threatened Australian freshwater cod species. *Heredity.* 2016;116:506–15.
- Morales HE, Pavlova A, Joseph L, Sunnucks P. Positive and purifying selection in mitochondrial genomes of a bird with mitonuclear discordance. *Mol Ecol.* 2015;24:2820–37.
- Ribeiro AM, Lloyd P, Bowie RCK. A tight balance between natural selection and gene flow in a southern African arid-zone endemic bird: spatial heterogeneous environments and gene flow. *Evolution.* 2011;65:3499–514.
- Hancock AM, Clark VJ, Qian Y, Di Rienzo A. Population genetic analysis of the uncoupling proteins supports a role for UCP3 in human cold resistance. *Mol Biol Evol.* 2011;28:601–14.
- Nunes VL, Beaumont MA, Butlin RK, Paulo OS. Multiple approaches to detect outliers in a genome scan for selection in ocellated lizards (*Lacerta Lepida*) along an environmental gradient: SELECTION IN OCELLATED LIZARDS. *Mol Ecol.* 2011;20:193–205.
- Li Y, Ren Z, Shedlock AM, Wu J, Sang L, Tersing T, et al. High altitude adaptation of the schizothoracine fishes (Cyprinidae) revealed by the mitochondrial genome analyses. *Gene.* 2013;517:169–78.
- Cheviron ZA, Brumfield RT. Migration-Selection Balance And Local Adaptation Of Mitochondrial Haplotypes In Rufous-Collared Sparrows (*Zonotrichia Capensis*) Along An Elevational Gradient. *Evolution.* 2009;63:1593–605.
- Xu S, Luosang J, Hua S, He J, Ciren A, Wang W, et al. High altitude adaptation and phylogenetic analysis of Tibetan horse based on the mitochondrial genome. *J Genet Genomics.* 2007;34:720–9.
- Zhuang X, CHC C. ND6 Gene "Lost" and Found: Evolution of Mitochondrial Gene Rearrangement in Antarctic Notothenioids. *Mol Biol Evol.* 2010;27:1391–403.
- Menezes AN, Viana MC, Furtado C, Schrago CG, Seuánez HN. Positive selection along the evolution of primate mitogenomes. *Mitochondrion.* 2013;13:846–51.
- Tomasco IH, Lessa EP. The evolution of mitochondrial genomes in subterranean caviomorph rodents: adaptation against a background of purifying selection. *Mol Phylogenet Evol.* 2011;61:64–70.
- Foote AD, Morin PA, Durban JW, Pitman RL, Wade P, Willerslev E, et al. Positive selection on the killer whale mitogenome. *Biol Lett.* 2011;7:116–8.
- Welch AJ, Bedoya-Reina OC, Carretero-Paulet L, Miller W, Rode KD, Lindqvist C. Polar bears exhibit genome-wide signatures of bioenergetic adaptation to life in the Arctic environment. *Genome Biol Evol.* 2014;6:433–50.
- Kerr KCR. Searching for evidence of selection in avian DNA barcodes: SEARCHING FOR SELECTION IN AVIAN DNA BARCODES. *Mol Ecol Resour.* 2011;11:1045–55.
- Marshall HD, Baker AJ, Grant AR. Complete mitochondrial genomes from four subspecies of common chaffinch (*Fringilla Coelebs*): new inferences about mitochondrial rate heterogeneity, neutral theory, and phylogenetic relationships within the order Passeriformes. *Gene.* 2013;517:37–45.
- 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;19:1760–5.
- Baker AJ, Pereira SL, Haddrath OP, Edge K-A. Multiple gene evidence for expansion of extant penguins out of Antarctica due to global cooling. *Proc R Soc B Biol Sci.* 2006;273:11–7.

36. Clarke JA, Ksepka DT, Stucchi M, Urbina M, Giannini N, Bertelli S, et al. Paleogene equatorial penguins challenge the proposed relationship between biogeography, diversity, and Cenozoic climate change. *Proc Natl Acad Sci*. 2007;104:11545–50.
37. Ksepka DT, Bertelli S, Giannini NP. The phylogeny of the living and fossil Sphenisciformes (penguins). *Cladistics*. 2006;22:412–41.
38. Subramanian S, Beans-Picon G, Swaminathan SK, Millar CD, Lambert DM. Evidence for a recent origin of penguins. *Biol Lett*. 2013;9:20130748–8.
39. Hays C. Effects of the el Niño 1982–83 on Humboldt penguin colonies in Perú. *Biol Conserv*. 1986;36:169–80.
40. Simeone A, Araya B, Bernal M, Diebold E, Grzybowski K, Michaels M, et al. Oceanographic and climatic factors influencing breeding and colony attendance patterns of Humboldt penguins *Spheniscus Humboldtii* in central Chile. *Mar Ecol Prog Ser*. 2002;227:43–50.
41. Paredes R, Zavalaga CB, Battistini G, Majluf P, McGill P. Status of the Humboldt penguin in Peru, 1999–2000. *Waterbirds*. 2003;26:129.
42. Vianna JA, Cortes M, Ramos B, Sallaberry-Pincheira D, Gonzalez-Acuna D, Dantas GPM, et al. Changes in abundance and distribution of Humboldt penguin *Spheniscus Humboldtii*. *Mar Ornithol*. 2014;42:153–9.
43. Ainley DG. *The Adélie penguin: bellwether of climate change*. New York: Columbia University Press; 2002.
44. Forcada J, Trathan PN. Penguin responses to climate change in the Southern Ocean. *Glob. Change Biol*. 2009;15:1618–30.
45. Forcada J, Trathan PN, Reid K, Murphy EJ, Croxall JP. Contrasting population changes in sympatric penguin species in association with climate warming. *Glob. Change Biol*. 2006;12:411–23.
46. Lynch H, Fagan W, Naveen R, Trivelpiece S, Trivelpiece W. Differential advancement of breeding phenology in response to climate may alter staggered breeding among sympatric pygoscelid penguins. *Mar Ecol Prog Ser*. 2012;454:135–45.
47. Slack KE. Early penguin fossils, plus mitochondrial genomes, calibrate avian evolution. *Mol Biol Evol*. 2006;23:1144–55.
48. Slack KE, Janke A, Penny D, Arnason U. Two new avian mitochondrial genomes (penguin and goose) and a summary of bird and reptile mitogenomic features. *Gene*. 2003;302:43–52.
49. Subramanian S, Denver DR, Millar CD, Heupink T, Aschrafi A, Emslie SD, et al. High mitogenomic evolutionary rates and time dependency. *Trends Genet*. 2009;25:482–6.
50. Subramanian S, Lingala SG, Swaminathan S, Huynen L, Lambert D. Second generation DNA sequencing of the mitogenome of the chinstrap penguin and comparative genomics of Antarctic penguins. *Mitochondrial DNA*. 2014; 25:271–2.
51. Li C, Zhang Y, Li J, Kong L, Hu H, Pan H, et al. Two Antarctic penguin genomes reveal insights into their evolutionary history and molecular changes related to the Antarctic environment. *GigaScience* [Internet]. 2014 [cited 2017 Mar 11];3. Available from: <https://academic.oup.com/gigascience/article-lookup/doi/10.1186/2047-217X-3-27>
52. Ritchie PA. Ancient DNA. Enables timing of the Pleistocene origin and Holocene expansion of two Adélie penguin lineages in Antarctica. *Mol Biol Evol*. 2003;21:240–8.
53. Aljanabi SM, Martinez I. Universal and rapid salt-extraction of high quality genomic DNA for PCR-based techniques. *Nucleic Acids Res*. 1997;25:4692–3.
54. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al. The sequence alignment/map format and SAMtools. *Bioinformatics*. 2009;25: 2078–9.
55. Zerbino DR, Birney E. Velvet: algorithms for de novo short read assembly using de Bruijn graphs. *Genome Res*. 2008;18:821–9.
56. Labuschagne C, Kotzé A, Grobler JP, Dalton DL. The complete sequence of the mitochondrial genome of the African penguin (*Spheniscus Demersus*). *Gene*. 2014;534:113–8.
57. Watanabe M, Nikaido M, Tsuda TT, Kobayashi T, Mindell D, Cao Y, et al. New candidate species most closely related to penguins. *Gene*. 2006;378:65–73.
58. Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG. The CLUSTAL\_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res*. 1997;25: 4876–82.
59. Librado P, Rozas J. DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics*. 2009;25:1451–2.
60. Mindell DP, Sorenson MD, Dimcheff DE. An extra nucleotide is not translated in mitochondrial ND3 of some birds and turtles. *Mol Biol Evol*. 1998;15:1568–71.
61. Tamura K, Nei M, Kumar S. Prospects for inferring very large phylogenies by using the neighbor-joining method. *Proc Natl Acad Sci*. 2004;101:11030–5.
62. Tamura K, Stecher G, Peterson D, Filipski A, Kumar S. MEGA6: Molecular Evolutionary Genetics Analysis Version 6.0. *Mol Biol Evol*. 2013;30:2725–9.
63. Excoffier L, Lischer HEL. Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and windows. *Mol Ecol Resour*. 2010;10:564–7.
64. R Core Team. R: A language and environment for statistical computing. Vienna: R Foundation for Statistical Computing; 2013. Available from: <http://www.R-project.org>.
65. Gibson AA. Comprehensive analysis of mammalian mitochondrial Genome Base composition and improved phylogenetic methods. *Mol Biol Evol*. 2004; 22:251–64.
66. Huelsenbeck JP, Ronquist F. MRBAYES: Bayesian inference of phylogenetic trees. *Bioinforma Oxf Engl*. 2001;17:754–5.
67. Ronquist F, Huelsenbeck JP. MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics*. 2003;19:1572–4.
68. Guindon S, Gascuel OA. Simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Syst Biol*. 2003;52:696–704.
69. Rambaut A. FigTree 1.3.1. 2009. <http://tree.bio.ed.ac.uk/software/figtree>.
70. Meiklejohn CD, Montooth KL, Rand DM. Positive and negative selection on the mitochondrial genome. *Trends Genet*. 2007;23:259–63.
71. Kosakovsky Pond SL, Frost SDW. Not so different after all: a comparison of methods for detecting amino acid sites under selection. *Mol Biol Evol*. 2005; 22:1208–22.
72. Kosakovsky Pond SL, Murrell B, Fourment M, Frost SDW, Delpont W, Scheffler KA. Random effects branch-site model for detecting episodic diversifying selection. *Mol Biol Evol*. 2011;28:3033–43.
73. Kosakovsky Pond SLK, Frost SDW, Muse SV. HyPhy: hypothesis testing using phylogenies. *Bioinformatics*. 2005;21:676–9.
74. Feldman GC, McClain CR. Ocean Color Web, SeaWiFS Reprocessing # 4 NASA Goddard Space Flight Center. Kuring N, Bailey SW, Eds. 2014. <http://oceancolor.gsfc.nasa.gov>. Accessed 27 Nov 2014.
75. Siegel DA, Behrenfeld MJ, Maritorena S, McClain CR, Antoine D, Bailey SW, et al. Regional to global assessments of phytoplankton dynamics from the SeaWiFS mission. *Remote Sens Environ*. 2013;135:77–91.
76. OceanColor Biology Processing Group. OceanColor [Internet]. Available from: <http://oceancolor.gsfc.nasa.gov/REPROCESSING/R2009/ocv6/>
77. Parkinson CL, Greenstone R. *EOS Data Products Handbook. V. 2*. NASA Goddard Space Flight Center; Greenbelt, MD United States. 266p. 2000.
78. Williams GN, Dogliotti AI, Zaidman P, Solis M, Narvarte MA, González RC, et al. Assessment of remotely-sensed sea-surface temperature and chlorophyll-a concentration in san Matías gulf (Patagonia, Argentina). *Cont Shelf Res*. 2013;52:159–71.
79. Borboroglu PG, Boersma PD, editors. *Penguins: natural history and conservation*. Seattle: University of Washington Press; 2013.
80. QGIS Development Team. QGIS Geographic Information System. Open Source Geospatial Foundation [Internet]. 2009. Available from: <http://qgis.osgeo.org>
81. Diniz-Filho JAF, Soares TN, Lima JS, Dobrovolski R, Landeiro VL, Telles MP d C, et al. Mantel test in population genetics. *Genet Mol Biol*. 2013;36:475–85.
82. Storfer A, Murphy MA, Evans JS, Goldberg CS, Robinson S, Spear SF, et al. Putting the “landscape” in landscape genetics. *Heredity*. 2007;98:128–42.
83. Pereira SL. Mitochondrial genome organization and vertebrate phylogenetics. *Genet Mol Biol*. 2000;23:745–52.
84. Suzuki Y, Nei M. False-positive selection identified by ML-based methods: examples from the Sig1 gene of the diatom *Thalassiosira weissflogii* and the tax gene of a human T-cell lymphotropic virus. *Mol Biol Evol*. 2004;21:914–21.
85. Burger G, Lang BF, Reith M, Gray MW. Genes encoding the same three subunits of respiratory complex II are present in the mitochondrial DNA of two phylogenetically distant eukaryotes. *Proc Natl Acad Sci*. 1996;93:2328–32.
86. Brandt U. Energy Converting NADH. Quinone Oxidoreductase (Complex I). *Annu Rev Biochem*. 2006;75:69–92.
87. Sazanov LA, Walker JE. Cryo-electron crystallography of two sub-complexes of bovine complex I reveals the relationship between the membrane and peripheral arms. *J Mol Biol*. 2000;302:455–64.
88. da Fonseca RR, Johnson WE, O'Brien SJ, Ramos M, Antunes A. The adaptive evolution of the mammalian mitochondrial genome. *BMC Genomics*. 2008;9:119.
89. Garvin MR, Bielawski JP, Sazanov LA, Gharrett AJ. Review and meta-analysis of natural selection in mitochondrial complex I in metazoans. *J Zool Syst Evol Res*. 2015;53:1–17.

90. Zink RM. Natural selection on mitochondrial DNA in Parus and its relevance for phylogeographic studies. *Proc R Soc B Biol Sci.* 2005;272:71–8.
91. Yu L, Wang X, Ting N, Zhang Y. Mitogenomic analysis of Chinese snub-nosed monkeys: evidence of positive selection in NADH dehydrogenase genes in high-altitude adaptation. *Mitochondrion.* 2011;11:497–503.
92. Teacher AG, André C, Merilä J, Wheat CW. Whole mitochondrial genome scan for population structure and selection in the Atlantic herring. *BMC Evol Biol.* 2012;12:248.
93. Kadenbach B. Intrinsic and extrinsic uncoupling of oxidative phosphorylation. *Biochim Biophys Acta.* 2003;1604:77–94.
94. Zhang J, Yang J-R. Determinants of the rate of protein sequence evolution. *Nat Rev Genet.* 2015;16:409–20.
95. Thiel M, Macaya E, Acuna E, Arntz W, Bastias H, Brokordt K, et al. The Humboldt Current System of Northern and Central Chile: Oceanographic Processes, Ecological Interactions And Socioeconomic Feedback. In: Gibson R, Atkinson R, Gordon J, editors. *Oceanogr. Mar. Biol.* [Internet]. CRC Press; 2007 [cited 2017 Mar 11]. p. 195–344. Available from: <http://www.crcnetbase.com/doi/abs/10.1201/9781420050943.ch6>
96. Vargas FH, Harrison S, Rea S, Macdonald DW. Biological effects of el Niño on the Galápagos penguin. *Biol Conserv.* 2006;127:107–14.
97. Barriga GP, Boric-Bargetto D, San Martin MC, Neira V, van Bakel H, Thompsom M, et al. Avian influenza virus H5 strain with north American and Eurasian lineage genes in an Antarctic penguin. *Emerg Infect Dis.* 2016; 22:2221–3.
98. Hurt AC, Vijaykrishna D, Butler J, Baas C, Maurer-Stroh S, Silva-de-la-Fuente MC, et al. Detection of Evolutionarily Distinct Avian Influenza A Viruses in Antarctica. *mBio.* 2014;5 e01098–14-e01098–14
99. Hurt AC, Su YCF, Aban M, Peck H, Lau H, Baas C, et al. Evidence for the introduction, Reassortment, and persistence of diverse influenza A viruses in Antarctica. Schultz-cherry S, editor. *J Virol* 2016;90:9674–9682.
100. Ainley D, Russell J, Jenouvrier S, Woehler E, Lyver PO, Fraser WR, et al. Antarctic penguin response to habitat change as Earth's troposphere reaches 2°C above preindustrial levels. *Ecol Monogr.* 2010;80:49–66.
101. Trathan PN, Fretwell PT, Stonehouse B. First recorded loss of an emperor penguin Colony in the recent period of Antarctic regional warming: implications for other colonies. Briffa M, editor. *PLoS One* 2011;6:e14738.
102. Boersma PD, Rebstock GA. Climate change increases reproductive failure in Magellanic penguins. *Chiaradia a. PLoS One.* 2014;9:e85602.
103. Boersma PD. Population trends of the Galápagos penguin: impacts of el Niño and la Niña. *condor.* 1998;100:245–53.
104. Cobb KM, Charles CD, Cheng H, Edwards RL. El Niño/southern oscillation and tropical Pacific climate during the last millennium. *Nature.* 2003;424:271–6.
105. Latif M, Kleeman R, Eckert C. Greenhouse warming, decadal variability, or el Niño? An attempt to understand the anomalous 1990s. *J Clim.* 1997;10: 2221–39.
106. Cimino MA, Fraser WR, Irwin AJ, Oliver MJ. Satellite data identify decadal trends in the quality of *Pygoscelis* penguin chick-rearing habitat. *Glob Change Biol.* 2013;19:136–48.
107. Kogut MH, Klasing K. An immunologist's perspective on nutrition, immunity, and infectious diseases: introduction and overview. *J Appl Poult Res.* 2009; 18:103–10.
108. Grimaldi WW, Seddon PJ, Lyver PO, Nakagawa S, Tompkins DM. Infectious diseases of Antarctic penguins: current status and future threats. *Polar Biol.* 2015;38:591–606.
109. McClintock J, Ducklow H, Fraser W. Ecological responses to climate change on the Antarctic peninsula. *Am Sci.* 2008;96:302.
110. Williams TD. *The penguins: Spheniscidae.* Oxford. New York: Oxford University Press; 1995.

Submit your next manuscript to BioMed Central and we will help you at every step:

- We accept pre-submission inquiries
- Our selector tool helps you to find the most relevant journal
- We provide round the clock customer support
- Convenient online submission
- Thorough peer review
- Inclusion in PubMed and all major indexing services
- Maximum visibility for your research

Submit your manuscript at  
[www.biomedcentral.com/submit](http://www.biomedcentral.com/submit)

