

ARTICLE



Selection-driven adaptation to the extreme Antarctic environment in the Emperor penguin

Federica Pirri^{1,2}, Lino Ometto³, Silvia Fuselli⁴, Flávia A. N. Fernandes^{1,5}, Lorena Ancona¹, Nunzio Perta¹, Daniele Di Marino¹, Céline Le Bohec^{5,6}, Lorenzo Zane² and Emiliano Trucchi¹✉

© The Author(s), under exclusive licence to The Genetics Society 2022

The eco-evolutionary history of penguins is characterised by shifting from temperate to cold environments. Breeding in Antarctica, the Emperor penguin appears as an extreme outcome of this process, with unique features related to insulation, heat production and energy management. However, whether this species actually diverged from a less cold-adapted ancestor, more ecologically similar to its sister species, the King penguin, is still an open question. As the Antarctic colonisation likely resulted in vast changes in selective pressure experienced by the Emperor penguin, the relative quantification of the genomic signatures of selection, unique to each sister species, could answer this question. Applying phylogeny-based selection tests on 7651 orthologous genes, we identified a more pervasive selection shift in the Emperor penguin than in the King penguin, supporting the hypothesis that its extreme cold adaptation is a derived state. Furthermore, among candidate genes under selection, four (*TRPM8*, *LEPR*, *CRB1*, and *SFI1*) were identified before in other cold-adapted homeotherms, like the woolly Mammoth, while other 161 genes can be assigned to biological functions relevant to cold adaptation identified in previous studies. Location and structural effects of *TRPM8* substitutions in Emperor and King penguin lineages support their functional role with putative divergent effects on thermal adaptation. We conclude that extreme cold adaptation in the Emperor penguin largely involved unique genetic options which, however, affect metabolic and physiological traits common to other cold-adapted homeotherms.

Heredity (2022) 129:317–326; <https://doi.org/10.1038/s41437-022-00564-8>

INTRODUCTION

Adaptation to severe Arctic and Antarctic temperatures is rare among terrestrial vertebrates and restricted to warm-blooded lineages (Storey and Storey 1992; Blix 2016). The main challenge for adaptation to extreme cold is keeping adequately high core body temperature, which in homeotherms can be obtained by the combination of several physiological, morphological, and behavioural responses (Allen 1877; Scholander 1955; Cannon and Nedergaard 2010; Tattersall et al. 2012; Blix 2016; Roussel et al. 2020): *i*) Minimising heat dissipation from the body surface *e.g.*, by reducing the surface area-to-volume ratio, by raising insulating hair or feathers, by decreasing peripheral circulation, and/or by balling up or huddling; *ii*) Increasing heat production via shivering and nonshivering thermogenesis through brown adipocytes; *iii*) Temporarily hibernating (*i.e.*, torpor in birds).

In contrast, the genetic basis of cold adaptation in homeotherms is not well understood. Only few candidate genes were found in common across cold-adapted mammals and birds (Yudin et al. 2017; Wollenberg Valero et al. 2014), suggesting that cold adaptation is likely the result of selection on different genes, which are nevertheless relevant to the same set of physiological and metabolic functions. For example, only four candidate genes, which are related to cardiovascular function, were found in

common in three out of six mammals dwelling in/near the Arctic or Antarctica (Yudin et al. 2017). However, different candidate genes could be involved in heart and vascular development and regulation (Liu et al. 2014; Vianna et al. 2020). Different genes acting in fatty-acid metabolism have been identified to be under positive selection in the polar bear and the Arctic fox (Liu et al. 2014; Castruita et al. 2020; Kumar et al. 2015). In mammals, the molecular mechanisms underpinning heat production via non-shivering thermogenesis have been extensively investigated and have been mainly linked to UCP-1, a mitochondrial uncoupling protein involved in the generation of heat in brown adipose tissue (Lowell and Spiegelman 2000). Non-shivering thermogenesis is instead not well understood in birds, where it is not clear whether the avian homologue of UCP-1 (*i.e.*, avian uncoupling protein, avUCP) has a similar thermogenic role or is mainly required against oxidative stress (Talbot et al. 2004). However, the same non-shivering thermogenic pathways could be used in birds (Tigano et al. 2018), and even in ectothermic vertebrates like reptiles (Akashi et al. 2016).

Birds are known for their fast adaptation rate, which allowed the colonisation of a huge diversity of environments, including polar ones (Zhang et al. 2014). In particular, the clade of penguins, which likely originated in temperate environments, successfully

¹Department of Life and Environmental Sciences, Marche Polytechnic University, Ancona, Italy. ²Department of Biology, University of Padova, Padova, Italy. ³Department of Biology and Biotechnology, University of Pavia, Pavia, Italy. ⁴Department of Life Sciences and Biotechnology, University of Ferrara, Ferrara, Italy. ⁵Université de Strasbourg, CNRS, IPHC UMR 7178, F-67000 Strasbourg, France. ⁶Département de Biologie Polaire, Centre Scientifique de Monaco, Monaco, Monaco. Associate editor: Rowan Barrett.

✉email: e.trucchi@univpm.it

diversified in the cold Antarctic and sub-Antarctic ecosystems (Pan et al. 2019; Vianna et al. 2020), featuring unique adaptations for insulation, heat production and energy management (Scholander 1955; Rowland et al. 2015). Our understanding of the underlying genetic determinants of such adaptations is still rather scarce. Testing about one third of the total genes across all penguin genomes, Vianna et al. (2020) identified blood pressure, cardiovascular regulation, and oxygen metabolism in muscles as functions, potentially involved in thermoregulation, which have been target of positive selection. Analyses of *Spheniscus* and *Pygoscelis* mitochondrial genomes also revealed a correlation between the pattern of diversity of the ND4 gene and sea surface temperature, suggesting this gene is involved in environmental temperature adaptation (Ramos et al. 2018). Moreover, Emperor and Adelle penguin genomes show the highest rate of duplication of β -keratin genes as compared to non-penguin birds, suggesting important changes in feathers and skin structure during their evolution to increase core body insulation (Li et al. 2014). Alternative gene pathways related to lipid metabolism and phototransduction appeared as under selection in these two species (Li et al. 2014).

The Emperor penguin (*Aptenodytes forsteri*) is the only warm-blooded vertebrate thriving and breeding during the harshest Antarctic winter, facing profound seasonal changes in daylight length as well as severe cold and wind conditions (Blix 2016; Goldsmith and Sladen 1961). To withstand such hostile environment, the Emperor penguin shows multiple morphological, physiological, behavioural adaptations, like improved thermoregulation systems in the head, wings, and legs (Frost et al. 1975; Thomas and Fordyce 2008), and efficient energy storage management system for long-term fasting (Groscolas 1990; Cherel et al. 1994; Groscolas and Robin 2001). Conversely, its sister species, the King penguin (*A. patagonicus*), breeds exclusively in year-round ice-free sub-Antarctic islands and in Tierra del Fuego. Extreme cold adaptation in the Emperor penguin has been suggested to be a derived feature from a less cold-adapted ancestor, likely more ecologically similar to the King penguin (Vianna et al. 2020). Such marked ecological transition should have left a clear signature of selection change across the genome of the Emperor penguin, with some genes becoming the novel targets of positive selection while others getting released from previous selective pressures associated with the ancestral habitat.

Here, we apply phylogeny-based tests to identify genes that markedly changed in their selection regime during the evolutionary history of the Emperor penguin, using its less cold-adapted sister species, the King penguin, as a control. If the common ancestor ecology was similar to the King penguin one, we expect a more intense selection shift across the Emperor penguin genome, with genes under positive selection related to adaptations to cold. By using a phylogenetic framework including seven species of penguins and 13 other birds, we compare the pattern of molecular evolution (Yang 2007; Wertheim et al. 2015) between Emperor and King penguins across 7651 orthologous genes and explore the gene ontology (GO) terms to identify the molecular functions that may have undergone positive selection in the Emperor penguin. To allow for a broader comparison with other cold-adapted vertebrates, we also investigated the overlap between the biological functions of the candidate genes identified in the Emperor penguin and the metabolic and physiological functions related to cold adaptation found in previous studies.

METHODS

Orthologous coding sequences identification

In order to test for selective signatures in coding sequences of Emperor and King penguins, we implemented a comparative phylogenomic analysis. We selected at least one species for each extant penguin genus based on the phylogeny of Pan et al. (2019), and other bird species to

evenly represent the clade of Core Waterbirds (which includes penguins), the tropicbirds (one species only), and, less densely, the Core Landbirds according to Jarvis et al. (2014). The resulting dataset included seven penguin species (*Eudyptula minor minor*, *Spheniscus magellanicus*, *Eudyptes chrysolophus*, *Pygoscelis papua*, *Pygoscelis adeliae*, *Aptenodytes patagonicus* and *Aptenodytes forsteri*) and 13 additional bird species (*Phaethon lepturus*, *Eurypyga helias*, *Gavia stellata*, *Fulmarus glacialis*, *Phalacrocorax carbo*, *Nipponia nippon*, *Egretta garzetta*, *Pelecanus crispus*, *Haliaeetus leucocephalus*, *Tyto alba*, *Cariama cristata*, *Corvus brachyrhynchos* and, as a more distant outgroup, *Opisthocomus hoazin*). We reconstructed the topology of the phylogenetic relationships among the selected species on the basis of the high-quality whole-genome phylogenies in Jarvis et al. (2014) and Pan et al. (2019). All coding sequences (CDS) of each of these twenty bird species were downloaded from GigaDB and Genbank (Supplementary Table S1).

One-to-one orthologs were identified by applying a reciprocal best-hit approach using pairwise BLAST searches with an e-value cut off of $1e-15$, a nucleotide sequence identity of at least 70%, and a fraction of aligned CDS of at least 60% (Savini et al. 2021). Only CDS longer than 150 bp that were a reciprocal best-hit between the Emperor penguin and the other species were retained. The orthologous gene sequences were then aligned with MAFFT (Madeira et al. 2019) and the alignments trimmed to maintain the open reading frame using a custom perl script. The resulting nucleotide alignments were re-aligned using the PRANK algorithm (Löytynoja 2013) in TranslatorX (Abascal et al. 2010), which aligns protein-coding sequences based on their corresponding amino acid translations. Since the results of dN/dS analysis could have been affected by poorly aligned regions or by regions that were too different to be considered reliably as orthologous, we removed CDS alignments that included internal stop codons and used a custom perl script to remove problematic regions as in Han et al. (2009) (see also Ramasamy et al. 2016). In fact, small inversions, especially when considering many species spanning a broad phylogenetic space, or sequencing artefacts could result in very “diverged” regions. Such differences would then be interpreted by a phylogeny-based selection test as the result of a high substitution rate (especially concerning dN), when in fact they are not. Furthermore, as phylogenetic-based selection tests are not able to properly deal with alignment gaps (Yang 2007), we filtered all the alignments with a custom perl script and kept only sites that were unambiguously present in at least 16 of the 20 sequences and always present in both our species of interest (i.e., King and Emperor penguins). The minimum length of an alignment for subsequent analyses was set to 150 bp (50 codons). Our full pipeline for one-to-one orthologs finding, nucleotide and amino acid alignment, and gene alignments cleaning is available at <https://github.com/evolinus/penguins>, with a step-by-step tutorial.

Identification of selection regime shifts

We first used CODEML in the PAML package (Yang 2007) to estimate synonymous (dS) and nonsynonymous (dN) substitution rates and to identify genes which are characterised by lineage-specific ω values (i.e., dN/dS , which is a proxy for the level of past selective pressure in the gene). In particular, we separately investigated the two scenarios of different ω in the King or in the Emperor lineage. If one branch in the phylogeny shows a value of ω (ω_i) significantly larger than in the other branches (i.e., ω_b : background ω), such a foreground lineage may have been targeted by positive (Darwinian) or relaxed selection (allowing the accumulation of nonsynonymous substitutions), whereas when ω_i is lower than ω_b , the foreground lineage may have evolved under stronger selective constraints (e.g., purifying selection). To identify potential signatures of selection started before the divergence of the Emperor and the King lineages, we also tested their ancestral branch for lineage-specific ω values.

We first ran the two-ratio *branch* model (one ω for the foreground branch, another ω for the background branches; set parameters “model = 2, NSsites = 0, fix_omega = 0”) and the one-ratio *branch* model (one ω estimate for all branches, as null model; set parameters “model = 0, NSsites = 0, fix_omega = 0”) on the unrooted phylogenetic tree of the species of interest. We determined the topology of such a tree (Fig. 1A) by manually combining the total evidence nucleotide tree of the avian family (Jarvis et al. 2014) and the phylogenomic reconstruction of penguins (Pan et al. 2019). The two models (two-ratios vs. one-ratio) were compared by likelihood ratio tests (LRTs). False discovery rates (FDR) were computed using the qvalue package (Storey et al. 2017) and the p.adjust function in R (R Core Team 2013) using the Benjamini-Hochberg procedure to adjust for

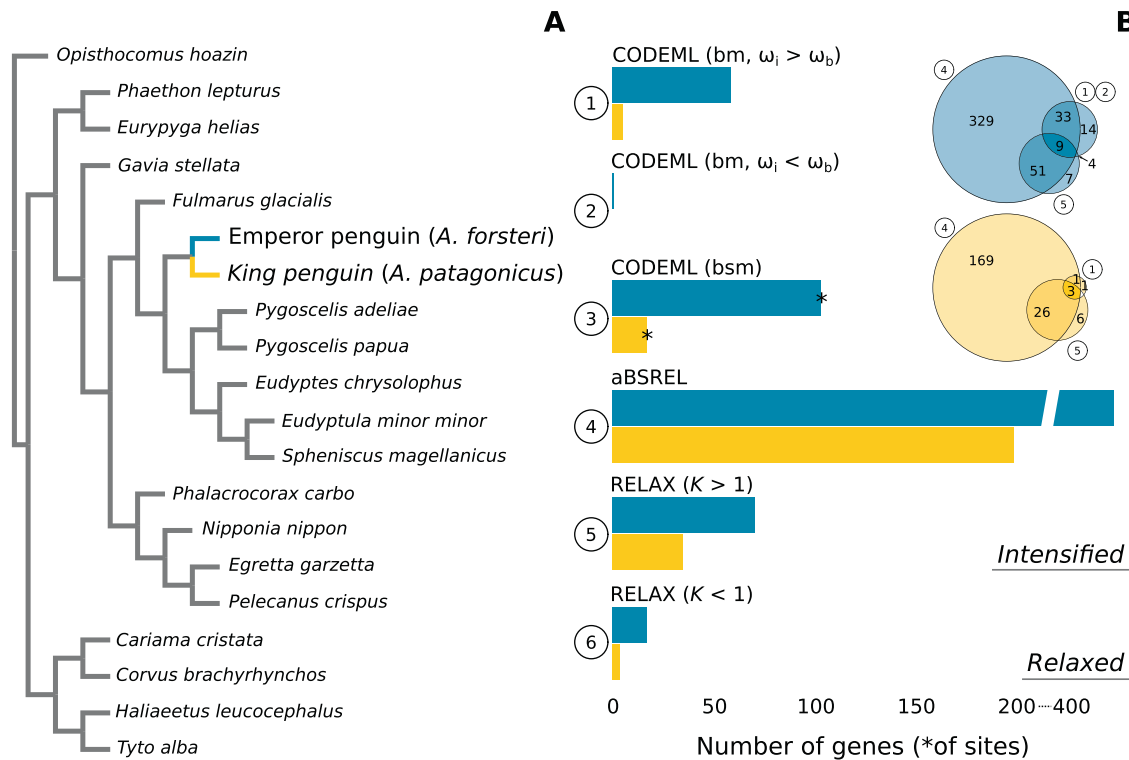


Fig. 1 Selection regime shifts in Emperor and King penguins genomes. **A** Phylogenetic tree used in the selection tests based on Jarvis et al. 2014 and Pan et al. 2019 (refer to the original phylogenetic trees for nodes support). The Emperor and the King penguin are highlighted in blue and yellow, respectively. Note that branch length is not to scale. **B** Comparison between Emperor and King penguins for genes with FDR > 0.05 in each of the tests performed; bm: branch model; bsm: branch-site model; ω_i : ω in the target species; ω_b : background ω ; For sake of completeness, we also show the genes putatively under selection according to RELAX (with $K > 1$). **Inset**. Venn diagram showing the overlap among CODEML (bm), aBSREL, and RELAX ($K > 1$). Note that the total number of genes is different between the Emperor and the King penguins and the size of the circles scales to the maximum in each of the two graphs. The overlap between CODEML (bm) or/and aBSREL with RELAX ($K < 1$) is 3, 1, and 1 gene, respectively (not shown).

multiple testing. An FDR or adjusted p -values significant threshold of 0.05 was used.

We then used a branch-site model to test for sites under selection in the candidate genes from the previous test. The parameters for the null model were set as “model = 2, NSsites = 2, fix_omega = 1”, while the parameters for the alternative model were set as “model = 2, NSsites = 2, fix_omega = 0”. LRT and FDR were computed as for the branch model tests. To control for misalignments that could have biased the results, we visually checked the sequence alignment of all candidate genes under selection using MEGA7 (Kumar et al. 2016).

We also applied aBSREL (adaptive Branch-Site Random Effects Likelihood) from the HyPhy package to our set of orthologous coding sequences. Jointly modelling site-level and branch-level ω heterogeneity, aBSREL tests for a branch of interest in the phylogeny if a proportion of sites evolved under positive selection, without testing for selection at specific sites. (Smith et al. 2015; Pond et al. 2005). In contrast to the branch-site test in CODEML, which assumes 4 ω rate classes for each branch and assigns each site to one of these classes, the aBSREL test uses AIC_c (Akaike Information Criterion with correction for small sample size) to infer the optimal number of ω rate categories per branch, not making the assumption that all branches exhibit the same degree of substitution rate heterogeneity (Smith et al. 2015). Although we expect a broad overlap between CODEML and aBSREL results, a higher sensitivity should characterise the latter approach (Smith et al. 2015). Signatures of positive selection were searched by setting a priori the King and the Emperor penguin lineages as test branches in the phylogeny. LRT was performed by comparing the full model to a null model where branches were not allowed to have rate classes of $\omega > 1$. A Benjamini-Hochberg correction was used to control the probability of making false discoveries and only tests with adjusted p -values < 0.05 were considered significant.

Beside presenting novel drivers of selection, the major ecological shift occurring in the Emperor penguin lineage should have also released some of the selective constraints characterising the ancestral ecological niche. As a consequence, some genes could show a signature of relaxed selection in

this lineage, a higher number than in the King penguin. To test for relaxation of selective constraints on a specific lineage we used RELAX from the HyPhy package, a general hypothesis testing framework that determines whether the strength of natural selection has been relaxed or intensified along a set of test branches defined a priori in a phylogenetic tree (Wertheim et al. 2015; Pond et al. 2005). It estimates a selection intensity parameter K , in which a significant $K > 1$ indicates intensification in the selection strength, whereas a significant $K < 1$ indicates relaxation in the strength of selection in the test branches (Wertheim et al. 2015). We tested whether selection pressure has increased or decreased in either the King or in the Emperor penguin lineage as compared to the rest of the phylogeny. In the null model, the selection intensity parameter K was set to 1 for all the branches of the phylogenetic tree, whereas in the alternative model the parameter K was inferred for every tested branch. The increase or relaxation of selection was validated by a LRT with 1 degree of freedom. Again, the Benjamini-Hochberg procedure was used to adjust for multiple testing with adjusted p -values < 0.05 considered significant.

Functional characterization of candidate genes under selection

To test whether the set of candidate genes for positive selection in the Emperor penguin lineage are functionally involved in cold adaptation, we tested these genes for functional GO terms enrichment by using the g:GOST function in g:Profiler (Raudvere et al. 2019). GO terms were assigned to candidate genes based on the Ensembl GO predictions for the flycatcher (*Ficedula albicollis*). Significantly enriched categories included at least two genes, and the Benjamini-Hochberg method was used for multiple testing correction to estimate significance (at adjusted p -values < 0.05). We used REVIGO (Supek et al. 2011) to summarise the resulting lists of GO terms in order to obtain a non-redundant and more easily interpretable set of GO terms. GO terms enrichment was performed on two lists of genes: i) candidate genes supported by both CODEML and aBSREL and ii) candidate genes supported by either CODEML or aBSREL.

Table 1. Candidate genes for positive/relaxed selection inferred by CODEML (*branch model* – *bm*, *branch-site model* – *bsm*), aBSREL, and RELAX ($K > 1$) which we could assign to biological functions suggested as related to cold adaptation in previous studies (see References within the table) on the basis of their GeneCards description.

Function/ metabolism of	CODEML (bm)	CODEML (bsm)	aBSREL	Relax ($K > 1$)	Relax ($K < 1$)	GO ID	Previously identified genes	References
Cardiovascular activity and regulation			DNASE1, ADAMTS13, EIF2AK1, MYBPC3, UBA5, ENTPD1, NRP1, PKDCC, ROCK2, LMOD2, PLEKHM2, ATP11C, HGSNAT, TFRC, UFL1, SPIC, MYOCD, STXBP4, FBLN5, MYLK3, AGXT2, FLVCR1	AGXT2, MYLK3	FLVCR1, EDRF1	GO:0072359 (adj <i>p</i> - val = 0.01); GO:0060047 (adj <i>p</i> - val = 0.04); GO:0003306 (adj <i>p</i> - val = 0.02)*; GO:1901863 (adj <i>p</i> - val = 0.03)*	SPTBN5, KNG1, ICAM4, DSP, TNN, ARID5B, CUL7, EHD3, VCL, ALPK3, XIRP1, ADORA2B, SOD1, POMC, CD36, ENPEP, MME, BDKRB2	Yudin et al. 2017; Liu et al. 2014; Wollenberg Valero et al. 2014; Vianna et al. 2020
Skin thickness			PRDM5, KIAA1211, EGF, FAM83G, MATN3, DSC1, SGSMB, GAB1, RAB11FIP1, ASAH1	ASAHI			DSP, EXPH5, EVPL, DSG1	Yudin et al. 2017; Li et al. 2014
Feathers development	WNT3A, RPL31**	WNT3A	SHOC2, RPL31**, WNT3A	RPL31**			(beta-keratin gene families)	Li et al. 2014
Immunity	DHX36, ANKRD17, TEC, SLIT2, TRIM29	DHX36, ANKRD17	MILR1, TRAF3IP2, SH2D4B, SPLL2B, MAP3K5, TBK1, IFIH1, RASGRP1, IFIT5, PLD4, CD101, LGALS8, ASB9, CMAS, AOAH, C7, MRC1, JCHAIN, HERC5, ATG5, LTA4H, ANKRD17, TRIM29	HERC5, ATG5, LTA4H, DHX36, TEC	ANKRD17	GO:0002376 (adj <i>p</i> - val = 0.0002); GO:0006952 (adj <i>p</i> - val = 0.01); GO:0032607 (adj <i>p</i> - val = 0.02)	SELPLG, ZDBF2, MAPK1, MAPK14	Yudin et al. 2017
Lipid and fatty acid	ASAP1	ASAP1	FAAH, VTG2, LEPR , PLCXD1, OSBPL2, LMAN1, PPARA, FADS1, ACSBG2, FNDC3B, GK5, ARSK, PDZRN3, DECR2, ACOT7, ACOX3, KIAA1468, PCCA, ACOT12, VPS13A	HMGCL, PCCA, ACOT12, VPS13A	ASAP1	GO:0006631 (adj <i>p</i> - val = 0.04); GO:0006071 (adj <i>p</i> - val = 0.05); GO:0033993 (adj <i>p</i> - val = 0.03)*	NRIID2, LPL, CETP, SOD1, LEPR , APOB, ABCC6, UCP3, FASN, GLPTD1, AKT2, SLC38A4, SF11	Kumar et al. 2015; Yudin et al. 2017; Castruita et al. 2020; Liu et al. 2014; Tigano et al. 2018; Wollenberg Valero et al. 2014; Li et al. 2014; Lin et al. 2019
Glucose (including insulin)	ADCY5, IDE, SLC2A1, PAK1	IDE, PAK1	AAED1, ALG6, SLC17A9, ABCA10, TTYH3, ABCC8, GRB14, GAB1, PHKB, STXBP4, NRG4, SLC7A6, AGL, PGM2L1, IDE, ADCY5, SLC2A1	SLC7A6, AGL, PGM2L1, SF11	PAK1	GO:0071333 (adj <i>p</i> - val = 0.05)*	SF11 , SLC38A4, AKT2, GLPTD1	Kumar et al. 2015; Yudin et al. 2017
Thyroid hormones			NCOA3, MARCH6			GO:0009725 (adj <i>p</i> -val = 0.03)	THRB, RARB	Tigano et al. 2018

Table 1. continued

Function/ metabolism of	CODEML (bm)	CODEML (bsm)	aBSREL	Relax ($K > 1$)	Relax ($K < 1$)	GO ID	Previously identified genes	References
Non-shivering thermogenesis			PRDM16, L2HGDH, CACNA1D, ATP1A1				UCPs, avUCP, THRB, RARB, NR1D2	Lowell and Spiegelman 2000; Talbot et al. 2004; Tigano et al. 2018
Shivering thermogenesis							NEB, ADORA2B, SOD1, HSPB2	Yudin et al. 2017; Wollenberg Valero et al. 2014
Response to oxidative stress	GMEB1, ATOX1	GMEB1	MAP3K5, PTGR2, SELENOP, COQ6, GMEB1, PEX16	COQ6	VRK2		SOD1, UCP3	Wollenberg Valero et al. 2014
Stress response			EIF2AK1, CUL4B, MAP3K3, ULK2, SLC9A8, MAP4K3, TH	HSP40		GO:0006950 (adj p - val = 4.0E-05)	HSP47, HSPA8, HSPB2, STUB1, MAPK1, UCP3, HSF1, MAPK14, UNG, SOD1	Wollenberg Valero et al. 2014
Homeostasis	AQP2		CHERP, AQP11, SLC9A8, ATP1A1, AQP2			GO:0006874 (adj p - val = 0.03); GO:0072503 (adj p - val = 0.04)	LPL, CETP; SLC6A19, ACE2, AGT	Wollenberg Valero et al. 2014; Vienna et al. 2020
Circadian rhythm	CREB1, PHLPP1		PROX1, ROCK2, GPR176, ASMT, PTCHD1		PHLPP1		ADCY2, ADCY8, CALML14, CAMK2	Lin et al. 2019
Phototransduction			RGS9, CRB2, MREG, ARR3, TLE4, ABCB5, FOXN4, GRIFIN, SPATA7, ABCA4, CRB1	CRB1		GO:0042462 (adj p - val = 0.04)	CNGB1, MYO3A, UACA, CRB1 , CRY2, MYO3B	Li et al. 2014
Mitochondrial activity	GLS	GLS	CHCHD6, PTCD3, SLC25A4, VPS13D, SPG20, SERAC1, OPA1, SLC25A3, PNPT1, LRPPRC, GLS	SLC25A3, PNPT1, LRPPRC, GLS			COX1, ATP8, ND4L, ND4	Ramos et al. 2018
Temperature sensing	TRPM8	TRPM8	TRPM8				TRPM8 , TRPV3, TRPV4, TRPA1, TRPM4	Lynch et al. 2015
Membrane fluidity	SGMS2	SGMS2	SGMS2			GO:1905371 (adj p - val = 0.04)*	–	This study

Candidate genes under selection in the Emperor penguins which were also found in previous studies are in bold (four genes).

Note that, in the GO ID column, we associated the Gene Ontology Biological Process supported by gProfiler in candidate genes inferred by aBSREL or CODEML (bm), or in the subset of genes inferred by both (marked with *), with relevant biological functions; the GOs in this column do not only refer to the genes indicated in the previous columns. **Weak evidence of gene expression change in a study on feathers development (Ng et al. 2015). Biological function *membrane fluidity* is proposed herein as related to cold adaptation.

To compare our results with the recent literature on genetics of cold adaptation, we compiled a list of biological/molecular functions characterising the candidate genes for cold adaptation retrieved in previous studies in vertebrates (Table 1). The list included: *cardiovascular activity and regulation, skin thickness, immunity, lipid and fatty acid metabolism, glucose (including insulin) metabolism, thyroid hormones, non-shivering thermogenesis, shivering thermogenesis, response to oxidative stress, stress response, homeostasis, circadian rhythm, phototransduction, mitochondrial activity, feathers development, temperature sensing*. We checked whether any of the GO terms enriched in candidate genes supported by either CODEML or aBSREL could be assigned to any of the 15 biological/molecular functions listed above. In addition, we assigned, whenever possible, genes from this list to the same biological/molecular functions, using the gene function description from the human gene database GeneCards (Stelzer et al. 2016).

RESULTS

Signature of selection shift in the Emperor penguin lineage

We identified 7651 orthologous coding sequences across seven penguin species and 13 other birds, corresponding to about 50% of the total number of genes in an avian genome (Zhang et al. 2014). Across all of the tests, by applying a significance threshold of $FDR < 0.05$, we consistently identified more candidate genes which underwent a shift in their selection regime (intensified or relaxed) in the Emperor penguin lineage than in the King penguin one (Fig. 1B). Even though we found a much larger number of genes putatively under selection using the aBSREL model (Supplementary Tables S2, S3), the overlap between the CODEML (with $\omega_i > \omega_b$) *branch* model (Supplementary Table S4) or RELAX (with $K > 1$; Supplementary Tables S5, S6) with aBSREL was on average 80% of the former ones (Fig. 1B, inset), which is significantly greater than expected by chance (Supplementary Fig. S1).

Using the CODEML *branch* model, we found 59 candidate genes with signals of positive or relaxed selection in the Emperor penguin, showing a ω significantly greater than the background, and one candidate gene under purifying selection showing a lower ω value than the background. In comparison, only five genes, with ω greater than the background, were retained as candidates of positive or relaxed selection in the King penguin lineage. Although greater than those of the background branches, the ω values of most of the candidate genes in the Emperor or the King penguin, still remain lower than one, making it difficult to distinguish between positive selection and relaxation of purifying selection. The CODEML *branch-site* test indicated a total of 104 sites in the 60 candidate genes in the Emperor lineage, whereas 17 were suggested in the five genes in the King lineage (Fig. 1B; Supplementary Table S7). Five genes were identified as candidate under selection in the branch ancestral to Emperor and King penguins, two of which appear as candidate genes also in the Emperor lineage (Supplementary Table S4). On the other hand, aBSREL identified a much larger number of candidate genes under positive selection than CODEML *branch* model (422 in the Emperor lineage and 199 in the King lineage). Between 70 and 80% of the CODEML candidate genes (42 and 4 genes considering the Emperor and the King penguin lineage, respectively) were also significant in aBSREL results (Fig. 1B). According to RELAX, 17 genes in the Emperor penguin lineage and four genes in the King penguin lineage bear a significant signature of relaxed selection ($K < 1$; Fig. 1B; Supplementary Tables S8, S9). Concerning the Emperor penguin lineage, five of the 17 genes (i.e., *FLVCR1*, *ANKRD17*, *ASAP1*, *PAK1*, *PHLPP1*) are also candidates in either CODEML (*branch* model, $\omega_i > \omega_b$), aBSREL, or both, further supporting the signal of relaxed purifying selection. Observed overlap of candidate genes among different tests within each species is greater than random expectation (Supplementary Fig. S1). Considering the joint results from all selection tests, 16 genes showed signature of positive selection in both the Emperor and the King lineage. In this case, however, the observed overlap is within the 95% intervals of the null (random) distribution.

Biological functions enrichment in Emperor penguin candidate genes

After correcting for multiple tests and using REVIGO's redundancy elimination algorithm, we found 16 enriched GO biological process terms in the candidate genes for positive selection suggested by both CODEML and aBSREL (Supplementary Table S10) in the Emperor penguin lineage. Some of these GOs are related to heart and muscle development (GO:0003306, GO:1901863), and to metabolism of lipid (GO:0033993), glucose (GO:0071333) and sphingolipid, like for example ceramide, (GO:1905371). When considering all candidate genes supported by either CODEML or aBSREL in the Emperor penguin lineage, we retrieved 34 enriched GO biological process terms (Supplementary Table S11), 12 of which could be assigned to one of the biological/molecular functions putatively related to cold adaptation from previous studies (Table 1). In addition, manually screening the biological functions of all candidate genes using the human database GeneCards, we identified 161 genes which could be assigned to the biological/molecular functions identified in previous studies (Table 1). Four genes identified as under selection in the Emperor penguin were already found in previous studies: *TRPM8* (*temperature sensing*), indicated as under selection by both CODEML (*branch* and *branch-site* models) and aBSREL; *LEPR* (*lipid and fatty acid metabolism*) and *CRB1* (*phototransduction*) were suggested as candidate genes under selection by aBSREL; *SFI1* (*glucose metabolism*) which showed a significant signal of intensified selection in the RELAX test ($K > 1$). In addition, the sphingomyelin synthase2 (*SGMS2*), seems markedly related to cold adaptation, as it regulates biological membrane fluidity at low temperatures (Wang et al. 2014). We propose to add this function (named e.g., as *membrane fluidity*) to those already suggested as related to cold adaptation in vertebrates (Table 1), in order to be considered for investigation in other biological systems.

DISCUSSION

Signature of major novel genetic changes in the Emperor penguin adaptation to Antarctica

A larger fraction of the genes tested in our analyses shows signatures of novel selection regimes in the Emperor penguin as compared to the King penguin, either as intensification or relaxation of selection pressure (Fig. 1B). One possible explanation of this pattern is that the ancestor of both species had ecological preferences which were more similar to the King penguin, while the Antarctic ecology of the Emperor penguin is a derived, though rather recent (ca. 1–2 Mya; Gavryushkina et al. 2017), adaptation. Interestingly, two of the five genes identified as under selection in the ancestral lineage of Emperor and King penguins also showed a significant signature of selection in the Emperor penguin lineage and one of the two (*TRPM8*) appear as related to cold adaptation, opening further questions on the ecological pressure that led to the following species divergence. A general shift from warmer to colder habitats has been suggested for penguins in general (Vianna et al. 2020). Rapid adaptation to polar lifestyle is not new in homeotherms, as suggested for the recent divergence (less than 0.5 Mya) of polar bear from brown bear (Liu et al. 2014; Castruita et al. 2020), or of Arctic fox from its common ancestor with red fox (ca. 2.9 Mya; Kumar et al. 2015). King and Emperor penguins critically differ in their breeding range, as they reproduce in temperate-cold sub-Antarctic islands with average winter temperature of ca. 3 °C, or on the Antarctic sea-ice featuring average winter temperature of –25 °C, respectively. Upon colonisation of Antarctica, novel selective pressures are expected to have appeared while others, characterising the former ecology, should have relaxed. Relaxed purifying selection can in fact be an additional source of evolutionary novelties (Hunt et al. 2011), as it can have non-linear consequences on a trait, including stabilising

or balancing selection, pseudogenization or, on the contrary, recruitment for a different function (Lahti et al. 2009). An alternative explanation for the observed pattern is that selection was more efficient along the Emperor penguin lineage due to its larger and constant population size through time, as revealed in previous studies (Cristofari et al. 2016), when compared to the King penguin markedly oscillating demographic trajectory (Trucchi et al. 2014; Cristofari et al. 2018; Trucchi et al. 2019). In fact, at low population size genetic drift overwhelms selection, leading to fixation of both synonymous and nonsynonymous variants, and potentially blurring the dN/dS ratios. However, the higher signature of relaxed purifying selection in the Emperor penguin contrasts with such an alternative explanation, thus leaning towards extreme cold adaptation as a derived ecology in this species.

Looking for a common genetic basis of cold adaptation in homeotherms

Many of the candidate genes under selection in the Emperor penguin are involved in functional pathways relevant to cold adaptation identified in previous studies (Table 1). Nonetheless, we discovered only four genes in common with those previously identified in other cold-adapted vertebrates (notably, two of which were found in the woolly mammoth). We found no overlap between our set of candidate genes under selection and those discovered in previous studies on penguins (Li et al. 2014; Vianna et al. 2020). Unlike our test which explicitly focuses on the contrast between the two sister *Aptenodytes* species, previous studies were either testing the differentiation between one penguin species (*A. forsteri* or *Pygoscelis papua*) and 48 other bird species (Li et al. 2014) or the general signal of adaptation in all of the penguin species (Vianna et al. 2020). Such lack of overlap is not surprising given the suggested polygenic basis of most phenotypic traits (Barghi et al. 2020) shaped by the contribution of many genes (Boyle et al. 2017). As also emerged in other vertebrates (see references in Table 1), traits related to cardiovascular function, lipid and fatty acid metabolism, glucose metabolism, oxidative stress and stress response, insulation (including skin thickness and feathers development), phototransduction and mitochondrial activity show several candidate genes under selection in the Emperor penguin (Table 1).

Genes involved in fatty-acid metabolism have been identified to be under positive selection both in polar bear and Arctic fox, indicating similar evolutionary constraints on fat metabolism in these two cold-adapted mammals (Liu et al. 2014; Kumar et al. 2015). The storage of subcutaneous fat is also crucial in the Emperor penguin, both because it represents the main source of energy during the long fasting periods (Blem 1990; Cherel et al. 1994; Groscolas 1990) and because it provides thermal insulation (Kooyman et al. 1976). Moreover, fatty-acids can stimulate muscle thermogenic processes in birds and therefore may be a very important component of the adaptive response to cold temperatures in penguins (Duchamp et al. 1999; Toyomizu et al. 2002; Talbot et al. 2004; Rey et al. 2010). Accordingly, our results revealed candidate genes under selection in the Emperor penguin such as *PPARA*, regulating eating behaviour (Fu et al. 2003), controlling lipid absorption in the intestine (Poirier et al. 2001) and fatty-acid oxidation (Lemberger et al. 1996), *ADCY5*, associated with body weight (Li and Li 2019), and *LEPR*, the leptin receptor, involved in fat and glucose metabolism, in appetite regulation through its effects on food intake and energy consumption (Zhang et al. 1994; Halaas et al. 1995; Pelleymounter et al. 1995), and in adaptive thermogenesis (Yang et al. 2011).

While shivering thermogenesis might be the main thermogenic mechanism in birds following short-term cold exposure (Teulier et al. 2010), cold acclimated birds show non-shivering thermogenesis mediated by *avUCP* expression within skeletal muscles (Talbot et al. 2004). According to our analyses, some candidate genes could

be assigned to the non-shivering thermogenesis category (Table 1), but none of them could be unambiguously associated with shivering thermogenesis. Among the former, *Na,K-ATPase (ATP1A1)* is a membrane enzyme that utilises energy derived from the hydrolysis of ATP to pump Na^+ , wasting energy as heat, thus playing a significant role in thermal tolerance and energy balance (Geering et al. 1987; Iannello et al. 2007). It was demonstrated that its expression is affected by heat stress (Sonna et al. 2002) and consistently increases during mammal hibernation (Vermillion et al. 2015). *L2HGDH* was found to be associated with the TCA cycle, electron transport and glycolysis (Oldham et al. 2006) and it was identified as one of the candidate genes under positive selection in three high-altitude passerine birds (Hao et al. 2019). *PRDM16* is a zinc-finger protein that activates brown fat-selective genes responsible for mitochondrial biogenesis and oxidative metabolism, while repressing the expression of a wide range of genes in white fat cells (Seale et al. 2007; Kajimura et al. 2015). This protein appears to play a role also in the development and function of beige cells (Ohno et al. 2012; Seale et al. 2011). Retinoic acid and thyroid hormones, whose candidate binding sites have been found in a mammal *UCP-1* enhancer (Lowell and Spiegelman 2000), and the nuclear receptor (*NR1D2*) have been suggested to be involved in thermal adaptation in birds (Tigano et al. 2018). *NR1D2* as well as *MARCH6* and *NCOA3*, also involved in thyroid hormone regulation and action (Zelcer et al. 2014; Ishii et al. 2021), influencing baseline temperature (Elliott et al. 2013) and thermoregulation in response to cold stimuli in birds (Vézina et al. 2015), were also present among Emperor penguin candidate genes.

One of the most promising candidate genes under selection is *TRPM8*, which encodes the sensor for noxious cold temperature (Yin et al. 2018), showing signatures of selection in both the ancestral and the Emperor lineages characterised by different codons (Fig. 2; Supplementary Information S1). Setting the physiological range of temperature tolerance (Matos-Cruz et al. 2017), any biological thermosensory apparatus should be under strong evolutionary pressures to noxious high or low temperatures (Myers et al. 2009). Indeed, evolutionary tuning of five temperature-sensitive transient receptor potential channels, including *TRPM8*, has been likely key in the adaptation of the woolly mammoth to the Arctic (Lynch et al. 2015). A previous study demonstrated the crucial role of a single-point mutation located at site 906 (as per *Gallus gallus* coordinates in Yin et al. 2018; 919 as per coordinates in Yang et al. 2020) for the activation of *TRPM8* pore domain channel in the Emperor penguin (Yang et al. 2020). Interestingly, our comparative selection scan suggested, instead, positive selection at two other sites (i.e., I1058T, M1069Y; Supplementary Table S7) in *Aptenodytes* lineages. After aligning 541 vertebrate *TRPM8* ortholog sequences available in GenBank (accessed on 22/11/2021), we found that the substitution at site 906 is not unique to Emperor penguins but it is instead widespread in birds with different ecology and habitat preference, including warm tropical regions (see Supplementary Information S1). Our candidate substitution M1069Y is also common in penguins and other birds, where methionine likely represents the ancestral amino acid state whereas tyrosine, found in the King penguin lineage, is the derived one. Conversely, the substitution I1058T (as per *Gallus gallus* coordinates in Yin et al. 2018), which characterises the Emperor penguin sequence, is extremely rare in birds, being present in one other species only (*Sitta europaea*). This substitution is also rare in mammals and reptiles where it has been detected so far in seven (only bats) and one species only, respectively. Both mutations are exposed toward the receptor channel in the distal carboxyl terminus domain (Fig. 2; Supplementary Information S1), a key structural element of *TRPM8* whose cold-induced folding determines cold-driven gating of the sensor channel (Diaz-Franulic et al. 2020): I1058T is located just after the last ultra-conserved residue present in the CTDH2 domain (Yin et al. 2018) in the helices

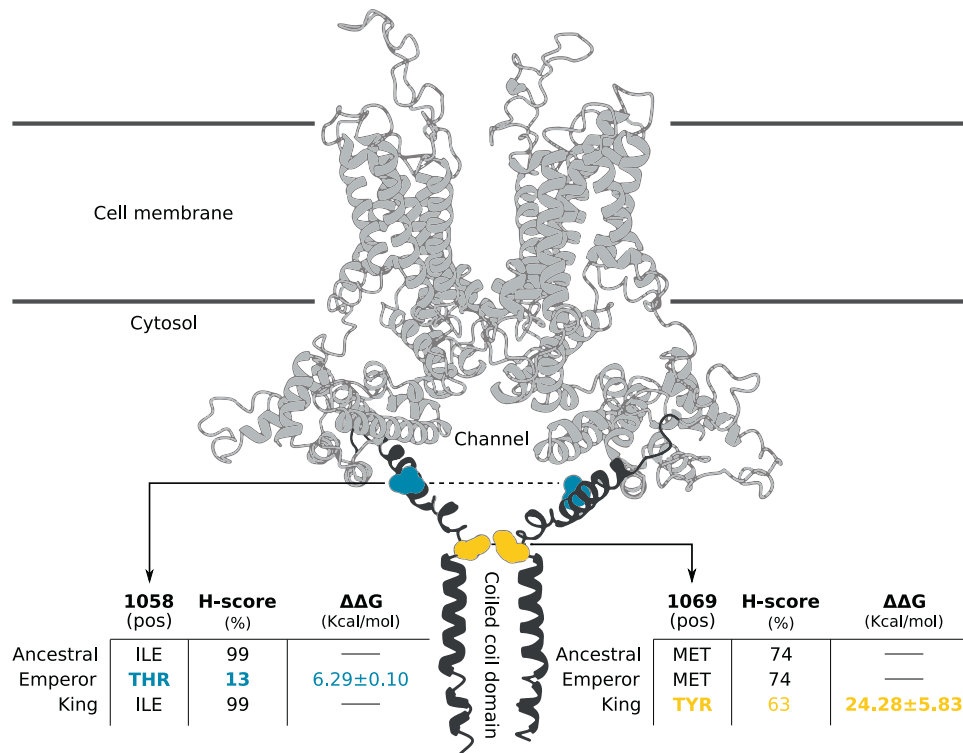


Fig. 2 Ligand-free TRPM8 structure obtained by comparative modelling (four subunits in total, only two shown for clarity). Carboxyl terminus domain, including the coiled coil, is shown in dark grey, while part of the rest of the protein is in light grey (MHR1/2/3 domains are not shown; see Supplementary Fig. S2 for completeness). Location of the substitutions found in Emperor (blue) and King (yellow) penguin lineages are also shown. H-score: hydrophobicity score (Monera et al. 1995); $\Delta\Delta G$: free energy difference as estimated by FoldX (Schymkowitz et al. 2005).

delimiting the channel, whereas M1069Y is right at the entrance of the coiled coil domain. Both amino acid changes (I1058T, M1069Y) alter the hydrophobic profile of the receptor channel (Fig. 2). Yang et al. (2020) suggests that decreased TRPM8 sensitivity (i.e., activation at lower temperature) correlates with lower total hydrophobicity of the channel side chains, hypothesizing this could be a fine-tuning molecular mechanism for thermal adaptation in vertebrates. According to this model, substitution I1058T, found almost uniquely in the Emperor penguin, has a strong impact on noxious cold temperature sensing in this species as it represents a change from very high (isoleucine hydrophobicity score is 99, with the score ranging from 0 to 100; Monera et al. 1995) to almost neutral (threonine hydrophobicity score is 13) hydrophobicity. The substitution M1069Y, found in the King penguin lineage, causes only a slight decrease in the hydrophobicity score (from 74 of the methionine to 63 of the tyrosine). In addition, preliminary estimates applying FoldX approach (Schymkowitz et al. 2005), suggest that substitution M1069Y results in a very intense structural destabilization ($\Delta\Delta G$ of $+24.28 \pm 5.83$ Kcal/mol; Supplementary Information S1) of the receptor while in the closed state, potentially increasing its temperature of activation. Of note, I1058T has a lower predicted destabilizing effect ($\Delta\Delta G$ of 6.29 ± 0.10 Kcal/mol). As it appeared multiple times during TRPM8 evolution across distantly related bird lineages, M1069Y seems like a common (although rather coarse as compared to the total channel hydrophobicity fine regulation mentioned above) switch for adaptation to (likely higher) temperature. All considered, we speculate that King and Emperor penguins adaptively moved in different directions from the thermal niche occupied by their common ancestor. Further molecular dynamics analyses, also comparing the effects of these mutations on different background sequences of TRPM8 from different species with different thermal niches, could help testing our evolutionary hypothesis and, more

generally, help understanding the mechanism of fine and coarse regulation of this fundamental noxious cold receptor.

CONCLUSION

The genetic basis of extreme cold adaptation in homeotherm vertebrates belongs to that subset of the evolutionary diversity of life which is both very rare (Blix 2016) and currently the most challenged by climate change (Gilg et al. 2012; Descamps et al. 2017). Beside a few target genes in common across species (Table 1), most of the molecular pathways at the basis of colonisation of Arctic and Antarctic environments appear as species-specific so that losing each of these uniquely adapted organisms corresponds to the loss of a whole evolutionary trajectory. Ecological specialisation, in particular for extreme environments, could be a higher risk condition for extinction (Colles et al. 2009) when environmental changes are as fast as in the case of the current global warming. Examples from the past are not encouraging. In fact, none of the extant rhino species, which are all adapted to warm climate, descends from any of the three extinct cold-adapted rhino species (Liu et al. 2021). A similar evolutionary endpoint could have characterized the diversification of elephants, with the extinction of the cold-adapted mammoths after the last glaciation (Lynch et al. 2015). Extreme cold adaptation in the Emperor penguin appears to be a derived trait based on a large set of unique genetic changes making this species a candidate evolutionary *cul-de-sac* in the contemporary climate change scenario.

DATA AVAILABILITY

All genomic data and protein structure used in this work were publicly available. Novel custom bioinformatic scripts used in this work are made publicly available at <https://github.com/evolinus/penguins>.

REFERENCES

- Abascal F, Zardoya R, Telford MJ (2010) TranslatorX: Multiple alignment of nucleotide sequences guided by amino acid translations. *Nucleic Acids Res* 38:7–13
- Akashi HD, Cádiz Díaz A, Shigenobu S, Makino T, Kawata M (2016) Differentially expressed genes associated with adaptation to different thermal environments in three sympatric Cuban *Anolis* lizards. *Mol Ecol* 25:2273–2285
- Allen JA (1877) The influence of Physical conditions in the genesis of species. *Radic Rev* 1:108–140
- Barghi N, Hermisson J, Schlötterer C (2020) Polygenic adaptation: A unifying framework to understand positive selection. *Nat Rev Genet* 21:769–781
- Blem CR (1990) Avian energy storage. *Curr Ornithol* 7:59–113
- Blix AS (2016) Adaptations to polar life in mammals and birds. *J Exp Biol* 219:1093–1105
- Boyle EA, Li Yi, Pritchard JK (2017) An expanded view of complex traits: From polygenic to omnigenic. *Cell* 169:1177–1186
- Cannon B, Nedergaard J (2010) Thyroid hormones: Igniting brown fat via the brain. *Nat Med* 16:965–967
- Castruita JAS, Westbury MV, Lorenzen ED (2020) Analyses of key genes involved in Arctic adaptation in polar bears suggest selection on both standing variation and de novo mutations played an important role. *BMC Genom* 21:1–8
- Cherel Y, Gilles J, Handrich Y, Le Maho Y (1994) Nutrient reserve dynamics and energetics during long-term fasting in the king penguin (*Aptenodytes patagonicus*). *J Zool* 234:1–12
- Colles A, Liow LH, Prizing A (2009) Are specialists at risk under environmental change? Neoeological, paleoecological and phylogenetic approaches. *Ecol Lett* 12:849–863
- Cristofari R, Bertorelle G, Ancel A, Benazzo A, Le Maho Y, Pongonis PJ et al. (2016) Full circumpolar migration ensures evolutionary utility in the Emperor penguin. *Nat Commun* 7:1–9
- Cristofari R, Liu X, Bonadonna F, Cherel Y, Pistorius P, Le Maho Y et al. (2018) Climate-driven range shifts of the king penguin in a fragmented ecosystem. *Nat Clim Change* 8:245–251
- Descamps S, Aars J, Fuglei E, Kovacs KM, Lydersen C, Pavlova O et al. (2017) Climate change impacts on wildlife in a High Arctic archipelago-Svalbard, Norway. *Glob Change Biol* 23:490–502
- Diaz-Franulic I, Raddatz N, Castillo K, González-Nilo FD, Latorre R (2020) A folding reaction at the C-terminal domain drives temperature sensing in TRPM8 channels. *PNAS* 117:20298–20304
- Duchamp C, Marmonier F, Denjean F, Lachuer J, Eldershaw TPD, Rouanet JL et al. (1999) Regulatory, cellular and molecular aspects of avian muscle non-shivering thermogenesis. *Ornis Fennica* 76:151–165
- Elliott KH, Welcker J, Gaston AJ, Hatch SA, Palace V, Hare JF et al. (2013) Thyroid hormones correlate with resting metabolic rate, not daily energy expenditure, in two charadriiform seabirds. *Biol Open* 2:580–586
- Frost PGH, Siegfried WR, Greenwood PJ (1975) Arterio-venous heat exchange systems in the Jackass penguin *Spheniscus demersus*. *J Zool* 175:231–241
- Fu J, Gaetani S, Oveisi F, Lo Verme J, Serrano A, Rodríguez de Fonseca F et al. (2003) Oleyethanolamide regulates feeding and body weight through activation of the nuclear receptor PPAR-alpha. *Nature* 425:90–93
- Gavryushkina A, Heath TA, Ksepka DT, Stadler T, Welch D, Drummond AJ (2017) Bayesian total-evidence dating reveals the recent crown radiation of penguins. *Syst Biol* 66:57–73
- Geering K, Kraehenbuhl JP, Rossier BC (1987) Maturation of the catalytic alpha unit of Na, K-ATPase during intracellular transport. *J Cell Biol* 105:2613–2619
- Gilg O, Kovacs KM, Aars J, Fort J, Gauthier G, Grémillet D et al. (2012) Climate change and the ecology and evolution of Arctic vertebrates. *Ann N. Y Acad Sci* 1249:166–190
- Goldsmith R, Sladen WJ (1961) Temperature regulation of some Antarctic penguins. *J Physiol* 157:251–262
- Groscolas R (1990) Metabolic adaptations to fasting in emperor and king penguins. In *Penguin Biology* (LS Davis & JT Darby ed.), 269–296 San Diego: Academic Press
- Groscolas R, Robin JP (2001) Long-term fasting and re-feeding in penguins. *Comp Biochem Physiol Part A Mol Integr Physiol* 128:645–655
- Halaas JL, Gajiwala KS, Maffei M, Cohen SL, Chait BT, Rabinowitz D et al. (1995) Weight-reducing effects of the plasma protein encoded by the obese gene. *Science* 269:543–546
- Han MV, Demuth JP, McGrath CL, Casola C, Hahn MW (2009) Adaptive evolution of young gene duplicates in mammals. *Genome Res* 19:859–867
- Hao Y, Xiong Y, Cheng Y, Song G, Jia C, Qu Y et al. (2019) Comparative transcriptomics of 3 high-altitude passerine birds and their low-altitude relatives. *PNAS* 116:11851–11856
- Hunt BG, Ometto L, Wurm Y, Shoemaker D, Soojin VY, Keller L et al. (2011) Relaxed selection is a precursor to the evolution of phenotypic plasticity. *PNAS* 108:15936–15941
- Iannello S, Milazzo P, Belfiore F (2007) Animal and human tissue Na,K-ATPase in normal and insulin-resistant states: regulation, behaviour and interpretative hypothesis on NEFA effects. *Obes Rev* 8:231–251
- Ishii S, Amano I, Koibuchi N (2021) The role of thyroid hormone in the regulation of cerebellar development. *Endocrinol Metab* 36:703–716
- Jarvis ED, Mirarab S, Aberer AJ, Li B, Houde P, Li C et al. (2014) Whole genome analyses resolve early branches in the tree of life of modern birds. *Science* 346:1320–1331
- Kajimura S, Spiegelman BM, Seale P (2015) Brown and beige fat: Physiological roles beyond heat generation. *Cell Metab* 22:546–559
- Kooyman GL, Gentry RL, Bergman WP, Hammel HT (1976) Heat loss in penguins during immersion and compression. *Comp Biochem Physiol Part A* 54:75–80
- Kumar V, Kutschera VE, Nilsson MA (2015) Genetic signatures of adaptation revealed from transcriptome sequencing of Arctic and red foxes. *BMC Genom* 16:1–13
- Kumar S, Stecher G, Tamura K (2016) MEGA7: Molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Mol Biol Evol* 33:1870–1874
- Lahti DC, Johnson NA, Ajie BC, Otto SP, Hendry AP, Blumstein DT et al. (2009) Relaxed selection in the wild. *Trends Ecol Evol* 24:487–496
- Lemberger T, Saladin R, Vazquez M, Assimacopoulos F, Staels B, Desvergne B et al. (1996) Expression of the peroxisome proliferator-activated receptor alpha gene is stimulated by stress and follows a diurnal rhythm. *J Biol Chem* 271:1764–1769
- Li FG, Li H (2019) A time-dependent genome-wide SNP-SNP interaction analysis of chicken body weight. *BMC Genom* 20:1–9
- Li C, Zhang Y, Li J, Kong L, Hu H, Pan H et al. (2014) Two Antarctic penguin genomes reveal insights into their evolutionary history and molecular changes related to the Antarctic environment. *Gigascience* 3:2047–2217X
- Lin Z, Chen L, Chen X, Zhong Y, Yang Y, Xia W et al. (2019) Biological adaptations in the Arctic cervid, the reindeer (*Rangifer tarandus*). *Science* 364:eaav6312
- Liu S, Lorenzen ED, Fumagalli M, Li B, Harris K, Xiong Z et al. (2014) Population genomics reveal recent speciation and rapid evolutionary adaptation in polar bears. *Cell* 157:785–794
- Liu S, Westbury MV, Dussex N, Mitchell KJ, Sinding MH, Heintzman PD et al. (2021) Ancient and modern genomes unravel the evolutionary history of the rhinoceros family. *Cell* 184:4874–4885
- Lowell BB, Spiegelman BM (2000) Towards a molecular understanding of adaptive thermogenesis. *Nature* 404:652–660
- Löytynoja A (2013) Phylogeny-aware alignment with PRANK. *Methods Mol Biol* 1079:155–170
- Lynch VJ, Bedoya-Reina OC, Ratan A, Sulak M, Drautz-Moses DI, Perry GH et al. (2015) Elephantid genomes reveal the molecular bases of woolly mammoth adaptations to the Arctic. *Cell Rep* 12:217–228
- Madeira F, Park YM, Lee J, Buso N, Gur T, Madhusoodanan N et al. (2019) The EMBL-EBI search and sequence analysis tools APIs in 2019. *Nucleic Acids Res* 47:636–641
- Matos-Cruz V, Schneider ER, Mastrotto M, Merriman DK, Bagriantsev SN, Gracheva EO (2017) Molecular prerequisites for diminished cold sensitivity in ground squirrels and hamsters. *Cell Rep* 21:3329–3337
- Monera OD, Sereda TJ, Zhou NE, Kay CM, Hodges RS (1995) Relationship of sidechain hydrophobicity and alpha-helical propensity on the stability of the single-stranded amphipathic alpha-helix. *J Pept Sci* 1:319–329
- Myers BR, Sigal YM, Julius D (2009) Evolution of Thermal Response Properties in a Cold-Activated TRP Channel. *PLoS one* 4:e5741
- Ng CS, Chen CK, Fan WL, Wu P, Wu SM, Chen JJ et al. (2015) Transcriptomic analyses of regenerating adult feathers in chicken. *BMC Genom* 16:1–16
- Ohno H, Shinoda K, Spiegelman BM, Kajimura S (2012) PPARγ agonists induce a white-to-brown fat conversion through stabilisation of PRDM16 protein. *Cell Metab* 15:395–404
- Oldham MC, Horvath S, Geschwind DH (2006) Conservation and evolution of gene coexpression networks in human and chimpanzee brains. *PNAS* 103:17973–17978
- Pan H, Cole TL, Bi X, Fang M, Zhou C, Yang Z et al. (2019) High-coverage genomes to elucidate the evolution of penguins. *GigaScience* 8:1–17
- Pelleymounter MA, Cullen MJ, Baker MB, Hecht R, Winters D, Boone T et al. (1995) Effects of the obese gene product on body weight regulation in ob/ob mice. *Science* 269:540–543
- Poirier H, Niot I, Monnot MC, Braissant O, Meunier-Durmort C, Costet P et al. (2001) Differential involvement of peroxisome-proliferator-activated receptors alpha and delta in fibrate and fatty-acid-mediated inductions of the gene encoding liver fatty-acid-binding protein in the liver and the small intestine. *Biochem J* 355:481–488
- Pond SK, Frost S, Muse SV (2005) HyPhy: hypothesis testing using phylogenies. *Bioinform* 21:676–679
- R Core Team (2013) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria
- Ramasamy S, Ometto L, Crava CM, Revadi S, Kaur R, Horner DS et al. (2016) The evolution of olfactory gene families in *Drosophila* and the genomic basis of chemical-ecological adaptation in *Drosophila suzukii*. *Genome Biol Evol* 8:2297–2311
- Ramos B, González-Acuña D, Loyola DE, Johnson WE, Parker PG, Massaro M et al. (2018) Landscape genomics: natural selection drives the evolution of mitochondrial genomes in penguins. *BMC Genom* 19:1–17

- Raudvere U, Kolberg L, Kuzmin I, Arak T, Adler P, Peterson H et al. (2019) g:Profiler: a web server for functional enrichment analysis and conversions of gene lists. *Nucleic Acids Res* 47:191–198
- Rey B, Roussel D, Romestaing C, Belouze M, Rouanet JL, Desplanches D et al. (2010) Up-regulation of avian uncoupling protein in cold-acclimated and hyperthyroid ducklings prevents reactive oxygen species production by skeletal muscle mitochondria. *BMC Physiol* 10:1–12
- Roussel D, Le Coadic M, Rouanet JL, Duchamp C (2020) Skeletal muscle metabolism in sea-acclimated king penguins I Thermogenic mechanisms. *J Exp Biol* 223:pjeb233668
- Rowland LA, Bal NC, Periasamy M (2015) The role of skeletal-muscle-based thermogenic mechanisms in vertebrate endothermy. *Biol Rev* 90:1279–1297
- Savini G, Scolari F, Ometto L, Rota-Stabelli O, Carraretto D, Gomulski LM et al. (2021) Viviparity and habitat restrictions may influence the evolution of male reproductive genes in tsetse fly (*Glossina*) species. *BMC Biol* 19:1–13
- Scholander PF (1955) Evolution of climatic adaptation in homeotherms. *Evolution* 9:15–26
- Schymkowitz J, Borg J, Stricher F, Nys R, Rousseau F, Serrano L (2005) The FoldX web server: an online force field. *Nuc Acids Res* 33:382–388
- Seale P, Kajimura S, Yang W, Chin S, Rohas LM, Uldry M et al. (2007) Transcriptional control of brown fat determination by PRDM16. *Cell Metab* 6:38–54
- Seale P, Conroe HM, Estall J, Kajimura S, Frontini A, Ishibashi J et al. (2011) Prdm16 determines the thermogenic program of subcutaneous white adipose tissue in mice. *J Clin Invest* 121:96–105
- Smith MD, Wertheim JO, Weaver S (2015) Less is more: An adaptive branch-site random effects model for efficient detection of episodic diversifying selection. *Mol Biol Evol* 32:1342–1353
- Sonna LA, Fujita J, Gaffin SL, Lilly CM (2002) Effects of heat and cold stress on mammalian gene expression. *J Appl Physiol* 92:1725–1742
- Stelzer G, Rosen N, Plaschkes I, Zimmerman S, Twik M, Fishilevich S et al. (2016) The geneCards suite: From gene data mining to disease genome sequence analysis. *Curr Protoc Bioinform* 54:1–30
- Storey KB, Storey JM (1992) Natural freeze tolerance in ectothermic vertebrates. *Annu Rev Physiol* 54:619–637
- Storey JD, Bass AJ, Dabney A, Robinson D (2017) qvalue: Q-value estimation for false discovery rate control R package version 2.15.0
- Supek F, Bošnjak M, Škunca N, Šmuc T (2011) REVIGO summarises and visualises long lists of gene ontology terms. *PLoS one* 6:e21800
- Talbot DA, Duchamp C, Rey B, Hanuise N, Rouanet JL, Sibille B et al. (2004) Uncoupling protein and ATP/ADP carrier increase mitochondrial proton conductance after cold adaptation of king penguins. *J Physiol* 558:123–135
- Tattersall GJ, Sinclair BJ, Withers PC, Fields PA, Seebacher F, Cooper CE et al. (2012) Coping with thermal challenges: Physiological adaptations to environmental temperatures. *Compr Physiol* 2:2151–2202
- Teulier L, Rouanet JL, Letexier D, Romestaing C (2010) Cold-acclimation-induced non-shivering thermogenesis in birds is associated with upregulation of avian UCP but not with innate uncoupling or altered ATP efficiency. *J Exp Biol* 213:2476–2482
- Thomas DB, Fordyce RE (2008) The heterothermic loophole exploited by penguins. *Aust J Zool* 55:317–321
- Tigano A, Reiertsen TK, Walters JR, Friesen VL (2018) A complex copy number variant underlies differences in both colour plumage and cold adaptation in a dimorphic seabird. *BioRxiv* 507384. <https://doi.org/10.1101/507384>
- Toyomizu M, Ueda M, Sato S, Seki Y, Sato K, Akiba Y (2002) Cold-induced mitochondrial uncoupling and expression of chicken UCP and ANT mRNA in chicken skeletal muscle. *FEBS Lett* 529:313–318
- Trucchi E, Gratton P, Whittington JD, Cristofari R, Le Maho Y, Stenseth NC et al. (2014) King penguin demography since the last glaciation inferred from genome-wide data. *Proc R Soc B* 281:20140528
- Trucchi E, Cristofari R, Le Bohec C (2019) Reply to: The role of ocean dynamics in king penguin range estimation. *Nat Clim Change* 9:122–122
- Vermillion KL, Anderson KJ, Hampton M, Andrews MT (2015) Gene expression changes controlling distinct adaptations in the heart and skeletal muscle of a hibernating mammal. *Physiol Genom* 47:58–74
- Vézina F, Gustowska A, Jalvingh KM, Chastel O, Piersma T (2015) Hormonal correlates and thermoregulatory consequences of moulting on metabolic rate in a northerly wintering shorebird. *Physiol Biochem Zool* 82:129–142
- Vianna JA, Fernandes FA, Frugone MJ, Figueiró HV, Pertierra LR, Noll D et al. (2020) Genome-wide analyses reveal drivers of penguin diversification. *PNAS* 117:22303–22310
- Wang Q, Tan X, Jiao S (2014) Analyzing cold tolerance mechanism in transgenic Zebrafish (*Danio rerio*). *PLoS one* 9:e102492
- Wertheim JO, Murrell B, Smith MD, Kosakovsky Pond SL, Scheffler K (2015) RELAX: Detecting relaxed selection in a phylogenetic framework. *Mol Biol Evol* 32:820–832
- Wollenberg Valero KC, Pathak R, Prajapati I, Bankston S, Thompson A, Usher J et al. (2014) A candidate multimodal functional genetic network for thermal adaptation. *PeerJ* 2:e578
- Yang Z (2007) PAML 4: A program package for phylogenetic analysis by maximum likelihood. *Mol Biol Evol* 24:1586–1591
- Yang J, Bromage TG, Zhao Q, Xu BH, Gao WL, Tian HF et al. (2011) Functional evolution of leptin of *Ochotona curzoniae* in adaptive thermogenesis driven by cold environmental stress. *PLoS one* 6:e19833
- Yang S, Lu X, Wang Y, Xu L, Chen X, Yang F et al. (2020) A paradigm of thermal adaptation in penguins and elephants by tuning cold activation in TRPM8. *PNAS* 117:8633–8638
- Yin Y, Wu M, Zubcevic L, Borschel WF, Lander GC, Lee SY (2018) Structure of the cold- and menthol-sensing ion channel TRPM8. *Science* 359:237–241
- Yudin NS, Larkin DM, Ignatieva EV (2017) A compendium and functional characterization of mammalian genes involved in adaptation to Arctic or Antarctic environments. *BMC Genet* 18:33–43
- Zelcer N, Sharpe LJ, Loregger A, Kristiana I, Cook EC, Phan L et al. (2014) The E3 ubiquitin ligase MARCH6 degrades squalene monooxygenase and affects 3-hydroxy-3-methyl-glutaryl coenzyme A reductase and the cholesterol synthesis pathway. *Mol Cell Biol* 34:1262–1270
- Zhang Y, Proenca R, Maffei M, Barone M, Leopold L, Friedman JM (1994) Positional cloning of the mouse obese gene and its human homologue. *Nature* 372:425–432
- Zhang G, Li C, Li Q, Li B, Larkin DM, Lee C et al. (2014) Comparative genomics reveals insights into avian genome evolution and adaptation. *Science* 346:1311–1320

ACKNOWLEDGEMENTS

ET was supported by PNRA16_00164 (“Programma Nazionale di Ricerca in Antartide. Bando PNRA 5 aprile 2016, n. 651. – Linea B “Genomica degli adattamenti estremi alla vita in Antartide”).

AUTHOR CONTRIBUTIONS

FP performed all of the genomic analyses from orthologous gene search to phylogenetic-based selection scans, and drafted the manuscript; LO designed a custom pipeline for the orthologous gene search, supervised the genomic analyses, discussed the results and commented on the manuscript; SF supervised the genomic analyses, discussed the results and commented on the manuscript; FANF and LA contributed to the candidate genes functional annotation and commented on the manuscript; NP and DDM contributed the protein structural description and functional annotation of newly discovered genetic variants; CLB provided insights on the ecological determinants of *Aptenodytes* species divergence, discussed the results and commented on the manuscript; LZ discussed the study design and the results, and commented on the manuscript; ET conceived the idea, designed the study, supervised the analyses, discussed the results, and wrote the final version of the manuscript.

COMPETING INTERESTS

The authors declare no competing interests.

ADDITIONAL INFORMATION

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41437-022-00564-8>.

Correspondence and requests for materials should be addressed to Emiliano Trucchi.

Reprints and permission information is available at <http://www.nature.com/reprints>

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

Springer Nature or its licensor holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.