April 22, 2013

Second generation DNA sequencing of the mitogenome of the Chinstrap penguin and comparative genomics of Antarctic penguins

Sankar Subramanian^{*}, Syamala Gowri Lingala, Siva Swaminathan, Leon Huynen and David Lambert^{*}

Environmental Futures Centre, Griffith University, Nathan Qld 4111, Australia

Keywords: Chinstrap penguin, mitochondrial genome, Adélie penguin, phylogenetics, natural selection and neutral evolution

Running head: Mitogenome of the Chinstrap penguin

^{*} Corresponding authors

Correspondence

Sankar Subramanian / David Lambert

170 Kessels Road

Nathan

QLD 4105, Australia

Tel: +61 737357495

Fax: +61 737357459

Email: s.subramanian@griffith.edu.au or d.lambert@griffith.edu.au

Abstract

The complete mitochondrial genome of the Chinstrap penguin (*Pygoscelis antarcticus*) was sequenced and compared with other penguin mitogenomes. The genome is 15,972 bp in length with the number and order of protein coding genes and RNAs being very similar to that of other known penguin mitogenomes. Comparative nucleotide analysis showed the Chinstrap mitogenome shares 94% homology with the mitogenome of its sister species, *Pygoscelis adeliae* (Adélie penguin). Divergence at nonsynonymous nucleotide positions was found to be up to 23 times less than that observed in synonymous positions of protein coding genes, suggesting high selection constraints. The complete mitogenome data will be useful for genetic and evolutionary studies of penguins.

1 The Chinstrap penguin is one of the three species in the genus *Pygoscelis*, which predominantly
2 inhabits in the Antarctic and sub-Antarctic islands (Merchant and Higgins 1990). Here we report
3 the complete mitochondrial genome sequence of this penguin. The Chinstrap mitogenome was
4 amplified from total genomic DNA as three overlapping fragments of 9650bp, 4133bp and
5 approximately 6000bp using primers designed using the complete mitochondrial genome of
6 Adélie penguin (Subramanian et al., 2009). The PCR products were nebulized and sequenced
7 using a GS Junior FLX Sequencer. After filtering for contaminants and short reads,
8 approximately 9 million base pairs were used for mitogenome reconstruction, to give an average
9 coverage of 37.5X.

11 The complete genome of the Chinstrap penguin is 15,972 bp in length. The mitogenome consists
12 of 13 protein-coding genes, 22 tRNAs, two rRNAs (12S and 16S) and a D-loop region. The
13 length and order of these genes are similar to those reported for other penguin species. We
14 recovered the full length of all genes except for the D-loop region, which is missing ~382 bp.
15 However, D-loop regions have been sequenced previously and are available in GenBank. The
16 genomic base composition of the heavy strand is 30.2%, 22.3%, 14.2% and 33.3% for A, T, G
17 and C respectively. The high As and Cs compared to their complementary Ts and Gs suggests
18 strand asymmetry, which is a common phenomenon in vertebrate mitogenomes (Reyes et al.
19 1998). While ATG was the start codon for 11 genes, COXI and ND3 genes start with GTG and
20 ATC codons respectively.

22 Comparative nucleotide analysis of the Chinstrap mitogenome with that of its closely related
23 sister species *Pygoscelis adeliae* showed that these two genomes share 93.7% homology. Analysis
24 of protein coding genes revealed that divergence at evolutionarily constrained sites such as

nonsynonymous sites (dN) and RNAs (dR) was 0.013 and 0.02 substitutions per site (s/s) respectively. In contrast, the divergence at synonymous positions (dS) was 0.30 s/s, which is 23 and 15 times higher than that at nonsynonymous and RNA sites. Hence the dN/dS and dR/dS ratios are 0.04 and 0.07 respectively. Assuming neutral evolution at synonymous sites the above ratios suggest that approximately 96% of the mutations in nonsynonymous positions and 93% of the mutations in RNA sites have been removed by purifying selection (Li 1997). Because of hyper-mutation and saturation effects we excluded the D-loop regions in all analyses.

To infer the phylogenetic relationship of Chinstrap to other penguins we compared the mitogenomes of Adélie, rock-hopper and little blue penguins. We also obtained the mitogenomes of loon, albatross, zebra finch, chicken, ostrich and tinamou. We used the algorithm *Modeltest* (Posada and Crandall 1998) implemented in the software *MEGA* (Tamura et al. 2011) to determine the best model of evolution and constructed a maximum likelihood tree using a GTR+Gamma+Invariant model (Figure 1). Tree branch lengths suggest that the rates of molecular evolution in Chinstrap and Adélie penguins are very similar. The complete mitochondrial genome data reported here could be useful to further examine the evolutionary patterns of penguin mitogenomes.

Acknowledgement

We thank Oliver Haddrath and Allan Baker (Royal Ontario Museum, Canada) for kindly providing the Chinstrap DNA.

Declaration of interests

This work was supported by Australia India Strategic Research Fund, Griffith University and the Australian Research Council. The authors report no conflict of interest. The authors alone are responsible for the content and writing of the paper.

References

- Li, W-H. 1997. *Molecular Evolution*. Sinauer Associates, Sunderland, Massachusetts.
- Merchant S and Higgins PJ. 1990. *The Handbook of Australian, New Zealand and Antarctic Birds*. Oxford University Press, Melbourne.
- Posada, D and Crandall KA. 1998. MODELTEST: testing the model of DNA substitution. *Bioinformatics* 14:817-818.
- Reyes A, Gissi C, Pesole G and Saccone C. 1998. Asymmetrical directional mutation pressure in the mitochondrial genome of mammals. *Mol Biol Evol* 15:957-966.
- Subramanian S, Denver DR, Millar CD, Heupink T, Aschrafi A, Emslie SD, Baroni C, and Lambert DM. 2009. High mitogenomic evolutionary rates and time dependency. *Trends in Genetics* 25:482-486.
- Tamura K, Peterson D, Peterson N, Stecher G, Nei M, and Kumar S. 2011. MEGA5: Molecular Evolutionary Genetics Analysis Using Maximum Likelihood, Evolutionary Distance, and Maximum Parsimony Methods. *Mol Biol Evol* 28:2731-2739.

Figure Legend

Figure 1. Maximum likelihood tree based on complete mitochondrial genomes of four penguins and other bird species. The numbers shown above the nodes are the bootstrap supports based on 100 replications.