

RESEARCH ARTICLE

Quantitative Phylogenomics of Within-Species Mitogenome Variation: Monte Carlo and Non-Parametric Analysis of Phylogeographic Structure among Discrete Transatlantic Breeding Areas of Harp Seals (*Pagophilus groenlandicus*)

Steven M. Carr^{1*}, Ana T. Duggan^{1‡}, Garry B. Stenson², H. Dawn Marshall³

1 Genetics, Evolution, and Molecular Systematics Laboratory, Department of Biology, Memorial University of Newfoundland, St John's, NL A1B 3X9, Canada, **2** Wildlife Genetics and Genomics Laboratory, Department of Biology, Memorial University of Newfoundland, St John's, NL A1B 3X9, Canada, **3** Marine Mammals Section, Science Branch, Dept. of Fisheries and Oceans, PO Box 5667, St. John's, Nfld., A1C 5X1, Canada

‡ Current address: McMaster Ancient DNA Centre, McMaster University, Hamilton ON, L8S 4L8, Canada

* scarr@mun.ca



CrossMark
click for updates

OPEN ACCESS

Citation: Carr SM, Duggan AT, Stenson GB, Marshall HD (2015) Quantitative Phylogenomics of Within-Species Mitogenome Variation: Monte Carlo and Non-Parametric Analysis of Phylogeographic Structure among Discrete Transatlantic Breeding Areas of Harp Seals (*Pagophilus groenlandicus*). PLoS ONE 10(8): e0134207. doi:10.1371/journal.pone.0134207

Editor: Axel Janke, BiK-F Biodiversity and Climate Research Center, GERMANY

Received: June 26, 2014

Accepted: July 7, 2015

Published: August 24, 2015

Copyright: © 2015 Carr et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: Lab work was supported by an NSERC Discovery Grant [RGPIN-2010-05265] (<http://www.nserc-crsng.gc.ca>) and DFO University Partnership grants (including funding from the Northern Cod Research Program) to SMC and GBS. Funding for field collection was provided by the Department of Fisheries and Oceans (Canada), the Institute of Marine Resources (Norway), and the Polar Research of Marine Fisheries and Oceanography (Russia).

Abstract

Phylogenomic analysis of highly-resolved intraspecific phylogenies obtained from complete mitochondrial DNA genomes has had great success in clarifying relationships within and among human populations, but has found limited application in other wild species. Analytical challenges include assessment of random versus non-random phylogeographic distributions, and quantification of differences in tree topologies among populations. Harp Seals (*Pagophilus groenlandicus* Erleben, 1777) have a biogeographic distribution based on four discrete trans-Atlantic breeding and whelping populations located on “fast ice” attached to land in the White Sea, Greenland Sea, the Labrador ice Front, and Southern Gulf of St Lawrence. This East to West distribution provides a set of *a priori* phylogeographic hypotheses. Outstanding biogeographic questions include the degree of genetic distinctiveness among these populations, in particular between the Greenland Sea and White Sea grounds. We obtained complete coding-region DNA sequences (15,825 bp) for 53 seals. Each seal has a unique mtDNA genome sequence, which differ by 6 ~ 107 substitutions. Six major clades / groups are detectable by parsimony, neighbor-joining, and Bayesian methods, all of which are found in breeding populations on either side of the Atlantic. The species coalescent is at 180 KYA; the most recent clade, which accounts for 66% of the diversity, reflects an expansion during the mid-Wisconsinan glaciation 40 ~ 60 KYA. F_{ST} is significant only between the White Sea and Greenland Sea or Ice Front populations. Hierarchical AMOVA of 2-, 3-, or 4-island models identifies small but significant Φ_{SC} among populations within groups, but not among groups. A novel Monte-Carlo simulation indicates that the observed distribution of individuals within breeding populations over the phylogenetic tree requires

Competing Interests: The authors have declared that no competing interests exist.

significantly fewer dispersal events than random expectation, consistent with island or *a priori* East to West 2- or 3-stepping-stone biogeographic models, but not a simple 1-step trans-Atlantic model. Plots of the cumulative pairwise sequence difference curves among seals in each of the four populations provide continuous proxies for phylogenetic diversification within each. Non-parametric Kolmogorov-Smirnov (K-S) tests of maximum pairwise differences between these curves indicates that the Greenland Sea population has a markedly younger phylogenetic structure than either the White Sea population or the two Northwest Atlantic populations, which are of intermediate age and homogeneous structure. The Monte Carlo and K-S assessments provide sensitive quantitative tests of within-species mitogenomic phylogeography. This is the first study to indicate that the White Sea and Greenland Sea populations have different population genetic histories. The analysis supports the hypothesis that Harp Seals comprises three genetically distinguishable breeding populations, in the White Sea, Greenland Sea, and Northwest Atlantic. Implications for an ice-dependent species during ongoing climate change are discussed.

Introduction

Analyses of multiple complete intraspecific mtDNA genomes were first applied to humans to clarify the historical emergence of modern humans and their subsequent migrations Out of Africa [1]. The accumulation of many thousands of such genomes has clarified more recent population movements in great detail, including those influenced by successive Holocene glaciations [2]. For example, high-resolution mitogenomic sampling of Iberian refugial lineages effectively discriminates postglacial dispersal and ecological events [3]. Only a few other wild species have been the subject of extensive mitogenomic phylogeographic analysis. Atlantic Cod (*Gadus morhua*) have a much more ancient species structure than previously suspected, with coalescence during the Wisconsinan glacial, and major clade differentiation at the peak of the Sangamon / Würm interglacial rather than subsequent to the last Recent glacial [4]. The major clades are phylogeographically mixed, and well-separated geographic samples coalesce only towards the base of the gene tree. The sister species of Atlantic Cod, the Walleye Pollock (*Gadus chalcogrammus*), includes geographically isolated populations originally supposed to be of recent origin that are shown by mitogenomic analysis to be ancient [5]. In contrast, separate geographic populations of freshwater whitefish (*Coregonus*) are discrete genomic clades [6]. Major phylogenetic lineages in Killer Whales (*Orcinus*) correspond to discrete geographic populations including some assignable to subspecies [7], and in Fin Whales (*Balaenoptera*) to what appear to be recognizable species [8].

Standard methods of phylogenetic analysis and (or) quantitative assessment of within-species mtDNA population structure founder when every individual is a distinct branch. Hierarchical analysis of molecular variance (AMOVA) among haplotypes *per se* is uninformative, as the entire variance occurs among individuals with respect to the total ($F_{ST} = 1.0$). AMOVA based on nucleotide divergences between haplotypes ($\phi_{ST} < 1.0$) may partition variance among populations ($\phi_{SC} > 0.0$), or groups ($\phi_{CT} > 0.0$), but does not capture phylogenetic structure within and among populations. Similarly, row by column tests of the relative abundances of major clades across populations cannot take into account phylogenetic relationships either within or among those clades [9]. Evaluation of hypotheses that phylogeographic sub-components within species have different phylogenetic tree topologies is elusive, not least