

The Nature of Neanderthal Introgression Revealed by 27,566 Icelandic Genomes

Laurits Skov et al

22 April 2020

Overview

Used a 2-state hidden Markov model (not conditioned on archaic genomes) to search for archaic haplotypes in 55,132 Icelandic haploid genomes sequenced to 30X.

Found 112,709 archaic haplotypes covering 1.179 Gb, or 48.2% of the callable genome.

These introgressed fragments cover 1.34% of the average Icelandic genome—an underestimate, because the method can detect introgressed fragments only if they are sufficiently long and sufficiently divergent. Estimated false-negative rate suggests a true fraction closer to 1.9%.

Multiple archaic sources

Neanderthal	84.5%
Denisovan	3.3%
Unknown	12.2%

The Denisovan fraction is larger than expected through incomplete lineage sorting. Suggests Denisovan gene flow either via Neanderthals or directly into moderns.

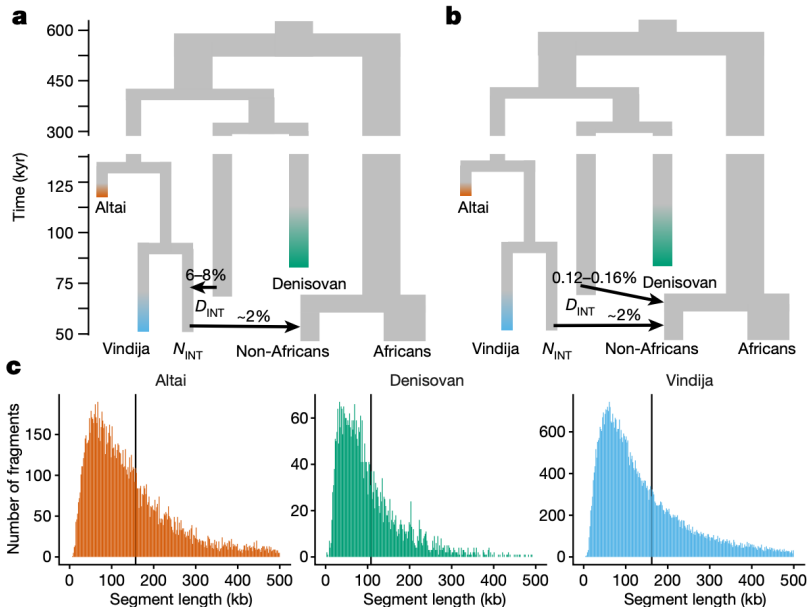
Mosaic haplotypes

18.9% of archaic haplotypes were “mosaic,” including Denisovan as well as Neanderthal variants.

This is remarkable. The authors discuss two hypotheses: (1) Denisovans interbred with Neanderthals, who later interbred with moderns, or (2) Denisovans interbred directly with moderns.

The first hypothesis seems more plausible, because we know there was interbreeding between N and D. Furthermore, the fraction of archaic DNA in moderns is low, and this would have reduced the opportunity for recombinant N-D haplotypes to form.

Models of history



Population size estimates

Archaic population sizes were estimated by comparing introgressed haplotypes with the sequences of archaic genomes.

Estimate small N_e for both archaics: 2000–3000 individuals.

I don't fully understand what they did and have been emailing with the authors.

Distribution of archaic DNA w/i the modern genome

I don't entirely understand the prose, but I think it says that archaic DNA is more common where recombination rate is high.

This is interesting, because archaic haplotypes should be shorter in these regions and thus more difficult to detect.

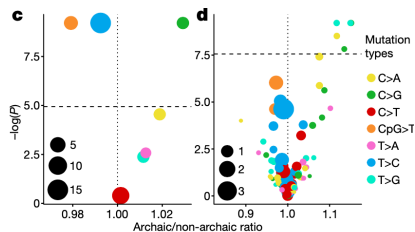
Implies that archaic variants are more likely to survive if they can be uncoupled from deleterious archaic variants.

Mutational bias

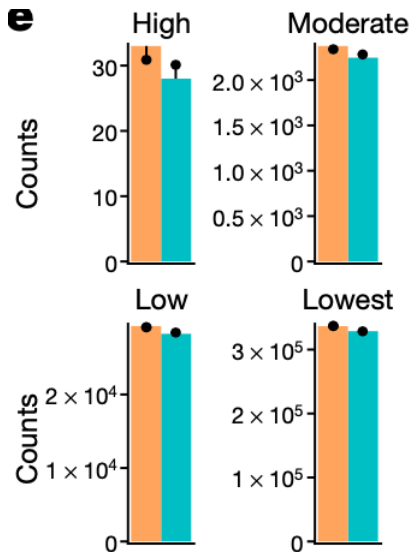
Archaiics and moderns have similar mutation rates.

But they differ in the frequencies of different categories of mutation.

The largest differences may reflect differences between the two populations in the average ages of mothers and fathers.



No evidence of selection against archaic DNA



No excess of deleterious variants in archaic haplotypes.

Yet abundant evidence shows that archaic genomes had lots of deleterious variants.

Presumably, these were removed in moderns by purifying selection.

Effects on function

Looked for associations between archaic haplotypes and 271 phenotypes.

Removed associations that were better explained by a nearby non-archaic site.

Pooled linked archaic haplotypes

After filtering, only 5 archaic associations remained.

Verify only 3 of 26 previously-reported archaic associations.

Neither was there a significant association between any of the phenotypes and the *number* of archaic haplotypes.

Discussion questions

1. How might we distinguish between the two hypotheses to explain the presence of Denisovan DNA and Neanderthal-Denisovan mosaics?
2. Are we convinced by the argument that archaic DNA doesn't affect function?
3. What might explain the apparent difference in average ages of mothers and fathers?