



HAL
open science

Archaic hominin traits through the splicing lens

Maxime Rotival

► **To cite this version:**

Maxime Rotival. Archaic hominin traits through the splicing lens. 2023, pp.800-801. 10.1038/s41559-023-02045-5 . pasteur-04605514

HAL Id: pasteur-04605514

<https://pasteur.hal.science/pasteur-04605514>

Submitted on 10 Jun 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

Human evolution: Archaic hominin traits through the splicing lens.

Machine-learning-based prediction of splicing in extinct hominin species highlights the effect of natural selection on splice altering variants and reveals phenotypic differences with modern humans.

Maxime Rotival¹

¹Institut Pasteur, Université Paris Cité, CNRS UMR2000, Human Evolutionary Genetics Unit, F-75015 Paris, France. Email: maxime.rotival@pasteur.fr

The sequencing of archaic hominin genomes has fostered a renewed interest in the identity of our extinct relatives and their legacy in the genome of modern humans. However, characterizing archaic hominins' phenotypes is limited by the scarcity of remains and the rapid degradation of soft tissues after death. Various attempts have been made to infer archaic phenotypes based on their DNA methylation patterns¹, or through the study of the regulatory alleles that they share with modern humans². Yet, the extensive purge of archaic DNA from the genome of modern humans³, and the possibility of regulatory divergence between Archaic and modern humans⁴ limit the effectiveness of these approaches. Writing in *Nature Ecology and Evolution*, Brand et al⁵ implemented an innovative solution to this issue by focusing on alternative splicing, informing us on archaic hominins' phenotypes and revealing how splice altering variants that modern humans inherited from their archaic relatives helped them adapt to their environment but also contributed to disease.

Alternative splicing, the process by which exons of a gene are joined in different combinations to form alternative mRNA molecules, is a major contributor to the heritability of complex traits, on a par with variants that affect gene expression levels⁶. Because molecular mechanisms underlying alternative splicing are deeply conserved across eukaryotes⁷, Brand et al reasoned that the same rules should govern alternative splicing in humans and archaic hominins. They thus used a deep learning algorithm called SpliceAI⁸ to predict allele-specific effects on splicing at >1.5 million loci that differ between modern humans and four archaic hominins (three Neanderthals and one Denisovan). The authors identified thousands of splice altering variants (SAVs) present in archaic hominins, 37% of which are archaic-specific. The remainder included both ancient alleles that predate the split between modern humans and their archaic relatives and have survived in modern humans to this day (59%, referred to as ancient SAVs), or alleles of archaic origin that were introgressed into modern humans through admixture (4%, referred to as introgressed SAVs).

The authors showed that archaic alleles that are specific to a single lineage (e.g., specific to the Vindija-cave Neanderthals) are enriched for SAVs compared to archaic variants that are shared across multiple archaic lineages. This result, which they attribute to the purging of SAVs that appeared early in archaic hominin evolution, is consistent with SAVs being more frequently found within Neanderthal lineages. Indeed, Neanderthals had a smaller effective population size relative to Denisovans and natural selection was thus less efficient at purging deleterious alleles from their genomes. Brand et al. further showed that archaic-specific SAVs are more likely than ancient SAVs to alter conserved sites and disrupt protein function. Finally, the authors analysed the phenotypic impact of the genetic burden carried by our late archaic relatives, by intersecting archaic-specific SAVs with known disease genes, either identified by genome-wide association studies or associated with rare mendelian disorders. In doing so, they

pinpointed the specific phenotypes that are associated with SAV-containing genes in each archaic lineage such as fragile skin in Neanderthals, or muscular abnormalities in Denisovans.

Next, Brand et al, zoomed in on SAVs that were introgressed in modern humans. The authors found that introgressed SAVs are enriched near genes that are expressed in a tissue-specific manner and show that the vast majority of SAVs observed in humans today are older than the split between Denisovans and Neanderthals. These results suggest that natural selection has purged more recent SAVs from human genomes shortly after their introgression, leaving only SAVs with more localized effects on gene expression. Yet, the authors further showed that introgressed SAVs are enriched near genes associated with multiple phenotypes, including hay fever and allergies, *Helicobacter pylori* serologic status or systemic sclerosis, suggesting that SAVs inherited from archaic hominins still contribute to disease today.

Finally, the authors explored how splice altering variants that modern humans inherited from their archaic relatives helped them adapt to their environment. Aside from previously described SAVs associated with COVID-19 susceptibility at the *OAS1* locus, and rhinitis at the *TLRI* locus, the authors report novel splice altering variants at the *EPAS1* locus, where an Denisovan-introgressed haplotype has contributed to the adaptation of Tibetans peoples to high altitudes⁹. The reported SAV leads to nonsense-mediated decay of *EPAS1* and correlates with lower haemoglobin levels in Tibetans¹⁰, consistent with the positive effect of *EPAS1* on altitude-induced erythropoiesis¹¹. At the *ERAP2* locus, the authors identify three introgressed SAVs of potential evolutionary importance. First, they identify an ancient SAV evolving under balancing selection associated with stimulation-induced splicing of *ERAP2*¹² and increased survival during the back death¹³. In addition, they report two novel, Neanderthal-specific SAVs of unknown functional relevance, one of which is predicted to induce nonsense-mediated decay of *ERAP2*. Together, these results highlight how SAVs inherited from archaic hominins contributed to human adaptation.

The work presented by Brand et al. shows us how prediction of molecular phenotypes from DNA sequence⁸, can teach us about extinct species and their impact on the evolutionary history of their living relatives. Yet, it also has some limitations. First, it relies heavily on the accuracy of the underlying predictions. Its conclusions could thus suffer from the poor performance of current prediction methods when applied to deep intronic variants and distal enhancers. Second, Brand et al only consider variants with a widespread effect on splicing. Yet, context-specific splicing has been shown to play an essential role in the development of the brain and testis¹⁴, and may have played a role in the purging of archaic variants observed in these tissues⁴. While the general splicing machinery is extremely conserved, such dynamic and tissue-specific splicing may be much more variable across species, which may complicate the portability of predictions across hominins. More accurate models powered by massively parallel splicing assays¹⁵, and the mapping of splice quantitative trait loci⁶ are needed to embrace the complexity of such context-specific splicing. Finally, structural variants (eg. indels) and variants located on sexual chromosomes have so far been left aside, despite their relevance for human phenotypes and hybrid infertility. Future efforts must assess the role of these variants in the purge of archaic alleles from modern human genomes.

The work of Brand et al is a powerful proof of concept that lays the foundations for future studies aiming to resurrect the alternative splicing landscape of extinct species.

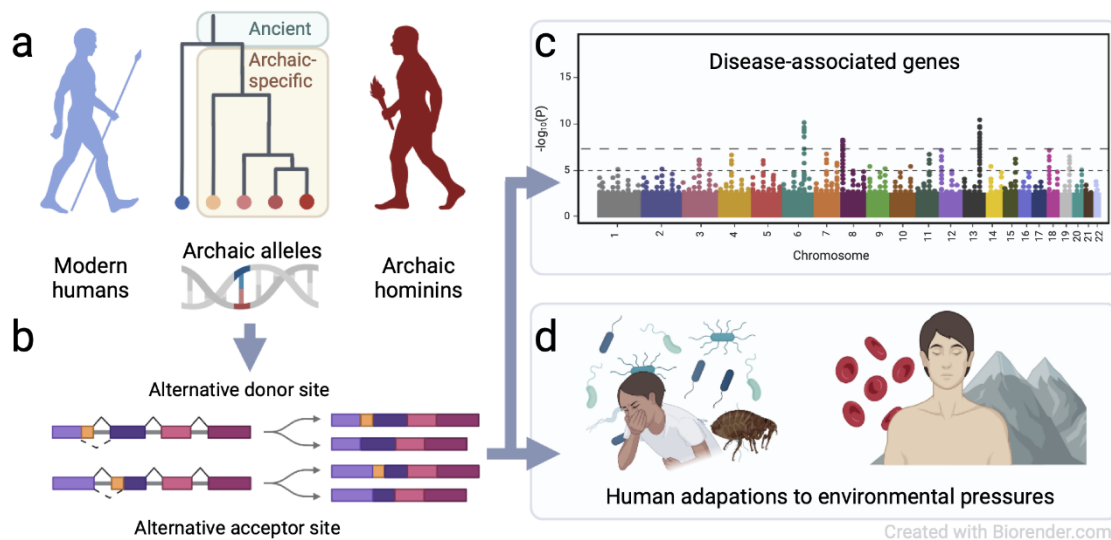


Figure 1: Defining the effect of Archaic alleles on splicing and human evolution.

a) Brand et al infer genetic differences between modern humans and four archaic hominins lineages. b) They evaluate the impact on splicing of thousands of archaic alleles and show an increased deleteriousness of splice altering variants (SAVs) that are specific to one or more archaic lineages relative to those that are shared with modern humans. c) They further overlap these SAVs with disease genes to reveal specific traits of archaic humans, and d) explore the impact of introgressed SAVs on human phenotypes and adaptation to environmental pressures.

Competing interests

The authors declare no competing interest.

References

- 1 Gokhman, D. *et al.* Reconstructing the DNA methylation maps of the Neandertal and the Denisovan. *Science* **344**, 523-527, doi:10.1126/science.1250368 (2014).
- 2 Colbran, L. L. *et al.* Inferred divergent gene regulation in archaic hominins reveals potential phenotypic differences. *Nat Ecol Evol* **3**, 1598-1606, doi:10.1038/s41559-019-0996-x (2019).
- 3 Harris, K. & Nielsen, R. The Genetic Cost of Neanderthal Introgression. *Genetics* **203**, 881-891, doi:10.1534/genetics.116.186890 (2016).
- 4 McCoy, R. C., Wakefield, J. & Akey, J. M. Impacts of Neanderthal-Introgressed Sequences on the Landscape of Human Gene Expression. *Cell* **168**, 916-927 e912, doi:10.1016/j.cell.2017.01.038 (2017).
- 5 Colin M. Brand, L. L. C., and John A. Capra. Resurrecting the Alternative Splicing Landscape of Archaic Hominins using Machine Learning. *Nature Ecology and Evolution* XXXXXX (2023).
- 6 Li, Y. I. *et al.* RNA splicing is a primary link between genetic variation and disease. *Science* **352**, 600-604, doi:10.1126/science.aad9417 (2016).
- 7 Collins, L. & Penny, D. Complex spliceosomal organization ancestral to extant eukaryotes. *Mol Biol Evol* **22**, 1053-1066, doi:10.1093/molbev/msi091 (2005).

- 8 Jaganathan, K. *et al.* Predicting Splicing from Primary Sequence with Deep Learning. *Cell* **176**, 535-548 e524, doi:10.1016/j.cell.2018.12.015 (2019).
- 9 Huerta-Sanchez, E. *et al.* Altitude adaptation in Tibetans caused by introgression of Denisovan-like DNA. *Nature* **512**, 194-197, doi:10.1038/nature13408 (2014).
- 10 Beall, C. M. *et al.* Natural selection on EPAS1 (HIF2alpha) associated with low hemoglobin concentration in Tibetan highlanders. *Proc Natl Acad Sci U S A* **107**, 11459-11464, doi:10.1073/pnas.1002443107 (2010).
- 11 Liu, H. *et al.* EPAS1 regulates proliferation of erythroblasts in chronic mountain sickness. *Blood Cells Mol Dis* **84**, 102446, doi:10.1016/j.bcmd.2020.102446 (2020).
- 12 Ye, C. J. *et al.* Genetic analysis of isoform usage in the human anti-viral response reveals influenza-specific regulation of ERAP2 transcripts under balancing selection. *Genome Res* **28**, 1812-1825, doi:10.1101/gr.240390.118 (2018).
- 13 Klunk, J. *et al.* Evolution of immune genes is associated with the Black Death. *Nature* **611**, 312-319, doi:10.1038/s41586-022-05349-x (2022).
- 14 Mazin, P. V., Khaitovich, P., Cardoso-Moreira, M. & Kaessmann, H. Alternative splicing during mammalian organ development. *Nat Genet* **53**, 925-934, doi:10.1038/s41588-021-00851-w (2021).
- 15 Rong, S. *et al.* Large scale functional screen identifies genetic variants with splicing effects in modern and archaic humans. *bioRxiv*, 2022.2011.2020.515225, doi:10.1101/2022.11.20.515225 (2022).