The evolutionary history of Neanderthal and Denisovan Y chromosomes

Martin Petr, Mateja Hajdinjak, Qiaomei Fu, Elena Essel, Hélène Rougier, Isabelle Crevecoeur, Patrick Semal, V Liubov, Vladimir B Doronichev, Carles Lalueza-Fox, et al.

REPORT

Y EVOLUTION

The evolutionary history of Neanderthal and Denisovan Y chromosomes

Martin Petr[1]*, Mateja Hajdinjak[1,2], Qiaomei Fu[3,4,5], Elena Essel[1], Hélène Rougier[6], Isabelle Crevecoeur[7], Patrick Semal[8], Liubov V. Golovanova[9], Vladimir B. Doronichev[9], Carles Lalueza-Fox[10], Marco de la Rasilla[11], Antonio Rosas[12], Michael V. Shunkov[13], Maxim B. Kozlikin[13], Anatoli P. Derevianko[13], Benjamin Vernot[1], Matthias Meyer[1], Janet Kelso[1]*

Ancient DNA has allowed the study of human history in previously unseen detail. However, we lack comprehensive studies of the Y chromosomes of Denisovans and Neanderthals. Sequencing Y chromosomes from two Denisovans and three Neanderthals shows that the Y chromosomes of Denisovans split around 700 thousand years ago from a lineage shared by Neanderthals and modern human Y chromosomes, which diverged from each other around 370 thousand years ago. The phylogenetic relationships of archaic and modern human Y chromosomes differ from the population relationships inferred from the autosomal genomes and mirror mitochondrial DNA phylogenies, indicating replacement of both the mitochondrial and Y chromosomal gene pools in late Neanderthals. This replacement is plausible if the low effective population size of Neanderthals resulted in an increased genetic load in Neanderthals relative to modern humans.

Ancient DNA (aDNA) has transformed our understanding of human evolutionary history, revealing complex patterns of population migration and gene flow, including admixture from archaic humans into modern humans. Particularly important have been analyses of autosomal sequences (*1*, *2*), which represent a composite of genealogies of any individual's ancestors. Although mitochondrial DNA (mtDNA) and Y chromosomes only provide information about single maternal and paternal lineages, they offer a distinctive perspective on various aspects of population history such as sex-specific migration, matrilocality and patrilocality, and variance in reproductive success between individuals (*3*–*5*). Furthermore, because of their lower effective population size (N_e) compared with that of autosomal loci, coalescent times of mtDNA and Y chromosomes sampled from two populations provide an upper bound for the last time they experienced gene flow.

The mtDNA and autosomal sequences of Neanderthals, Denisovans, and modern humans have revealed puzzling phylogenetic discrepancies. Autosomal genomes show that Neanderthals and Denisovans are sister groups that split from modern humans between 550 thousand and 765 thousand years (ka) ago (*6*). By contrast, the mtDNAs of Neanderthals and modern humans are more similar to one another [time to the most recent common ancestor (TMRCA) of 360 to 468 ka ago] than to the mtDNAs of Denisovans (*7*). Notably, ~400-ka-old early Neanderthals from Sima de los Huesos were shown to carry mitochondrial genomes related to Denisovan mtDNAs (*8*, *9*). This suggests that Neanderthals originally carried a Denisovan-like mtDNA, which was later completely replaced through ancient gene flow from an early lineage related to modern humans (*7*, *9*).

The Y chromosomes of Neanderthals and Denisovans should provide an additional source of information about population splits and gene flow events between archaic and modern humans or populations related to them. However, with the exception of a small amount of Neanderthal Y chromosome coding sequence (118 kb) (*10*), none of the male Neanderthals or Denisovans studied to date have yielded sufficient amounts of endogenous DNA to allow comprehensive studies of archaic human Y chromosomes.

Previous genetic studies identified two male Denisovans, Denisova 4 (55 to 84 ka old) and Denisova 8 (106 to 136 ka old) (*11*, *12*), and two male late Neanderthals, Spy 94a (38 to 39 ka old) and Mezmaiskaya 2 (43 to 45 ka old) (*13*) (Fig. 1A). To enrich for Y chromosome DNA from these individuals, we performed hybridization capture using probes we designed to target ~6.9 Mb of the nonrecombining portion of the human Y chromosome (Fig. 1B) (*14*). This yielded sequence coverage of 1.4× for Denisova 4, 3.5× for Denisova 8, 0.8× for Spy 94a, and 14.3× for Mezmaiskaya 2 (Fig. 1C and table S2). In addition, we used a capture array designed for modern human Y chromosomes (*3*) to obtain 7.9× coverage of ~560 kb of the Y chromosome from the ~46- to 53-ka-old El Sidrón 1253 Neanderthal (Fig. 1C and table S2), which has been analyzed previously (*15*, *16*).

To call genotypes of the captured archaic human and previously published modern human Y chromosomes (*4*, *17*, *18*), we leveraged the haploid nature of the human Y chromosome and implemented a consensus approach that requires at least 90% of the reads observed at each site covered by at least three reads to agree on a single allele (*14*). This minimizes the impact of aDNA damage on genotyping accuracy while allowing for a small amount of sequencing error or contamination (fig. S8) (*14*).

To determine the relationships between Denisovan, Neanderthal, and modern human Y chromosomes, we constructed a neighbor-joining tree from the Y chromosome genotype calls (*14*). Unlike the rest of the nuclear genome, which puts Denisovans and Neanderthals as sister groups to modern humans (*2*), the Denisovan Y chromosomes form a separate lineage that split before Neanderthal and modern human Y chromosomes diverged from each other (Fig. 2A). Notably, all three late Neanderthal Y chromosomes cluster together and fall outside of the variation of present-day human Y chromosomes [Fig. 2A; (*10*)].

To estimate the TMRCA of archaic and modern human Y chromosomes, we adapted a previously published method that calculates the archaic-modern human TMRCA as a proportion of the deepest known split in present-day human Y variation (*4*, *10*, *14*) and is therefore robust to low coverage and aDNA damage (*10*) (fig. S8 and table S2). We first calculated the mutation rate in the 6.9-Mb target region to be 7.34×10^{-10} per base pair per year [bootstrap confidence interval (CI) 6.27×10^{-10} to 8.46×10^{-10}] (fig. S11 and table S11) (*14*) and used it to estimate a TMRCA of ~249 ka ago (bootstrap CI 213 to 293 ka ago) (fig. S11 and table S11) (*14*) for the African A00 lineage and a set of non-African Y chromosomes (*4*, *18*). This is consistent with other studies of present-day human Y chromosomes (*4*, *17*), suggesting that the Y chromosomal regions we sequenced are not unusual in terms of their mutation rate. We then used this A00 divergence time of 249 ka ago to infer TMRCAs between archaic

[1]Department of Evolutionary Genetics, Max Planck Institute for Evolutionary Anthropology, D-04103 Leipzig, Germany. [2]The Francis Crick Institute, NW1 1AT London, UK. [3]Key Laboratory of Vertebrate Evolution and Human Origins of Chinese Academy of Sciences, IVPP, CAS, Beijing 100044, China. [4]CAS Center for Excellence in Life and Paleoenvironment, Beijing 100044, China. [5]University of Chinese Academy of Sciences, Beijing 100049, China. [6]Department of Anthropology, California State University, Northridge, Northridge, CA 91330-8244, USA. [7]Université de Bordeaux, CNRS, UMR 5199-PACEA, 33615 Pessac Cedex, France. [8]Royal Belgian Institute of Natural Sciences, 1000 Brussels, Belgium. [9]ANO Laboratory of Prehistory 14 Linia 3-11, St. Petersburg 1990 34, Russia. [10]Institute of Evolutionary Biology, Consejo Superior de Investigaciones Científicas, Universitat Pompeu Fabra, 08003 Barcelona, Spain. [11]Área de Prehistoria, Departamento de Historia, Universidad de Oviedo, 33011 Oviedo, Spain. [12]Departamento de Paleobiología, Museo Nacional de Ciencias Naturales, Consejo Superior de Investigaciones Científicas, 28006 Madrid, Spain. [13]Institute of Archaeology and Ethnography, Siberian Branch, Russian Academy of Sciences, Novosibirsk, Russia.
*Corresponding author. Email: mp@bodkan.net (M.P.); kelso@eva.mpg.de (J.K.)

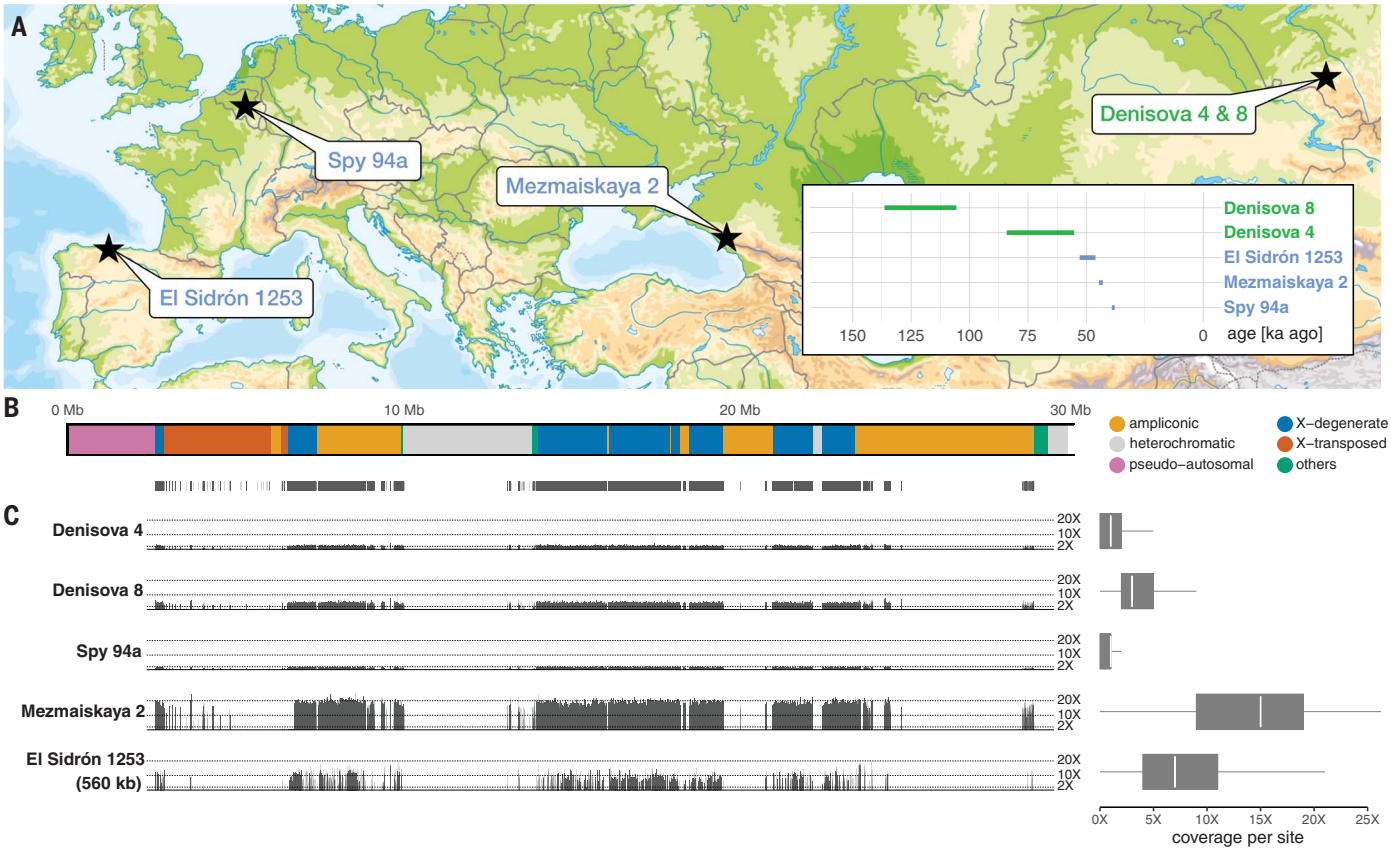

Fig. 1. Overview of male archaic humans in our study. (**A**) Archaeological site locations. Ages of specimens are shown as an inset (*12*, *13*, *15*). (**B**) Portion of the human Y chromosome targeted for capture [legend on right, coordinates of genomic regions are from (*30*)]. Thin black vertical lines show individual target capture regions. (**C**) (Left) Spatial distribution of sequencing coverage along the ~6.9 Mb of capture target regions. The heights of the thin vertical bars represent average coverage in each target region. Coordinates are aligned to match the chromosome shown in (B). (Right) Coverage across all target sites for each individual to the left.

Y chromosomes and present-day non-African Y chromosomes for each archaic individual (fig. S14 and table S12) (*14*). The two Denisovan Y chromosomes split from the modern human lineage around 700 ka ago (Denisova 8: 707 ka ago, CI 607 to 835 ka ago; Denisova 4: 708 ka ago, CI 550 to 932 ka ago) (Fig. 2B and table S12). By contrast, the three Neanderthal Y chromosomes split from the modern human lineage about 350 ka ago: 353 ka ago for Spy 94a (CI 287 to 450 ka ago), 370 ka ago for Mezmaiskaya 2 (CI 326 to 420 ka ago), and 339 ka ago for El Sidrón 1253 (CI 275 to 408 ka ago) (Fig. 2B and table S12). Additionally, we used the proportion of sharing of derived alleles with the high-coverage Mezmaiskaya 2 to estimate the TMRCA of the three Neanderthal Y chromosomes to around 100 ka ago (figs. S25 and S26). We validated the robustness of all TMRCA estimates using filters of varying levels of stringency and different genotype calling methods and also by comparing capture and shotgun sequence results (figs. S19, S21, and S23). Although there was some evidence of capture bias in the data (fig. S7), we observed no consistent differences between capture data and shotgun sequences or between individuals showing different read length distributions, indicating that technical biases do not affect our inferences (fig. S21).

Our estimates of the Neanderthal–modern human TMRCA (Fig. 2B) are younger than the previous estimate of ~588 ka ago from the El Sidrón 1253 individual (*10*). This older estimate was calculated from ~3× coverage of 118 kb of nuclear exome capture sequence and, because of the limited amount of data, used single-nucleotide polymorphisms supported even by single reads (*10*, *16*). However, this is problematic because it can result in an increased rate of erroneously called genotypes, leading to some shared alleles derived from Neanderthal and modern human being converted to the ancestral state, increasing the apparent TMRCA. When we applied filtering designed to mitigate errors (*14*) to the original El Sidrón 1253 data, we arrived at TMRCA estimates for El Sidrón 1253 consistent with all other Neanderthals in our study (fig. S22).

The estimates of Denisovan–modern human Y chromosome TMRCA agree with population split times inferred from autosomal sequences, suggesting that the differentiation of Denisovan Y chromosomes from modern humans occurred through a simple population split (*19*). By contrast, the young TMRCA of Neanderthal and modern human Y chromosomes and mtDNAs suggest that these loci have been replaced in Neanderthals through gene flow from an early lineage closely related to modern humans (Fig. 3A) (*7*). Previous work indicates that the rate of gene flow from modern humans into Neanderthals was on the order of only a few percent (*20*, *21*). Because the fixation probability of a locus is equal to its initial frequency in a population (*22*), the joint probability of both Neanderthal mtDNA and Y chromosomes being replaced by their introgressed modern human counterparts starting from a low initial frequency is even lower. However, owing to their low N_e and reduced efficacy of purifying selection, Neanderthals have been shown to have accumulated an excess of deleterious variation compared with modern humans (*16*), and it has been suggested that introgressed DNA was not neutral (*23*, *24*).

To explore the dynamics of modern human Y chromosomes introgressed into Neanderthals, we simulated introgression of a nonrecombining, uniparental locus under purifying selection (*14*, *25*). We considered a range of values

for the following parameters: Neanderthal and modern human N_e, the time that both populations evolved independently after their split, and the amount of sequence under selection, all of which affect the amount of deleterious variation that accumulated in Neanderthal and modern human populations before introgression (*14*). We simulated introgression of modern human Y chromosomes into the Neanderthal population in a single pulse and varied the contribution between 1 and 10%. We then traced the frequency of the introgressed modern human Y chromosomes in Neandertals over 100 ka. For each combination of parameters, we calculated how much lower the fitness of an average Neanderthal Y chromosome is compared with an average modern human Y chromosome using all linked

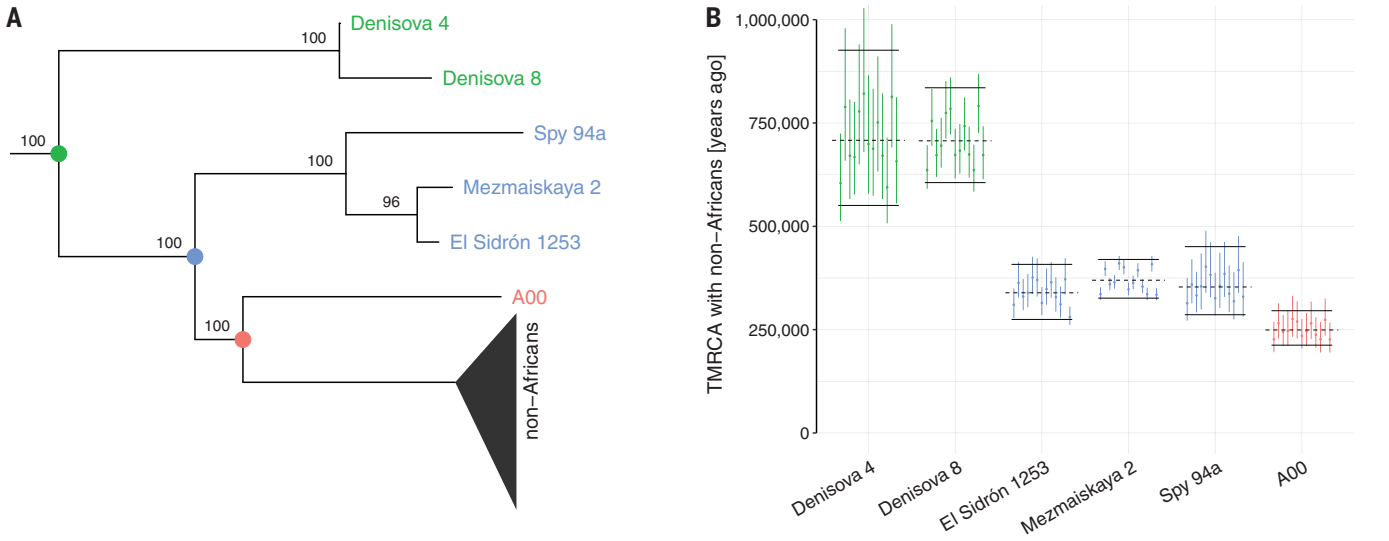

Fig. 2. Phylogenetic relationships between archaic and modern human Y chromosomes. (**A**) Neighbor-joining tree estimated from the Y chromosome genotype calls, excluding C-to-T and G-to-A polymorphisms, rooted with a chimpanzee as the outgroup (*14*). Numbers show bootstrap support out of 100 bootstrap replicates. Terminal branch lengths are not informative about the ages of specimens (Fig. 1A), owing to differences in sequence quality. (**B**) Estimates of TMRCA between Y chromosomes along the *x* axis and a panel of 13 non-African Y chromosomes. Each dot represents a TMRCA with a single non-African Y chromosome, with error bars showing 95% CI from a resampling of branch counts (*14*). Black horizontal lines show the mean TMRCA calculated across the full non-African panel (dotted lines) with resampling-based 95% CI (solid lines) (*14*).

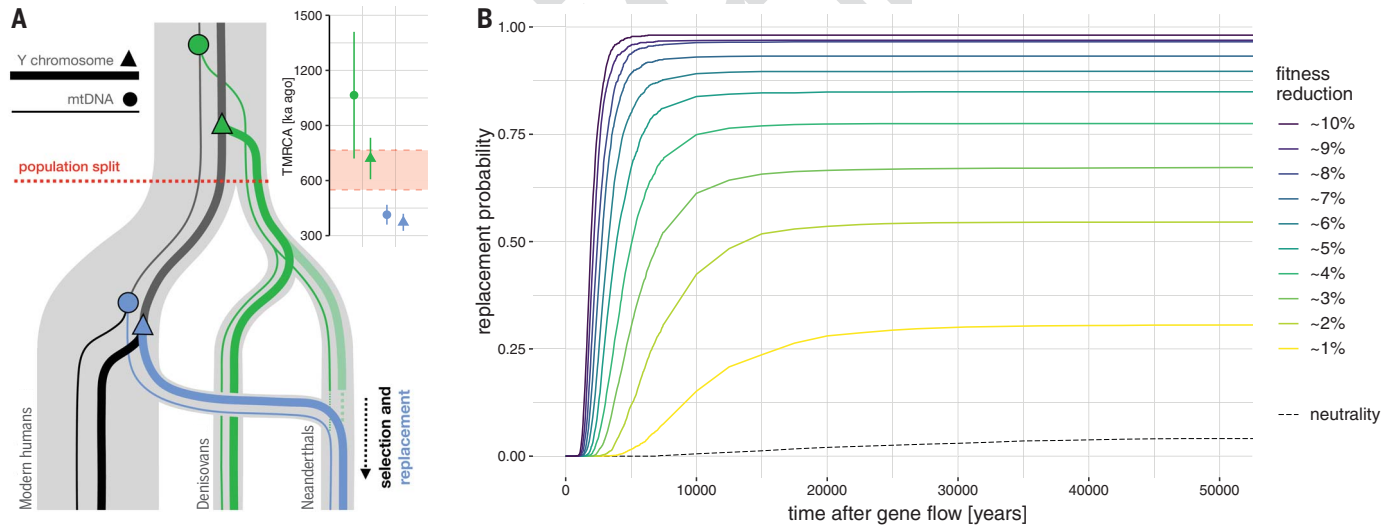

Fig. 3. Proposed model for the replacement of Neanderthal Y chromosomes and mtDNA. (**A**) Relationships between archaic and modern human mtDNA and Y chromosomes. The semitransparent Neanderthal lineage indicates a (as yet unsampled) hypothetical Y chromosome replaced by an early lineage related to modern human Y chromosomes. Most recent common ancestors with modern human lineages are shown for mtDNA (circles) and Y chromosomes (triangles). The inset shows TMRCAs for the four nodes in the diagram: Y chromosome TMRCAs as estimated by our study and mtDNA TMRCA estimates from the literature (*7*, *8*). The red shaded area highlights the 95% CI for the population split time between archaic and modern humans, shown as the dotted red horizontal line (*6*). (**B**) Probability of replacement of a non-recombining, uniparental Neanderthal locus over time, assuming a given level of fitness burden relative to its modern human counterpart. Trajectories are based on forward simulations across a grid of parameters (figs. S27 to S29) (*14*), with N_e of modern humans and Neanderthals fixed at 10,000 and 1000, respectively. Modern human introgression was simulated in a single pulse at 5%. Replacement probabilities from a wider range of model parameters are shown in fig. S31.

deleterious mutations on each simulated chromosome (*14*). This allows us to make a general statement about the probability of replacement in terms of the difference in fitness between Neanderthal and modern human Y chromosomes while abstracting over other factors that affect reproductive fitness but are currently impossible to simulate accurately (*26*).

For example, assuming 5% gene flow from modern humans, we found that even a 1% reduction in Neanderthal Y chromosome fitness increases the probability of replacement after 50 ka to ~25%, and a 2% reduction in fitness increases this probability to ~50% (Fig. 3B). However, the rate of gene flow as well as any factor that contributes to the difference in fitness between Neanderthal and modern human Y chromosomes will have a pronounced effect on the replacement probability (figs. S27 to S32). Given the crucial role of the Y chromosome in reproduction and fertility and its haploid nature, it is possible that deleterious mutations or structural variants on the Y chromosome have a larger impact on fitness than considered in our simulations. We therefore refrain from making predictions about the specific process of replacement, because we lack information about the frequencies of introgressed Y chromosomes in older Neanderthals, potential sex bias in the gene flow, and the fitness effects of single-nucleotide and structural variants on the Y chromosome (*26*). Nevertheless, our models are a proof-of-principle demonstration that even a simple difference in the efficacy of purifying selection between two lineages can markedly affect introgression dynamics of nonrecombining, uniparental DNA.

We conclude that the Y chromosomes of late Neandertals represent an extinct lineage closely related to modern human Y chromosomes that introgressed into Neanderthals between ~370 and ~100 ka ago. The presence of this Y chromosome lineage in all late Neanderthals makes it unlikely that genetic changes that accumulated in Neanderthal and modern human Y chromosomes before the introgression led to incompatibilities between these groups (*10*). Furthermore, we predict that the ~400-ka-old Sima de los Huesos Neanderthals should carry a Y chromosome lineage more similar to that of Denisovans than to that of later Neanderthals (*8*, *9*). Although the amount of modern human gene flow into Neanderthals appears to have been limited (*13*, *20*, *21*), we demonstrate that the replacement of mtDNA and Y chromosomes in Neanderthals is highly plausible, given the higher genetic load in Neanderthals compared with that in modern humans. Our results imply that differences in genetic load in uniparental loci between two hybridizing populations is a plausible driver for the replacements observed in other hybridization events (*27*–*29*).

REFERENCES AND NOTES

1. R. E. Green *et al*., *Science* **328**, 710–722 (2010).
2. M. Meyer *et al*., *Science* **338**, 222–226 (2012).
3. S. Lippold *et al*., *Investig. Genet.* **5**, 13 (2014).
4. M. Karmin *et al*., *Genome Res.* **25**, 459–466 (2015).
5. I. Olalde *et al*., *Science* **363**, 1230–1234 (2019).
6. K. Prüfer *et al*., *Nature* **505**, 43–49 (2014).
7. C. Posth *et al*., *Nat. Commun.* **8**, 16046 (2017).
8. M. Meyer *et al*., *Nature* **505**, 403–406 (2014).
9. M. Meyer *et al*., *Nature* **531**, 504–507 (2016).
10. F. L. Mendez, G. D. Poznik, S. Castellano, C. D. Bustamante, *Am. J. Hum. Genet.* **98**, 728–734 (2016).
11. S. Sawyer *et al*., *Proc. Natl. Acad. Sci. U.S.A.* **112**, 15696–15700 (2015).
12. K. Douka *et al*., *Nature* **565**, 640–644 (2019).
13. M. Hajdinjak *et al*., *Nature* **555**, 652–656 (2018).
14. See supplementary materials.
15. R. E. Wood *et al*., *Archaeometry* **55**, 148–158 (2013).
16. S. Castellano *et al*., *Proc. Natl. Acad. Sci. U.S.A.* **111**, 6666–6671 (2014).
17. Q. Fu *et al*., *Nature* **514**, 445–449 (2014).
18. S. Mallick *et al*., *Nature* **538**, 201–206 (2016).
19. K. Prüfer *et al*., *Science* **358**, 655–658 (2017).
20. M. Kuhlwilm *et al*., *Nature* **530**, 429–433 (2016).
21. M. J. Hubisz, A. L. Williams, A. Siepel, *PLOS Genet.* **16**, e1008895 (2020).
22. J. F. Crow, M. Kimura, *An Introduction to Population Genetics Theory* (The Blackburn Press, 1970).
23. K. Harris, R. Nielsen, *Genetics* **203**, 881–891 (2016).
24. I. Juric, S. Aeschbacher, G. Coop, *PLOS Genet.* **12**, e1006340 (2016).
25. B. C. Haller, P. W. Messer, *Mol. Biol. Evol.* **36**, 632–637 (2019).
26. S. Colaco, D. Modi, *Reprod. Biol. Endocrinol.* **16**, 14 (2018).
27. J. W. Ballard, M. C. Whitlock, *Mol. Ecol.* **13**, 729–744 (2004).
28. T. Bonnet, R. Leblois, F. Rousset, P.-A. Crochet, *Evolution* **71**, 2140–2158 (2017).
29. F. A. Seixas, P. Boursot, J. Melo-Ferreira, *Genome Biol.* **19**, 91 (2018).
30. L. Skov, M. H. Schierup; Danish Pan Genome Consortium, *PLOS Genet.* **13**, e1006834 (2017).
31. M. Petr, "The evolutionary history of Neandertal and Denisovan Y chromosomes" - code and Jupyter notebooks. Zenodo (2020); doi:10.5281/zenodo.3941654.
32. M. Petr, "The evolutionary history of Neandertal and Denisovan Y chromosomes" - Y chromosome capture designs. Zenodo (2020); doi: 10.5281/zenodo.3940568.

ACKNOWLEDGMENTS

We thank S. Pääbo, M. Stoneking, B. Peter, M. Slatkin, L. Skov, and E. Zavala for helpful discussions and comments on the manuscript. **Funding:** Q.F. was supported by funding from the Chinese Academy of Sciences (XDB26000000) and the National Natural Science Foundation of China (91731303, 41925009, 41630102). A.R. was funded by Spanish government (MICINN/FEDER) (grant number CGL2016-75109-P). The reassessment of the Spy collection by H.R., I.C., and P.S. was supported by the Belgian Science Policy Office (BELSPO 2004-2007, MO/36/0112). M.V.S., M.B.K., and A.P.D. were supported by the Russian Foundation for Basic Research (RFBR 17-29-04206). This study was funded by the Max Planck Society and the European Research Council (grant agreement number 694707). **Author contributions:** M.P. and J.K. analyzed data. M.H., Q.F., and E.E. performed laboratory experiments. H.R., I.C., P.S., L.V.G., V.B.D., C.L.-F., M.d.l.R., A.R., M.V.S., M.B.K., and A.P.D. provided samples. B.V., M.M., and J.K. supervised the project. M.P. and J.K. wrote and edited the manuscript with input from all co-authors. **Competing interests:** The authors declare no competing interests. **Data and materials availability:** Complete source code for data processing and simulations, as well as Jupyter notebooks with all analyses and results, can be found at Zenodo (*31*). Coordinates of the capture target regions and sequences of the capture probes are available at Zenodo (*32*). All sequence data are available from the European Nucleotide Archive under the accession number PRJEB39390.

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/369/issue/page/suppl/DC1
Materials and Methods
Figs. S1 to S32
Tables S1 to S14
References (*33*–*67*)
MDAR Reproducibility Checklist

View/request a protocol for this paper from *Bio-protocol*.

9 March 2020; accepted 6 August 2020
10.1126/science.abb6460