bioRxiv preprint doi: https://doi.org/10.1101/657247; this version posted June 2, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

Neanderthal-Denisovan ancestors interbred with a distantly-related hominin

Alan R. Rogers[1*], Nathan S. Harris[1], Alan A. Achenbach[1]

1 Dept. of Anthropology, University of Utah, Salt Lake City, USA

* rogers@anthro.utah.edu

Abstract

Previous research has shown that modern Eurasians interbred with their Neanderthal and Denisovan predecessors. We show here that hundreds of thousands of years earlier, the ancestors of Neanderthals and Denisovans interbred with their own Eurasian predecessors—members of a "superarchaic" population that separated from other humans about 2 mya. The superarchaic population was large, with an effective size between 10 and 46 thousand individuals. We confirm previous findings that:
(1) Denisovans also interbred with superarchaics, (2) Neanderthals and Denisovans separated early in the middle Pleistocene, (3) their ancestors endured a bottleneck of population size, and (4) the Neanderthal population was large at first but then declined in size. We provide qualified support for the view that (5) Neanderthals interbred with the ancestors of modern humans.

Author summary

We show that early in the middle Pleistocene, long before the expansion of modern humans into Eurasia, the "neandersovan" ancestors of Neanderthals and Denisovans undertook a very similar expansion. In both cases, an African population expanded into Eurasia, endured a narrow bottleneck of population size, interbred with indigeneous Eurasians, largely replaced them, and split into eastern and western sub-populations. In the earlier expansion, neandersovans interbred with a "superarchaic" population that had been separate since about 2 mya and may represent the original expansion of humans into Eurasia.

Introduction

We used genetic data to study the history of human populations during the middle Pleistocene. Early in this period, the ancestors of modern humans separated from those of Neanderthals and Denisovans. Somewhat later, Neanderthals and Denisovans separated from each other. The paleontology and archeology of this period also record important changes, as large-brained hominins appear in Europe and Asia, and Acheulean tools appear in Europe [1, 2]. We studied this period using genetic data from modern Africans and Europeans, and from the two archaic populations, Neanderthals and Denisovans.

Fig. 1 illustrates our notation. Upper-case letters refer to populations, and combinations such as XY refer to the population ancestral to X and Y. X represents an African population (the Yorubans), Y a European population, N Neanderthals, and

bioRxiv preprint doi: https://doi.org/10.1101/657247; this version posted June 2, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

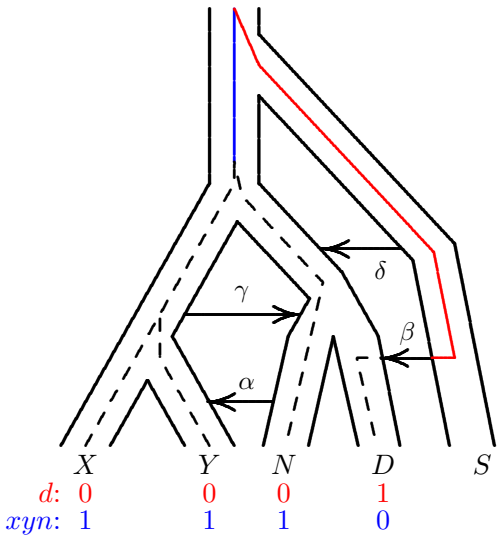

Fig 1. A population network including four episodes of gene flow, with an embedded gene genealogy. Upper case letters (X, Y, N, D, and S) represent populations (Africa, Europe, Neanderthal, Denisovan, and superarchaic). Greek letters label episodes of admixture. d and xyn illustrate two nucleotide site patterns, in which 0 and 1 represent the ancestral and derived alleles. A mutation on the red branch would generate site pattern d. One on the blue branch would generate xyn. For simplicity, this figure refers to Neanderthals with a single letter. Elsewhere, we use two letters to distinguish between the Altai and Vindija Neanderthals.

D Denisovans. S is an unsampled "superarchaic" population that is distantly related to other humans. Lower-case letters at the bottom of Fig. 1 label "nucleotide site patterns." A nucleotide site exhibits site pattern xyn if random nucleotides from populations X, Y, and N carry the derived allele, but those sampled from other populations are ancestral. Site pattern probabilities can be calculated from models of population history, and their frequencies can be estimated from data. Our Legofit [3] software estimates parameters by fitting models to these relative frequencies.

The current data include two high-coverage Neanderthal genomes: one from the Altai Mountains of Siberia [6] and the other from Vindija Cave in Croatia [8]. These appear in site pattern labels as "a" and "v". Thus, av is the site pattern in which the derived allele appears only in nucleotides sampled from the two Neanderthal genomes. Fig. 2 shows the site pattern frequencies studied here.

Greek letters in Fig. 1 label episodes of admixture. We label models by concatenating greek letters, to indicate the episodes of admixture they include. For example, model "$\alpha\beta$" includes only episodes α and β. Our model does not include gene flow from Denisovans into moderns, because there is no evidence of such gene flow into Europeans. Two years ago we studied a model that included only one episode of admixture: α, which refers to gene flow from Neanderthals into Europeans [9]. The left panel of Fig. 3 shows the residuals from this model, using the new data. Several are far from zero, suggesting that something is missing from the model [10].

Recent literature suggests some of what might be missing. There is evidence for admixture into Denisovans from a "superarchaic" population, which was distantly related to other humans [6, 8, 11–13] and also for admixture from early moderns into Neanderthals [13]. These episodes of admixture appear as β and γ in Fig. 1. Adding β and/or γ to the model improved the fit, yet none of the resulting models were satisfactory. For example, model $\alpha\beta\gamma$ implied (implausibly) that superarchaics

bioRxiv preprint doi: https://doi.org/10.1101/657247; this version posted June 2, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

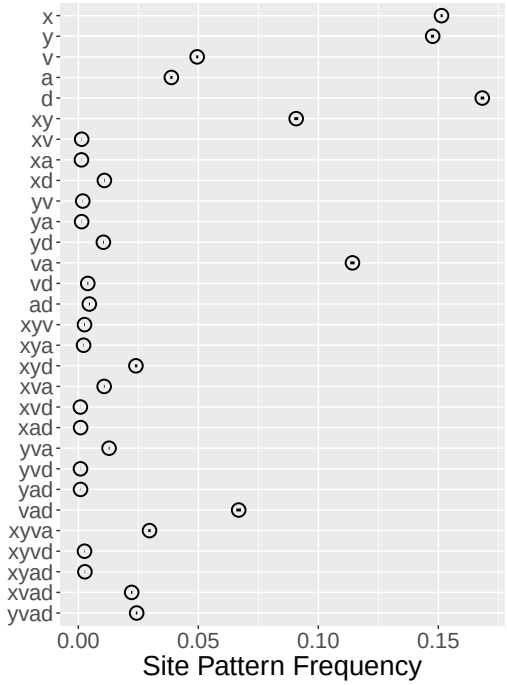

Fig 2. Observed site pattern frequencies. Horizontal axis shows the relative frequency of each site pattern in random samples consisting of a single haploid genome from each of X, Y, V, A, and D, representing Africa, Europe, Vindija Neanderthal, Altai Neanderthal, Denisovan, and superarchaic. Horizontal lines (which look like dots) are 95% confidence intervals estimated by a moving-blocks bootstrap [4]. Data: Simons Genome Diversity Project [5] and Max Planck Institute for Evolutionary Anthropology [6–8].

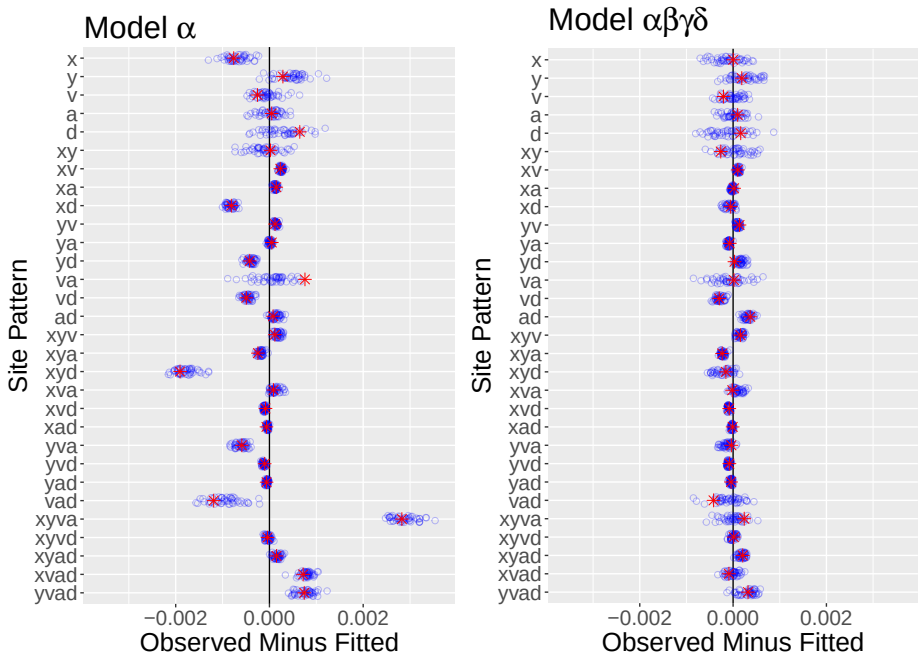

Fig 3. Residuals from models α and $\alpha\beta\gamma\delta$. Key: red asterisks, real data; blue circles, bootstrap replicates.

bioRxiv preprint doi: https://doi.org/10.1101/657247; this version posted June 2, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

Table 1. Bootstrap estimate of predictive error (bepe) values and bootstrap model average (booma) weights

Model	bepe	weight
α	1.14×10^{-6}	0
$\alpha\delta$	0.81×10^{-6}	0
$\alpha\gamma$	0.62×10^{-6}	0
$\alpha\gamma\delta$	0.39×10^{-6}	0
$\alpha\beta$	0.17×10^{-6}	0
$\alpha\beta\gamma$	0.17×10^{-6}	0
$\alpha\beta\delta$	0.14×10^{-6}	0.04
$\alpha\beta\gamma\delta$	0.11×10^{-6}	0.96

separated from other hominins eight million years ago.

To understand what might still be missing, consider what we know about the early middle Pleistocene, around 600 kya. At this time, large-brained hominins appear in Europe along with Acheulean stone tools [1, 2]. They were probably African immigrants, because similar fossils and tools occur earlier in Africa. According to one hypothesis, these early Europeans were Neanderthal ancestors [14, 15]. Somewhat earlier—perhaps 750 kya [16, table S12.2, p. 90]—the "neandersovan" ancestors of Neanderthals and Denisovans separated from the lineage leading to modern humans. Neandersovans may have separated from an African population and then expanded into Eurasia. If so, they would not have been expanding into an empty continent, for Eurasia had been inhabited since 1.85 mya [17]. Neandersovan immigrants would have met the indigenous superarchaic population of Eurasia. This suggests a fourth episode of admixture—from superarchaics into neandersovans—which appears as δ in Fig. 1.

Results

We considered eight models, all of which include α, and including all combinations of β, γ, and/or δ. In choosing among complex models, it is important to avoid overfitting. Conventional methods such as AIC [18] are not available, because we don't have access to the full likelihood function. Instead, we use the *bootstrap estimate of predictive error* (bepe) [3, 19, 20]. The best model is the one with the lowest value of bepe. When no model is clearly superior, it is better to average across several than to choose just one. For this purpose, we used *bootstrap model averaging* (booma) [3, 21]. The booma weight of the ith model is the fraction of data sets (including the real data and 50 bootstrap replicates) in which that model "wins," i.e. has the lowest value of bepe. The bepe values and booma weights of all models are in table 1.

The best model is $\alpha\beta\gamma\delta$, which includes all four episodes of admixture. It has smaller residuals (Fig. 3, right), the lowest bepe value, and the largest booma weight. One other model—$\alpha\beta\delta$—has a positive booma weight, but all others have weight zero. To understand what this means, recall that bootstrap replicates approximate repeated sampling from the process that generated the data. The models with zero weight lose in *all* replicates, implying that their disadvantage is large compared with variation in repeated sampling. On this basis, we can reject these models. Neither of the two remaining models can be rejected. These results provide strong support for two episodes of admixture (β and δ) and qualified support for a third (γ). Not only does this support previously-reported episodes of gene flow, it also provides evidence of another, in which superarchaics contributed genes to neandersovans. Model-averaged parameter estimates, which use the weights in table 1, are listed in SI table 1 and graphed in Fig. 4.

The superarchaic separation time, T_{XYNDS}, has a point estimate of 2.4 mya. This

bioRxiv preprint doi: https://doi.org/10.1101/657247; this version posted June 2, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

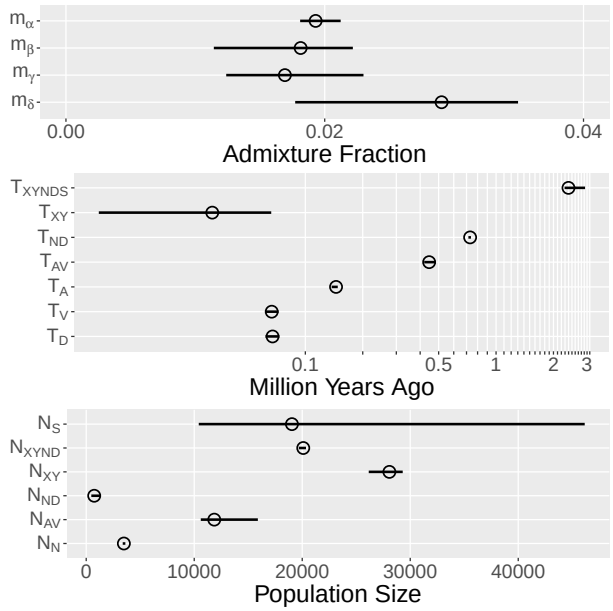

Fig 4. Model-averaged parameter estimates with 95% confidence intervals estimated by moving-blocks bootstrap [4]. Key: m_α, fraction of Y introgressed from N; m_β, fraction of D introgressed from S; m_γ, fraction of N introgressed from XY; m_δ; fraction of ND introgressed from S; T_{XYNDS}, superarchaic separation time; T_{XY}, separation time of X and Y; T_{ND}, separation time of N and D; T_{AV}, end of early epoch of Neanderthal history; T_A, age of Altai Neanderthal fossil; T_V, age of Vindija Neanderthal fossil; T_D, age of Denisovan fossil; N_S, size of superarchaic population; N_{XYND}, size of populations $XYND$ and $XYNDS$; N_{XY}, size of population XY; N_{ND}, size of population ND; N_{AV}, size of early Neanderthal population; N_N, size of late Neanderthal population. Parameters that exist in only one model are not averaged.

estimate may be biased upward, because our molecular clock assumes a fairly low mutation rate of 0.38×10^{-9} per nucleotide site per year. Other authors prefer slightly higher rates [22]. Although the mutation rate is apparently insensitive to generation time among the great apes, it is sensitive to the age of male puberty. If the average age of puberty during the past two million years were half-way between that of modern humans and chimpanzees, the yearly mutation rate would be close to 0.45×10^{-9} [23, Fig. 2B], and our estimate of T_{XYNDS} would drop to 2.0 mya—just at the origin of the genus *Homo*. Under this clock, the 95% confidence interval is 1.9–2.5 mya.

The lower end of this interval hardly differs from the 1.85 mya date of the earliest Eurasian archaeological remains at Dmanisi [17]. It is possible that superarchaics separated from an African population 1.9 mya, expanded into Eurasia, and left those remains at Dmanisi. If so, then superarchaics descend from the earliest human dispersal into Eurasia. On the other hand, some authors prefer a higher mutation rate of 0.5×10^{-9} per year [6]. Under this clock, the lower end of our confidence interval would be 1.7 mya. Thus, our results are also consistent with the view that superarchaics entered Eurasia after the earliest remains at Dmanisi.

Parameter N_S is the effective size of the superarchaic population. This parameter can be estimated because there are two sources of superarchaic DNA in our sample (β and δ), and this implies that coalescence time within the superarchaic population affects site pattern frequencies. Although this parameter has a broad confidence interval, even the low end implies a fairly large population of about 10,000. This does not require

bioRxiv preprint doi: https://doi.org/10.1101/657247; this version posted June 2, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

large numbers of superarchaic humans, because effective size can be inflated by geographic population structure [24]. Our large estimate may mean that neandersovans and Denisovans received gene flow from two different superarchaic populations.

Parameter T_{ND} is the separation time of Neanderthals and Denisovans. Our point estimate—731 kya—is remarkably old. Furthermore, the neandersovan population that preceded this split was remarkably small: $N_{ND} \approx 700$. This supports our previous results, which indicated an early separation of Neanderthals and Denisovans and a bottleneck among their ancestors [9].

Because our analysis includes two Neanderthal genomes, we can estimate the effective size of the Neanderthal population in two separate epochs. The early epoch extends from $T_{AV} = 446$ kya to $T_{ND} = 731$ kya, and within this epoch the effective size was large: $N_{AV} \approx 12,000$. It was smaller during the later epoch: $N_N \approx 3500$. These results support previous findings that the Neanderthal population was large at first but then declined in size [6, 8].

Discussion

Early in the middle Pleistocene—about 600 kya—large-brained hominins appear in the fossil record of Europe along with Acheulean stone tools. There is disagreement about how these early Europeans should be interpreted. Some see them as the common ancestors of modern humans and Neanderthals [25], others as an evolutionary dead end, later replaced by immigrants from Africa [26, 27], and others as early representatives of the Neanderthal lineage [14, 15]. Our estimates are most consistent with the last of these views. They imply that by 600 kya Neanderthals were already a distinct lineage, separate not only from the modern lineage but also from Denisovans.

These results resolve a discrepancy involving human fossils from Sima de los Huesos (SH). Those fossils had been dated to at least 350 kya and perhaps 400–500 kya [28]. Genetic evidence showed that they were from a population ancestral to Neanderthals and therefore more recent than the separation of Neanderthals and Denisovans [29]. However, genetic evidence also indicated that this split occurred about 381 kya [6, table S12.2]. This was hard to reconcile with the estimated age of the SH fossils. To make matters worse, improved dating methods later showed that the SH fossils are even older—about 600 ky, and much older than the molecular date of the Neanderthal-Denisovan split [30]. Our estimates resolve this conflict, because they push the date of the split back well beyond the age of the SH fossils.

Our estimate of the Neanderthal-Denisovan separation time conflicts with 381 kya estimate discussed above [6]. This discrepancy results in part from differing calibrations of the molecular clock. Under our clock, the 381 ky date becomes 502 ky [9], but this is still far from our own 731 ky estimate. The remaining discrepancy may reflect differences in our models of history. Misspecified models often generate biased parameter estimates.

Our new results on Neanderthal population size differ from those we published in 2017 [9]. At that time, we argued that the Neanderthal population was substantially larger than others had estimated. Our new estimates are more in line with those published by others [6, 8]. The difference does not result from our new and more elaborate model, because we get similar results from model α, which (like our 2017 model) allows only one episode of gene flow. Instead, it was including the Vindija Neanderthal genome that made the difference. Without this genome, we still get a large estimate ($N_N \approx 12,000$), even using model $\alpha\beta\gamma\delta$. This implies that the Neanderthals who contributed DNA to modern Europeans were more similar to the Vindija Neanderthal than to the Altai Neanderthal, as others have also shown [8].

Our results revise the date at which superarchaics separated from other humans.

bioRxiv preprint doi: https://doi.org/10.1101/657247; this version posted June 2, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

One previous estimate put this date between 0.9 and 1.4 mya [6, p. 47], which implied that superarchaics arrived well after the initial human dispersal into Eurasia, around 1.9 mya. This required a complex series of population movements between Africa and Eurasia [31, pp. 66-71]. Our new estimates do not refute this reconstruction, but they do allow a simpler one, which involves only three expansions of humans from Africa into Eurasia: an expansion of early *Homo* at about 1.9 mya, an expansion of neandersovans at about 700 kya, and an expansion of modern humans at about 50 kya.

Conclusions

It seems likely that superarchaics descend from the initial human settlement of Eurasia. As discussed above, the large effective size of the superarchaic population hints that it comprised at least two deeply-divided subpopulations, of which one mixed with neandersovans and another with Denisovans. We suggest that around 700 kya, neandersovans expanded from Africa into Eurasia, endured a bottleneck of population size, interbred with indigenous Eurasians, largely replaced them, and separated into eastern and western subpopulations—Denisovans and Neanderthals. These same events unfolded once again around 50 kya as modern humans expanded out of Africa and into Eurasia, largely replacing the Neanderthals and Denisovans.

Methods and materials

Samples

The modern human genomes in our sample are from the Simons Genome Diversity Project (SGDP) [5]. These include 3 Yorubans, 5 French, and 2 English, as detailed in Supplementary Information. The ancient genomes in our sample are the Altai Neanderthal genome [6], the Vindija Neanderthal genome [8], and the Denisovan genome [7]. Our Eurasian sample is entirely European, because there is no evidence of Denisovan admixture into Europe, and this allows us to avoid modeling Denisovan gene flow.

Outgroup

We used the chimpanzee [32] and gorilla [33, 34] reference genomes to call ancestral alleles. These genomes, aligned to human hg19, are available at the Santa Cruz Genome Browser. We downloaded .axt files aligned to human autosomes and generated a single .raf file for the chimpanzee and gorilla genomes, as described in Supporting Information.

Quality Control (QC)

We excluded sex chromosomes and normalized all variants at a given nucleotide site using the human reference genome. We excluded sites within 7 bases of the nearest INDEL and included sites only if they are monomorphic or are biallelic SNPs. For the SGDP genomes, we excluded nucleotide sites at which FL equals 0 or N. For the three archaic genomes, we excluded sites at which genotype quality (GQ) is less than 30. For the Denisovan and Altai Neanderthal genomes, we also excluded sites at which mapping quality (MQ) is less than 37. These QC filters are justified in Supporting Information.

bioRxiv preprint doi: https://doi.org/10.1101/657247; this version posted June 2, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

Molecular clock calibration

We assume that the modern and neandersovan lineages separated $T_{XYND} = 25920$ generations before the present [9]. This date assumes a mutation rate of 1.1×10^{-8} per site per generation [35] and a generation time of 29 y—a yearly rate of 0.38×10^{-9}.

Legofit analyses

The program `sitepat`, a component of the Legofit package [3], was used to call ancestral alleles, tabulate site patterns, and generate bootstrap replicates. Estimation then used the `legofit` program. These pipelines are detailed in Supporting Information.

Acknowledgments

We thank Ryan Bohlender, Elizabeth Cashdan, Jon Seger, and Timothy Webster for comments. This work was supported by NSF BCS 1638840, NSF GRF 1747505, and the Center for High Performance Computing at the University of Utah.

References

1. Klein RG. Anatomy, Behavior, and Modern Human Origins. Journal of World Prehistory. 1995;9(2):167–198.
2. Klein RG. The Human Career: Human Biological and Cultural Origins. 3rd ed. Chicago: University of Chicago Press; 2009.
3. Rogers AR. Legofit: Estimating Population History from Genetic Data. bioRxiv. 2019;doi:10.1101/613067.
4. Liu RY, Singh K. Moving Blocks Jacknife and Bootstrap Capture Weak Dependence. In: LePage R, Billard L, editors. Exploring the "Limits" of the Bootstrap. New York: Wiley; 1992. p. 225–248.
5. Mallick S, Li H, Lipson M, Mathieson I, Gymrek M, Racimo F, et al. The Simons Genome Diversity Project: 300 Genomes from 142 Diverse Populations. Nature. 2016;538:201–206. doi:10.1038/nature18964.
6. Prüfer K, Racimo F, Patterson N, Jay F, Sankararaman S, Sawyer S, et al. The Complete Genome Sequence of a Neanderthal from the Altai Mountains. Nature. 2014;505(7481):43–49. doi:10.1038/nature12886.
7. Meyer M, Kircher M, Gansauge MT, Li H, Racimo F, Mallick S, et al. A High-Coverage Genome Sequence from an Archaic Denisovan Individual. Science. 2012;338(6104):222–226.
8. Prüfer K, de Filippo C, Grote S, Mafessoni F, Korlević P, Hajdinjak M, et al. A High-Coverage Neandertal Genome from Vindija Cave in Croatia. Science. 2017;doi:10.1126/science.aao1887.
9. Rogers AR, Bohlender RJ, Huff CD. Early History of Neanderthals and Denisovans. Proceedings of the National Academy of Sciences, USA. 2017;114(37):9859–9863. doi:10.1073/pnas.1706426114.
10. Rogers AR, Bohlender RJ, Huff CD. Reply to Mafessoni and Prüfer: Inferences with and without Singleton Site Patterns. Proceedings of the National Academy of Sciences, USA. 2017;114(48):E10258–E10260. doi:10.1073/pnas.1717085114.

bioRxiv preprint doi: https://doi.org/10.1101/657247; this version posted June 2, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

11. Waddell PJ, Ramos J, Tan X. *Homo denisova*, Correspondence Spectral Analysis, Finite Sites Reticulate Hierarchical Coalescent Models and the Ron Jeremy Hypothesis. ArXiv. 2011;.

12. Waddell PJ. Happy New Year *Homo erectus*? More Evidence for Interbreeding with Archaics Predating the Modern Human/Neanderthal Split. ArXiv. 2013;.

13. Kuhlwilm M, Gronau I, Hubisz MJ, de Filippo C, Prado-Martinez J, Kircher M, et al. Ancient Gene Flow from Early Modern Humans into Eastern Neanderthals. Nature. 2016;doi:10.1038/nature16544.

14. Hublin JJ. The Origin of Neandertals. Proceedings of the National Academy of Sciences, USA. 2009;106(38):16022–16027.

15. Hublin JJ. Climatic Changes, Paleogeography, and the Evolution of the Neandertals. In: Akazawa T, Aoki K, Bar-Yosef O, editors. Neandertals and Modern Humans in Western Asia. Kluwer; 1998. p. 295–310.

16. Li H, Mallick S, Reich D. Population Size Changes and Split Times; 2014. Supplementary Information 12 of Prüfer et al. [6].

17. Ferring R, Oms O, Agustí J, Berna F, Nioradze M, Shelia T, et al. Earliest Human Occupations at Dmanisi (Georgian Caucasus) Dated to 1.85–1.78 Ma. Proceedings of the National Academy of Sciences, USA. 2011;108(26):10432–10436.

18. Akaike H. A New Look at the Statistical Model Identification. IEEE Transactions on Automatic Control. 1974;19(6):716–723.

19. Efron B. Estimating the Error Rate of a Prediction Rule: Improvement on Cross-Validation. Journal of the American Statistical Association. 1983;78(382):316–331.

20. Efron B, Tibshirani RJ. An Introduction to the Bootstrap. New York: Chapman and Hall; 1993.

21. Buckland ST, Burnham KP, Augustin NH. Model Selection: an Integral Part of Inference. Biometrics. 1997;53(2):603–618.

22. Kong A, Frigge ML, Masson G, Besenbacher S, Sulem P, Magnusson G, et al. Rate of *de Novo* Mutations and the Importance of Father/'s Age to Disease Risk. Nature. 2012;488(7412):471–475.

23. Amster G, Sella G. Life History Effects on the Molecular Clock of Autosomes and Sex Chromosomes. Proceedings of the National Academy of Sciences. 2016;113(6):1588–1593. doi:10.1073/pnas.1515798113.

24. Nei M, Takahata N. Effective Population Size, Genetic Diversity, and Coalescence Time in Subdivided Populations. Journal of Molecular Evolution. 1993;37:240–244.

25. Rightmire GP. Human Evolution in the Middle Pleistocene: The Role of Homo heidelbergensis. Evolutionary Anthropology. 1998;6(6):218–227.

26. Foley R, Lahr MM. Mode 3 Technologies and the Evolution of Modern Humans. Cambridge Archaeological Journal. 1997;7(1):3–36.

bioRxiv preprint doi: https://doi.org/10.1101/657247; this version posted June 2, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

27. Endicott P, Ho SYW, Stringer C. Using Genetic Evidence to Evaluate Four Palaeoanthropological Hypotheses for the Timing of Neanderthal and Modern Human Origins. Journal of Human Evolution. 2010;59(1):87–95.

28. Bischoff JL, Shamp DD, Aramburu A, Arsuaga JL, Carbonell E, De Castro JB. The Sima de los Huesos Hominids Date to Beyond U/Th Equilibrium (> 350kyr) and Perhaps to 400–500kyr: New Radiometric Dates. Journal of Archaeological Science. 2003;30(3):275–280.

29. Meyer M, Arsuaga JL, de Filippo C, Nagel S, Aximu-Petri A, Nickel B, et al. Nuclear DNA Sequences from the Middle Pleistocene Sima de los Huesos Hominins. Nature. 2016;531:504–507. doi:10.1038/nature17405.

30. Bischoff JL, Williams RW, Rosenbauer RJ, Aramburu A, Arsuaga JL, García N, et al. High-resolution U-series Dates from the Sima de los Huesos Hominids Yields: Implications for the Evolution of the Early Neanderthal Lineage. Journal of Archaeological Science. 2007;34(5):763–770.

31. Reich D. Who We Are and How We Got Here: Ancient DNA and the New Science of the Human Past. New York: Pantheon Books; 2018.

32. Chimpanzee Sequencing and Analysis Consortium. Initial Sequence of the Chimpanzee Genome and Comparison with the Human Genome. Nature. 2005;437(7055):69–87.

33. Scally A, Dutheil JY, Hillier LW, Jordan GE, Goodhead I, Herrero J, et al. Insights into Hominid Evolution from the Gorilla Genome Sequence. Nature. 2012;483(7388):169–175.

34. Gordon D, Huddleston J, Chaisson MJ, Hill CM, Kronenberg ZN, Munson KM, et al. Long-Read Sequence Assembly of the Gorilla Genome. Science. 2016;352(6281):aae0344. doi:10.1126/science.aae0344.

35. Veeramah KR, Hammer MF. The Impact of Whole-Genome Sequencing on the Reconstruction of Human Population History. Nature Reviews Genetics. 2014;15(3):149–162.

Figure legends

1. A population network including four episodes of gene flow, with an embedded gene genealogy. Upper case letters (X, Y, N, D, and S) represent populations (Africa, Europe, Neanderthal, Denisovan, and superarchaic). Greek letters label episodes of admixture. d and xyn illustrate two nucleotide site patterns, in which 0 and 1 represent the ancestral and derived alleles. A mutation on the red branch would generate site pattern d. One on the blue branch would generate xyn. For simplicity, this figure refers to Neanderthals with a single letter. Elsewhere, we use two letters to distinguish between the Altai and Vindija Neanderthals.

2. Observed site pattern frequencies. Horizontal axis shows the relative frequency of each site pattern in random samples consisting of a single haploid genome from each of X, Y, V, A, and D, representing Africa, Europe, Vindija Neanderthal, Altai Neanderthal, Denisovan, and superarchaic. Horizontal lines (which look like dots) are 95% confidence intervals estimated by a moving-blocks bootstrap [4]. Data: Simons Genome Diversity Project [5] and Max Planck Institute for Evolutionary Anthropology [6–8].

bioRxiv preprint doi: https://doi.org/10.1101/657247; this version posted June 2, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

3. Residuals from models α and $\alpha\beta\gamma\delta$. Key: red asterisks, real data; blue circles, bootstrap replicates.

4. Model-averaged parameter estimates with 95% confidence intervals estimated by moving-blocks bootstrap [4]. Key: m_α, fraction of Y introgressed from N; m_β, fraction of D introgressed from S; m_γ, fraction of N introgressed from XY; m_δ; fraction of ND introgressed from S; T_{XYNDS}, superarchaic separation time; T_{XY}, separation time of X and Y; T_{ND}, separation time of N and D; T_{AV}, end of early epoch of Neanderthal history; T_A, age of Altai Neanderthal fossil; T_V, age of Vindija Neanderthal fossil; T_D, age of Denisovan fossil; N_S, size of superarchaic population; N_{XYND}, size of populations $XYND$ and $XYNDS$; N_{XY}, size of population XY; N_{ND}, size of population ND; N_{AV}, size of early Neanderthal population; N_N, size of late Neanderthal population. Parameters that exist in only one model are not averaged.

Supporting information legends

1. Supplementary Methods