

The Derived *FOXP2* Variant of Modern Humans Was Shared with Neandertals

Johannes Krause,^{1,*} Carles Lalueza-Fox,²
Ludovic Orlando,^{3,4} Wolfgang Enard,¹
Richard E. Green,¹ Hernán A. Burbano,¹
Jean-Jacques Hublin,¹ Catherine Hänni,^{3,4}
Javier Fortea,⁵ Marco de la Rasilla,⁵
Jaume Bertranpetit,⁶ Antonio Rosas,⁷
and Svante Pääbo¹

¹Max Planck Institute for Evolutionary Anthropology
Deutscher Platz 6
04103 Leipzig
Germany

²Departament de Biologia Animal
Facultat de Biologia
Universitat de Barcelona
Avenida Diagonal 645
08028 Barcelona
Spain

³Paléogénétique et Évolution Moléculaire
Institut de Génomique Fonctionnelle de Lyon
Université de Lyon
CNRS, INRA
École Normale Supérieure de Lyon, 46 Allée d'Italie
69364 Lyon, Cedex 07
France

⁴Institut Fédératif Biosciences Gerland Lyon Sud
Université Lyon 1
CNRS, INRA
École Normale Supérieure de Lyon, 46 Allée d'Italie
69364 Lyon, Cedex 07
France

⁵Área de Prehistoria Departamento de Historia
Universidad de Oviedo
Calle Teniente Alfonso Martínez s/n
33011 Oviedo
Spain

⁶Unitat de Biologia Evolutiva, CEXS
Universitat Pompeu Fabra
c. Dr. Aiguader 88
08003 Barcelona
Spain

⁷Departamento de Paleobiología
Museo Nacional de Ciencias Naturales (CSIC)
Calle José Gutierrez Abascal 2
28006 Madrid
Spain

Summary

Although many animals communicate vocally, no extant creature rivals modern humans in language ability. Therefore, knowing when and under what evolutionary pressures our capacity for language evolved is of great interest. Here, we find that our closest extinct relatives, the Neandertals, share with modern humans two

evolutionary changes in *FOXP2*, a gene that has been implicated in the development of speech and language. We furthermore find that in Neandertals, these changes lie on the common modern human haplotype, which previously was shown to have been subject to a selective sweep. These results suggest that these genetic changes and the selective sweep predate the common ancestor (which existed about 300,000–400,000 years ago) of modern human and Neandertal populations. This is in contrast to more recent age estimates of the selective sweep based on extant human diversity data. Thus, these results illustrate the usefulness of retrieving direct genetic information from ancient remains for understanding recent human evolution.

Results and Discussion

One approach to understanding the origin and evolution of language is to study the evolution of genes necessary for language acquisition. Although language and speech are clearly genetically complex phenomena, the only gene currently known that has a specific role in the development of language and speech is *FOXP2* [1, 2]. The inactivation of one *FOXP2* copy leads primarily to deficits in orofacial movements and linguistic processing similar to those in individuals with adult-onset Broca's aphasia [3]. Although *FOXP2* is among the 5% most conserved proteins among mammals [4], two amino acid substitutions have fixed in the human lineage since our split from the chimpanzee common ancestor. These amino acid substitutions are caused by nucleotide substitutions at positions 911 and 977 in exon 7 of the *FOXP2* gene and change threonine to aspartic acid and arginine to serine residues, respectively. Coalescent simulations using extant human diversity data in the region around exon 7 suggested that a recent selective sweep occurred [4, 5] and ended within the last 200,000 years [4]. Taken together, these results are compatible with the notion that these two amino acid substitutions are associated with the emergence of fully modern language ability [4, 5]. We have thus undertaken a targeted approach to determine the genotype of Neandertals at these *FOXP2* positions.

The retrieval of nuclear DNA sequences from ancient remains by the polymerase chain reaction (PCR) is fraught with difficulty, primarily because most remains contain extremely low quantities of endogenous DNA. An additional problem when Neandertal remains are studied is that modern human DNA is a frequent contaminant of ancient remains and laboratory reagents [6] and that most human DNA sequences cannot be distinguished from Neandertal sequences [7]. Thus, in addition to designing six primers that in four combinations amplify the nucleotide positions 911 and 977 (Figure 1), we designed three types of controls in order to as far as possible ensure the authenticity of our results.

First, we analyzed the ratio of Neandertal to modern human mitochondrial (mt) DNA in an mtDNA segment

*Correspondence: krause@eva.mpg.de

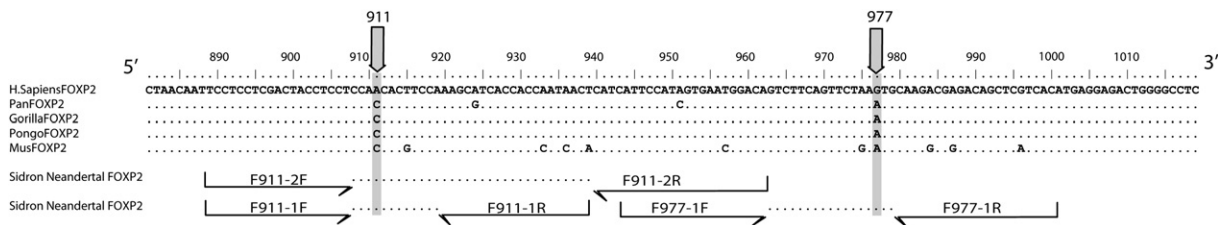


Figure 1. Sequence Alignment of Nucleotide Positions 880–1020 from the *FOXP2* Gene

The two nonsynonymous nucleotide substitutions on the human lineage are indicated by arrows. Identical positions in the alignment are given as dots. The three primer pairs used to retrieve the two substitutions from the El Sidrón Neandertals are indicated by arrows.

where Neandertals and contemporary humans can be distinguished on the basis of several substitutions [8] fixed in each group. Among 46 DNA extracts prepared from 22 Neandertal bones, we identified two bones where approximately 98%–99% of the hominid mtDNA in the extracts are of the Neandertal type and where DNA is abundant enough for PCR. Both bones originate from El Sidrón Cave in Asturias (north of Spain) [9]. These bones (numbers 1253 and 1351c) were removed under sterile conditions directly from the excavation in 2006, immediately frozen, and transported to our clean room facility where extractions were performed. Because additional DNA sequences from the mtDNA hypervariable region show that they stem from different individuals (data not shown), they are referred to as the first and the second Neandertal, respectively, below.

Second, we designed a number of additional controls for detecting modern human nuclear DNA contamination. To this end, we used genomic sequence data produced from a 38,000-year-old Neandertal from Vindija Cave, Croatia [8] to identify seven sequence positions on autosomes and the X chromosome that are ancestral (i.e., identical to the chimpanzee sequence) in the Vindija Neandertal but derived (i.e., different from the chimpanzee sequence) and not known to vary among current humans. We avoided C to T or G to A differences because these are common artifacts of ancient DNA sequences [10–14]. We chose several control positions based on the Vindija Neandertal because it is not certain that any particular one would be shared with the El Sidrón Neandertals. Furthermore, given that most extracts contain very low amounts of Neandertal DNA, we expect only a fraction of amplifications of nuclear target sequences in any one experiment to yield products.

Third, we took advantage of the detailed modern human Y chromosome phylogeny [15] and the fact that the time to the most recent common ancestor for modern human Y chromosomes, approximately 90,000 years ago [16], is more recent than the estimated population split time with Neandertals [17]. A priori we therefore expect Neandertal Y chromosomes to fall outside the variation of modern human Y chromosomes unless there was male gene flow from modern humans into Neandertals. This allowed us to design primers that amplify five positions that define the deepest nodes in the tree of current Y chromosomes as well as subgroups including almost all European and Asian and most African Y chromosomes (Figure S1 in the Supplemental Data available online).

We combined all control primer pairs with different primer sets for the *FOXP2* gene in four different primer

mixes and performed two-step multiplex PCRs [18]. All primers were amplified for 30 cycles in a first PCR from which aliquots were removed and then used to amplify each specific target individually in a second PCR. For each primer mix and individual, we also performed mock amplifications containing no template DNA. None of 108 secondary PCRs from such negative controls yielded any specific product. The results for the two Neandertals are summarized in Figure 2.

For the autosomal controls, nine out of 20 secondary PCRs yielded the relevant products. The fact that not all primer pairs yield products shows that the extracts contain such small amounts of nuclear DNA that amplification success is only sporadic and likely to often start from single-template molecules [10]. For two of the three autosomal positions analyzed, both individuals carry the ancestral allele. Out of the five products retrieved for the third autosomal position (A1, Table S1), both the ancestral and derived alleles are found in both El Sidrón Neandertals, suggesting that they were polymorphic at this position. The second individual in addition carried an allele not seen in modern humans or chimpanzees. Given that this sequence variant is a G to T substitution, not a substitution commonly associated with ancient DNA damage [10], it is likely to represent a genuine allele present in the Neandertals. However, further work would be necessary to verify this. Out of four X chromosomal positions, only one yielded products in one individual. It carried the ancestral state. Both Neandertals yielded products for Y chromosomal primer pairs, indicating that they were males. Strikingly, all 15 Y chromosomal products for the five assayed positions show the ancestral allele. This includes two polymorphisms that define the deepest split among current human Y chromosomes (Y2 and Y4, Figure S1) as well as two polymorphisms that cover less common African Y chromosomes (Y3 and Y5, Figure S1). These Y chromosome results must derive, then, either from Y chromosomes that fall outside the variation of modern humans or from the very rare African lineages not covered by the assay (Figure S1). For our purposes, this result shows that neither the maternally inherited mtDNA nor the paternally inherited Y chromosome shows evidence of gene flow from modern humans into Neandertals or of subsequent contamination of their mortal remains. From these results, we infer that the multiplex-PCR products derive from Neandertal nuclear-DNA templates.

Of the two *FOXP2* substitutions in exon 7, position 911 was retrieved four times from the first individual and twice from the second. In all cases, the recovered allele was the derived allele seen in modern humans. Position

FOXP2 Gene of Neanderthals

3

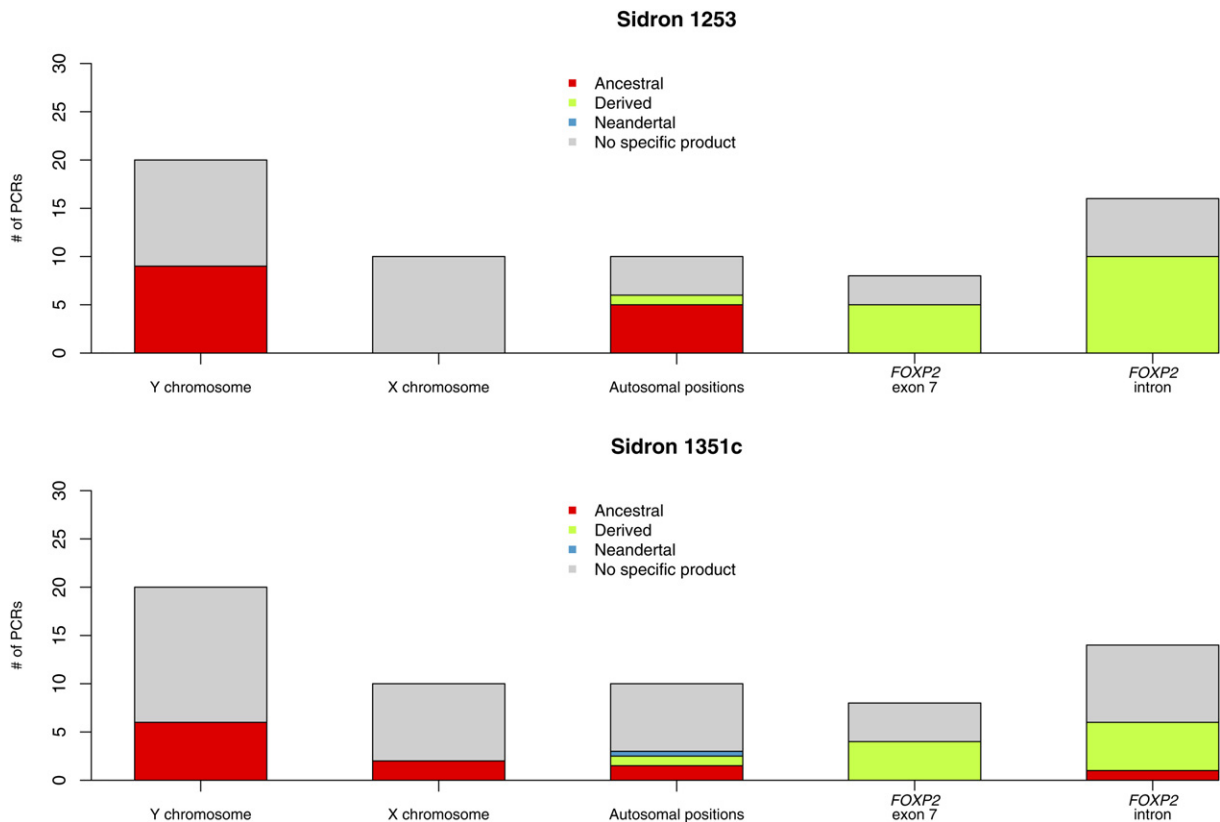


Figure 2. Summary of the Results of the Four Multiplex PCRs Carried out in Leipzig on the Two El Sidrón Neanderthals. PCRs that yielded products carrying ancestral allelic states are in red, derived allelic states are in green, blue indicates a derived Neandertal-specific state, and gray indicates PCRs where no specific products were obtained. Detailed results are given in Tables S1 and S2.

977 was retrieved once from the first individual and twice from the second individual. Again, in all cases the alleles were the derived variant that is fixed in modern humans. To test the reproducibility of these results, we sent bone fragments of the first Neandertal individual as well as information about the control primers to one laboratory in Barcelona, Spain and one laboratory in Lyon, France. Both laboratories confirmed the results by consistently finding the derived state for the *FOXP2* substitutions and the ancestral states in the controls (see Supplemental Experimental Procedures and Tables S3 and S4 for details). This not only corroborates the results obtained in Leipzig but also shows that at least the first individual was likely to be homozygous for the two *FOXP2* substitutions in exon 7, even under the conservative assumption that each successful amplification reflects a single starting molecule (individual 1; position 911, 13 independent amplifications, $p = 1.2 \times 10^{-4}$; position 977, five independent amplifications, $p = 3.1 \times 10^{-2}$) [19].

In order to address the question of whether the signal of a selective sweep detected by the allele-frequency spectrum close to exon 7 in the *FOXP2* gene among present-day humans was present also in the Neanderthals, we included in multiplex PCR 3 and 4, described above, nine primer pairs that encompass nucleotide polymorphisms in the intronic region upstream of exon 7 and that carry high-frequency alleles that are derived among humans [4]. For seven of these nine positions,

which are crucial for detecting and dating the selective sweep signal (Supplemental Experimental Procedures), at least one product was obtained from each of the two Neanderthals (Figure 2). For six of the seven sites, all Neandertal products showed the derived allele (Table S1). For one site, one product carried the derived allele in the first Neandertal and the second Neandertal carried the ancestral allele, indicating that this position (S5, Table S1) was polymorphic among Neanderthals. It should be noted that the frequency of the derived allele for the later position in contemporary humans is 0.88, whereas it is > 0.97 for the other six positions [4], suggesting that this derived allele may be younger than the others. Thus, not only was the derived form of the *FOXP2* protein present in Neanderthals, but it is also linked to the haplotype that is common among modern humans and appears to have been subject to a selective sweep.

If the selective sweep that affected this genomic region was in fact associated with one or both of the amino acid substitutions in *FOXP2*, then one of three scenarios must apply to the evolution of *FOXP2* in hominids. The first scenario is that the positively selected *FOXP2* haplotype was transferred into Neanderthals from modern humans or vice versa through gene flow. This seems to be an unlikely possibility. Neither mitochondrial nor—as we show here—Y chromosomal gene flow between the two hominid groups can be detected. Furthermore, other tests for admixture that rely on autosomal variation have hitherto failed to detect any gene flow

into Neandertals (J. Mullikin and D. Reich, personal communication). The second scenario is that the relevant *FOXP2* haplotype was present in the ancestral population of modern humans and Neandertals and was later positively selected in humans after their divergence from Neandertals. For this scenario to be likely, the relevant haplotype would have needed to be at a considerable frequency in the ancestral population in order to obtain the relatively high frequencies in Neandertals. However, the higher the frequency of a variant is before it becomes positively selected, the less likely it is to detect a signature of a selective sweep [20]. Hence, this scenario might also be considered relatively unlikely. The third scenario is that the selective sweep started before the divergence of the ancestral populations of Neandertals and modern humans around 300,000–400,000 years ago [17]. Note that given a divergence time of chimpanzees and humans of 6.5 million years [21, 22], the fixation of the sweep would have occurred within the last 260,000 years, and that a sweep driven by a selective advantage of 1% takes 50,000–80,000 years on average to come to completion [23]. This, in conjunction with the fact that deviations from model assumptions about generation times, selection coefficients, population sizes, and panmixia could further increase the variance on these time estimates, makes it possible that the sweep started or occurred in the common ancestral population of Neandertals and modern humans. In fact, estimates based on the variation in current human populations by necessity rest on simplistic models about human population history and parameters not known with certainty, making these estimates inherently unreliable with respect to the absolute timing of events. Therefore, if the third scenario is correct—i.e., the substitutions in *FOXP2* and the selective sweep originated in the common ancestor of Neandertals and modern humans—this work demonstrates the usefulness of directly testing for modern human substitutions in Neandertals. A Neandertal genome sequence will thus provide invaluable direct information on the timing of genetic changes on the human lineage relative to the divergence to Neandertals [8].

It is worth noting that we have analyzed only the two changes that are known to have occurred in exon 7 of the *FOXP2* gene after the divergence of humans and chimpanzees. Although unlikely given the high degree of conservation of this gene, it is possible that other Neandertal-specific substitutions existed elsewhere in *FOXP2*. Only the full sequence of the entire gene, as will eventually be provided by a Neandertal genome sequence [8], will therefore allow a comprehensive evaluation of the function of *FOXP2* in Neandertals.

In conclusion, the current results show that the Neandertals carried a *FOXP2* protein that was identical to that of present-day humans in the only two positions that differ between human and chimpanzee. Leaving out the unlikely scenario of gene flow, this establishes that these changes were present in the common ancestor of modern humans and Neandertals. The date of the emergence of these genetic changes therefore must be older than that estimated with only extant human diversity data, thus demonstrating the utility of direct evidence from Neandertal DNA sequences for understanding recent modern human evolution. Whatever

function the two amino acid substitutions might have for human language ability, it was present not only in modern humans but also in late Neandertals. Ongoing *in vivo* and *in vitro* experiments should help to delineate these functions.

Experimental Procedures

DNA Extraction and Amplification

The El Sidrón bone samples were removed under sterile conditions with laboratory coveralls, sterile gloves, face masks, and sterile blades directly from the excavation in 2006 and were immediately frozen and transported to Leipzig. In Leipzig, DNA sampling and extractions were performed in a laboratory dedicated exclusively to ancient DNA work. For minimizing contamination with modern human DNA, the surface of the bone was first removed, and then ~350 mg of bone was sampled, powdered, and extracted as described [24]. Two-step multiplex PCRs [17] in a total volume of 20 μ l containing up to 14 primer pairs were set up and incubated for 10 min at room temperature with 0.5 units of Shrimp Nuclease (Biotec Pharmacon) to degrade double-stranded DNA that may contaminate reagents. After inactivation of the nuclease for 30 min at 65°C, 5 μ l of extract was added. Thirty cycles of PCR were performed in the first and second PCR, respectively. Reaction conditions were as described [18] except for the annealing temperature, which was 55°C. Amplification products of the correct size were cloned with the TOPO TA cloning kit (Invitrogen), and three to 16 clones from each product were sequenced with an ABI3730 capillary sequencer (Applied Biosystems) (Table S1). Primer sequences, extraction and amplification conditions, and detailed results from the Barcelona and Lyon laboratories are described in the Supplemental Data.

Supplemental Data

Experimental Procedures, one figure, and five tables are available at <http://www.current-biology.com/cgi/content/full/17/21/4> [■ ■ ■ ■ / DC1/].

Acknowledgments

We are indebted to Adrian Briggs, Michael Hofreiter, and Tomislav Maricic for helpful comments; especially to Susan Ptak for help with statistical analyses and Matthias Meyer for contributing laboratory expertise; and to the Max Planck Society for funding. C.L.-F. and A.R. were supported by a grant from the Ministry of Education and Science of Spain. R.E.G. was supported by a postdoctoral fellowship from the National Sciences Foundation. The El Sidrón excavation project is funded by the Principado de Asturias (Spain). A.R., J.F., and M.R. provided Neandertal samples and palaeontological information; J.J.H. provided paleontological and archaeological information; J.K., L.O., and C.L.-F. extracted, amplified, sequenced, and analyzed ancient DNA; E.G. and H.B. helped design the control experiment; W.E. assisted in the interpretation of the results; C.H. coordinated the work in Lyon; J.B. provided laboratory space in Barcelona; S.P. initiated, planned, and coordinated the study; J.K., E.G. and S.P. wrote the paper.

Received: September 15, 2007

Revised: September 25, 2007

Accepted: September 26, 2007

Published online: October 18, 2007

References

1. Lai, C.S., Fisher, S.E., Hurst, J.A., Vargha-Khadem, F., and Monaco, A.P. (2001). A forkhead-domain gene is mutated in a severe speech and language disorder. *Nature* 413, 519–523.
2. MacDermot, K.D., Bonora, E., Sykes, N., Coupe, A.M., Lai, C.S., Vernes, S.C., Vargha-Khadem, F., McKenzie, F., Smith, R.L., Monaco, A.P., and Fisher, S.E. (2005). Identification of *FOXP2* truncation as a novel cause of developmental speech and language deficits. *Am. J. Hum. Genet.* 76, 1074–1080.

3. Vargha-Khadem, F., Gadian, D.G., Copp, A., and Mishkin, M. (2005). *FOXP2* and the neuroanatomy of speech and language. *Nat. Rev. Neurosci.* 6, 131–138.
4. Enard, W., Przeworski, M., Fisher, S.E., Lai, C.S., Wiebe, V., Kitano, T., Monaco, A.P., and Pääbo, S. (2002). Molecular evolution of *FOXP2*, a gene involved in speech and language. *Nature* 418, 869–872.
5. Zhang, J., Webb, D.M., and Podlaha, O. (2002). Accelerated protein evolution and origins of human-specific features: *Foxp2* as an example. *Genetics* 162, 1825–1835.
6. Hofreiter, M., Serre, D., Poinar, H.N., Kuch, M., and Pääbo, S. (2001). Ancient DNA. *Nat. Rev. Genet.* 2, 353–359.
7. Pääbo, S. (1999). Human evolution. *Trends Cell Biol.* 9, M13–M16.
8. Green, R.E., Krause, J., Ptak, S.E., Briggs, A.W., Ronan, M.T., Simons, J.F., Du, L., Egholm, M., Rothberg, J.M., Paunovic, M., and Pääbo, S. (2006). Analysis of one million base pairs of Neanderthal DNA. *Nature* 444, 330–336.
9. Rosas, A., Martínez-Maza, C., Bastir, M., García-Tabernero, A., Lalueza-Fox, C., Huguet, R., Ortiz, J.E., Julià, R., Soler, V., de Torres, T., et al. (2006). Paleobiology and comparative morphology of a late Neanderthal sample from El Sidron, Asturias, Spain. *Proc. Natl. Acad. Sci. USA* 103, 19266–19271.
10. Hofreiter, M., Jaenicke, V., Serre, D., Haeseler v., A., and Pääbo, S. (2001). DNA sequences from multiple amplifications reveal artifacts induced by cytosine deamination in ancient DNA. *Nucleic Acids Res.* 29, 4793–4799.
11. Stiller, M., Green, R.E., Ronan, M., Simons, J.F., Du, L., He, W., Egholm, M., Rothberg, J.M., Keates, S.G., Ovodov, N.D., et al. (2006). Patterns of nucleotide misincorporations during enzymatic amplification and direct large-scale sequencing of ancient DNA. *Proc. Natl. Acad. Sci. USA* 103, 13578–13584.
12. Gilbert, M.T., Binladen, J., Miller, W., Wiuf, C., Willerslev, E., Poinar, H., Carlson, J.E., Leebens-Mack, J.H., and Schuster, S.C. (2007). Recharacterization of ancient DNA miscoding lesions: insights in the era of sequencing-by-synthesis. *Nucleic Acids Res.* 35, 1–10.
13. Brotherton, P., Endicott, P., Sanchez, J.J., Beaumont, M., Barnett, R., Austin, J., and Cooper, A. (2007). Novel high-resolution characterization of ancient DNA reveals C > U-type base modification events as the sole cause of post mortem miscoding lesions. *Nucleic Acids Res.* 35, 5717–5728.
14. Briggs, A.W., Stenzel, U., Johnson, P.L., Green, R.E., Kelso, J., Prüfer, K., Meyer, M., Krause, J., Ronan, M.T., Lachmann, M., and Pääbo, S. (2007). Patterns of damage in genomic DNA sequences from a Neanderthal. *Proc. Natl. Acad. Sci. USA* 104, 14616–14621.
15. Underhill, P.A., Shen, P., Lin, A.A., Jin, L., Passarino, G., Yang, W.H., Kauffman, E., Bonnè-Tamir, B., Bertranpetit, J., Francalacci, P., et al. (2000). Y chromosome sequence variation and the history of human populations. *Nat. Genet.* 26, 358–361.
16. Thomson, R., Pritchard, J.K., Shen, P., Oefner, P.J., and Feldman, M.W. (2000). Recent common ancestry of human Y chromosomes: evidence from DNA sequence data. *Proc. Natl. Acad. Sci. USA* 97, 7360–7365.
17. Noonan, J.P., Coop, G., Kudaravalli, S., Smith, D., Krause, J., Alessi, J., Chen, F., Platt, D., Pääbo, S., Pritchard, J.K., and Rubin, E.M. (2006). Sequencing and analysis of Neanderthal genomic DNA. *Science* 314, 1113–1118.
18. Krause, J., Dear, P.H., Pollack, J.L., Slatkin, M., Spriggs, H., Barnes, I., Lister, A.M., Ebersberger, I., Pääbo, S., and Hofreiter, M. (2006). Multiplex amplification of the mammoth mitochondrial genome and the evolution of Elephantidae. *Nature* 439, 724–727.
19. Morin, P.A., Chambers, K.E., Boesch, C., and Vigilant, L. (2001). Quantitative polymerase chain reaction analysis of DNA from noninvasive samples for accurate microsatellite genotyping of wild chimpanzees (*Pan troglodytes verus*). *Mol. Ecol.* 10, 1835–1844.
20. Przeworski, M., Coop, G., and Wall, J.D. (2005). The signature of positive selection on standing genetic variation. *Evolution Int. J. Org. Evolution* 59, 2312–2323.
21. Kumar, S., Filipowski, A., Swarna, V., Walker, A., and Hedges, S.B. (2005). Placing confidence limits on the molecular age of the human-chimpanzee divergence. *Proc. Natl. Acad. Sci. USA* 102, 18842–18847.
22. Patterson, N., Richter, D.J., Gnerre, S., Lander, E.S., and Reich, D. (2006). Genetic evidence for complex speciation of humans and chimpanzees. *Nature* 441, 1103–1108.
23. Teshima, K.M., and Przeworski, M. (2006). Directional positive selection on an allele of arbitrary dominance. *Genetics* 172, 713–718.
24. Rohland, N., and Hofreiter, M. (2007). Comparison and optimization of ancient DNA extraction. *Biotechniques* 42, 343–352.