

# A Complete Neandertal Mitochondrial Genome Sequence Determined by High-Throughput Sequencing

Richard E. Green,<sup>1,\*</sup> Anna-Sapfo Malaspinas,<sup>2</sup> Johannes Krause,<sup>1</sup> Adrian W. Briggs,<sup>1</sup> Philip L.F. Johnson,<sup>3</sup> Caroline Uhler,<sup>4</sup> Matthias Meyer,<sup>1</sup> Jeffrey M. Good,<sup>1</sup> Tomislav Maricic,<sup>1</sup> Udo Stenzel,<sup>1</sup> Kay Prüfer,<sup>1</sup> Michael Siebauer,<sup>1</sup> Hernán A. Burbano,<sup>1</sup> Michael Ronan,<sup>5</sup> Jonathan M. Rothberg,<sup>6</sup> Michael Egholm,<sup>5</sup> Pavao Rudan,<sup>7</sup> Dejana Brajković,<sup>8</sup> Željko Kućan,<sup>7</sup> Ivan Gušić,<sup>7</sup> Mårten Wikström,<sup>9</sup> Liisa Laakkonen,<sup>10</sup> Janet Kelso,<sup>1</sup> Montgomery Slatkin,<sup>2</sup> and Svante Pääbo<sup>1</sup>

<sup>1</sup>Max-Planck Institute for Evolutionary Anthropology, D-04103 Leipzig, Germany

<sup>2</sup>Department of Integrative Biology

<sup>3</sup>Biophysics Graduate Group

<sup>4</sup>Department of Statistics

University of California, Berkeley, CA 94720, USA

<sup>5</sup>454 Life Sciences, Branford, CT 06405, USA

<sup>6</sup>The Rothberg Institute for Childhood Diseases, Guilford, CT 06437, USA

<sup>7</sup>Croatian Academy of Sciences and Arts, Zrinski trg 11, HR-10000 Zagreb, Croatia

<sup>8</sup>Croatian Academy of Sciences and Arts, Institute for Quaternary Paleontology and Geology, Ante Kovačića 5, HR-10000 Zagreb, Croatia

<sup>9</sup>Helsinki Bioenergetics Group, Program for Structural Biology and Biophysics, Institute of Biotechnology

<sup>10</sup>Division of Biochemistry, Department of Biological and Environmental Sciences, Faculty of Biosciences, University of Helsinki

FIN-00014 Helsinki, Finland

\*Correspondence: [green@eva.mpg.de](mailto:green@eva.mpg.de)

DOI 10.1016/j.cell.2008.06.021

## SUMMARY

A complete mitochondrial (mt) genome sequence was reconstructed from a 38,000 year-old Neandertal individual with 8341 mtDNA sequences identified among 4.8 Gb of DNA generated from ~0.3 g of bone. Analysis of the assembled sequence unequivocally establishes that the Neandertal mtDNA falls outside the variation of extant human mtDNAs, and allows an estimate of the divergence date between the two mtDNA lineages of  $660,000 \pm 140,000$  years. Of the 13 proteins encoded in the mtDNA, subunit 2 of cytochrome c oxidase of the mitochondrial electron transport chain has experienced the largest number of amino acid substitutions in human ancestors since the separation from Neandertals. There is evidence that purifying selection in the Neandertal mtDNA was reduced compared with other primate lineages, suggesting that the effective population size of Neandertals was small.

## INTRODUCTION

Although it is well established that Neandertals are the hominid form most closely related to present-day humans, their exact relationship with modern humans remains a topic of debate (Hublin and Pääbo, 2006; Soficaru et al., 2006; Harvati et al., 2007). Molecular genetic data first spoke to this issue in 1997, when a 379

base pair section of the hypervariable region I (*HVRI*) of the mitochondrial genome (mtDNA) was determined from the Neandertal-type specimen found in 1856 in Neander Valley, near Düsseldorf, Germany (Krings et al., 1997). Since then, a total of 15 complete or partial Neandertal *HVRI* sequences, as well as two *HVRII* sequences (Krings et al., 1999; Krings et al., 2000), have been described. Phylogenetic analyses of these suggest that Neandertal mtDNA falls outside the variation of modern human mtDNA. Since the mtDNA genome is maternally inherited without recombination, these results indicate that Neandertals made no lasting contribution to the modern human mtDNA gene pool (Krings et al., 1997; Currat and Excoffier, 2004; Serre et al., 2004).

High-throughput 454 sequencing techniques have recently been applied to ancient DNA (Green et al., 2006; Poinar et al., 2006; Stiller et al., 2006). These methods open new possibilities for the retrieval of ancient DNA that has hitherto relied either on the cloning of random molecules in bacteria (Higuchi et al., 1984; Pääbo, 1985; Noonan et al., 2005, 2006) or on the PCR amplification of individual DNA sequences of interest (Pääbo and Wilson, 1988; Pääbo et al., 2004). The main benefit of the 454 sequencing technique is the sheer volume of sequence data that make it practical to undertake genome-scale ancient DNA sequencing projects. This is particularly feasible for mitochondrial genomes (Gilbert et al., 2007), given their smaller size relative to the nuclear genome and their abundance in cells, where, typically, several hundred mtDNAs per nuclear genome exist.

The 454 sequence data from ancient DNA have also allowed an increased understanding of DNA diagenesis (i.e., how DNA

is modified during deposition in a burial context). In particular, they have allowed a quantitative model of how DNA degradation and chemical modification occurs, and how the effects of these processes interact with the molecular manipulations used to generate sequencing libraries (Briggs et al., 2007). Notably, although it was previously known that a high rate of cytosine deamination occurs in ancient DNA (Hofreiter et al., 2001), it has become clear that this is particularly prevalent in the ends of the ancient molecules, presumably because these are often single stranded (Briggs et al., 2007). Deamination of cytosine residues results in uracil residues that are read as thymine by DNA polymerases, leading to a high rate of C-to-T transitions. A high rate of G-to-A transitions observed near the 3' ends of sequence reads is thought to be caused by deaminated cytosine residues on the complementary strands used as templates during the fill-in reaction to create blunt ends when sequencing libraries are constructed (Briggs et al., 2007).

By 454 sequencing, we have generated 34.9-fold coverage of the Neandertal mtDNA genome from a Neandertal bone (Vindija bone 33.16) excavated in 1980 from Vindija Cave, Croatia (Malez and Ullrich, 1982). It has been dated to  $38,310 \pm 2130$  years before present (Serre et al., 2004). Previously, the mtDNA *HVRI* sequence of this bone has been determined (Serre et al., 2004), as well as 2414 bp of mtDNA sequences by 454 sequencing (Green et al., 2006). Here, we present its complete mtDNA sequence, as well as the insights it allows into recent human and Neandertal mtDNA evolution.

## RESULTS

### DNA Sequence Determination

Three DNA extracts, each from 100–200 mg of a Neandertal bone (Vindija 33.16) were prepared in our cleanroom facility where several precautions against DNA contamination are implemented (Experimental Procedures). These include complete separation from other parts of the laboratories, direct delivery of all equipment and reagents to the facility, positive pressure generated with filtered air that excludes particles larger than 0.2  $\mu\text{m}$ , and UV irradiation and bleach treatment of all surfaces. The bone surface was removed prior to extraction. However, the interior of bones is also often contaminated with modern human DNA, presumably due to past washing and other treatments of Neandertal bones. Thus, we analyzed each extract for contamination by extant human mtDNA by PCR with primers flanking positions in the *HVRI* that distinguish extant humans from Neandertals (Green et al., 2006), and amplify both types of mtDNA with similar efficiencies. Following amplification, we cloned the PCR product and sequenced 103–112 clones to determine the ratio of Neandertal to extant human mtDNA. The contamination rate in the three extracts ranged from 0%–0.9% (see Figure S1 available online).

From these extracts, we generated a total of nine 454 libraries in the cleanroom facility with 454 adapters with a Neandertal-specific sequence key that is unique to this project, and thus unequivocally identifies each sequence determined as derived from the extract of a Neandertal bone (Briggs et al., 2007). This allows detection of any contamination that may be introduced in subsequent handling and sequencing steps outside the clean-

room. To maximize the library and sequence yield, we incorporate two modifications to the standard 454 protocol that reduce the need to perform titration runs of libraries (Meyer et al., 2008) and allow more molecules to be retrieved during library preparation (Maricic and Pääbo, unpublished results). From these libraries, we generated a total of 39 million sequence reads by 147 runs on the GS FLX sequencing platform. Bases were called with the standard 454 signal threshold and filtering criteria with minor modifications tailored for short, ancient sequence reads (Meyer et al., 2008).

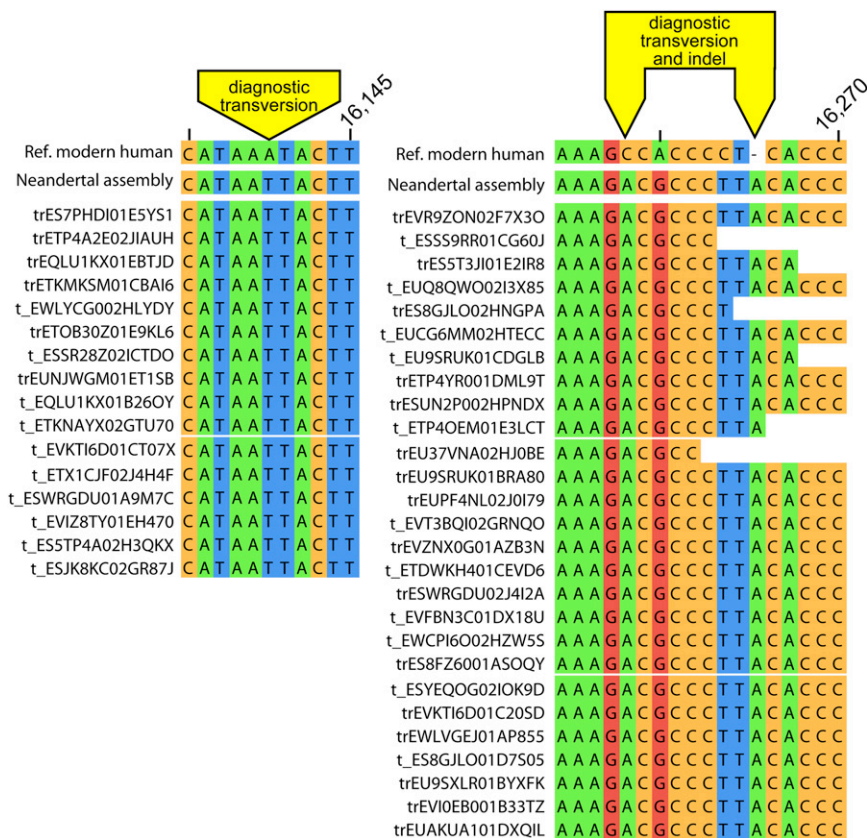
Neandertal sequences were identified within each run as described previously (Green et al., 2006), with the chief criterion being sequence similarity to a primate genome. mtDNA sequences were identified from these with further criteria (see Experimental Procedures), including similarity to the human mtDNA at least as great as to any nuclear DNA sequence. While the total fraction of sequences that are identified as Neandertal varied across libraries and library pools, the ratio of putative Neandertal nuclear DNA sequences to mtDNA sequences varied little among these libraries, and averaged 171. This corresponds to  $\sim 2100$  mtDNA molecules per cell. In total, 8341 mtDNA sequences were identified. They are of an average length of 69.4 bp (SD = 26.4), with the shortest fragment identified being 30 bp (limited by the length cut-off in the analysis pipeline), and the longest fragment being 278 bp (limited by the flow cycles performed on the GS FLX instrument).

### mtDNA Genome Assembly

Ancient DNA sequences present a challenge for DNA sequence assembly since they are typically short and exhibit high rates of nucleotide misincorporation. A further complication is that pyrosequencing (Ronaghi et al., 1998), for example, as performed on the GS FLX platform, calls long polymers of the same base with reduced accuracy. With these issues in mind, we designed an assembly procedure for ancient DNA. In short, each sequence identified as mtDNA was aligned over its entire length to the human reference mtDNA sequence (UCSC build hg18). These alignments were then merged, and each alignment column was examined to determine the majority base, yielding an assembled mtDNA sequence. Homopolymer lengths at positions where the reference human carries  $\geq 5$  identical bases were determined by analysis of the raw signal distributions as described in the Supplemental Data.

Following this examination, some problematic regions remained in the assembly. These include four regions of a total length of 20 bp, where no sequence coverage existed, eight other regions amounting to a total of 117 bp covered by only single reads, nine positions where no majority base existed due to low coverage, and 31 homopolymers for which the data were not sufficient to determine their length. These regions were amplified by PCR from a Vindija 33.16 bone extract in two-step multiplex PCRs (Krause et al., 2006), cloned, and sequenced by Sanger technology in order to complete the assembly (Experimental Procedures).

We then reapplied our mtDNA fragment detection pipeline to all Neandertal DNA sequences determined from the bone, but compared them to the assembled Neandertal mtDNA instead of the reference human mtDNA. This resulted in the detection



**Figure 1. Sequences Overlapping HVRI Positions Carrying Diagnostic Differences between Neandertal and Extant Humans**

All 43 sequences overlapping the three diagnostic positions are shown.

of an additional 721 mtDNA sequences that were initially missed, since they have a higher similarity to the human nuclear genome than human mtDNA, and thus could only be identified with the assembled Neandertal mtDNA. Interestingly, 522 of these sequences were similar to a single nuclear mtDNA insertion on chromosome 1 (hg18, position 554,327–560,165). In total, 8341 mtDNA fragments totaling 578,733 nt yielded a final assembly of 16,565 nt, where no position has less than 9-fold sequence coverage.

### Estimates of Contamination

In order to estimate the level of human contamination among the mtDNA sequences determined from these libraries, we realigned the individual sequences to the assembled mtDNA and used a number of differences between the putative Neandertal and human mtDNAs as diagnostic markers. We then examined each alignment to determine if it contained a position that allowed it to be classified as of either Neandertal or extant human origin.

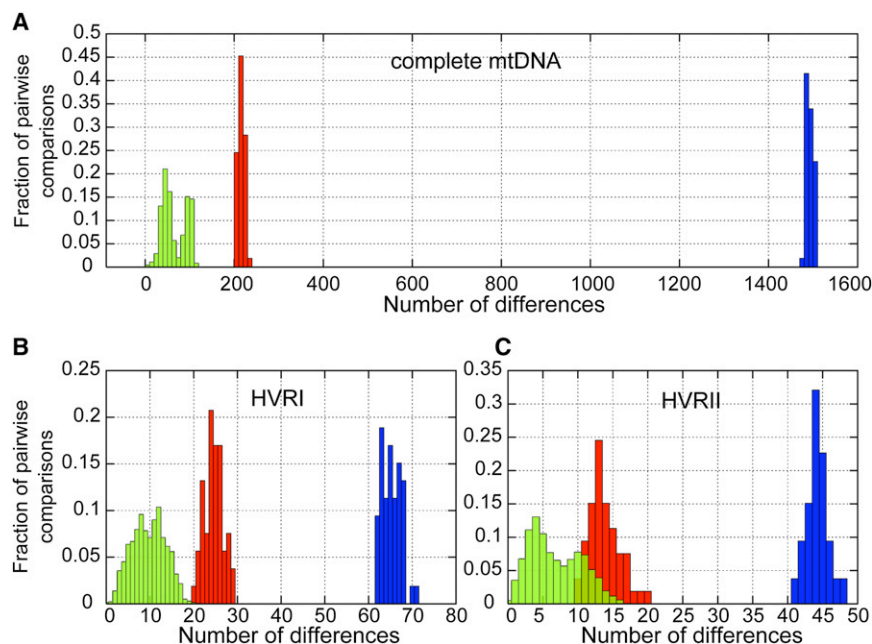
First, we used three human-Neandertal differences in the *HVRI*: an A-to-T and a C-to-A transversion, and an insertion of an A where all Neandertal sequences determined to date, including Vindija 33.16 (Serre et al., 2004), differ from all 1865 contemporary human *HVRI* sequences available in the mtDB database (Ingman and Gyllensten, 2006). Forty three sequences overlapped these *HVRI* positions, and all were of the Neandertal type (Figure 1). Second, we used four transversions between the putative Neandertal mtDNA and human mtDNAs that occur outside the *HVRI* and are neither adjacent to homopolymers nor poly-

morphic among the 1865 human mtDNA sequences. Among 192 sequences that overlapped these positions, 186 carried the base of the Neandertal assembly (Figure S2). Of the six sequences that did not, five were cases in which the observed base was T, the assembly base was C, and the modern human base was G or A. Thus, these are likely to be the result of deaminated cytosine residues that are common in ancient DNA (Hofreiter et al., 2001; Briggs et al., 2007; Brotherton et al., 2007). The alternative that these are contamination by a previously unknown modern human mtDNA sequence is unlikely because of the large number of modern human sequences available. One single sequence matches the human base at a diagnostic position. Interestingly, this sequence shows features commonly associated with ancient DNA (Figure S2) in that it begins with a C-to-T mismatch to the assembly, ends with two G-to-A mismatches, and is 65 nucleotides long, and thus relatively short. In spite of this, it may represent a contaminant, since modern DNA sequences retrieved from ancient specimens can carry such features (Sampietro et al., 2006). In order to estimate contamination levels, we thus disregard the five C-to-T mismatches as uninformative and count the last sequence as a contaminant. This yields a total of 229 out of 230 sequences carrying Neandertal diagnostic positions, and an estimate of the contamination rate of 0.4% with a 95% confidence interval (CI) of 0.01%–2.4%.

In order to maximize our power to detect contamination, we also used all positions in the Neandertal assembly that differ from extant humans and for which at least 99% of a panel of 311 worldwide modern human mtDNA sequences do not differ, ignoring positions that could be due to cytosine deamination. A total of 133 positions fulfilled these criteria and allowed 1963 fragments to be classified. Nine were of extant human origin, yielding a contamination estimate of 0.5% (95% CI = 0.21%–0.87%).

We conclude that so few of the mtDNA sequences determined derive from extant humans that they will not compromise the assembly, which has a 35-fold average coverage. The assembled mtDNA sequence, therefore, represents a reliable reconstruction of the mtDNA that this Neandertal individual carried when alive.

**MtDNA Sequence Analyses**  
Alignment of the 16,565 nt Neandertal mtDNA to the 16,568 nt human revised Cambridge reference mtDNA sequence (rCRS) (Andrews et al., 1999) revealed 206 differences (195 transitions



**Figure 2. Distribution of Pairwise mtDNA Sequence Differences**

Sequence differences among 53 humans (green), between humans and the Neandertal mtDNA (red), and between humans and chimpanzee (blue): (A) the complete mtDNA; (B) the HVRI (Neandertal position 16,044–16,411); and (C) the HVRII (position 57–372).

and 11 transversions). In the noncoding control region, the Neandertal sequence contains a deletion of four base pairs (CACA) at rCRS position 514 and a previously known insertion of one base pair (Serre et al., 2004) following position 16,263. The 13 protein-coding genes, the 22 tRNA genes, and two rRNA genes in the Neandertal mtDNA lack notable structural differences when compared to the human and chimpanzee mtDNAs (Table S2).

Figure 2A shows the distribution of pairwise sequence differences among 53 humans from around the world (Ingman et al., 2000), between these and the Neandertal, and between the modern humans and the chimpanzee mtDNAs. Among the humans, the number of differences ranges from 2 to 118, and is bimodal. The peak with a mode around 99 differences contains comparisons that involve at least one member of deep clades containing sub-Saharan African mtDNA. The peak with a mode around 44 differences involves comparisons between individuals outside these clades. By contrast, the number of differences between human mtDNAs and the Neandertal mtDNA ranges from 201 to 234, and is unimodal. Thus, the Neandertal mtDNA falls outside the variation of extant humans, even when nucleotide differences, uncorrected for multiple substitutions, are counted. Notably, when the comparisons are restricted to only the *HVRI* or *HVRII*, the distributions of pairwise differences among extant humans and between humans and the Neandertal overlap, illustrating that multiple substitutions complicate the reconstruction of mtDNA relationships when only these regions are studied (Figures 2B and 2C).

### MtDNA Divergence

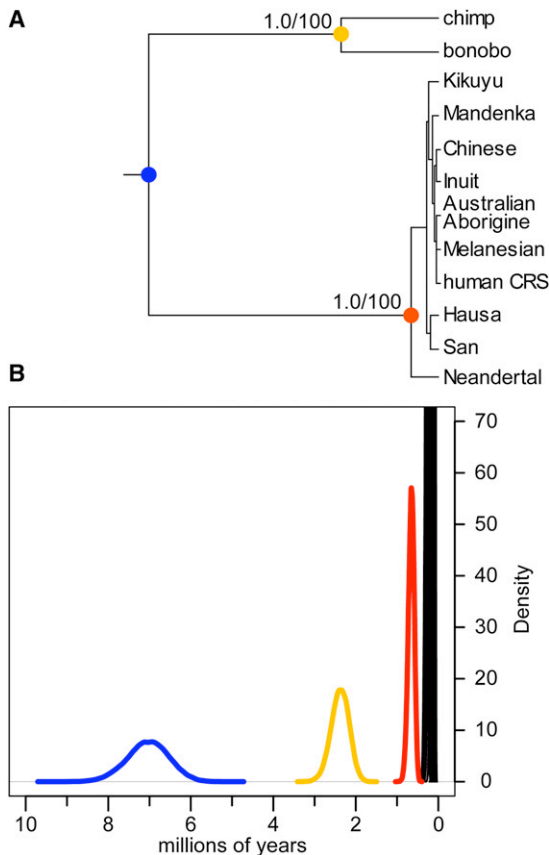
To further explore the evolutionary relationship between Neandertal and extant human mtDNAs, we estimated a phylogenetic tree with the complete mtDNA of Neandertal, 53 human mtDNAs, the human rCRS, and chimpanzee and bonobo mtDNAs as outgroups. Neighbor-joining, maximum likelihood,

parsimony, and a Bayesian approach (Supplemental Data) all indicated that the 54 extant humans formed a monophyletic group to the exclusion of the Neandertal, with complete support for each tree-building measure (100% bootstrap support; 1.0 posterior probability). Unsurprisingly, this topology is also found when analysis is restricted to various subsets of the sequence, such as protein-coding sequences and RNA-coding sequences (Supplemental Data).

In order to estimate the date of the divergence of the Neandertal and extant human mtDNAs, we compared the Neandertal mtDNA to 10 divergent human mtDNAs (Figure 3A), and tested if we could reject a molecular clock assuming the topology of the Bayesian tree (see Supplemental Data). This was not the case, indicating no significant heterogeneity in evolutionary rates among the sequences. To allow divergences of mtDNA sequences to be transformed to calendar years, we assumed—based on the fossil record—that humans and chimpanzees diverged between six (Galik et al., 2004) and eight million years ago (Brunet et al., 2002; Lebatard et al., 2008). This results in an estimate of the mean divergence time between Neandertal and extant human mtDNAs of 660,000, with a 95% credibility interval of 520,000–800,000 years ago (Figure 3B) (see also Supplemental Data), which overlaps with previous estimates based on *HVRI* and *I1* sequences (Kriings et al., 1997; Kriings et al., 1999). Since the width of the divergence credibility interval increases almost linearly with the posterior mean of the divergence estimate (Figure S5), further sequence data would be unlikely to decrease the width of the credibility interval (Yang and Rannala, 2006). However, if the estimated date of the divergence between humans and chimpanzees, or current assumptions about how the mtDNA evolves, were found to be incorrect, the estimates in calendar years of the divergence of the Neandertal and human mtDNAs would need to be revised.

### Mitochondrial Protein Evolution

Next, we estimated the substitutions that occurred in each of the 13 mtDNA protein-coding genes on the Neandertal and the human lineages with parsimony as an outgroup (Table 1). The total number of silent substitutions assigned to the Neandertal lineage, 44, is lower than the number assigned to the human lineage, 57. An apparent shortening of the Neandertal lineage relative to the human branch is also observed when all substitutions in the mtDNA are analyzed (Figure S3).



**Figure 3. Phylogenetic Tree and Divergence Time Estimate of mtDNA Sequences**

(A) Bayesian phylogenetic tree of complete mtDNA sequences of the Neanderthal, 10 extant humans, one chimpanzee, and one bonobo. Identical topologies for the Neanderthal and chimpanzee/bonobo split are produced by each tree-building method. The Bayesian posterior probability and the bootstrap support values are shown for two internal nodes.

(B) Posterior distribution of divergence times at each internal node using a 6–8 Mya for the ape/hominid divergence (blue node). The extant human divergences are shown in black, the Neanderthal/human divergence in red, the chimpanzee/bonobo divergence in yellow, and the ape/hominid in blue.

This dearth of Neanderthal lineage differences is not clearly concentrated in any particular region of the mtDNA (Figure S4A) or class of substitutions (Figures S4B and S4C), and is not statistically significant (Supplemental Data). It may reflect stochastic variation in the amount of substitutions on the two lineages, perhaps in combination with the fact that the Neanderthal mtDNA is ~38,000 years older than the human mtDNA sequence. However, for amino acid substitutions, we observe a similar number on the Neanderthal (20) and human (18) lineages. Thus, relative to silent substitutions, more amino acid replacements occurred on the Neanderthal lineage.

To explicitly test for an increase in amino acid replacements on the Neanderthal lineage, we evaluated the evolution of the 12 protein-coding genes (excluding *ND6*; see Supplemental Data) across seven primates (human, Neanderthal, chimpanzee, bonobo, gorilla, orangutan, and baboon) with a maximum likelihood framework. The ratio of amino acid replacements per replace-

**Table 1. Number of Synonymous and Nonsynonymous Substitutions in Each Protein-Coding mtDNA Gene Assigned to the Neanderthal or Extant Human Lineage by Parsimony with the Chimpanzee as an Outgroup**

Gene	Neanderthal		Extant Human	
	Synonymous	Nonsynonymous	Synonymous	Nonsynonymous
<i>ND1</i>	4	2	5	2
<i>ND2</i>	6	1	3	1
<i>COX1</i>	8	0	8	0
<i>COX2</i>	0	0	3	4
<i>ATP8</i>	2	1	3	0
<i>ATP6</i>	3	3	1	2
<i>COX3</i>	1	1	3	1
<i>ND3</i>	1	1	5	1
<i>ND4L</i>	3	1	1	0
<i>ND4</i>	5	0	9	0
<i>ND5</i>	7	5	6	4
<i>ND6</i>	1	1	1	0
<i>CYTB</i>	3	4	9	3
Total	44	20	57	18

ment site ( $d_N$ ) to silent substitutions per silent site ( $d_S$ ) can reveal the influence of natural selection on protein evolution with low  $d_N/d_S$  ( $\ll 1$ ) indicating purifying selection, and high  $d_N/d_S$  ( $> 1$ ) indicating positive selection. Overall,  $d_N/d_S$  was quite low across the entire phylogeny (0.051–0.087; Table S3), consistent with strong purifying selection on mitochondrial proteins in primates (Hasegawa et al., 1998). However,  $d_N/d_S$  varied significantly among branches in the phylogeny (all  $p < 0.0001$ ; Table S3). While lineage-specific estimates of  $d_N/d_S$  were sensitive to both the underlying model of sequence evolution and the number of outgroups considered (see Figure S6 and Table S3), we consistently observed the highest  $d_N/d_S$  on the Neanderthal lineage, and this value was significantly higher than background rates in all comparisons (all  $p < 0.01$ ; Table S3).

To determine if any of the 13 proteins encoded in the mtDNA show an unusual pattern of evolution in humans since the divergence from the Neanderthal, we contrasted the ratio of nucleotide polymorphisms in the 54 humans to fixed differences to Neanderthal at synonymous and nonsynonymous sites (Table 2). Under a standard neutral model, the ratio of diversity to divergence should be the same for both classes of sites (McDonald and Kreitman, 1991). Whereas this is the case for 12 of the protein-coding genes, *COX2* has an excess of amino acid divergence ( $p = 0.021$ ), consistent with the action of positive directional selection. However, when corrected for multiple tests of 13 genes, this yields only suggestive evidence for adaptive evolution of *COX2* (Bonferroni  $\alpha = 0.0038$ ).

*COX2* has experienced four amino acid substitutions on the human mtDNA lineage after its divergence from the Neanderthal lineage about 660,000 years ago. In order to see if any of these amino acids vary among humans today, we analyzed human mtDNA sequences in mtDB (Ingman and Gyllensten, 2006). At one of the four positions (rCRS position 7868), one of 2704 individuals carries the same base as the Neanderthal. Because

**Table 2. Polymorphism within Humans and Divergence to Neandertal at Synonymous and Nonsynonymous Sites in Protein-Coding mtDNA Genes**

Gene	Fixed Synonymous Differences	Human Synonymous polymorphic Sites	Fixed Nonsynonymous Differences	Human Nonsynonymous Polymorphic Sites	Neutrality Index	p Value, Uncorrected <sup>a</sup>
<i>ND1</i>	7	19	2	8	1.47	1
<i>ND2</i>	7	30	1	13	3.03	0.419
<i>COX1</i>	13	40	0	8	—	0.184
<i>COX2</i>	2	19	4	3	0.08	0.021
<i>ATP8</i>	4	6	1	2	1.33	1
<i>ATP6</i>	2	22	2	9	0.41	0.575
<i>COX3</i>	3	25	2	4	0.24	0.205
<i>ND3</i>	5	10	1	3	1.50	1
<i>ND4L</i>	2	9	1	1	0.22	0.423
<i>ND4</i>	7	43	0	12	—	0.328
<i>ND5</i>	11	51	7	23	0.71	0.58
<i>ND6</i>	1	20	0	5	—	1
<i>CYTB</i>	10	26	5	17	1.31	0.764

The neutrality index value (Rand and Kann, 1996) is the ratio of polymorphism to divergence at nonsynonymous sites versus this ratio at synonymous sites. Values smaller than 1 indicate an excess of nonsynonymous divergence. Values greater than 1 indicate excess nonsynonymous polymorphism within humans.

<sup>a</sup> Fisher's exact test.

mtDNA is inherited without recombination, and because the Neandertal mtDNA falls outside the variation of modern human mtDNA, this single modern human observation represents a reversion to the ancestral state seen in Neandertals and chimpanzees. Thus, these four amino acid substitutions occurred in the relatively short period after the divergence of Neandertal and extant human mtDNAs and before the most recent common ancestor of current human mtDNAs. The observation of four nonsynonymous substitutions on the modern human lineage, and no amino acid changes on the Neandertal lineage, stands in contrast to the overall trend of more nonsynonymous evolution in Neandertal protein-coding genes (Table 1), and deserves consideration.

### Subunit 2 of Cytochrome c Oxidase

COX2 encodes subunit 2 of the cytochrome c oxidase, or complex IV, of the inner mitochondrial membrane. Complex IV catalyzes the reduction of molecular oxygen to water with electrons from cytochrome c and protons from within the mitochondrial matrix. By acting as a proton pump (Belevich et al., 2006), it maintains a proton gradient across the mitochondrial inner membrane, which drives the phosphorylation of ADP to ATP (Mitchell, 1966; Reid et al., 1966). It has previously been noted that genes encoding various components of the electron transport chain evolve quickly in primates (Doan et al., 2004; Grossman et al., 2004).

To determine what functional effects, if any, these amino acid substitutions may have, we examined the crystal structure of the bovine cytochrome c oxidase complex (Figure 4) (Muramoto et al., 2007). All four substitutions are spatially far from functionally critical parts of the protein, such as the cytochrome c binding area and the bimetallic Cu<sub>A</sub> center that passes electrons from cytochrome c to COX1. However, three of them interact with the

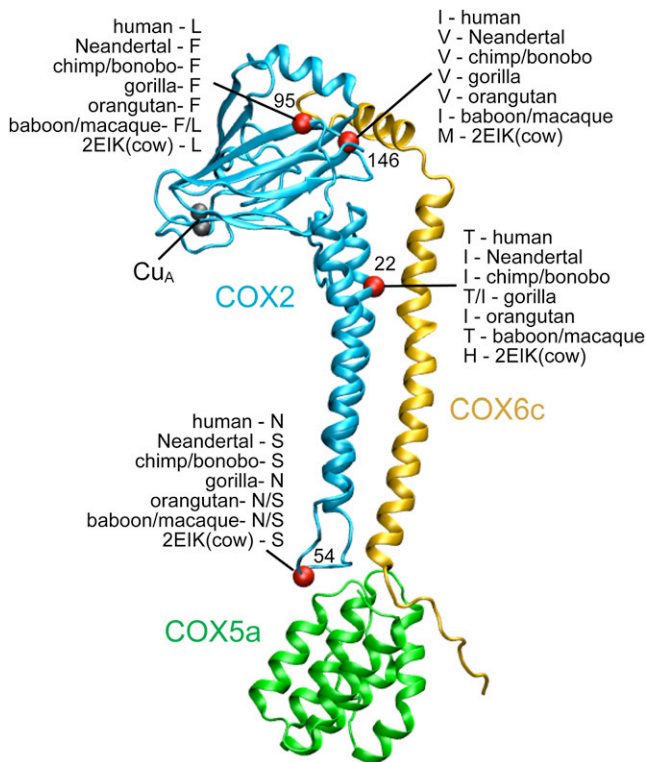
nuclear-encoded COX5a and COX6c subunits, which are thought to perform regulatory functions. It has been noted that COX5a evolves fast in primates (Uddin et al., 2008). However, the two COX5a sites that differ between humans and chimpanzees are not in close proximity to COX2. It may also be of relevance that two of the four amino acid residues that are specific to modern humans relative to Neandertals are variable among apes, and all four are variable when Old World monkeys are considered (Figure 4).

### DNA Diagenesis

The 8341 mtDNA sequences comprising the Neandertal mtDNA assembly allow us to gauge a number of factors of relevance to the state of preservation of mtDNA in this late Pleistocene individual. These estimates have an advantage over previous ones (e.g., Briggs et al., 2007) that were based on inferences from alignments to human and chimpanzee genome sequences, and were thus potentially affected by alignment uncertainties or uncertain inferences about evolutionary processes.

One feature of interest is that the retrieval of DNA across the mtDNA genome is more uneven than expected under a model of random sampling (Lander and Waterman, 1988) (Figure 5A). The variation in DNA retrieval is largely explained by a positive correlation ( $r = 0.49$ ) between GC content and the number of recovered fragments (Figure 5C). Furthermore, for shorter fragments, GC content is negatively correlated with length (Figure 5B). This may be due to denaturation of short, AT-rich fragments during library preparation.

Because the GS FLX read length of about 250 nucleotides is much longer than the average sequence length of 69 nucleotides, fragmentation points can be inferred for nearly all sequences. As previously reported (Briggs et al., 2007), purines



**Figure 4. COX2 Protein Sequence Differences between Neandertal and Modern Humans in Structural Context**

The amino acid positions of the four differences are shown in red. The copper center of COX2 (Cu<sub>A</sub>) is also shown. The amino acid at each position in some primate and the cow sequences (from which the structure [PDB identifier: 2EIK] is derived) are indicated.

are overrepresented before, and pyrimidines after, fragmentation points (Table S4). This is consistent with depurination in ancient DNA, which would induce strand breaks. When dinucleotides made up of the bases immediately 5' and 3' of fragmentation points are analyzed (Table S4), the most overrepresented bases are G 5' of breaks and T 3' of breaks, where breaks occur ~2.9 times more often than would be expected if fragmentation was random, and ~2.1 times more often than predicted by the individual contributions of G and T separately. Further investigations are necessary to understand the biochemical basis for this.

When we compare the observed rates of nucleotide misincorporations to maximum likelihood estimates for the expected numbers of misincorporations based on a model that we previously developed (Briggs et al., 2007), we find the model fits the overall data well (Figure S7). However, longer fragments have higher misincorporation rates (and, by inference, higher deamination rates), and shorter fragments have lower misincorporation rates than predicted by the model. This could be due to a reduced power to detect and align fragments that are both short and contain nucleotide misincorporations. Alternatively, it could represent a difference between shorter and longer fragments, for example, if longer fragments have, on average, longer single-stranded overhanging ends.

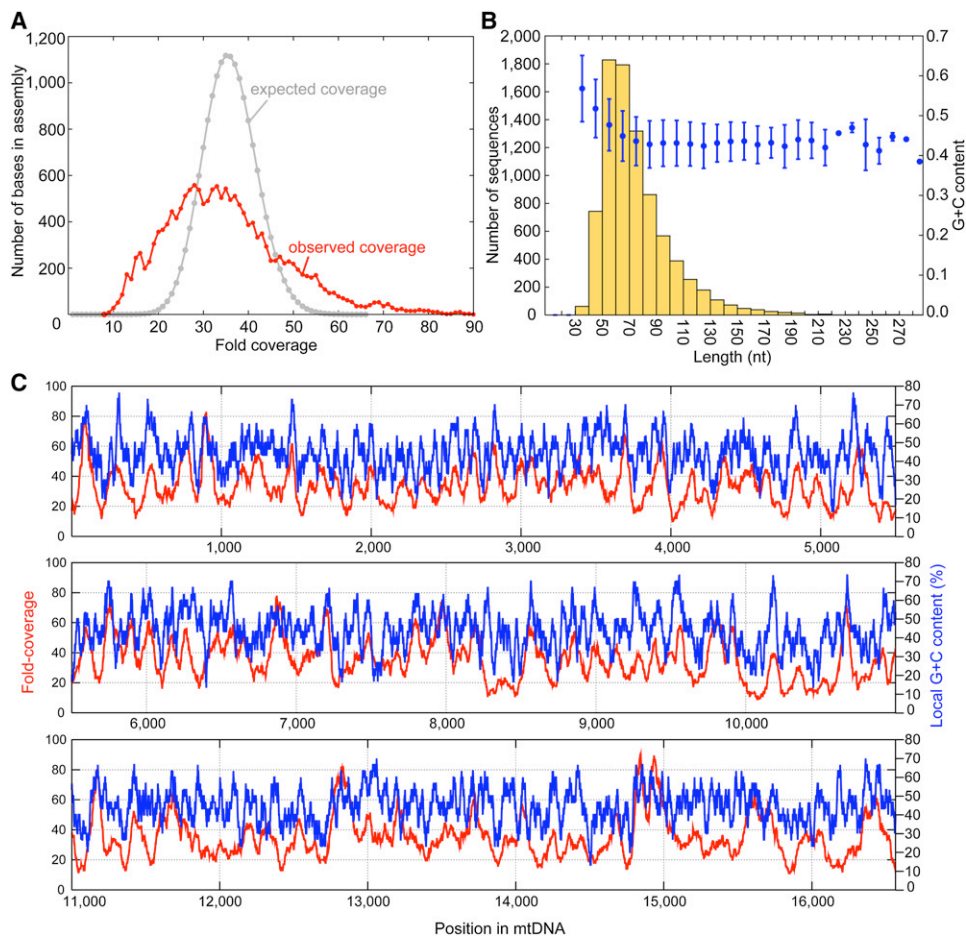
## DISCUSSION

### Neandertal Genetic History

The complete Neandertal mtDNA genome confirms and extends previous insights into the genetic history of Neandertals. First, it confirms that the Neandertal mtDNA falls outside the variation of extant human mtDNA. Second, it shows that the Neandertal mtDNA diverged from the extant human mtDNA lineage on the order of 660,000 years ago. Thus, the most recent common ancestor of human and Neandertal mtDNA lived more than two or three times as long ago as the most recent common ancestor of extant human mtDNAs.

A striking observation from the analysis of the 13 protein-coding genes in the mtDNA is that the ratio of nonsynonymous to synonymous evolutionary rates is significantly higher on the Neandertal lineage. One plausible explanation is that Neandertals had a smaller effective population size, and thus less effective purifying selection than humans. Previous studies have reported that mtDNA  $d_N/d_S$  ratios tend to be higher within than between species of great apes, including extant humans (Nachman et al., 1996; Templeton, 1996; Hasegawa et al., 1998; Kivisild et al., 2006). These results suggest that slightly deleterious amino acid variants segregate within populations, and that differences in the intensity of purifying selection may affect mtDNA  $d_N/d_S$  ratios. Previous estimates based on mean pairwise differences (MPD) within the mtDNA *HVRI* suggested that Neandertals (MPD = 5.5) had an effective population size similar to that of modern Europeans (MPD = 4.0) or Asians (MPD = 6.3), but lower than that of modern Africans (MPD = 8.1) (Krause et al., 2007b). Recent population genetic analyses have revealed a higher mtDNA amino acid substitution rate (Elson et al., 2004) and relatively more deleterious autosomal nuclear variants (Lohmueller et al., 2008) in Europeans than in Africans, presumably due to the smaller effective population size of Europeans. Thus, it seems plausible that Neandertals had a long-term effective population size smaller than that of modern humans. Population reductions caused by recurrent glaciations in Eurasia during Neandertals' ~400,000 years of existence may have contributed to this. Further work will reveal if a small effective population size is compatible with the extent of nucleotide diversity seen in the Neandertal nuclear genome.

A tantalizing finding is that, when all substitutions in the mtDNA are analyzed, the Neandertal mtDNA lineage is shorter than the human lineage by approximately 20% (Figure S3B). It is tempting to speculate that this is due to the fact that the Neandertal mtDNA is ~38,000 years older than the extant human mtDNA sequences to which it is compared. However, under the assumption that the evolutionary dates are reasonably accurate, the reduction in length is about three times as large as would be expected if it was entirely due to the age of the fossil. It should be noted that, for small evolutionary distances such as these, there is a large stochastic component to phylogenetic branch length. Thus, although the evolutionary dates are clearly dependent on many tenuous assumptions, it seems reasonable to assume that the majority of the discrepancy in length between the Neandertal and extant human mtDNA lineages is due to stochastic differences in the amounts of substitutions that have come to fixation on the two lineages.



**Figure 5. Sequence Coverage and Base Composition of the Neandertal mtDNA**

(A) The expected (gray) and observed (red) distribution of sequence depths in the Neandertal mtDNA assembly at 34.9-fold over all coverage. (B) The length distribution of sequences (yellow) and, for each length bin, the mean  $\pm$  SD of GC content (blue). (C) G + C content within a sliding window 30 bp 5' and 3' of each position (blue) and the observed coverage (red) at each position.

Another interesting observation is that COX2 stands out among proteins encoded in the mitochondrial genome as having experienced four amino acid substitutions on the modern human mtDNA lineage. Further work is warranted to elucidate the functional consequences of these amino acid substitutions. However, all these substitutions are in regions of the protein that, based on the crystal structure, do not have any obvious function, and they are variable among primates. Hence, they may represent either minor adaptive advantages, perhaps of regulatory relevance, or have no significant functional consequences for mitochondrial function. Unless other evidence for their importance becomes available, we see no need to invoke positive selection to account for the evolution of COX2 on the human lineage.

#### DNA Sequence Authenticity

Two features of ancient DNA in general, and Neandertal DNA in particular, may cause errors in nucleotide sequences in excess of what is expected from modern DNA. First, chemical modification in the ancient DNA may cause nucleotide misincorporations, and, second, the unintended presence in the experiments

of DNA from extant humans may be mistaken for Neandertal DNA. In the case of the Neandertal mtDNA genome sequences presented here, high average coverage of the random sequence reads in combination with amplification and sequencing of positions where coverage is low, or where longer nucleotide homopolymers may cause base calling problems, make us confident that the error rates from both these sources are low. Thus, the amount of data now accumulated for the Neandertal mtDNA allows extrapolations to what will be required to arrive at a reasonably reliable nuclear DNA sequences from the Neandertal.

With respect to nucleotide misincorporations, the current data show that errors causing a transition occur at positions carrying either a C or a G base in  $\sim 2.7\%$  of sequence reads. By contrast, all other errors occur in less than 0.1% of reads (Figure S8). Thus, C-to-T and G-to-A misincorporations are the predominant sequencing errors that need to be considered in terms of an ancient genome assembly. Since these misincorporations are drastically accumulated toward the 5' and 3' ends of molecules (Figure S7), respectively, and since there is a tendency for strand breaks in

the ancient DNA to occur 3' of purines and, in particular, at GT dinucleotides (Table S4), this might conceivably result in an accumulation of misincorporations at certain nucleotide positions, where even multiple retrieval of the same sequence may then not always allow the correct base to be determined. However, since GT dinucleotides cause only an ~3-fold increase in the likelihood of a break compared with random positions in the genome, already rather low sequence coverage makes this effect unlikely to cause errors in the assembled sequence. In fact, although it has been argued that C-to-T misincorporations are accumulated at particular sites in ancient mtDNA molecules (Gilbert et al., 2003), no positions in the mtDNA assembly have a mixture of observed bases that would confuse assembly (Figure S9), and higher relative amounts of mismatches correlate with lower sequence coverage. This suggests that stochastic events among few recovered sequences rather than hotspots for DNA modifications are the basis of these observations. For example, the least supported position of the Neandertal mtDNA assembly (position 5476) is supported by eight C observations and three T observations (Figure S9). Thus, we conclude that nucleotide misincorporations will not impede the determination of reliable Pleistocene genome sequences, provided that reasonable sequence coverage is achieved.

Contamination with extant human DNA is the other dominant source of erroneous Neandertal sequences. Given the high coverage and the fact that the best estimate of the contamination rate here is 0.5% (with an upper 95% confidence limit of 0.87%), we do not expect contamination to affect the mtDNA sequence assembly to any appreciable level. Under the assumption that the Neandertal mtDNA sequence is reliable, it is a useful tool for gauging contamination when sequencing the Neandertal nuclear genome. Previously, assays to determine contamination within Neandertal fossil extracts were limited to the *HVRI*, which carry few positions where extant humans differ from Neandertals. By contrast, the complete Neandertal mtDNA now offers 133 such positions. This enables a reliable estimation of mtDNA contamination by analyzing sequence reads from 454 libraries, rather than by PCR-based assays of the DNA extracts. For example, when we do this in a small preliminary data set initially published from this fossil (Green et al., 2006), 10 of 10 sequences are classified as Neandertal. However, in further unpublished sequencing runs from that library, 8 out of 75 diagnostic sequences derive from extant human mtDNA, suggesting a contamination rate of ~11% (CI = 4.7%–20%). This is in agreement with the suggestion (Wall and Kim, 2007) that contamination occurred in that experiment. That library was constructed outside our cleanroom facility and before the introduction of the Neandertal-specific key, which is crucial for the detection of contamination by other 454 libraries, and was therefore not used for the subsequent Neandertal genome sequencing project (Briggs et al., 2007). However, with the help of the mtDNA presented here, such levels of contamination are now easily detectable from 454 sequencing runs. Nevertheless, since the ratios of nuclear to mtDNA may differ between endogenous and contaminating DNA, it is of importance to identify nuclear DNA sequences diagnostic for Neandertals. Data collected from Vindija bone 33.16 now begin to make this possible (Krause et al., 2007a).

### Outlook for Nuclear Genome Sequencing

Irrespective of the fact that the amino acid substitutions in the human COX2 protein may not be of functional significance, they illustrate the power of Neandertal DNA sequences for finding accumulations of recent evolutionary changes in human genes. A genome-wide analysis of human and Neandertal genes holds the promise to identify many such cases for nuclear genes. However, whereas the mtDNAs of extant humans are monophyletic with respect to Neandertals, this is not expected to be the case for most nuclear genes (Pääbo, 1999). This fact will be helpful for the identification of recent positive selection in humans. If, for example, a completed selective sweep affected a gene in human ancestors after their separation from their common ancestors with Neandertals, extant human diversity in that region of the genome would coalesce to the exclusion of Neandertal alleles. Consequently, the combination of the accumulation of recent human substitutions in coding genes or conserved sequence elements along with reduced human diversity that excludes Neandertals may be a hallmark for genes and genomic regions that have been important in the emergence of fully modern humans.

Technically, the Neandertal mtDNA presented here is a useful forerunner for the sequencing of the Neandertal nuclear genome. First, the mtDNA assembly and the sequences of which it is comprised can be used to tune the parameters of models of DNA diagenesis. These parameters, in turn, can then be used to properly detect and align ancient Neandertal DNA sequence to the human genome. This will be particularly crucial before multifold coverage of the Neandertal nuclear genome is achieved. Second, the high coverage allows a first tentative estimate of how deeply a Neandertal genome would need to be sequenced in order to arrive at a reasonable accuracy of the sequence, given the biased retrieval of sequences (Figure 5), the misincorporation patterns (Figure S8), and other features of the DNA extracted from Neandertal bones. By resampling our sequences to simulate different depths of coverage, we find that, to achieve an error-rate of 1 in 10,000, 12-fold coverage would be required (Figure S10). Although all factors that effect ancient DNA error rates are unlikely to be identical between mitochondrial and nuclear DNA (e.g., homopolymer lengths and DNA methylation), this lends us confidence that a reliable Neandertal genome sequence will be achievable.

### EXPERIMENTAL PROCEDURES

#### DNA Extraction

DNA was extracted as previously described (Serre et al., 2004). After the first DNA extraction from 100–200 mg of bone, a second DNA extraction (“re-extract”) was carried out from the remaining bone pellet, as previously described (Rohland and Hofreiter, 2007a, 2007b), except that incubation was overnight at 56°C. One extract of the Vindija 33.16 bone (Vi80 [33.16] 120402; Figure S1) as well as the two re-extracts (Vi80 [33.16] 120402re, Vi80 [33.16] 031201re; Figure S1) were used for 454 library preparation and high-throughput sequencing. For PCR verification of the initial sequence assembly, DNA was extracted from 90 mg of the bone, as previously described (Rohland and Hofreiter, 2007b).

#### Assembly of the mtDNA Sequence

MtDNA sequences were identified with *megablast* (Zhang et al., 2000) and was required to be  $\geq 30$  nucleotides long and show  $\geq 90\%$  identity to the reference human mtDNA (GI: 17981852) or a version of this sequence carrying the Vindija

*HVRI*. mtDNA alignments, which had to be at least as high scoring as the best alignment against the human nuclear genome, were merged, and regions where the assembly was problematic were amplified by PCR. While nuclear insertions of mtDNA could, in theory, confound the assembly, this is unlikely for several reasons (see [Supplemental Data](#)). The Neandertal mtDNA sequence is available under accession number AM948965.

### Evolutionary Analyses

We aligned Neandertal, human, ape, and monkey mtDNAs (identifiers in [Supplemental Data](#)), and estimated trees by using neighbor-joining, maximum parsimony, maximum-likelihood, and Bayesian approaches. Support for nodes was assessed with bootstrap replicates, and a likelihood ratio test was used to test the clock assumption. In order to date mtDNA divergences, we estimated the posterior distribution of divergence times assuming that chimpanzees and humans diverged 6–8 million years ago. For protein-coding analyses, we concatenated 12 genes, excluding *ND6* and regions with overlapping reading frames and codons with insertion-deletion variation, and performed *dN/dS* analyses using an ML framework (Yang, 2007).

### ACCESSION NUMBERS

The Neandertal mtDNA sequence has been deposited in the EMBL database with accession number AM948965.

### SUPPLEMENTAL DATA

Supplemental Data include Supplemental Experimental Procedures, ten figures, and five tables and are available with this article online at <http://www.cell.com/cgi/content/full/134/3/416/DC1/>.

### ACKNOWLEDGMENTS

We thank the Croatian Academy of Sciences and Arts, the Berlin-Brandenburg Academy of Sciences, and the Presidential Innovation Fund of the Max Plank Society for making this work possible. We are indebted to Mark Stoneking, Weiwei Zhai, John Huelsenbeck, Jérôme Fuchs, Mario Mörl, Nick Patterson, David Reich, and Timothy White for helpful discussions. We also thank Bruce Rannala and Ziheng Yang for help with *mcmctree*, James Hudson Bullard for help with figures, and Christine Green for editing. A.-S.M., P.L.F.J., and M.S. are supported by NIH grant R01-GM40282. P.L.F.J. is also supported by a Chang-Lin Tien scholarship. M.E. is an employee of 454 Life Sciences, a Roche company.

Received: April 16, 2008

Revised: May 29, 2008

Accepted: June 13, 2008

Published: August 7, 2008

### REFERENCES

Andrews, R.M., Kubacka, I., Chinnery, P.F., Lightowlers, R.N., Turnbull, D.M., and Howell, N. (1999). Reanalysis and revision of the Cambridge reference sequence for human mitochondrial DNA. *Nat. Genet.* 23, 147.

Belevich, I., Verkhovskiy, M.I., and Wikstrom, M. (2006). Proton-coupled electron transfer drives the proton pump of cytochrome c oxidase. *Nature* 440, 829–832.

Briggs, A.W., Stenzel, U., Johnson, P.L., Green, R.E., Kelso, J., Pruffer, K., Meyer, M., Krause, J., Ronan, M.T., Lachmann, M., and Pääbo, S. (2007). Patterns of damage in genomic DNA sequences from a Neandertal. *Proc. Natl. Acad. Sci. USA* 104, 14616–14621.

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.

Brunet, M., Guy, F., Pilbeam, D., Mackaye, H.T., Likius, A., Ahounta, D., Beauvilain, A., Blondel, C., Bocherens, H., Boisserie, J.R., et al. (2002). A new hominid from the Upper Miocene of Chad, Central Africa. *Nature* 418, 145–151.

Currat, M., and Excoffier, L. (2004). Modern humans did not admix with Neanderthals during their range expansion into Europe. *PLoS Biol.* 2, e421. 10.1371/journal.pbio.0020421.

Doan, J.W., Schmidt, T.R., Wildman, D.E., Uddin, M., Goldberg, A., Huttemann, M., Goodman, M., Weiss, M.L., and Grossman, L.I. (2004). Coadaptive evolution in cytochrome c oxidase: 9 of 13 subunits show accelerated rates of nonsynonymous substitution in anthropoid primates. *Mol. Phylogenet. Evol.* 33, 944–950.

Elson, J.L., Turnbull, D.M., and Howell, N. (2004). Comparative genomics and the evolution of human mitochondrial DNA: assessing the effects of selection. *Am. J. Hum. Genet.* 74, 229–238.

Galik, K., Senut, B., Pickford, M., Gommery, D., Treil, J., Kuperavage, A.J., and Eckhardt, R.B. (2004). External and internal morphology of the BAR 1002'00 Orrorin tugenensis femur. *Science* 305, 1450–1453.

Gilbert, M.T., Tomsho, L.P., Rendulic, S., Packard, M., Drautz, D.I., Sher, A., Tikhonov, A., Dalen, L., Kuznetsova, T., Kosintsev, P., et al. (2007). Whole-genome shotgun sequencing of mitochondria from ancient hair shafts. *Science* 317, 1927–1930.

Gilbert, M.T., Willerslev, E., Hansen, A.J., Barnes, I., Rudbeck, L., Lynnerup, N., and Cooper, A. (2003). Distribution patterns of postmortem damage in human mitochondrial DNA. *Am. J. Hum. Genet.* 72, 32–47.

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.

Grossman, L.I., Wildman, D.E., Schmidt, T.R., and Goodman, M. (2004). Accelerated evolution of the electron transport chain in anthropoid primates. *Trends Genet.* 20, 578–585.

Harvati, K., Gunz, P., and Grigorescu, D. (2007). Cioclovina (Romania): affinities of an early modern European. *J. Hum. Evol.* 53, 732–746.

Hasegawa, M., Cao, Y., and Yang, Z. (1998). Preponderance of slightly deleterious polymorphism in mitochondrial DNA: nonsynonymous/synonymous rate ratio is much higher within species than between species. *Mol. Biol. Evol.* 15, 1499–1505.

Higuchi, R., Bowman, B., Freiberger, M., Ryder, O.A., and Wilson, A.C. (1984). DNA sequences from the quagga, an extinct member of the horse family. *Nature* 312, 282–284.

Hofreiter, M., Jaenicke, V., Serre, D., Haeseler Av, 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.

Hublin, J.J., and Pääbo, S. (2006). Neandertals. *Curr. Biol.* 16, R113–R114.

Ingman, M., and Gyllensten, U. (2006). mtDB: Human Mitochondrial Genome Database, a resource for population genetics and medical sciences. *Nucleic Acids Res.* 34, D749–D751.

Ingman, M., Kaessmann, H., Pääbo, S., and Gyllensten, U. (2000). Mitochondrial genome variation and the origin of modern humans. *Nature* 408, 708–713.

Kivisild, T., Shen, P., Wall, D.P., Do, B., Sung, R., Davis, K., Passarino, G., Underhill, P.A., Scharfe, C., Torroni, A., et al. (2006). The role of selection in the evolution of human mitochondrial genomes. *Genetics* 172, 373–387.

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.

Krause, J., Lalueza-Fox, C., Orlando, L., Enard, W., Green, R.E., Burbano, H.A., Hublin, J.J., Hanni, C., Fortea, J., de la Rasilla, M., et al. (2007a). The derived FOXP2 variant of modern humans was shared with Neandertals. *Curr. Biol.* 17, 1908–1912.

Krause, J., Orlando, L., Serre, D., Viola, B., Pruffer, K., Richards, M.P., Hublin, J.J., Hanni, C., Derevianko, A.P., and Pääbo, S. (2007b). Neandertals in central Asia and Siberia. *Nature* 449, 902–904.

- Krings, M., Capelli, C., Tschentscher, F., Geisert, H., Meyer, S., von Haeseler, A., Grossschmidt, K., Possnert, G., Paunovic, M., and Pääbo, S. (2000). A view of Neandertal genetic diversity. *Nat. Genet.* *26*, 144–146.
- Krings, M., Geisert, H., Schmitz, R.W., Krainitzki, H., and Pääbo, S. (1999). DNA sequence of the mitochondrial hypervariable region II from the neandertal type specimen. *Proc. Natl. Acad. Sci. USA* *96*, 5581–5585.
- Krings, M., Stone, A., Schmitz, R.W., Krainitzki, H., Stoneking, M., and Pääbo, S. (1997). Neandertal DNA sequences and the origin of modern humans. *Cell* *90*, 19–30.
- Lander, E.S., and Waterman, M.S. (1988). Genomic mapping by fingerprinting random clones: a mathematical analysis. *Genomics* *2*, 231–239.
- Lebataard, A.E., Bourles, D.L., Düringer, P., Jolivet, M., Braucher, R., Carcaillet, J., Schuster, M., Arnaud, N., Monie, P., Lihoreau, F., et al. (2008). Cosmogenic nuclide dating of *Sahelanthropus tchadensis* and *Australopithecus bahrelghazali*: Mio-Pliocene hominids from Chad. *Proc. Natl. Acad. Sci. USA* *105*, 3226–3231.
- Lohmueller, K.E., Indap, A.R., Schmidt, S., Boyko, A.R., Hernandez, R.D., Hubisz, M.J., Sninsky, J.J., White, T.J., Sunyaev, S.R., Nielsen, R., et al. (2008). Proportionally more deleterious genetic variation in European than in African populations. *Nature* *451*, 994–997.
- Malez, M., and Ullrich, H. (1982). Neuere pläoanthropologische untersuchungen am material aus der höhle vindija. *Palaeontologia Jugoslavia* *29*, 1–44.
- McDonald, J.H., and Kreitman, M. (1991). Adaptive protein evolution at the Adh locus in *Drosophila*. *Nature* *351*, 652–654.
- Meyer, M., Briggs, A.W., Maricic, T., Hober, B., Hoffner, B., Krause, J., Weihmann, A., Pääbo, S., and Hofreiter, M. (2008). From micrograms to picograms: quantitative PCR reduces the material demands of high-throughput sequencing. *Nucleic Acids Res.* *36*, e5. 10.1093/nar/gkm1095.
- Mitchell, P. (1966). Chemiosmotic coupling in oxidative and photosynthetic phosphorylation. *Biol. Rev. Camb. Philos. Soc.* *41*, 445–502.
- Muramoto, K., Hirata, K., Shinzawa-Itoh, K., Yoko-o, S., Yamashita, E., Aoyama, H., Tsukihara, T., and Yoshikawa, S. (2007). A histidine residue acting as a controlling site for dioxygen reduction and proton pumping by cytochrome c oxidase. *Proc. Natl. Acad. Sci. USA* *104*, 7881–7886.
- Nachman, M.W., Brown, W.M., Stoneking, M., and Aquadro, C.F. (1996). Non-neutral mitochondrial DNA variation in humans and chimpanzees. *Genetics* *142*, 953–963.
- Noonan, J.P., Coop, G., Kudravalli, 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.
- Noonan, J.P., Hofreiter, M., Smith, D., Priest, J.R., Rohland, N., Rabeder, G., Krause, J., Dettler, J.C., Pääbo, S., and Rubin, E.M. (2005). Genomic sequencing of Pleistocene cave bears. *Science* *309*, 597–599.
- Pääbo, S. (1985). Molecular cloning of ancient Egyptian mummy DNA. *Nature* *314*, 644–645.
- Pääbo, S. (1999). Human evolution. *Trends Cell Biol.* *9*, M13–M16.
- Pääbo, S., Poinar, H., Serre, D., Jaenicke-Despres, V., Hebler, J., Rohland, N., Kuch, M., Krause, J., Vigilant, L., and Hofreiter, M. (2004). Genetic analyses from ancient DNA. *Annu. Rev. Genet.* *38*, 645–679.
- Pääbo, S., and Wilson, A.C. (1988). Polymerase chain reaction reveals cloning artefacts. *Nature* *334*, 387–388.
- Poinar, H.N., Schwarz, C., Qi, J., Shapiro, B., Macphee, R.D., Buigues, B., Tikhonov, A., Huson, D.H., Tomsho, L.P., Auch, A., et al. (2006). Metagenomics to paleogenomics: large-scale sequencing of mammoth DNA. *Science* *311*, 392–394.
- Rand, D.M., and Kann, L.M. (1996). Excess amino acid polymorphism in mitochondrial DNA: contrasts among genes from *Drosophila*, mice, and humans. *Mol. Biol. Evol.* *13*, 735–748.
- Reid, R.A., Moyle, J., and Mitchell, P. (1966). Synthesis of adenosine triphosphate by a protonmotive force in rat liver mitochondria. *Nature* *212*, 257–258.
- Rohland, N., and Hofreiter, M. (2007a). Ancient DNA extraction from bones and teeth. *Nat. Protoc.* *2*, 1756–1762.
- Rohland, N., and Hofreiter, M. (2007b). Comparison and optimization of ancient DNA extraction. *Biotechniques* *42*, 343–352.
- Ronaghi, M., Uhlen, M., and Nyren, P. (1998). A sequencing method based on real-time pyrophosphate. *Science* *281*, 363–365.
- Sampietro, M.L., Gilbert, M.T., Lao, O., Caramelli, D., Lari, M., Bertranpetit, J., and Lalueza-Fox, C. (2006). Tracking down human contamination in ancient human teeth. *Mol. Biol. Evol.* *23*, 1801–1807.
- Serre, D., Langaney, A., Chech, M., Teschler-Nicola, M., Paunovic, M., Mennecier, P., Hofreiter, M., Possnert, G., and Pääbo, S. (2004). No evidence of Neandertal mtDNA contribution to early modern humans. *PLoS Biol.* *2*, e57. 10.1371/journal.pbio.0020057.
- Soficaru, A., Dobos, A., and Trinkaus, E. (2006). Early modern humans from the Pesteria Muierii, Baia de Fier, Romania. *Proc. Natl. Acad. Sci. USA* *103*, 17196–17201.
- 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.
- Templeton, A.R. (1996). Contingency tests of neutrality using intra/interspecific gene trees: the rejection of neutrality for the evolution of the mitochondrial cytochrome oxidase II gene in the hominoid primates. *Genetics* *144*, 1263–1270.
- Uddin, M., Opazo, J.C., Wildman, D.E., Sherwood, C.C., Hof, P.R., Goodman, M., and Grossman, L.I. (2008). Molecular evolution of the cytochrome c oxidase subunit 5A gene in primates. *BMC Evol. Biol.* *8*, 8. 10.1186/1471-2148-8-8.
- Wall, J.D., and Kim, S.K. (2007). Inconsistencies in Neanderthal genomic DNA sequences. *PLoS Genet.* *3*, 1862–1866.
- Yang, Z. (2007). PAML: Phylogenetic analysis by maximum likelihood. *Mol. Biol. Evol.* *24*, 1586–1591.
- Yang, Z., and Rannala, B. (2006). Bayesian estimation of species divergence times under a molecular clock using multiple fossil calibrations with soft bounds. *Mol. Biol. Evol.* *23*, 212–226.
- Zhang, Z., Schwartz, S., Wagner, L., and Miller, W. (2000). A greedy algorithm for aligning DNA sequences. *J. Comput. Biol.* *7*, 203–214.