

The American Journal of Human Genetics, Volume 91

Supplemental Data

**A Haplotype at *STAT2* Introgressed from Neanderthals
and Serves as a Candidate of Positive Selection**

in Papua New Guinea

Fernando L. Mendez, Joseph C. Watkins, and Michael F. Hammer

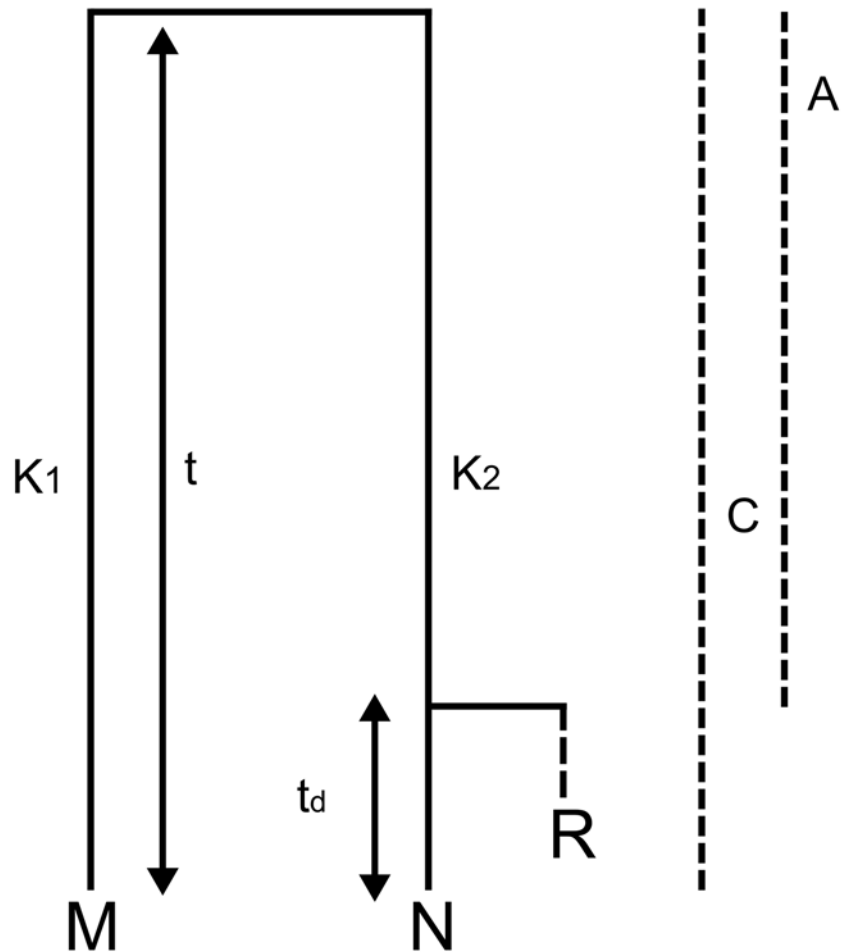


Figure S1. Representation of three lineages for which DNA sequence is available: The lineages **M** and **N** have complete sequence available, whereas the lineage **R** has incomplete sequence coverage. All mutations are polarized using the sequence of an outgroup. K_1 and K_2 mutations have occurred in the **M** and **N** lineages, respectively. The **R** lineage has sequence coverage for C of the K_2 mutations and shares the derived state at A of them. The lineages **M** and **N** diverged at time t in the past, whereas the lineage **R** branched from the lineage **N** at time t_d in the past.

Figure S2

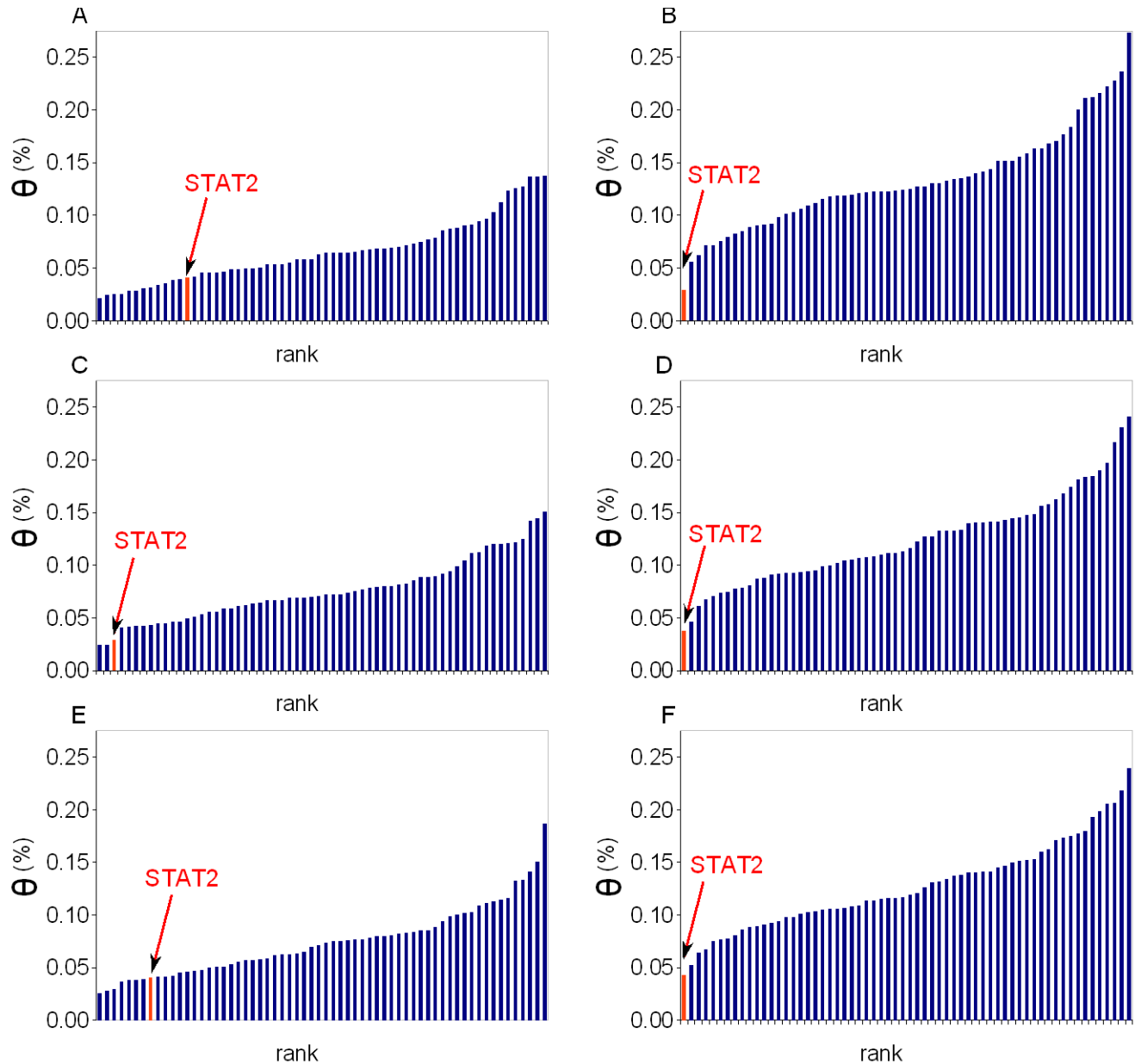


Figure S2. Empirical distribution of the value of θ (%) per site for 61 intergenic loci (blue) and the parts *STAT2* resequenced here (orange) in the populations of the resequence panel: (a) Biaka, (b) Mandenka, (c) San, (d) Papuans, (e) Han, and (f) French Basque populations.

Figure S3

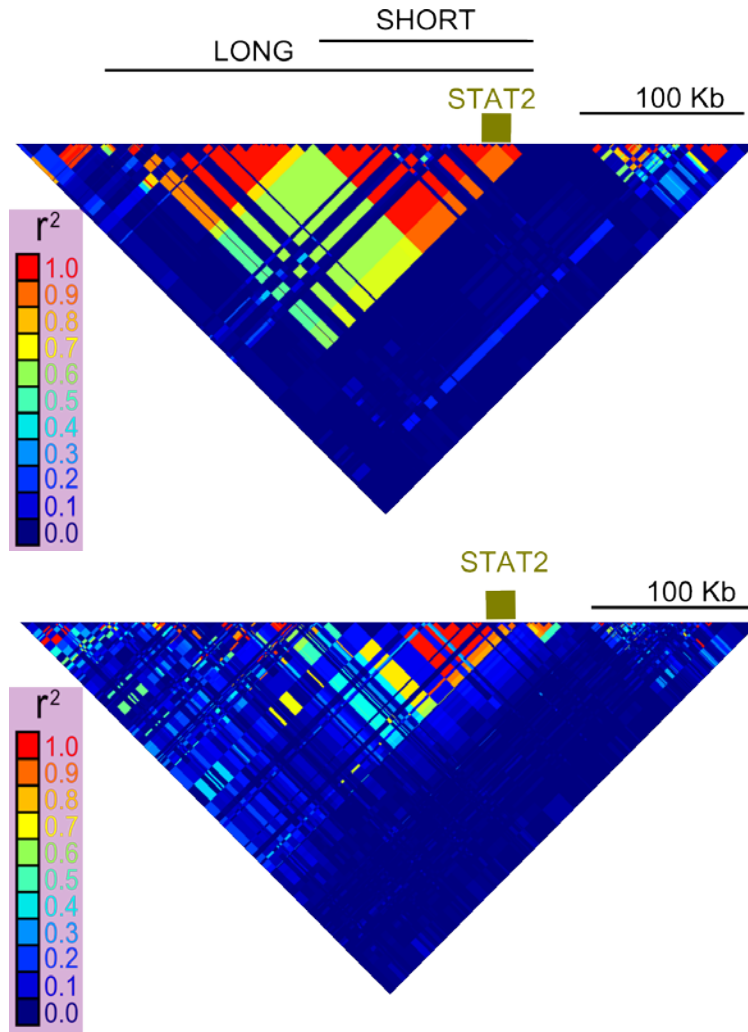


Figure S3. Comparison of pairwise linkage disequilibrium measured with r^2 for Japanese (upper) and Yoruba (lower) HapMap phase III populations for the sequence spanning from 54780000 to 55281000. Horizontal lines indicate the blocks in **Figure 1**.

Figure S4

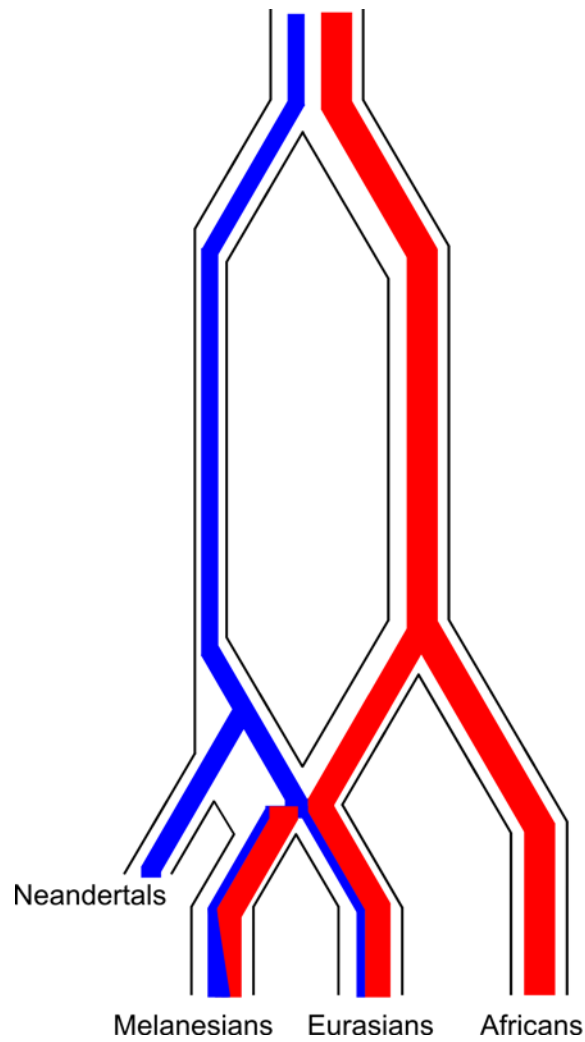


Figure S4. Schematic model for the trajectory of the alleles at *STAT2*. The red and blue lines represent lineages that evolved in the ancestors of AMH and Neanderthals, respectively. The Neanderthal lineage introgressed in modern humans after the out of Africa event, remaining originally at low frequency. Subsequently, the frequency of the introgressive lineage increased in Melanesians due to natural selection.

Table S1a. Coordinates of amplicons sequenced for 90 individuals of 6 populations (hg18)

Segment	Beginning	End	Length (bp)	Distance to the next amplicon (bp)
STAT2_A	55021597	55022185	589	984
STAT2_B	55023170	55024241	1072	1877
STAT2_C	55026119	55027026	908	1546
STAT2_D	55028573	55031762	3190	2650
STAT2_E	55034413	55036756	2344	3152
STAT2_F	55039909	55040411	503	

Table S1b. Primer information for amplicons sequenced in this work

Amplicon	Upper Name	Upper Seq	Lower Name	Lower Seq
STAT2_21466	STAT2U21466	TCTAACGGACAGGAGGGGCTAC	STAT2L23114	TCAGGGCAGTGGCTTCTTTC
STAT2_23089	STAT2U23089	AAAATGTTGAAATCCTTCCATACT	STAT2L24423	TAACCACCCTACTGCCTCTC
STAT2_26016	STAT2U26016	GGAGACTGGAATAGGGCAGC	STAT2L27037	AGGGTGTGGGGCGGTGTC
STAT2_28475	STAT2U28475	ACCTGTCTTTGCTCTCTCTCCC	STAT2L28860	TCTGTCTGTGGCTATCTGTCCC
STAT2_28749	STAT2U28749	CAGTATTTGGAATGTGGGCTA	STAT2L29409	GGCATAAGAAGGGGATAAACA
STAT2_29214	STAT2U29214	CTGAAGGTTTGGGCTGAGCA	STAT2L29782	GAACAGGGGTTGGGGGAAGT
STAT2_29638	STAT2U29638	ACCTTATTGCTGCCCTTTCC	STAT2L30956	TTGACCACTTTCCCCTATCC
STAT2_30514	STAT2U30514	GGGCAAAGAGTATGGAAG	STAT2L31822	ATCATCTGTAATGTGGGGAA
STAT2_34324	STAT2U34324	TGAAGGAGACTGGGCTCTG	STAT2L35461	GTGGTAGGAGTTAGGGTTGACAG
STAT2_35313	STAT2U35313	AATGTGCTTCCTACCTTTGG	STAT2L36582	CAGAATCTTGACAGCCCCTT
STAT2_36070	STAT2U36070	CACATTTGTTCCCGTCTCCC	STAT2L36848	ACTCAACTGGTAGCACTCCCTG
STAT2_39797	STAT2U39797	ACATCGTTTTTTTGGTGAAGA	STAT2L40437	GACTCCTCCCCTCGTAGAA

Table S2. SNP array information used to assess the frequency of the N lineage

Array	SNPs derived in Neandertal
Illumina 650	rs744051
Affy 250K Nsp	rs17118439 ^a
Affy 250K Sty	rs2371454
	rs871130
	rs744051
	rs2306694
	rs111718065

^a rs17118439 is found in different backgrounds in sub-Saharan Africans and non-Africans.

This SNP was used as diagnostic of the N lineage in non-Africans when Nsp was the only array genotyped in the population.

Table S3. Data from table 1 including NA19026 and additional sequence data (see text).

	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	
	4	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	
	7	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
	9	2	2	2	2	2	2	2	2	2	2	2	2	2	3	3	3	3	3	3	3	3	3	4	4	
	9	2	3	3	4	6	6	6	6	6	7	9	9	9	0	0	0	0	0	5	5	5	6	0	1	
	4	1	5	7	2	1	2	6	9	9	4	2	6	9	3	5	5	6	7	7	7	8	4	0	3	
	1	1	2	3	4	2	8	4	4	6	9	6	3	6	5	0	4	8	1	6	6	5	7	8	2	
	2	4	2	3	0	8	4	7	9	7	5	4	4	5	8	2	3	9	2	1	2	9	1	9	5	
	C	G	C	G	G	C	A	C	G	C	A	T	C	T	A	A	G	G	T	A	T	A	C	C	T	A
Neandertal	A	*	.	.	.	-	*	*	*	.	.	A	C	.	.	.	T	*	.	.
Denisova	A	.	.	*	.	.	.	*	T	*
N	A	-	-	A	C	.	.	.	T	A	.	.	.
D	*	*	.	.	.	T	*
S ^a	.	.	T	T	.	.	G	C	G
Mb-1	A	.	.	.	C	.	G	.	T	.	.	C
Mb-2	.	.	.	A	A	.	.	.	C	.	G	.	T	.	.	C
Mb-3	A	.	.	.	C	.	G	.	T	.	C	.	C
NA19026	C	.	G	G
Ma-1	C	.	G	G
Ma-2	C	.	G	.	.	C	G
Ma-3	C	.	G	.	.	C	G	G
Ma-4	.	A	C	.	G	.	.	C	G	G
Ma-5	C	T	G	.	.	C	G	G
Ma-6	C	.	G	C	.	C	G	G
Ma-7	C	.	G	.	.	C	G	G	G
Rec N-Mb	N	.	.	.	A	.	.	.	C	.	N	A	C	.	.	.	T	A	.	N	
Rec S-M	N	.	T	T	.	.	N	N

*, low sequenc coverage (ancient DNA) or missing data (lineage D)

^a: information for the positions shaded was obtained using KB1

Mandanka	HGDP00908	C	C	R	G
Mandanka	HGDP00911	A	.	.	.	C .	.	.	T .	.	.	C
Mandanka	HGDP00912	R	.	.	.	C .	.	.	Y Y .	Y .	.	S	K
Mandanka	HGDP00913	Y	C .	.	.	C .	C	G
Mandanka	HGDP00914	K	C	C	R	G
Mandanka	HGDP00915	R	.	.	.	C .	.	.	Y Y .	Y .	.	S	K	
Mandanka	HGDP01199	C	C	G	G	
Mandanka	HGDP01200	C	C	G	
Mandanka	HGDP01202	C	C	G	
Mandanka	HGDP01283	R	.	.	.	C .	.	.	Y .	Y .	.	S	K	
Mandanka	HGDP01284	.	.	W	.	A	.	.	.	C .	.	.	T .	.	.	C	
Mandanka	HGDP01286	C .	.	.	Y .	C	R .	.	.	Y .	.	G	
San	GM3043	.	Y	Y	.	S	Y	Y R	K	
San	JR013	.	Y	Y	.	S	Y	Y R	K	
San	JR020	.	Y	Y	.	S .	.	Y .	Y	Y R	K		
San	JR054	C .	.	S	Y .	.	.	S .	.	.	R	
San	JR077	.	R	C	Y .	Y	K	
San	JR301	.	Y	Y	.	S	Y	Y R	K	
San	JR305	C .	.	Y .	C	K	
San	JR321	.	A	C	C	G	
San	JR323	.	Y	Y	.	S	Y	Y R	K	
San	JR354	.	T	T	Y R	
Papuan	NG04	.	.	.	R	N	R	N	.	S .	.	Y .	.	S .	R	Y .	Y	.	Y .	M	
Papuan	NG06	.	.	.	R	N	.	N	.	S .	.	Y .	.	S .	R	Y .	.	Y .	M	
Papuan	NG07	.	.	.	R	N	.	N	.	S .	.	Y .	.	S .	R	Y .	.	Y .	M	
Papuan	NG13	A	C .	.	T .	A	
Papuan	NG14	A	C .	.	T .	A	
Papuan	NG15	.	.	.	R	N	.	N	.	S .	.	Y .	.	S .	R	Y .	.	Y .	M	
Papuan	NG17	N	.	N	W	.	R	Y .	Y .	M	
Papuan	NG18	A	C .	.	T .	A	
Papuan	NG20	.	.	.	R	.	.	.	S .	.	Y .	.	W	S	
Papuan	NG22	N	.	N	W	.	R	Y .	Y .	M	
Papuan	NG25	A	C .	.	T .	A	

Papuan	NG26	.	.	.	A	.	.	.	C	.	.	T	.	.	C
Papuan	NG29	.	.	.	R	N	N	.	S	.	.	Y	.	.	S	R	Y	.	W
Papuan	NG30	.	.	.	R	N	N	.	S	.	R	Y	.	.	S	R	Y	.	.
Papuan	NG34	A	C	.	.	T
Papuan	NG51	.	.	.	R	N	N	.	S	R	R	Y	.	.	S	R	Y	.	Y
Han	HGDP00774	.	.	.	A	.	.	.	C	.	.	T	.	.	C	R	.	.	.
Han	HGDP00775	.	.	.	A	.	.	.	C	.	.	T	.	.	C	.	R	.	.
Han	HGDP00777	.	.	.	A	.	.	.	C	.	.	T	.	.	C
Han	HGDP00778	.	.	.	A	.	.	.	C	.	.	T	.	.	C
Han	HGDP00780	.	.	.	A	.	.	.	C	.	.	T	.	.	C
Han	HGDP00785	.	.	.	A	.	.	.	C	.	.	T	.	.	C
Han	HGDP00786	.	.	.	A	.	W	.	C	.	.	T	.	.	C
Han	HGDP00815	.	.	.	A	.	.	.	C	.	.	T	.	.	C
Han	HGDP00819	.	.	.	A	.	.	.	C	.	.	T	.	.	C
Han	HGDP00977	.	.	.	A	.	.	.	C	.	.	T	.	.	C
Han	HGDP01288	.	.	.	A	.	.	.	C	.	.	T	.	.	C
Han	HGDP01290	.	.	.	A	.	.	.	C	.	.	T	.	.	C
Han	HGDP01293	.	.	.	A	.	.	.	C	.	.	T	.	.	C
Han	HGDP01294	.	S	.	A	.	.	.	C	.	.	Y	.	.	S	R	Y	.	Y
Han	HGDP01295	.	.	.	A	.	.	.	C	.	.	T	.	.	C
Han	HGDP01296	.	.	.	A	.	.	.	C	.	.	T	.	.	C
Basque	HGDP01357	.	.	.	A	.	.	.	C	.	.	T	.	.	C
Basque	HGDP01358	.	.	.	A	.	.	.	C	.	.	T	.	.	C
Basque	HGDP01359	.	.	.	R	N	N	.	S	.	.	Y	.	.	S	R	Y	.	Y
Basque	HGDP01360	.	.	.	R	A	.	.	C	.	.	T	.	.	C
Basque	HGDP01361	.	.	.	A	.	.	.	C	K	.	T	.	.	C
Basque	HGDP01362	.	.	M	R	A	.	.	C	.	.	T	.	.	C	.	.	.	R
Basque	HGDP01364	.	.	.	A	.	.	.	C	.	.	T	M	.	C
Basque	HGDP01370	.	.	.	A	.	.	.	C	.	.	T	.	.	C
Basque	HGDP01371	.	.	.	R	N	N	.	S	.	.	Y	.	.	S	R	Y	.	Y
Basque	HGDP01372	.	.	.	A	.	.	.	C	.	.	T	M	.	C
Basque	HGDP01374	.	.	.	A	.	.	.	C	.	.	T	.	.	C
Basque	HGDP01375	.	.	.	R	A	.	.	C	.	.	T	M	.	C

Basque	HGDP01376	.	.	.	R	A	C	.	.	.	T	C		
Basque	HGDP01377	A	C	.	.	.	T	C	S	.	.		
Basque	HGDP01378	R	N	.	N	.	S	.	.	.	Y	S	R	Y	Y	.	M	.	.
Basque	HGDP01379	A	C	.	.	.	T	C		

^a the ancestral state was inferred using chimpanzee

^b colors are used to represent function of the polymorphisms: yellow (non-synonymous), brown (synonymous) and green (UTR)

N, heterozygous indel; -, homozygous indel; * no sequence coverage

Table S5. Analysis of sites between 54770000 and 54913000 that differ between NA18956 and the reference sequence

SNPs present in NA18956	State in HGDP00542 ^{a,b,c}	State in Neandertal ^{a,b,c}	State in Denisova ^{a,b,c}	State in Reference ^a	Sequence coverage in Chimpanzee
54771772	1(3)[60]	1(5)[254]	0 (1)[37]	0	YES
54779342	1(3)[60]	-1	1(3)[37]	0	YES
54781290	1(6)[60]	-1	0(7)[37]	0	YES
54784285	0(2)[37]	0(1,1)[151,170]	0(1)[37]	1	YES
54786849	0(5)[60]	1(3)[254]	1(1)[37]	1	YES
54787283	0(1)[60]	-1	-1	1	YES
54791507	0(4)[60]	-1	0(6)[37]	1	YES
54791929	0(3)[60]	1(1)[19]	1(2)[0]	1	YES
54792771	1(4)[60]	1(2)[254]	0(2)[37]	0	YES
54793758	0(5)[60]	0(1)[21]	-1	1	YES
54794205	1(2)[60]	1(3,1)[254]	1(3)[37]	0	YES
54794749	0(2)[60]	0(2)[254]	0(2)[37]	0	YES
54795088	0(5)[60]	0(2)[254]	0(2)[37]	1	YES
54795802	1(5)[60]	1(3)[254]	0(1)[37]	0	YES
54796112	0(2)[60]	0(2,1)[55,13]	-1	1	YES
54796575	0(3)[60]	0(1)[254]	0(2)[37]	1	YES
54796904	0(3)[60]	0(3)[254]	0(1)[37]	1	YES
54797619	0(1)[37]	-1	0(11)[0]	0	YES
54798525	1(1,1)[37,37]	1(2)[254]	1(4)[37]	0	YES
54798855	0(3)[60]	-1	0(5)[37]	1	YES
54798938	0(4)[60]	0(3)[254]	0(4)[37]	1	YES
54799058	1(4)[60]	1(1)[254]	1(2)[37]	0	YES
54799412	1(4)[60]	1(3)[254]	1(2)[37]	0	YES
54799972	0(4)[60]	0(1)[157]	0(3)[37]	1	YES
54802423	0(6)[60]	0(4)[254]	0(10)[37]	1	YES
54802836	1(5)[60]	1(1)[165]	0(3)[37]	0	YES
54804669	1(6)[60]	-1	-1	0	YES
54804675	0(6)[60]	0(1)[254]	-1	0	YES

54806067	1(1)[37]	-1	0(8)[0]	0	YES
54806086	0(1)[37]	-1	0(2,4)[0,0]	1	YES
54806099	0(1)[37]	-1	0(2,1)[0,0]	1	YES
54806359	0(7)[60]	-1	0(5)[37]	1	YES
54808625	0(6)[60]	0(1)[254]	0(2)[37]	0	YES
54809324	1(3)[60]	-1	1(1)[37]	0	YES
54810243	0(5)[60]	-1	0(5,1)[37,0]	1	YES
54815898	0(3)[37]	-1	0(1,1)[23,0]	1	YES
54817021	0(3)[37]	2(1)[254]	0(1)[37]	1	YES
54817040	1(3)[60]	-1	1(1)[37]	0	YES
54817421	0(9)[60]	0(3)[254]	0(1)[37]	1	YES
54818387	1(8)[60]	1(2)[254]	0(3)[37]	0	YES
54818985	1(4)[60]	1(4)[254]	1(3)[37]	0	YES
54821519	0(1)[60]	-1	0(3,1)[37,0]	1	YES
54825834	-1	0(1)[20]	1(2)[0]	1	YES
54825865	0(1)[37]	0(2)[20]	0(1)[0]	0	YES
54825880	0(1)[37]	1(5)[20]	0(1,3)[37,0]	1	YES
54825998				-1	NO
54826091				-1	NO
54826861	0(10)[60]	-1	0(4)[37]	1	YES
54826902	0(7)[60]	0(1)[254]	0(4)[37]	0	YES
54827331	0(7)[60]	0(3,1)[153,133]	0(5)[37]	0	YES
54829254	0(3)[60]	-1	0(4)[37]	1	YES
54829965	0(4)[60]	0(1)[254]	-1	1	YES
54831174	1(2)[60]	-1	0(10)[37]	0	YES
54833077	1(3)[60]	1(3)[254]	0(3)[37]	0	YES
54834417	1(10)[60]	-1	1(1)[37]	0	YES
54835343	1(6)[60]	1(3)[254]	0(2)[37]	0	YES
54839093	1(8)[60]	1(3)[254]	0(2)[37]	0	YES
54839218	1(5)[60]	1(1)[254]	0(4)[37]	0	YES
54840207	1(10)[60]	-1	0(4)[37]	0	YES
54842900	1(15)[60]	1(2)[254]	1(3)[37]	0	YES
54846601				-1	NO

54847508	0(3)[60]	0(1)[5]	0(4)[37]	1	YES
54848757	0(3)[60]	-1	0(1)[37]	1	YES
54849129	1(3)[60]	0(1)[19]	0(2)[0]	0	YES
54851078	1(5)[60]	1(1)[254]	1(4)[37]	1	YES
54851607	0(4)[60]	0(4)[254]	0(3,1)[37,23]	1	YES
54851659	0(3)[60]	0(5)[254]	0(3)[37]	0	YES
54855674	1(6)[60]	-1	1(1)[37]	0	YES
54855891	0(8)[60]	-1	0(1)[37]	1	YES
54855905	0(8)[60]	-1	-1	1	YES
54857604	0(5)[60]	-1	0(2)[37]	0	YES
54862009	0(4)[60]	0(1)[254]	0(6)[37]	1	YES
54862562	1(4)[60]	-1	-1	0	YES
54863461	1(3)[37]	0(1)[6]	0(5)[37]	0	YES
54864585	0(6)[60]	0(1)[254]	0(1)[37]	1	YES
54865290	1(7)[60]	-1	1(3,2)[37,0]	0	YES
54867835	0(3)[60]	0(2)[215]	0(2)[37]	1	YES
54868459	1(3)[60]	1(1)[254]	0(1)[37]	0	YES
54870091	1(3)[60]	-1	1(1)[37]	1	YES
54870352	1(4)[60]	-1	1(3)[37]	0	YES
54872419	0(3)[60]	-1	0(2)[37]	0	YES
54872516	0(7)[60]	0(4)[254]	0(7)[37]	0	YES
54872565	1(5)[60]	1(1)[254]	1(4)[37]	1	YES
54872711	0(4)[60]	0(3)[254]	0(3)[37]	0	YES
54872732	1(6)[60]	1(1)[254]	0(4)[37]	0	YES
54873263	1(2)[37]	1(1)[254]	0(1)[37]	0	YES
54873843	1(5)[60]	-1	1(2)[37]	1	YES
54874720				-1	NO
54875251				-1	NO
54875494				-1	NO
54876082				-1	NO
54876735	1(1)[37]	-1	0(9)[0]	0	YES
54876817	0(1)[37]	-1	0(4)[0]	0	YES
54877170	1(2)[60]	-1	-1	0	YES

54878246	0(4)[60]	0(3)[76]	0(1)[37]	0	YES
54878600	0(2)[60]	-1	0(3)[37]	1	YES
54879298	0(4)[60]	0(2)[254]	0(3)[37]	0	YES
54880167	0(4)[60]	0(1)[249]	-1	1	YES
54880208	0(1)[60]	0(3)[163]	0(4)[37]	0	YES
54880433	0(2)[60]	0(1)[8]	1(1)[0]	1	YES
54880766				-1	NO
54880785				-1	NO
54881278	0(2)[60]	-1	0(2)[37]	0	YES
54882258	0(2)[60]	-1	-1	1	YES
54883246	0(6,3)[60,60]	0(1)[254]	0(1)[37]	0	YES
54883451	0(5)[37]	-1	0(1)[37]	1	YES
54883755	0(2,1)[60,37]	-1	-1	0	YES
54883787	0(3)[60]	-1	-1	1	YES
54883850	0(1)[60]	-1	0(11)[37]	0	YES
54884311	0(2,4)[60,60]	0(1)[254]	0(3)[37]	0	YES
54887504	0(1,1)[60,37]	-1	1(4)[37]	1	YES
54888006	0(3)[37]	-1	0(2)[37]	1	YES
54888776	1(4,2)[60,50]	-1	0(1)[0]	0	YES
54889163	1(1)[60]	0(1)[0]	0(2)[37]	0	YES
54889333	1(4,5)[60,60]	2(1)[254]	-1	0	YES
54889488	0(5,4)[60,37]	1(3)[38]	-1	1	YES
54890032	1(1)[60]	-1	0(3)[37]	0	YES
54890101	1(2)[60]	-1	-1	1	YES
54890687	1(2)[37]	-1	1(3)[37]	0	YES
54890740	0(5,2)[60,37]	0(1)[245]	1(5)[37]	1	YES
54890756	1(4,1)[60,37]	-1	1(3)[37]	0	YES
54890963	0(3,1)[60,37]	-1	0(2)[37]	0	YES
54891487	1(3,1)[60,60]	0(1)[254]	-1	0	YES
54893226	1(4,2)[60,60]	-1	0(1)[37]	0	YES
54894194	1(3)[60]	0(1)[245]	0(5)[37]	0	YES
54894974	0(7)[60]	2(1)[254]	0(2)[37]	1	YES
54895775	0(4)[60]	0(2)[100]	-1	0	YES

54896152	1(1)[60]	1(1)[18]	1(1)[37]	1	YES
54897521	1(1,2)[60,60]	1(1)[24]	1(2,1)[37,37]	0	YES
54898424	1(5,1)[60,37]	1(3,1)[248,2]	1(3)[37]	0	YES
54898578	0(3)[60]	-1	0(2)[37]	1	YES
54899250	0(2)[60]	-1	-1	0	YES
54900548				-1	NO
54900910	0(1)[37]	-1	0(3)[37]	0	YES
54901530	0(5)[60]	0(1)[254]	0(2)[37]	1	YES
54902174	0(2)[37]	0(1)[230]	0(2)[37]	1	YES
54904365	0(1)[60]	0(1)[40]	0(6)[37]	0	YES
54906867				-1	NO
54907064				-1	NO
54907309	1(1)[60]	0(2)[55]	0(2)[37]	0	YES
54907336	0(1)[37]	1(9)[55]	1(3)[37]	1	YES
54907404	0(7)[60]	0(1)[3]	-1	0	YES
54908240	0(2,1)[60,60]	1(1)[55]	-1	1	YES
54909614	0(1,2)[60,37]	0(2)[53]	1(1)[0]	1	YES
54911312	0(6)[60]	0(2)[254]	0(3)[37]	1	YES
54911960	0(10)[60]	2(1)[6]	-1	0	YES

^a 0,1,-1 and 2 indicate respectively, ancestral state, derived state, no call, and call differing from the alleles in NA18956

^b In parenthesis, the number of reads supporting the call; when both calls are present, the number of reads supporting the alternative call is given after the comma

^c In brackets, maximum alignment quality of a read supporting the call; maximum quality for a second call is given after the comma

Table S6. Frequency of the N lineage^a

N haplotype				Long Variant of N		
Geographic Region	Population	Frequency	N	N	N	
(source) ^b			(individuals)	(individuals)	(# chromosomes)	Freq
Western Europe						
Kayser	Rotterdam	0.068	280	0	NA	NA
	Helsinki	0.074	47	0	NA	NA
	Lyon	0.090	50	0	NA	NA
	Ngreece	0.069	51	0	NA	NA
	Ancona	0.031	49	0	NA	NA
	Førde	0.096	52	0	NA	NA
	Warsawa	0.122	49	0	NA	NA
	Uppsala	0.098	46	0	NA	NA
	Barcelona	0.021	47	0	NA	NA
	Prague	0.078	45	0	NA	NA
HGDP	Basque	0.063	24	24	0	0.00
	French	0.036	28	28	0	0.00
	Italian	0.038	13	12	0	0.00
	Russian	0.120	25	25	0	0.00
	Sardinian	0.018	28	28	0	0.00
	Tuscan	0.000	8	7	0	0.00
	Orcadian	0.100	15	15	0	0.00
Jorde	Slovenian	0.096	26	0	NA	NA
	N_european	0.065	23	0	NA	NA
HapMap	CEU	0.092	60	117	0	0.00

	TSI	0.100	25		88	0	0.00
Eastern Europe							
HGDP	Adygei	0.059	17		17	0	0.00
Jorde	Stalskoe	0.100	5		0	NA	NA
	Urkarah	0.111	18		0	NA	NA
Western Asia							
HGDP	Bedouin	0.000	46		45	0	0.00
	Druze	0.000	42		42	0	0.00
	Palestinian	0.032	47		46	0	0.00
Jorde	Kurd	0.000	24		0	NA	NA
South Asia							
HGDP	Balochi	0.021	24		24	0	0.00
	Brahui	0.060	25		25	0	0.00
	Burusho	0.000	25		25	0	0.00
	Hazara	0.091	22		22	0	0.00
	Kalash	0.000	23		23	0	0.00
	Makrani	0.020	25		25	0	0.00
	Pathan	0.000	22		22	0	0.00
	Sindhi	0.042	24		24	0	0.00
Jorde	Tamil-Nadu	0.000	27		0	NA	NA
	Andra Pradesh	0.022	46		0	NA	NA
	Pakistani	0.060	25		0	NA	NA
	Irula	0.063	24		0	NA	NA
	Nepalese	0.083	24		0	NA	NA
Central Asia							
Jorde	Kyrgystani	0.040	25		0	NA	NA
Southeast Asia							

HGDP	Cambodian	0.050	10		10	0	0.00
Jorde	Iban	0.140	25		0	NA	NA
	Vietnamese	0.000	7		0	NA	NA
	Cambodian	0.100	5		0	NA	NA
	Thai	0.083	24		0	NA	NA
East Asia							
HGDP	Dai	0.000	10		10	0	0.00
	Daur	0.167	9		9	1	0.06
	Han	0.015	34		34	0	0.00
	Han-NChina	0.000	10		10	0	0.00
	Hezhen	0.000	10		9	0	0.00
	Japanese	0.054	28		28	0	0.00
	Lahu	0.167	9		8	1	0.06
	Miao	0.050	10		10	0	0.00
	Mongola	0.000	10		10	0	0.00
	Naxi	0.167	9		8	1	0.06
	Oroqen	0.056	9		9	0	0.00
	She	0.000	10		10	0	0.00
	Tu	0.050	10		10	0	0.00
	Tujia	0.050	10		10	0	0.00
	Uyгур	0.000	10		10	0	0.00
	Xibo	0.111	9		9	0	0.00
	Yi	0.150	10		10	2	0.10
Jorde	Japanese	0.042	12		0	NA	NA
	Taiwan	0.000	3		0	NA	NA
	Chinese	0.000	7		0	NA	NA
HapMap	CHB	0.022	45		84	2	0.01

	JPT	0.089	45		86	7	0.04
North Asia							
HGDP	Yakut	0.040	25		25	0	0.00
Jorde	Buryat	0.000	25		0	NA	NA
America							
HGDP	Karitiana	0.000	13		13	0	0.00
	Maya	0.048	21		21	0	0.00
	Pima	0.111	18		14	0	0.00
	Surui	0.000	9		8	0	0.00
Jorde	Colombian	0.125	8		0	NA	NA
	Bolivian	0.022	23		0	NA	NA
	Totonac	0.042	24		0	NA	NA
Melanesia							
HGDP	Papuan	0.588	17		17	17	0.50
	Melanesian	0.438	16		11	5	0.23
This work	Papuan	0.641	32		0	NA	NA
Jorde	Samoa	0.423	13		0	NA	NA
	Tonga	0.278	9		0	NA	NA
This work	Vanuatu	0.652	23		0	NA	NA
	New-Britain	0.725	20		0	NA	NA
North Africa							
HGDP	Mozabite	0.111	27		27	0	0.00
Sub-Saharan Africa							
HGDP	Bantu Kenya	0.000	12		11	0	0.00
	Bantu SouthAfrica	0.000	8		8	0	0.00
	Biaka	0.000	25		22	0	0.00
	Mandenka	0.000	22		22	0	0.00

	Mbuti	0.000	14		13	0	0.00
	San	0.000	6		5	0	0.00
	Yoruba	0.000	21		21	0	0.00

^aThe Neandertal haplotype was determined by genotyping of one or several diagnostic SNPs

^bData sources are 14 and 15 for HGDP (supplemented in some cases), 16 for Kayser, and 17 for Jorde and HapMap