



Supporting Online Material for

Ancient DNA from the First European Farmers in 7500-Year-Old Neolithic Sites

Wolfgang Haak,* Peter Forster, Barbara Bramanti, Shuichi Matsumura, Guido Brandt, Marc Tänzer, Richard Villems, Colin Renfrew, Detlef Gronenborn, Kurt Werner Alt, Joachim Burger

*To whom correspondence should be addressed at the Molecular Archaeology Group, Institute of Anthropology, Colonel Kleinmann Weg 2, SBII 02, Johannes Gutenberg University, Mainz D-55128, Mainz, Germany. E-mail: haakw@uni-mainz.de

Published 11 November 2005, *Science* **310**, 1016 (2005)
DOI: 10.1126/science.1118725

This PDF file includes:

Materials and Methods
Figs. S1 and S2
Tables S1 to S5
References and Notes

Supporting Online Material

Materials and Methods

Samples.

Ancient samples were taken from 57 individuals from sites detailed in Table S3. The six N1a samples that yielded N1a haplotypes (Table S4) were found in four sites in the area of the *Linearbandkeramik* culture in Germany, namely Halberstadt, Unterwiederstedt, two from Derenburg Meerenstieg II and Flomborn. One N1a sample stems from the *Alföldi Vonaldisztes Kerámia (AVK)* site Ecsegfalva 23A. AVK is a contemporary LBK equivalent in the Eastern Great Hungarian Plain (Alföld) in Hungary (Fig. 1). Whenever possible, we sampled individuals in the unwashed and untreated state after excavation (e.g. Derenburg, Halberstadt and Unterwiederstedt).

Modern N1a sequences were consulted both from the literature and from databases (S1-S3). The modern N1a data are shown in Table S2.

Ancient DNA work. The aDNA work was conducted in two laboratories, located in separate buildings. One ancient DNA laboratory is dedicated solely to pre-PCR procedures and free of any molecular work, and a second laboratory accommodates all post-PCR analyses. The former laboratory follows strict clean-room conditions. Workers enter the clean room area only in the morning directly from home after a shower and wearing freshly washed clothes. All rooms and workbenches are routinely treated with bleach and UV-irradiated overnight. All workers wear clean-room overalls plus particular shoes, face masks, face shields and gloves. Every item entering is extensively washed with bleach and subsequently UV-irradiated. HPLC-water for PCR and filtered water for cleaning is UV-irradiated with a waterproof UV-bulb for at least 10h. All possible further methodical precautions were taken, i. e. total removal of sample surface, repeated irradiation with UV light, mock extractions, PCR blanks, independent replication of extraction, amplification and cloning of PCR products.

Sample preparation. Whenever possible, two anatomically distant samples (2cm x 2cm femoral compacta or intact molars) were taken from each Neolithic individual. Alternatively (DEB1 and FLO1) samples were divided and subsequently treated independently. The samples were UV-

irradiated (260nm) on all sides for at least 45 min. Teeth sample surfaces were gently wiped with bleach. The entire surface of every bone and tooth sample was removed by using dental drills, and the samples were cut into smaller pieces with diamond drills (Dremel, Germany and KaVO EWL K11 Type 4980, Germany). Tooth crowns were removed and were not used for the analysis. All samples were again UV-irradiated for 45 min. The samples were ground to fine powder by using a mixer mill (Retsch, Haan, Germany) and stored until use at 4 °C.

Ancient DNA extraction. Samples of 0.4-1 g powdered bone/teeth were incubated in 3.5 mL of extraction buffer (0.5M EDTA, pH 8.5; 0.5% N-lauryl sarcosine; 20 mg/μL proteinase K) on a rotary mixer over night at 37 °C. DNA was extracted with phenol/chloroform/isoamyl alcohol (25:24:1, Roth, Germany); the supernatant was transferred to Centricon 100 filter units (Amicon, Millipore) and was concentrated and washed several times with UV-treated-HPLC water.

PCR. The amplification of 413 bp (np 15997-16409) of the mtDNA control region was performed by using four overlapping primer pairs (Table S5). The amplification reaction was set up in a volume of 50 μL, containing 1x PCR Gold Buffer (Applied Biosystems, Darmstadt, Germany), 2.5 mM MgCl₂ solution (Applied Biosystems), 0.2 mM dNTP Mix (MBI, Fermentas GmbH, St. Leon-Rot, Germany), 2.5 U AmpliTaq Gold[®] DNA Polymerase (Applied Biosystems, Darmstadt, Germany), 0.2 μM each primer and 1-8 μL DNA from bone or teeth extracts. In cases where we used Uracil-N-Glycosylase (UNG) (Sigma-Aldrich), 1U was added directly to the PCR reaction and incubated at room temperature for 40 min. The cycle conditions using a Mastercycler gradient (Eppendorf, Hamburg, Germany) consisted of an initial denaturation at 92 °C for 6 min, 38 cycles of 30 s at 92 °C, 30 s at 58 °C and 30 s at 72 °C, followed by a final extension at 60 °C for 30 min. Amplicons were purified using the Invisorb Rapid PCR Purification Kit (Invitek, Berlin-Buch, Germany) according to the manufacturer's instructions. Amplification and subsequent digestion of specific mtDNA coding region sites were performed as described (S4-S6).

Cloning and sequencing.

Subsequent cloning and sequencing of PCR products was performed by using a pUC18 (T-vector, own production) transformed to an E. coli culture (RRI). Selected clones were picked directly into PCR (50 μl final volume), containing 1x PCR Buffer (ABgene[™], Hamburg,

Germany), 2.5 mM MgCl₂ solution (ABgeneTM, Hamburg, Germany), 0.2 mM dNTP Mix (MBI Fermentas GmbH, St. Leon-Rot, Germany), 1 U DNA Polymerase (ABgeneTM, Hamburg, Germany), 0.2 μM each of universal M13 forward and reverse primer. Cycle conditions consisted of an initial denaturation at 94 °C for 15 min, 30 cycles of 30 s at 94 °C, 30 s at 58 °C and 30 s at 72 °C. PCR products carrying the expected insert length (screened by 2% agarose gel electrophoresis) were purified as mentioned above. Cycle sequencing was carried out with the DNA Sequencing Kit (Terminator Cycle Sequencing Ready Reaction; Applied Biosystems, Darmstadt, Germany) using 25 cycles at 92°C denaturing for 30 s, 15 s annealing at 55°C with the universal M13 primer and elongation at 72°C for 2.5 min. The sequencing products were analysed by capillary electrophoresis on an ABI PRISMTM 310 Genetic Analyzer (Applied Biosystems, Darmstadt, Germany) using the Collection Software (ABI Prism Sequencing AnalysisTM 3.0; Applied Biosystems, Darmstadt, Germany). Further sequence analyses were carried out using the programs *Seqman II*TM and *MegAlign*TM from the DNA Star Software package (version 4.05). “Phantom mutations” as defined by Bandelt and colleagues (S7) were excluded through our approach of sequencing overlapping clones. DNA sequence data is deposited under accession numbers D9211974-D9211979 in Genbank (S2).

Phylogenetic analysis.

A median-joining network (Fig. 3) of N1a sequences was constructed with the Network 4.1.1.1 programme (S8) (Fluxus Technology Ltd., Clare, Suffolk, UK) according to instructions on the website and in the literature (S9). Epsilon was set to 0 and default weights were used. Four samples producing reticulations in the network were omitted from network analysis. These were all samples showing 16147T (one Indian (S10), one Armenian (S11), one Uzbek sample (S3), and one Turk (S12)).

Simulation Method.

A discrete-generation coalescent method (S13) was used to follow the change in the frequency of a particular mtDNA lineage in the course of 275 generations (7425 years assuming the generation time as 27 years). It allowed multiple coalescent and/or migration events per generation. We started with 2300 samples from the modern population and simulated genealogies backwards in time. When we reached 275 generations before the present, we allocated an mtDNA lineage to each ancestor randomly according to the assumed mtDNA frequencies (i.e., 8% versus 92%) in

the original population. By counting the number of descendants in the modern population for all the ancestors, we obtained the resulting frequencies of mtDNA lineages in the modern population. We firstly tested the simplest demographic scenario where the target population was genetically isolated (i.e. the migration rate was zero). Next, we allowed migration between the target population and a surrounding large population. For simplicity, we assumed that the surrounding population did not originally contain the mtDNA lineage in question, and that it was too large to allow back-migration of this lineage to the target population.

Simulation Results.

We simulated a scenario that would maximise the chance that the N1a has been lost by genetic drift in the course of approximately the past 7500 years (for convenience, we chose a time span of 7425 years resulting from 275 generations with a female generation time of 27 years). The generation time of 27 years is conservative, being at the low end of modern hunter-gatherer generation times. We assumed a small Neolithic population of only 3500 effective females (roughly equivalent to 5000 actual females) inhabiting the LBK area, and we assumed the lower confidence limit of 8% N1a in the Neolithic population, and a population growth of only 0.45 % (about 7 times lower than a documented agricultural expansion (*SI4*)) for the first 25 generations, and a subsequent population growth of 0.1% per year (*SI5*) for the next 250 generations. We ran the simulation 4000 times and recorded the number of N1a types in the modern sample of 2300 individuals. The simulation showed that we should observe at least 74 N1a out of the 2300 modern samples. 95% of the total runs ended between 119 and 259 N1a in the modern sample. Next, we allowed migration between the Neolithic population and the surrounding population per generation. The simulation showed that a migration rate of 1% per generation throughout 7425 years between the Neolithic population and the surrounding population is not enough to reduce the N1a percentage to the low value observed today, because only 5.5% of the total runs ended in <6 N1a in the modern sample.

These simulations reject the simple hypothesis in which modern Europeans are direct descendants from these first farmers and have lost N1a mainly by genetic drift. Hence the simulations confirm that the first farmers in Central Europe had limited success in leaving a genetic mark on the female lineages of modern Europeans.

Authentication criteria.

The internal evidence thus overwhelmingly suggests that the Neolithic N1a types are genuine, but nevertheless we also observed the inevitable effects of modern human contamination (*S16*, *S17*), which we kept to a minimum by progressively improving the laboratory procedures in our clean room facilities, and by applying quality control criteria. The details are as follows:

- All sequences were reproduced from at least two extracts and four PCRs and by subsequent cloning of at least 5 but usually 10 to 12 clones per PCR.
- Concerning the whole Neolithic sample set, reproducible amplification success was always restricted to certain archaeological sites. This would not have been the case for laboratory-internal background contamination.
- Simultaneously amplified DNA from a parallel study on specimens from outside the LBK area yielded no N1a-haplotype (data not shown)
- In those cases where the archaeologists and anthropologists who had worked on the skeletons prior to sampling for genetic analysis were known to us, we never were able to amplify the DNA of these investigators after removing and decontaminating sample surfaces following our protocols (see Methods section).
- The contamination rate of our clones and of PCR blanks varied with time, while the reproducibility of the sequences at hand varied only from sample to sample.
- Overlapping amplicons from independent PCRs complemented one another to haplotypes which make phylogenetic sense. Those sequences that we identified as contaminations hardly did so.
- Post-mortem DNA damage common and characteristic in ancient samples was observed in all samples, but could not be reproduced.
- The overall good preservation state of the bones and teeth which led to the finding of up to 450bp mtDNA in one case and nuclear DNA in 3 out of 24 LBK samples (12%).

Only a minor proportion of clones (0% in DEB1 and HAL2, 1.1% in DEB3, 5.2% in UWS5, 18.3% in FLO1 and 15.5% in ECS 1) showed additional divergent sequences. These were identified as contaminating DNAs as they were never reproducible in a second experiment (See sequence alignments Fig. S2). By following these strict technical procedures and a rigorous experimental design we can exclude contamination with a high level of confidence. Nevertheless,

from a theoretical point of view it is impossible to completely rule out a systematic contamination of a skeleton (see also (S18)). However, even if the extremely unlikely event of a systematic contamination of a whole skeleton did take place and in addition our decontamination procedures were not able to remove it, this should be very rare and will not affect our sample of 57 (of which 24 were successfully typed) skeletons as a whole and thence neither the overall interpretation of our data.

Table S1. Assigned haplogroups of the remaining LBK samples successfully analysed.

Country	Location/Site	ID	Haplogroup	12308 Hinf	7025 Alu I
Austria	Asparn Schletz	ASP2	H*	-12308 Hinf	-7025 Alu I
Germany	Derenburg/Meerenstieg II	DEB2	K		
Germany	Derenburg/Meerenstieg II	DEB4	HV	-12308 Hinf	+7025 Alu I
Germany	Derenburg/Meerenstieg II	DEB5	HV	-12308 Hinf	+7025 Alu I
Germany	Eilsleben	EIL1	H*	-12308 Hinf	-7025 Alu I
Germany	Flomborn	FLO2	H*	-12308 Hinf	-7025 Alu I
Germany	Flomborn	FLO4	T		
Germany	Flomborn	FLO5	K		
Germany	Flomborn	FLO6	K		
Germany	Halberstadt	HAL1	V	-12308 Hinf	+7025 Alu I
Germany	Halberstadt	HAL3	T		
Germany	Schwetzingen	SCHWE1	T		
Germany	Schwetzingen	SCHWE2	T		
Germany	Schwetzingen	SCHWE4	H*		-7025 Alu I
Germany	Schwetzingen	SCHWE5	T		
Germany	Seehausen	SEE1	J*		
Germany	Unterwiederstedt	UWS2	K		
Germany	Vaihingen an der Enz	VAI3	U3		

Table S2. Modern N1a data. For colouring see the legend below the table.

Individual	#	HVS I haplotype								Reference	Associated Nation, Culture, Language or Region	
BACO1	1	16086.C	16147.A	16223.T	16248.T	16320.T	16324.C	16355.T		(S19)	Açores	
BAIIRA	1	16147.G	16172.C	16223.T	16248.T	16295.T	16355.T			(S20)	Irani Caste, Irani (Persian) speaking, India	
BBAS16	6	16147.A	16172.C	16189.C	16223.T	16248.T	16320.T	16355.T		(S21)	Bashkir	
BBAS7	1	16147.A	16172.C	16187.A	16189.C	16223.T	16248.T	16320.T	16355.T	(S21)	Bashkir	
BCHU	1	16147.A	16172.C	16223.T	16248.T	16320.T	16355.T			(S21)	Chuvashi	
BCRO1	1	16147.A	16172.C	16209.C	16223.T	16248.T	16320.T	16355.T		(S22)	Croatian-Italian##	
BCRO2	1	16147.A	16172.C	16209.C	16223.T	16248.T	16320.T	16355.T		(S22)	Croatian-Italian##	
BKOM16	6	16086.C	16147.A	16223.T	16248.T	16278.T	16320.T	16355.T		(S21)	Komi	
BKOM710	4	16147.A	16172.C	16189.C	16223.T	16248.T	16272.G	16320.T	16355.T	(S21)	Komi	
BMOL1	1	16092.C	16129.A	16147.A	16154.C	16172.C	16223.T	16248.T	16320.T	16355.T	(S22)	Molise##
BUZB1	1	16147.A	16172.C	16189.C	16234.T	16290.T	16362.C			(S3)	Uzbek	
CALTUR	1	16129.A	16147.A	16248.T						(S12)	Turkey	
DALT1	1	16147.G	16172.C	16189.C	16223.T	16248.T	16320.T	16355.T		(S23)	Altaiian tribe, Turkic-speaker, Altai Republic	
DALT2	2	16147.A	16172.C	16189.C	16223.T	16248.T	16320.T	16355.T		(S23)	Altaiian tribe, Turkic-speaker, Altai Republic	
FSWI1	1	16086.C	16147.A	16223.T	16248.T	16320.T	16355.T			(S24)	Romansh-speak., Sent, Lower Engadin, Switzerland	
FSWI2	1	16086.C	16147.A	16223.T	16248.T	16320.T	16355.T			(S24)	Romansh-speak., Lower Engadin, Switzerland	
FSWI3	1	16086.C	16147.A	16223.T	16320.T	16355.T				(S24)	Romansh-speak., Sent, Lower Engadin, Switzerland	
GTAT1	1	16174.A	16172.C	16223.T	16248.T	16320.T				(S3)	Tatar	
HSCO	1	16086.C	16147.A	16223.T	16248.T	16319.A	16320.T	16355.T		(S25)	Scotland (Western Isles) EMBL:AY025314	
KAMH1	1	16147.G	16172.C	16223.T	16248.T	16260.T	16355.T			(S26)	Amhara, Ethiopia	
KAMH23	2	16147.G	16172.C	16223.T	16248.T	16355.T				(S26)	Amhara, Ethiopia	
KERI1	1	16147.G	16172.C	16223.T	16248.T	16355.T				(S26)	Eritrean, Eritrea	
KGUR1	1	16147.G	16172.C	16223.T	16248.T	16355.T				(S26)	Gurage, Ethiopia	
KINDL	1	16147.A	16172.C	16223.T	16248.T	16294.T	16320.T	16355.T		(S10)	Andhra Pradesh, India	
KINDR	1	16147.T	16189.C	16223.T	16243.C	16278.T	16355.T	16362.C		(S10)	Andhra Pradesh, India	
KOSLO1	1	16147.A	16172.C	16223.T	16248.T	16320.T	16355.T			(S27)	Slovakia	
KOSLO2	1	16147.A	16154.C	16172.C	16223.T	16320.T	16355.T			(S27)	Slovakia	
KRIEGC	1	16147.A	16172.C	16223.T	16248.T	16355.T				(S28)	Egyptian, Cairo	
KRIEGYA	1	16147.A	16172.C	16223.T	16248.T	16291.T	16320.T	16355.T		(S28)	Egypt, Assiut, Arabic	
KTIG1	1	16147.G	16172.C	16223.T	16248.T	16355.T				(S26)	Tigrai, Ethiopia	

Individual	#	HVS I haplotype								Reference	Associated Nation, Culture, Language or Region	
KTZD	1	16147.G	16172.C	16223.T	16248.T	16355.T				(S29)	Datoga tribe, Sudanic-speaking, Lake Eyasi, Tanzania	
KTZI	1	16147.G	16172.C	16223.T	16248.T	16355.T				(S29)	Iraqw tribe, Cushitic-speaking, Lake Eyasi, Tanzania	
KYEM1	1	16124.C	16147.G	16172.C	16213.A	16223.T	16248.T	16355.T		(S29)	Yemeni, Yemen	
KYEM2	1	16147.A	16154.C	16170.G	16172.C	16223.T	16248.T	16320.T	16355.T	(S26)	Yemeni, Yemen	
KYEM3	1	16147.A	16154.C	16170.G	16172.C	16223.T	16248.T	16320.T	16355.T	(S26)	Yemeni, Yemen	
KYEM45	2	16147.A	16172.C	16223.T	16248.T	16320.T	16355.T			(S26)	Yemeni, Yemen	
KYEM6	1	16147.G	16172.C	16223.T	16248.T	16263.C	16266.T	16355.T		(S26)	Yemeni, Yemen	
KYEM78	2	16147.G	16223.T	16248.T	16263.C	16266.T	16355.T			(S26)	Yemeni, Yemen	
LHUN7	1	16147.A	16172.C	16189.C	16223.T	16248.T	16320.T	16355.T		(S30)	Budapest, Hungary	
LHUN8	1	16086.C	16147.A	16164.G	16172.C	16223.T	16248.T	16320.T	16355.T	(S30)	Budapest, Hungary	
METIRA	1	16147.A	16172.C	16223.T	16248.T	16320.T	16355.T			(S31)	Mashhad \ Ostan-e-Khorasan	
MHAV26	1	16147.G	16172.C	16209.N	16223.T	16248.T	16295.T	16355.T		(S32)	Havik group, Brahmin caste, NW Karnataka, India	
MHAV29	1	16147.G	16172.C	16223.T	16248.T	16295.T	16355.T			(S32)	Havik group, Brahmin caste, NW Karnataka, India	
MHAV38	1	16147.G	16172.C	16223.T	16248.T	16295.T	16355.T			(S32)	Havik group, Brahmin caste, NW Karnataka, India	
MHAV45	1	16147.G	16172.C	16209.C	16223.T	16248.T	16295.T	16355.T		(S32)	Havik group, Brahmin caste, NW Karnataka, India	
ONORW	1	16092.C	16129.A	16147.A	16154.C	16172.C	16223.T	16248.T	16320.T	16355.T	(S33)	E Norway
PASNOR	1	16086.C	16147.A	16148.T	16214.T	16223.T	16320.T	16355.T		(S34)	central/south Norway	
PBRUNS	1	16092.C	16129.A	16147.A	16154.C	16172.C	16223.T	16248.T	16320.T	16355.T	(S35)	German or foreigner, Braunschweig, Germany
PBUR57	1	16147.A	16172.C	16223.T	16248.T	16295.T	16320.T	16355.T		(S36)	Buryat tribe, Ulan-Ude, Buryat Republic	
PBUR58	1	16147.A	16172.C	16183.C	16189.C	16223.T	16248.T	16320.T	16355.T	(S36)	Buryat tribe, Ulan-Ude, Buryat Republic	
PFRA1	1	16147.A	16223.T	16248.T	16320.T	16355.T				(S3)	Lyonnais, Rhone, France	
PFRA2	1	16147.A	16172.C	16206.G	16223.T	16248.T	16320.T	16355.T		(S3)	Maine-Anjou, Sarthe, France	
PFRA3	1	16086.C	16147.A	16164.G	16172.C	16223.T	16248.T	16320.T	16355.T	(S3)	Brittany, Finistère, France	
PGREIF	1	16114.A	16147.A?	16172.C	16223.T	16248.T	16320.T	16325.C	16355.T	(S37)	German, Vorpommern	
PMUEN	1	16147.A	16172.C	16195.C	16223.T	16248.T	16320.T	16355.T		(S38)	Munster area, Germany	
PPORT	1	16086.C	16147.A	16164.G	16172.C	16223.T	16248.T	16320.T	16355.T	(S39)	central Portugal between Douro and Tejo, 119	
QTUR	1	16147.A	16172.C	16223.T	16248.T	16320.T	16355.T			(S40)	Altaic, Turkmen, Turkmenistan	
RARM73	1	16147.T	16172.C	16223.T	16248.T	16278.T	16355.T			(S11)	Armenian, Armenia	
RARM92	1	16147.A	16172.C	16223.T	16248.T	16355.T				(S11)	Armenian, Armenia	
RCAUC	1	16147.G	16172.C	16223.T	16248.T	16295.T	16297.C	16355.T		(S11)	Kabardian, north Caucasus	
RDAN	1	16086.C	16147.A	16189.C	16223.T	16248.T	16320.T	16355.T		(S11)	Sonderborg, Denmark	

Individual	#	HVS I haplotype								Reference	Associated Nation, Culture, Language or Region
REST	1	16147.A	16172.C	16223.T	16248.T	16294.T	16320.T	16355.T		(S11)	Estonian
RGRE	1	16147.G	16172.C	16223.T	16224.C	16248.T	16355.T	16357.C		(S11)	Thessaloniki, Greece
RIRAN	1	16147.A	16172.C	16223.T	16248.T	16320.T	16355.T			(S11)	Iran
RKIZIL	1	16147.A	16172.C	16189.C	16223.T	16248.T	16320.T	16355.T		(S41)	Kizil, Altai Republic
SEST	1	16147.A	16172.C	16223.T	16248.T	16320.T	16355.T			(S42)	Estonia
TTURK13	3	16147.A	16172.C	16189.C	16223.T	16248.T	16320.T	16355.T		(S43)	Turkey
TTURK4	1	16147.G	16172.C	16223.T	16248.T					(S43)	Turkey
TTURK5	1	16147.G	16172.C	16223.T	16248.T	16295.T	16297.C	16344.T	16355.T	(S43)	Turkey
VEST1	1	16147.A	16172.C	16223.T	16248.T	16320.T				(S3)	Estonia
VEST25	4	16147.A	16172.C	16223.T	16248.T	16320.T	16355.T			(S3)	Estonia
VKAL1	1	16147.A	16172.C	16189.C	16223.T	16248.T	16320.T	16355.T		(S3)	Kalmyks
VKAZ15	5	16147.A	16172.C	16189.C	16223.T	16248.T	16320.T	16355.T		(S3)	Kazakhs
VPOL1	1	16147.A	16172.C	16248.T	16320.T	16355.T				(S3)	Poland
VRUS1	1	16147.A	16172.C	16223.T	16248.T	16320.T	16355.T			(S3)	Russia
VRUS2	1	16147.A	16189.C	16223.T	16248.T	16272.G	16320.T	16355.T		(S3)	Russia
VSVE1	1	16147.A	16154.C	16172.C	16223.T	16320.T	16355.T			(S3)	Sweden
VUDM1	1	16086.C	16147.A	16223.T	16248.T	16278.T	16320.T	16355.T		(S3)	Udmurts
WSOM	1	16147.G	16172.C	16213.A	16223.T	16248.T	16355.T			(S44)	Somali, E Africa

Legend:

White/Italic: Samples causing reticulations

Blue: European/Eurasian branch

Orange: Central Asian branch

Green: African/SouthAsian branch

Pink: Unclear sample sizes

Red: Unclear/reconstructed positions

Yellow: ignored position

Table S3. Neolithic LBK/AVK samples that have been analysed.

Location/Site	Number of Samples	Country
Asparn Schletz	1	Austria
Derenburg/Meerenstieg II	5	Germany
Halberstadt	3	Germany
Schwetzingen	4	Germany
Flomborn	6	Germany
Eilsleben	1	Germany
Seehausen	3	Germany
Sondershausen	4	Germany
Bruchstedt	5	Germany
Unterwiederstedt	10	Germany
Vaihingen a.d. Enz	1	Germany
Unseburg	3	Germany
Wiesbaden-Erbenheim	8	Germany
Viesenhäuserhof	1	Germany
Bruchenbrücken	1	Germany
Ecsegfalva	1	Hungary

Table S4. Additional information on the six N1a individuals.

Individual	Coordinates	Grave	Anthropological/Archaeological information	Source	Dating	Literature
Derenburg 1	51.86 N 10.9 E	33	Date of excavation: 1999 Skeleton: Nearly complete preservation, good condition, juvenile, presumably female, leftside flexed supine position, East-West orientation. Burial goods: 1 stone pearl	Landesamt für Archäologie Sachsen-Anhalt, Halle/Saale, Germany	LBK	(S45)
Derenburg 3	51.86 N 10.9 E	21	Date of excavation: 1999 Skeleton: Good condition, but disturbances from the lower spine to the femurs, adult, presumably female, rightside flexed position, East-West orientation. Burial goods: 2 vessels, 1 grinding stone	Landesamt für Archäologie Sachsen-Anhalt, Halle/Saale, Germany	LBK	(S45)
Halberstadt 2	51.54 N 11.04 E	35	Date of excavation: 2000 Skeleton: Complete preservation in excellent condition, leftside flexed position, West-East orientation Burial goods: 2 open bowl , 1 closed bowl, 1 Zipfelschale	Landesamt für Archäologie Sachsen-Anhalt, Halle/Saale, Germany	LBK	(S46)
Unterviederstedt 5	51.66 N 11.53 E	Mass grave, no. 7	Date of excavation: 1998 Skeleton: Nearly complete preservation, good condition, female, mature (35-45 years) Burial goods: none	Landesamt für Archäologie Sachsen-Anhalt, Halle/Saale, Germany	LBK	(S47)
Flomborn 1	49.69 N 8.15 E	13	Date of excavation: 1900-1903 Skeleton: Nearly complete preservation, partially secondarily lost, mint condition, presumably female, mature, leftside flexed position, East-West orientation Burial goods: 1 shoe-last adze, 1 grinding stone	Landesdenkmalamt Rheinland-Pfalz, Museum Worms, Germany	LBK	(S48)
Ecsegfalva 1	47.15 N 20.91 E	23A	Date of excavation: 1999 Skeleton: Nearly complete preservation, good condition, female, leftside flexed position Burial goods: none No sign of a grave pit as such, the body had presumably been lain on the soil or in a depression in the soil.	Hungarian Natural History Museum, Hungary	AVK later sixth millennium BC OxA-10678	(S49)

Table S5: Sequences of the four overlapping primer pairs used in this study.

Name	Sequence 5'-3'	Source
L15996	CTCCACCATTAGCACCCAAAGC	(S50)
H16142	ATGTACTACAGGTGGTCAAG	(S51)
L16117	TACATTACTGCCAGCCACCAT	this study
H16233	GCTTTGGAGTTGCAGTTGATGTGT	this study
L16209	CCCATGCTTACAAGCAAGT	(S52)
H16348	ATGGGGACGAGAAGGGATTTG	this study
L16287	CACTAGGATACCAACAAACC	(S52)
H16410	GCGGGATATTGATTCACGG	(S52)

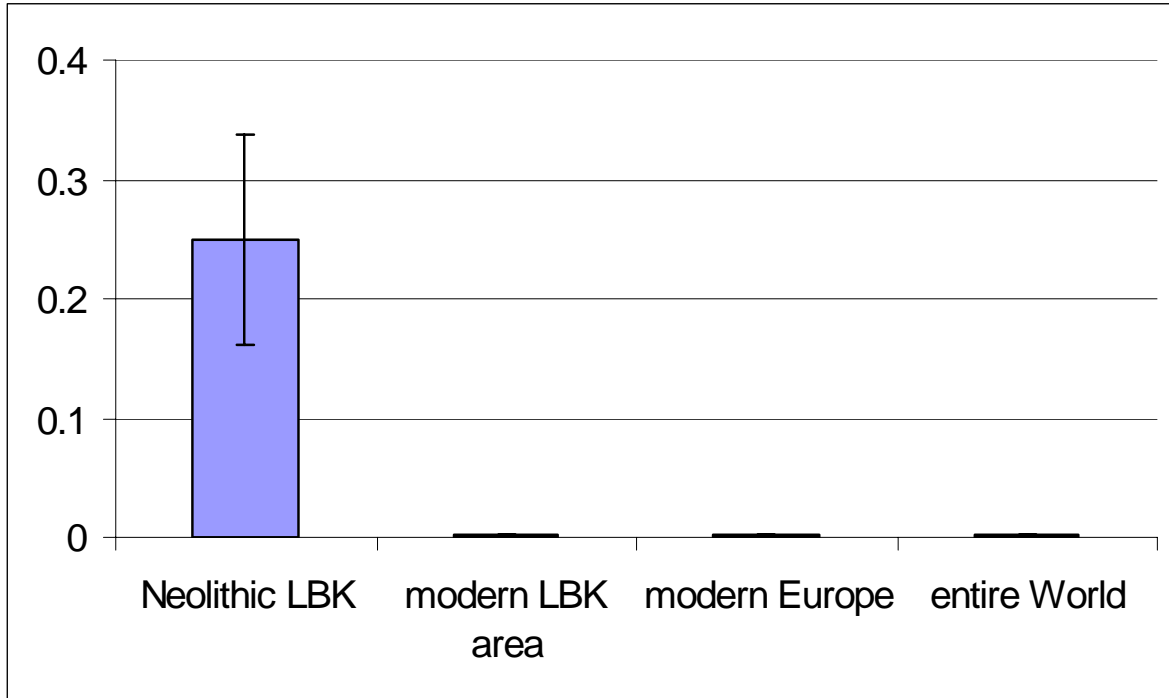


Fig. S1. Frequencies of N1a mtDNA lineages in Neolithic and in modern humans. In the Neolithic sample, the values are as follows: 6/24 N1a in the Neolithic sample (mean 0.250, SE 0.088), 5/2300 N1a in modern Central Europeans (mean 0.0022, SE 0.0010), 36/21384 N1a in all modern Europeans (mean 0.0014, SE 0.00026) and 105/32132 N1a in the entire World (mean 0.0021, SE 0.00026). Frequencies are taken from the literature (*S1*), Genbank (*S2*), and the EBC database (*S3*).

J01415

16000 16010 16020 16030 16040 16050 16060 16070 16080 16090 16100 16110 16120 16130 16140 16150 16160 16170 16180 16190 16200
TAA GATT CTA ATT TAA AACT ATT CT CT GT TCT TTT CAT GGG GAAG CA GATT TGG GT ACC ACG CA A GT ATT GACT CAC CCAT CA A CA A CCG GT AT GT ATT TCG TAC ATT ACT GGC A GCA CCA T GA AT ATT GT ACG GT ACC ATA AAT ACT TG ACC A CT GT AGT AC AT AAA AAC CCA AT CCA CAT CA AAA CCCC CT CCCC AT GC TT ACA A

DEB1_A.V1.1 A
DEB1_A.V1.2 T
DEB1_A.V1.3
DEB1_A.V1.4 TT T
DEB1_A.V1.5 A A
DEB1_A.V1.6
DEB1_A.V1.7
DEB1_A.V1.8 C
DEB1_A.V1.9
DEB1_A.V1.10
DEB1_A.V1.11 A T T T A A
DEB1_B.V1.1
DEB1_B.V1.2
DEB1_B.V1.3
DEB1_B.V1.4 A A A A A
DEB1_B.V1.5
DEB1_B.V1.6
DEB1_B.V1.7 T
DEB1_B.V1.8
DEB1_B.V1.9
DEB1_B.V1.10 A
DEB1_A.II1UNG.1 A C
DEB1_A.II1UNG.2 A C
DEB1_A.II1UNG.3 A C
DEB1_A.II1UNG.4 A C
DEB1_A.II1UNG.5 A C
DEB1_A.II1UNG.6 A C
DEB1_A.II1UNG.7 A C
DEB1_A.II1UNG.8 A C
DEB1_A.II1UNG.9 A C
DEB1_A.II1UNG.10 A C
DEB1_B.II1.1 A C
DEB1_B.II1.2 A C
DEB1_B.II1.3 A C
DEB1_B.II1.4 A C
DEB1_B.II1.5 A C
DEB1_B.II1.6 A C
DEB1_B.II1.7 A C
DEB1_B.II1.8 A C
DEB1_B.II1.9 T T A C T T
DEB1_B.II1.10 T T T A C T T
DEB1_B.II1.11 A C T
DEB1_A.III1UNG.1
DEB1_A.III1UNG.2
DEB1_A.III1UNG.3
DEB1_A.III1UNG.4
DEB1_A.III1UNG.5
DEB1_A.III1UNG.6
DEB1_A.III1UNG.7
DEB1_A.III1UNG.8
DEB1_A.III1UNG.9
DEB1_A.III1UNG.10
DEB1_A.III1UNG.11
DEB1_A.III1UNG.12
DEB1_B.III1.1
DEB1_B.III1.2
DEB1_B.III1.3
DEB1_B.III1.4
DEB1_B.III1.5
DEB1_B.III1.6
DEB1_B.III1.7
DEB1_B.III1.8
DEB1_B.III1.9
DEB1_B.III1.10
DEB1_B.III1.11
DEB1_A.IV1UNG.1
DEB1_A.IV1UNG.2
DEB1_A.IV1UNG.3
DEB1_A.IV1UNG.4
DEB1_A.IV1UNG.5
DEB1_A.IV1UNG.6
DEB1_A.IV1UNG.7
DEB1_A.IV1UNG.8
DEB1_A.IV1UNG.9
DEB1_A.IV1UNG.10
DEB1_A.IV1UNG.11
DEB1_B.IV1.1
DEB1_B.IV1.2
DEB1_B.IV1.3
DEB1_B.IV1.4
DEB1_B.IV1.5
DEB1_B.IV1.6
DEB1_B.IV1.7
DEB1_B.IV1.8
DEB1_B.IV1.9
DEB1_B.IV1.10
DEB1_B.IV1.11
DEB1_B.IV1.12

DEB1

16210 16220 16230 16240 16250 16260 16270 16280 16290 16300 16310 16320 16330 16340 16350 16360 16370 16380 16390 16400
 GCAAGTACAGCAATCAACCGCTGAACATACACACATCAACTGCAACTCCAAAGCCACCCCTCACCCCACTAGGATACCAAGCAACCTACCCACCCCTTAAACAGTACATAGTACATAAAGCCATTTACCGTACATAGCACATTACAGTCAATCCCTTCTCGTCCCCATGGATGACCCCCCTCAGATAGGGGTCCCTTGACCACCATCGT

J01415
 DEB3_A_V1_1
 DEB3_A_V1_2
 DEB3_A_V1_3
 DEB3_A_V1_4
 DEB3_A_V1_5
 DEB3_A_V1_6
 DEB3_A_V1_7
 DEB3_A_V1_8
 DEB3_A_V1_9
 DEB3_A_V1_10
 DEB3_B_V1UNG_1
 DEB3_B_V1UNG_2
 DEB3_B_V1UNG_3
 DEB3_B_V1UNG_4
 DEB3_B_V1UNG_5
 DEB3_B_V1UNG_6
 DEB3_B_V1UNG_7
 DEB3_B_V1UNG_8
 DEB3_B_V1UNG_9
 DEB3_B_V1UNG_10
 DEB3_B_V1UNG_11
 DEB3_B_V1UNG_12
 DEB3_B_V2UNG_1
 DEB3_B_V2UNG_2
 DEB3_B_V2UNG_3
 DEB3_B_V2UNG_4
 DEB3_B_V2UNG_5
 DEB3_B_V2UNG_6
 DEB3_B_V2UNG_7
 DEB3_B_V2UNG_8
 DEB3_B_V2UNG_9
 DEB3_B_V2UNG_10
 DEB3_B_V2UNG_11
 DEB3_A_IV1_1
 DEB3_A_IV1_2
 DEB3_A_IV1_3
 DEB3_A_IV1_4
 DEB3_A_IV1_5
 DEB3_A_IV1_6
 DEB3_A_IV1_7
 DEB3_A_IV1_8
 DEB3_A_IV1_9
 DEB3_B_IV1UNG_1
 DEB3_B_IV1UNG_2
 DEB3_B_IV1UNG_3
 DEB3_B_IV1UNG_4
 DEB3_B_IV1UNG_5
 DEB3_B_IV1UNG_6
 DEB3_B_IV1UNG_7
 DEB3_B_IV1UNG_8
 DEB3_A_III1_1
 DEB3_A_III1_2
 DEB3_A_III1_3
 DEB3_A_III1_4
 DEB3_A_III1_5
 DEB3_A_III1_6
 DEB3_A_III1_7
 DEB3_A_III1_8
 DEB3_A_III1_9
 DEB3_A_III1_10
 DEB3_A_III1_11
 DEB3_B_III1UNG_1
 DEB3_B_III1UNG_2
 DEB3_B_III1UNG_3
 DEB3_B_III1UNG_4
 DEB3_B_III1UNG_5
 DEB3_B_III1UNG_6
 DEB3_B_III1UNG_7
 DEB3_B_III1UNG_8
 DEB3_B_III1UNG_9
 DEB3_A_IV1_1
 DEB3_A_IV1_2
 DEB3_A_IV1_3
 DEB3_A_IV1_4
 DEB3_A_IV1_5
 DEB3_A_IV1_6
 DEB3_A_IV1_7
 DEB3_A_IV1_8
 DEB3_B_IV1UNG_1
 DEB3_B_IV1UNG_2
 DEB3_B_IV1UNG_3
 DEB3_B_IV1UNG_4
 DEB3_B_IV1UNG_5
 DEB3_B_IV1UNG_6
 DEB3_B_IV1UNG_7
 DEB3_B_IV1UNG_8
 DEB3_B_IV1UNG_9

```

..... T ..... T .....
DEB3_A_IV1_2 ..... T .....
DEB3_A_IV1_3 ..... T .....
DEB3_A_IV1_4 ..... T .....
DEB3_A_IV1_5 ..... T .....
DEB3_A_IV1_6 ..... T .....
DEB3_A_IV1_7 ..... T T .....
DEB3_A_IV1_8 ..... T .....
DEB3_A_IV1_9 ..... T T .....
DEB3_B_IV1UNG_1 ..... T .....
DEB3_B_IV1UNG_2 ..... T .....
DEB3_B_IV1UNG_3 ..... T .....
DEB3_B_IV1UNG_4 ..... T .....
DEB3_B_IV1UNG_5 ..... T .....
DEB3_B_IV1UNG_6 ..... T .....
DEB3_B_IV1UNG_7 ..... T .....
DEB3_B_IV1UNG_8 ..... T .....
DEB3_A_III1_1 ..... T .....
DEB3_A_III1_2 ..... T .....
DEB3_A_III1_3 ..... T .....
DEB3_A_III1_4 ..... T .....
DEB3_A_III1_5 ..... TT .....
DEB3_A_III1_6 ..... TT .....
DEB3_A_III1_7 ..... T .....
DEB3_A_III1_8 ..... C .....
DEB3_A_III1_9 ..... C .....
DEB3_A_III1_10 ..... C .....
DEB3_A_III1_11 ..... C .....
DEB3_B_III1UNG_1 ..... T .....
DEB3_B_III1UNG_2 ..... NN .....
DEB3_B_III1UNG_3 ..... T .....
DEB3_B_III1UNG_4 ..... T .....
DEB3_B_III1UNG_5 ..... T .....
DEB3_B_III1UNG_6 ..... T .....
DEB3_B_III1UNG_7 ..... T .....
DEB3_B_III1UNG_8 ..... T .....
DEB3_B_III1UNG_9 ..... T .....
DEB3_A_IV1_1 ..... T .....
DEB3_A_IV1_2 ..... T .....
DEB3_A_IV1_3 ..... T .....
DEB3_A_IV1_4 ..... T .....
DEB3_A_IV1_5 ..... T .....
DEB3_A_IV1_6 ..... T .....
DEB3_A_IV1_7 ..... T .....
DEB3_A_IV1_8 ..... T .....
DEB3_B_IV1UNG_1 ..... T .....
DEB3_B_IV1UNG_2 ..... T .....
DEB3_B_IV1UNG_3 ..... T .....
DEB3_B_IV1UNG_4 ..... T .....
DEB3_B_IV1UNG_5 ..... T .....
DEB3_B_IV1UNG_6 ..... T .....
DEB3_B_IV1UNG_7 ..... T .....
DEB3_B_IV1UNG_8 ..... T .....
DEB3_B_IV1UNG_9 ..... T .....

```

DEB3 continued

J01415
 16000 16010 16020 16030 16040 16050 16060 16070 16080 16090 16100 16110 16120 16130 16140 16150 16160 16170 16180 16190 16200
 TAA GATTCTAATTAAACTATTCTCTGTTCTTTTCATGGGGAAGCAGATTGGGGTACCGCAAGTATTGACTCACCCTCAACGACCGCTATGTATTGTTGACATTAAGTGGCAGCCACCATGAATATTGTACGGTACCATAAACTTGACCACTGTAGTACATAAAAACCAATCCACATCAAAAACCCCTCCCATGCTTACAA
 FLO1_A_II_1 T T
 FLO1_A_II_2 T T
 FLO1_A_II_7 T T
 FLO1_A_II_8 T T
 FLO1_A_II_12 T T
 FLO1_A_II_5 T T
 FLO1_A_II_3 T T
 FLO1_A_II_4 T T
 FLO1_A_II_9 T T
 FLO1_A_II_10 G T R
 FLO1_B_II1UNG_2 T A
 FLO1_B_II1UNG_4 T A
 FLO1_B_II1UNG_10 A
 FLO1_B_II1UNG_9 A
 FLO1_B_II1UNG_7 A
 FLO1_B_II1UNG_11 A
 FLO1_B_II1UNG_6 A
 FLO1_B_II1UNG_3 A
 FLO1_B_II1UNG_5 A
 FLO1_B_II1UNG_12 A
 FLO1_A_III_3 A T TT
 FLO1_A_III_5 A T TT
 FLO1_A_III_7 A T TT
 FLO1_A_III_2 T A C
 FLO1_A_III_6 A C
 FLO1_A_III_9 A C T T
 FLO1_A_III_8 T A C
 FLO1_A_III_10 A C
 FLO1_A_III_11 T A C
 FLO1_A_III_12 A C
 FLO1_B_III1UNG_2 A C
 FLO1_B_III1UNG_3 A C
 FLO1_B_III1UNG_4 A C
 FLO1_B_III1UNG_6 A C
 FLO1_B_III1UNG_7 A C
 FLO1_A_III4_1 A C
 FLO1_A_III4_2 A C
 FLO1_A_III4_4 A C
 FLO1_A_III4_5 A C
 FLO1_A_III4_7 A C
 FLO1_A_III4_8 A C
 FLO1_A_III4_9 A C
 FLO1_A_III4_10 A C
 FLO1_B_III4UNG_1 A C
 FLO1_B_III4UNG_2 A C
 FLO1_B_III4UNG_5 A C
 FLO1_B_III4UNG_7 A C
 FLO1_B_III4UNG_6 A C
 FLO1_B_III4UNG_8 A C
 FLO1_B_III4UNG_10 A C
 FLO1_A_IV1_1 A C
 FLO1_A_IV1_2 A C
 FLO1_A_IV1_4 A C
 FLO1_A_IV1_6 A C
 FLO1_A_IV1_7 A C
 FLO1_A_IV1_10 A C
 FLO1_A_IV1_8 A C
 FLO1_A_IV1_3 A C
 FLO1_B_IV1UNG_3 A C
 FLO1_B_IV1UNG_5 A C
 FLO1_B_IV1UNG_8 A C
 FLO1_B_IV1UNG_11 A C
 FLO1_B_IV1UNG_4 A C
 FLO1_B_IV1UNG_9 A C
 FLO1_B_IV1UNG_12 A C
 FLO1_B_IV1UNG_10 A C
 FLO1_B_IV1UNG_11 A C

FLO1

Fig. S2. Sequence alignments of the six Neolithic samples.

Names for the sequenced clones are given as follows:

SAMPLENAME_Extraction as letters_primer pair as Roman numbers/PCR number(use of UNG when mentioned)_number of clone.

References and Notes

- S1. A. Rohl, B. Brinkmann, L. Forster, P. Forster, *Int. J. Legal Med.* **115**, 29 (2001).
- S2. <http://www.ncbi.nlm.nih.gov/Genbank/index.html>. 2005.
- S3. Estonian Biocentre Database 2005. Richard Villems, unpublished data.
- S4. A. Torroni *et al.*, *Genetics* **144**, 1835 (1996).
- S5. N. Izagirre, R. C. de la, *Am. J. Hum. Genet.* **65**, 199 (1999).
- S6. C. Lalueza-Fox *et al.*, *Proc. R. Soc. Lond B Biol. Sci.* **271**, 941 (2004).
- S7. H. J. Bandelt, L. Quintana-Murci, A. Salas, V. Macaulay, *Am. J. Hum. Genet.* **71**, 1150 (2002).
- S8. <http://fluxus-engineering.com>. 2005.
- S9. H. J. Bandelt, P. Forster, A. Rohl, *Mol. Biol. Evol.* **16**, 37 (1999).
- S10. T. Kivisild *et al.*, *Curr. Biol.* **9**, 1331 (1999).
- S11. M. Richards *et al.*, *Am. J. Hum. Genet.* **67**, 1251 (2000).
- S12. F. Calafell, P. Underhill, A. Tolun, D. Angelicheva, L. Kalaydjieva, *Ann. Hum. Genet.* **60** (Pt 1), 35 (1996).
- S13. G. Laval, L. Excoffier, *Bioinformatics.* **20**, 2485 (2004).
- S14. M. Tremblay, H. Vezina, *Am. J. Hum. Genet.* **66**, 651 (2000).
- S15. C. Renfrew, P. Bahn, *Archaeology. Theory and Practice* (Thames and Hudson, London, 1996).
- S16. D. Serre *et al.*, *PLoS. Biol.* **2**, 313 (2004).
- S17. E. Willerslev, A. Cooper, *Proc. Biol. Sci.* **272**, 3 (2005).
- S18. M. T. P. Gilbert, H.-J. Bandelt, M. Hofreiter, I. Barnes, *Trends in Ecology & Evolution* In press (2005).
- S19. A. Brehm, L. Pereira, T. Kivisild, A. Amorim, *Hum. Genet.* **114**, 77 (2003).
- S20. M. M. Baig, A. A. Khan, K. M. Kulkarni, *Ann. Hum. Genet.* **68**, 453 (2004).
- S21. M. Bermisheva, K. Tambets, R. Villems, E. Khusnutdinova, *Mol. Biol. (Mosk)* **36**, 990 (2002).
- S22. C. Babalini *et al.*, *Eur. J. Hum. Genet.* **13**, 902 (2005).

- S23. M. V. Derenko *et al.*, *Ann. Hum. Genet.* **67**, 391 (2003).
- S24. P. Forster, University of Hamburg (1996).
- S25. A. Helgason *et al.*, *Am. J. Hum. Genet.* **68**, 723 (2001).
- S26. T. Kivisild *et al.*, *Am. J. Hum. Genet.* **75**, 752 (2004).
- S27. Z. Koledová, Comenius University, Bratislava (2005).
- S28. M. Krings *et al.*, *Am. J. Hum. Genet.* **64**, 1166 (1999).
- S29. A. Knight *et al.*, *Curr. Biol.* **13**, 464 (2003).
- S30. P. Lahermo *et al.*, *Hereditas* **132**, 35 (2000).
- S31. M. Metspalu *et al.*, *BMC. Genet.* **5**, 26 (2004).
- S32. J. L. Mountain *et al.*, *Am. J. Hum. Genet.* **56**, 979 (1995).
- S33. S. H. Opdal *et al.*, *Acta Paediatr.* **87**, 1039 (1998).
- S34. G. Passarino *et al.*, *Eur. J. Hum. Genet.* **10**, 521 (2002).
- S35. H. Pfeiffer, P. Forster, C. Ortmann, B. Brinkmann, *Int. J. Legal Med.* **114**, 169 (2001).
- S36. B. Pakendorf *et al.*, *Am. J. Phys. Anthropol.* **120**, 211 (2003).
- S37. M. Poetsch, H. Wittig, D. Krause, E. Lignitz, *Forensic Sci. Int.* **137**, 125 (2003).
- S38. H. Pfeiffer *et al.*, *Int. J. Legal Med.* **112**, 291 (1999).
- S39. L. Pereira, M. J. Prata, A. Amorim, *Ann. Hum. Genet.* **64**, 491 (2000).
- S40. L. Quintana-Murci *et al.*, *Am. J. Hum. Genet.* **74**, 827 (2004).
- S41. F. X. Ricaut, C. Keyser-Tracqui, J. Bourgeois, E. Crubezy, B. Ludes, *Hum. Biol.* **76**, 109 (2004).
- S42. A. Sajantila *et al.*, *Genome Res.* **5**, 42 (1995).
- S43. K. Tambets *et al.*, in *Archeogenetics: DNA and the population prehistory of Europe*, C. Renfrew, K. Boyle, Eds. (McDonald Institute for Archaeological Research, Cambridge, 2000), chap. 25, pp. 219-236.
- S44. E. Watson, P. Forster, M. Richards, H. J. Bandelt, *Am. J. Hum. Genet.* **61**, 691 (1997).
- S45. U. Müller, *Jahresschrift für mitteldeutsche Vorgeschichte* **85**, (2002).
- S46. T. Autze, *Archäologie in Sachsen-Anhalt* **3**, in press (2005).

- S47. C. Meyer, O. Kürbis, K. W. Alt, *Jahresschrift für mitteldeutsche Vorgeschichte* **88**, 31 (2004).
- S48. I. Richter, *Mainzer Zeitschrift* **63/64**, 158 (1968).
- S49. A. Whittle, L. Bartosiewicz, D. Boric, P. Pettitt, M. Richards, *Anteus* **25**, 63 (2002).
- S50. P. Endicott *et al.*, *Am. J. Hum. Genet.* **72**, 178 (2003).
- S51. A. C. Stone, M. Stoneking, *Am. J. Hum. Genet.* **62**, 1153 (1998).
- S52. O. Handt, M. Krings, R. H. Ward, S. Paabo, *Am. J. Hum. Genet.* **59**, 368 (1996).