www.sciencemag.org/cgi/content/full/1176869/DC1

Supporting Online Material for

Genetic Discontinuity Between Local Hunter-Gatherers and Central Europe's First Farmers

B. Bramanti,* M. G. Thomas, W. Haak, M. Unterlaender, P. Jores, K. Tambets, I. Antanaitis-Jacobs, M. N. Haidle, R. Jankauskas, C.-J. Kind, F. Lueth, T. Terberger, J. Hiller, S. Matsumura, P. Forster, J. Burger

*To whom correspondence should be addressed. E-mail: bramanti@uni-mainz.de

Published 3 September 2009 on *Science* Express
DOI: 10.1126/science.1176869

This PDF file includes:

Genetic discontinuity between local hunter-gatherers and Central Europe's first farmers

Bramanti, B.*, Thomas, M.G., Haak, W., Unterlaender, M., Jores, P., Tambets, K., Antanaitis-Jacobs, I., Haidle, M.N., Jankauskas, R., Kind, C.-J., Lueth, F., Terberger, T., Hiller, J., Matsumura, S., Forster, P., Burger J.

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

Samples. Samples were taken from 36 Stone Age non-farming individuals (Table S1) mainly from locations in Central, northern Europe and northeastern Europe (fig. 1). One Mesolithic individual is from the Netherlands (Hardinxveld –Giessendam, De Bruin G1; *S1*). Remains of four Central German individuals (Bad Dürrenberg 1 and 2, Unseberg and Bottendorf) were excavated in the 1930's (*S2*). 10 individuals are from the hunter-gatherer site Ostorf of Northern Germany. Ostorf is considered a Mesolithic enclave that is surrounded by the Neolithic "Trichterbecher" (Funnel Beaker) Culture (*S3*). Five samples are from three caves in South Germany: the samples from Hohlenstein-Stadel were taken from two skulls out of a skull burial, whereas the individuals from Hohle Fels and Falkensteiner Höhle are represented by isolated bones. Other non-farming samples are from Lithuania and North East Poland. Here the Neolithic transition process is rather complex (*S4-S6*). The actual transition to a farming economy based on animal husbandry and plant domestication appears not to have started before *ca*. 2,500 BC. Further, we included two pre-farming individuals from the Samara region of Russia, attributed to the Yelshanskaya Culture (8000-7000 calBC). Their attribution to the Neolithic by some authors is based solely upon the use of ceramics, and no other indication of farming economy is known for this culture.

Anti-contamination measurements, sample preparation and DNA extraction. Palaeogenetic analysis was conducted in the ancient DNA facilities in Mainz, Germany. The pre-PCR laboratories are situated in a humanities building located separately from the post-PCR area. Pre-PCR laboratories have dedicated equipment

and are subjected to overnight UV-light exposure. Within the pre-PCR laboratory suite, sample preparation is physically separated from DNA extraction and from PCR setup facilities. Workers approached the pre-PCR area only after having had a shower and wearing freshly laundered clothes. When entering the laboratory changing rooms, they stripped and changed into disposable overalls, gloves, over-shoes, surgical face masks and plastic face-shields and followed an irreversible sequence of work steps to avoid contamination. Frequent surface washing with UV-irradiated water and soap followed by bleaching was routinely performed after each work step (*S7*).

Samples were provided by various archaeologists/anthropologists from different universities/institutions. After informed consent, five of them (Table 1) additionally provided swab samples of the archaeologists and anthropologists who had access to the ancient specimens or prepared the samples for the supply. Generally, no special precautions had been taken to prevent DNA contamination during excavation and handling before DNA analysis. Nevertheless, in no case did the DNA sequence of the archaeologists and anthropologists (Table S2) match those obtained from the prehistoric individuals. After arrival in the laboratory, samples were submitted to decontamination procedures consisting of 45-60 min UV irradiation of each side, mechanically removing of the outer layer by using an electric drill (Dremel, Racine, WI) and again UV irradiation. The samples were then finely powdered by using a mixer mill (Retsch, Haan, Germany) and stored at 4 °C until use.

Aliquots of 0.3-1 g powder were incubated overnight on a rotary mixer at 37 °C in a decalcification and digestion solution consisting of 0.5 M EDTA (pH 8.3), 0.5 % N-Laurylsarcosine and 1-30 U Proteinase K. DNA was extracted by using phenol/chloroform/isoamyl alcohol (25:24:1, Roth, Germany) then desalted and concentrated using 50 or 100 kDa Centricon micro-concentrators (Amicon, Millipore). When possible, at least two independent extracts were obtained from two different samples of each individual (see Table S4). In addition to blank extraction controls, animal samples (*Bos taurus)* were processed as controls.

Diagenetic analysis. Eight samples were submitted to diagenetic analyses measuring the degree of mineral alteration in the bone sample (*S8-S10*; Table S4). Small-angle X-ray scattering (SAXS) analysis on a laboratory apparatus (Bruker AXS

NanoSTAR) at Cardiff University was conducted according to previously published procedures (*S10*). Samples were ground into fine powder using an agate mortar and pestle and loaded into the vacuum sample chamber of the NanoSTAR. Measurements were taken for 3 hours with an X-ray wavelength of 1.54 Å (CuKα) and a sample-to-detector distance of 1.25m. Data was spherically averaged into one-dimensional traces, a background image was subtracted, and crystallite thickness values were calculated as described elsewhere (*S8, S10*). Crystal thickness values for all samples but one fell in the physiological range, described as below 5 nm thick in the smallest dimension; the one exception was Unseburg which had a crystal thickness of 5.5 nm, outside the physiological range. Consistently, this sample did not yield unambiguous mitochondrial sequences. All other samples showed diagenetic characteristics in line with previously studied successful ancient DNA samples (*S10*), leading us to expect successful DNA amplification from the current samples. Exceptions were Donkalnis 4 and Spiginas 3. Although showing positive diagenesis results, results were not replicable, most probably due to the presence of contaminant molecules.

mtDNA amplification: Overlapping PCR-amplicons were combined in different sets to span the entire mitochondrial HVS-I region (np 15997-16409) (Table S3). The amplification reaction was set up in a final volume of 50 µL, containing 1x PCR Gold Buffer (Applied Biosystems, Darmstadt, Germany), 2.5-3.5 U of AmpliTaq Gold® DNA Polymerase (Applied Biosystems), 2.5 mM $MgCl_2$ (Applied Biosystems), 0.2 mM dNTP mix (MBI Fermentas GmbH, St. Leon-Rot, Germany; Qiagen GmbH, Hilden, Germany), 0.2 µM each primer (Biospring, Frankfurt am Main, Germany), 20 µG BSA (Roche, Germany), and 1-8 µL of ancient DNA extract. A few amplifications were carried out on extracts treated with *Escherichia coli* uracil-N-glycosylase (UNG) which cleaves deaminated cytosine residues (*S11*); 1 U was added directly to the PCR reaction mixture and incubated at room temperature for 40 min. The amplifications were carried out in a Mastercycler gradient (Eppendorf, Hamburg, Germany). The PCR cycle conditions consisted of an initial denaturation at 92-94 °C for 6 min., 38-55 cycles of 30 s for 1 min. at 92-94 °C, 30 s for 1 min. at 53-58 °C and 30 s for 1 min. at 72 °C, followed by a final extension at 60 °C for 30-60 min.

Amplifications of clade-defining mtDNA coding region polymorphisms and subsequent digestions (Table S4) were performed as described (*S12-S14*).

Additionally, several informative SNP sites (*S15)* were amplified in a two step multiplex approach (Table S3) in cases where only a small amount of sample material was available (Hohlenstein-Stadel 5830a and 5830b; Hohle Fels 49Ib1 66 and 10Ic 405; Falkensteiner Höhle).

Amplification sets resulting in one or more of the multiple controls (mock extraction, cattle bone, PCR blank) showing bands on an agarose gel were generally discarded. Only in a few cases where sample material was limited, controls were cloned and sequenced for comparison. Whenever the cloned sequences from the control PCRs showed a sequence observed in a specimen, all simultaneously produced PCR-products were discarded. We note that sequences that were finally identified as contamination could never be replicated in different extractions and amplifications.

Cloning and sequencing: Amplification products were purified using the Invisorb Rapid PCR Purification Kit (Invitek, Berlin-Buch, Germany) according to the manufacturer's instructions. Successfully amplified fragments were directly sequenced and/or cloned and sequenced (see Tables 1 and S4). Cloning was performed via ligation of the amplicons with a pUC18-T-vector (overhanging T, own modification) followed by transformation of the resulting constructs into an *E. coli* culture (RRI). A variable number of colonies (8-15) were picked directly into PCR (50 µL final volume), containing 1x PCR Buffer (ABgene™ , Hamburg, Germany), 0.5-1 U DNA Polymerase (ABgene™), 2.5 mM $MgCl_2$ (ABgene™), 0.2 mM dNTP mix (Fermentas), 0.2 µM each of universal M13 forward and reverse primer. Cycle conditions consisted of an initial denaturation at 94 °C for 5-15 min, 25-30 cycles of 30 s at 94 °C, 30 s at 58 °C and 30 s at 72 °C. Amplicons carrying the expected insert (detected by 2 % agarose gel electrophoresis) were purified as mentioned above and sequenced with Big Dye Terminator chemistry (v3.1, Applied Biosystems) using 25 cycles at 92 °C for 30 s, 15 s at 55 °C (universal M13 primer) and 2.5 min at 72 °C. The analysis of the sequencing products were performed by capillary electrophoresis on ABI PRISM™ 310 and ABI PRISM™ 3130 Genetic Analyzers (Applied Biosystems). Sequences were further analyzed using the programs *Seqman II*™ and *MegAlign*™ from the DNA Star Software package (version 4.05 and 7). The obtained clone sequences are reported in Table S7, whereas in Table S6 the sequences of Neolithic farmers (*S7*, *S16*) are listed for comparison.

Validation of ancient DNA data: Contamination of mitochondrial DNA cannot be completely ruled out when PCR protocols with high cycle numbers are applied. However, to reduce the number of contaminating molecules the above mentioned lab procedures were applied. Further, multiple independent extractions and PCR amplifications from different parts of each skeleton were carried out whenever possible, mostly followed by bacterial cloning to detect PCR product heterogeneity. Additionally, STR-profiles of 3 individuals were established, i.e. whenever amplifiable nuclear DNA was present (Table S5). We further STR genotyped archaeologists and geneticists who had been in contact with the specimens to monitor for post-excavation contamination. Moreover, results of molecular sexing of the skeletons were compared with morphological sex determination (Table S1), and showed to be consistent, adding confidence to the authenticity of our DNA data.

Eight samples were submitted to diagenesis testing (*S10*). The structure of the mineral crystallites as seen by SAXS was comparable with that of recently buried bone in all but one sample (Table S1). Thus, the values of mineral crystallite characteristics are within the range that according to experience makes ancient DNA preservation feasible.

For mtDNA-sequence determination, we considered those substitutions as consistent that were observed in clone sequences from different amplifications and, where possible, from independent extracts obtained from anatomically different positions in the skeleton (see Tables S1 and S4). As we already outlined, we cannot exclude that any single prehistoric human was systematically contaminated across the whole skeleton by an unknown modern contaminator. But we consider this extremely unlikely as we have never found sequences that match those of the co-workers or the archaeologists or the anthropologists, incidentally showing that our sample pre-treatment in combination with our cleaning procedures are highly effective. Contamination is present among the clones, but is recognizable since contaminating sequences were not reproducible from different extracts. However, even in the unlikely case of one or two skeletons being contaminated entirely – for example, by extensive handling or washing after the excavation – this would not significantly influence the results of our rather conservative simulations.

The high observed sequence diversity among the recovered U5-types makes contamination of the samples from a single source biologically impossible. An additional argument for the authenticity of our sequences from 20 archaeological specimens is the fact that their mtDNA clade distribution looks different to what we would expect from any random contaminating population. This is because 13 out of 20 types belong to the clade U5, with a total frequency of 65.0 % (95 % CI: 40.8-84.6 %). Even the lower end of this frequency estimate (40.8 %) is considerably higher than the modern frequency of U5 in Europe.

As an additional contamination control, primers with 3´-specificity for U5 sites (189c, 192t, and 270t; amplicons 2a, 2b, 3a, and 3c, Table S3), were used to amplify various PCR blank controls. In 5 % (0-11.75 %) those controls yielded amplicons. Although rather high, this contamination rate of 5 % excludes a systematic laboratory U5 background. Further, bones that previously yielded U5 types with universal HVS-I primers, yielded the same types using the U5-specific primers.

Lastly, a comparison of the ancient sequences obtained with the human genome and known nuclear insertions did not show chromosomal sequences sharing a MaxIdent score over 95% with our sequences. Thus, we can exclude, with high confidence, any nuclear insertion of mitochondrial DNA (numts) (*S17*).

Negative and discarded results

Two samples from Lithuania (Plinkaigalis and Sventoji) and one from the Netherlands (De Bruin G1) were excluded from further analysis after several initial PCR-attempts were negative. While for Plinkaigalis we could not find any taphonomic explanation, for Sventoji we can argue that DNA was completely destroyed by the conditions of humidity and acidity of the peat-bog where the specimen was found. The HVS-I sequences of nine specimens (Donkalnis 2, Donkalnis 4, Dudka 1, Bad Dürrenberg 1, Unseburg, Bottendorf, Ostorf SK11, Ostorf SK3, Ostorf SK20a) were excluded since multiple sequences were present indicating the presence of contaminating molecules (data not shown). For two samples (Hohle Fels 10Ic 405 and Falkensteiner Höhle FH) that were only sequenced directly and not cloned, we could not distinguish among original mutations and possible deaminations. Thus the HVS-I sequences were not determined. In the case of Spiginas 3, a low

average deamination rate of 0.9 ‰, which has never been observed in authentic ancient DNA sequences, clearly identified the amplicons as modern contamination. The Polish sample Drestwo 1 was also discarded since both the HVS-I sequences and the STR-genotype could not be amplified from a second extract.

Positively established mtDNA- and genotypes

Samples Kretuonas 1, Kretuonas 3 and Donkalnis 1, and Hohlenstein-Stadel 5830b, showed the same mtDNA type: 16192t 16270t. Thus, we had to consider the possibility of cross-contamination or contamination from the archaeologists. We conclude that both possibilities are highly unlikely, because: i) the samples were analyzed at different times with little possibility of cross-contamination between all of them; ii) various other samples with different mtDNA types were co-analyzed in varying order and never showed this mtDNA type; iii) in hundreds of animal and human specimens analyzed so far, we never experienced a consistent reciprocal contamination between individuals; iv) a contamination by a single archaeologist can be excluded for 3 out of four specimens, since they originate from different archaeological excavations; v) the two individuals from Kretuonas yielded consistent STR-genotypes that were different from each other (Table S5); vi) the amelogenin genotype showed that those two were of different sex, which is in accordance with a previous anthropological examination (Table S1); vii) one person working in the Mainz ancient DNA laboratories sharing the U5 mtDNA-type of the four samples (Kretuonas 1, Kretuonas 3, Donkalnis 1 and Hohlenstein-Stadel 5830b) can be excluded as a contaminant source for two of them because of differing STR-profiles (Table S5). As expected for highly degraded DNA, only the shorter STRs could be reproducibly amplified from different extracts and amplifications (*S18, S19*). Although a close relationship between individuals Kretuonas 1 and Kretuonas 3 could not be ruled out the two STR profiles differed from each other and from each modern STR profile typed in our lab. Contamination of Donkalnis 1, which yielded no STR results, cannot be excluded. But contamination is unlikely, since a recent contamination by laboratory personnel should yield both mitochondrial and nuclear DNA, which is not the case; viii) cross-contamination of Hohlenstein-Stadel 5830b is extremely unlikely because it has been analyzed 2-3 years after the other samples,

shows an extremely good preservation status, and its HVS I-type was confirmed by a series of 5 coding region SNPs.

The other two pairs of identical sequences (Ostorf SK8d and Ostorf SK35, showing 16270t; Ostorf SK12a and Ostorf SK18 showing 16093c 16126c 16153a 16294t) originate from the same archaeological site but were taken during different excavations (in 1904 and 1961 for the first pair and 1904 and 1935 for the second). Contamination from subsequent handling cannot be excluded, but is unlikely since the skeletons were not washed but covered by soil when they arrived in the lab. Hence, we consider biological kinship to be more likely in this case.

Phylogeography of mtDNA types. Two samples from Ostorf (SK8d and SK35) show HVS I - type 16270t which was confirmed by RFLP analysis to belong to clade U5. The sequence is found at low frequency throughout Europe and has been considered the ancestral sequence type of this clade, dated 45,000-52,000 YBP (95 % CI; *S20*).

The U5b type with the HVS-I motif 16192t 16270t was observed in three Eastern Baltic (Lithuanian) and one Central European (South-West German) prehistoric Specimen (Kretuonas 1, Kretuonas 3, Donkalnis 1 and Hohlenstein-Stadel 5830b).

The exact matches or close derivatives of that sequence type have also been identified in modern populations of the region (*e.g. S21, S22, S23*). This motif, which is confirmed by SNP analysis (14793A) to belong to the subclade U5b, is present both in the phylogenetic branches U5b1 and U5b2 (*S22, S24*; see also Figure S1A). While the type 16192t 16270t belonging to U5b1 seems to be evenly distributed throughout Europe, those belonging to U5b2 tend to be more characteristic of western and southern Europe and Scandinavia. However, at the present resolution, with the exception of Hohlenstein- Stadel 5830b, the exact location of those samples on a phylogenetic tree remains unresolved (Figure S1A). For this sample, the SNP 1721T indicates that it belongs to the subclade U5b2. 1721T was also observed in the Falkensteiner Höhle sample, whose HVS-I type was not determined.

The HVS-I sequence of Dudka 2 (Poland) is 16189c 16270t. It is classified here as U5b1. This subclade has been observed in several European populations, including

those of the study region of prehistoric samples (Figure S1A). The individual Dudka 3 (Poland) presents three consistent mutations at nps 16189c 16265g 16270t. It is a one-step derivative of the mtDNA type found in Dudka 2 (Figure S1) and most likely belongs to U5b1 as well.

Three out of five U5a sequences (Drestwo 2, Chekalino IVa and Hohlenstein-Stadel5830a) showed identical or close matches with those common among extant Europeans. The sequence of Drestwo 2 from North-Eastern Europe (Poland) with HVS-I motif 16192t 16256t 16270t has been observed among contemporary Polish and in surrounding populations (Figure S1A). The type of the skeletal individual Chekalino IV from Eastern Europe (Russia) with its additional transition at np16294 has been earlier described for example among modern-day Germans and Czechs; more derived variants of this sequence are present also among Latvians and Lithuanians (Figure S1A). Hohlenstein-Stadel 5830a belongs to a sub-branch U5a2a (*S25*). The clade is characterized among other diagnostic mutations by HVS-I transversion at np 16114 from C to A (Figure S1B). This particular aDNA sequence has lost the transition at np 16270 and is thus two steps away from a large node shared by many populations. Its closest relatives in modern populations can be found for example among Latvians and Tatars. The two samples Lebyazhinka IVa (Central Russia) and Ostorf SK19 (Germany) show specific U5a-types (16192t 16241c 16256t 16270t 16399g and 16168t 16192t 16256t 16270t 16302g). Due to the substitution at np 16399 Lebyazhinka IVa can be classified into U5a1.

Two samples (Hohle Fels 49Ib166 and 10Ic405) were attributed by SNPs analysis to clade U. For Hohle Fels 49Ib166 additionally, the HVS-I region was determined. It shows no difference from the reference sequence (rCRS). All remaining U types (Spiginas 4, Lithuania and Bad Dürrenberg 2, Eastern Germany), with a transition at np 16356 (clade U4) can be observed everywhere across Europe; While its frequency is high in Eastern Europe and Western Siberia (*S26-S28*), this lineage is also common in Northern Europe.

Two individuals (Ostorf SK12a, Ostorf SK 18) could be classified as clade T2e (*S25*). The two sequences showed an identical mutation pattern (16093c 16126c 16153a 16294t). Even though the same motif was found to be shared between Icelanders and Scottish Western Islanders (*S29*), the clade T2e seems generally to be more prevalent in the Southwestern fringe of Europe (*S30*).

The other sequences identified in the sample of Ostorf are 16069t 16126c (SK45a), belonging to clade J, and 16224c 16311c (SK28a) belonging to clade K. All DNA sequences are deposited under accession number bankit 1259945 in GenBank.

Summary Statistics: An analysis of the molecular variance was performed on 349bp of HVS-I sequence (np 16,052-16,400 using Arlequin 3.1 (*S31*). Sequences were assigned to one of 3 populations, hunter-gatherers (n=20), Neolithic farmers (n=25) (*S7, S16* and Table S6) and modern Central Europeans (n=484). Transformed pairwise genetic distances, F_{ST}, were estimated according to the formula of Reynolds (*S32*), and P-values were obtained by 10,000 permutations.

Further, we calculated F_{ST}s between the following modern populations to provide a comparison for F_{ST} values involving our ancient samples: Sicily, Scotland, Papua New Guinea, Australia, S-Germany, England & Wales, Spain, India, Syria, Greece, Czech Republic, China, Slovenia. For each geographic location mtDNA sequence homologous to the region used for our ancient DNA analysis was taken from 20 individuals chosen at random from a larger database (Peter Forster, Cambridge). F_{ST}s were calculated as for our ancient DNA comparisons using the pairwise method (see above).

Coalescent Simulation: To test if the observed genetic differences between our population samples, as measured by F_{ST}, could be explained under the null-hypothesis of genetic drift over time in a continuous population, we performed coalescent simulations using Bayesian Serial Sim Coal (*S33*). We assumed a single population of current size N=12 million, having grown exponentially over the last 300 generations (7,500 years assuming a generation time of 25 years) from a Neolithic population of unknown female effective size N_N. Values of N_N ranging from 1,000 to 100,000 were explored. The Neolithic population was further assumed to have grown exponentially over the proceeding 1500 generations (37,500 years assuming a generation time of 25 years) from an Upper Paleolithic European founding population of unknown female effective population size N_{UP}. Values of N_{UP} ranging from 10 to 5,000 were explored. The Upper Paleolithic European founding population was sampled from a constant ancestral African effective population of 5000 females. 349bp sequences were

evolved on the simulated coalescent trees assuming a mutation rate of 7.5 x 10^{-6} per site per generation (*S34*), a transition bias of 0.9841 and a continuous gamma distribution of mutation rates among sites, of parameter 0.205 (*S35*). 10,000 coalescent simulations were performed for each combination of ancestral population sizes (N_{UP} and N_N), giving a total of 100,000,000 simulations. Three sets of sequences were sampled from the coalescent simulations according to the numbers (hunter-gatherers = 20, farmers = 25, moderns = 484) and dates (see Table 1 and S1) of the observed sequence data, and pairwise F_{ST} values calculated as described above. Finally, the proportion of times where the simulated F_{ST} was greater than that estimated from the observed data ($P_{S>O}$) was recorded for each combination of N_{UP} and N_N (Figure 2). Execution of Bayesian Serial Sim Coal was controlled and simulation results were harvested using scripts written in the programming language Python (*S36*). Results were analyzed and plotted using the statistical package 'R' (*S37*). We recognize that the mutation rate we have used is high compared to others published but we reason that this assumption is conservative as higher mutation rate will generate greater differences between populations sampled at different time periods, thus easing the burden of explaining the high observed F_{ST} values.

Construction of a network. A phylogenetic network (Figure S1) of U5 sequences was constructed manually and checked by the use of the Network 4.1.1.1 program (*S38*). Further information is given in the caption of Figure S1.

Analysis of STRs. Nuclear DNA was investigated with the kit AmpFlSTR® Profiler Plus® (PE Applied Biosystems, Foster City, CA, USA) under previously described conditions (*S18, S19, S39, S40*). Raw data are shown in Table S5.

Figure S1 (A and B) Phylogenetic network of mtDNA HVS-I sequences of clade U5 and the location of studied prehistoric samples in it. Only the main nodes and some of their closest derivatives in regard to the aDNA sequences are given. Numbers on links indicate the mutated nucleotide positions relative to rCRS (*S42*), the region between np 16090 and 16365 is considered. Back-mutations are prefixed with a sign '@'. Diagnostic coding and non-coding region mutations outside HVS-I are given in square brackets. Possible alternative pathways are shown with broken lines. aDNA samples are marked in red, the ones reported in this study are underlined. Blue letters indicate the data taken from literature; their possible subclade affiliations are deduced from HVS-I sequences. Violet letters are those for whom the diagnostic coding region markers have been typed (*S22*). Samples with complete mtDNA sequences are dark green (*S24*) or light blue (*S45*). The position of three Lithuanian and two aDNA German samples is left unresolved. The branch U5a2a is given on panel B, all other branches of U5 are shown on panel A. Populations are indicated by the following abbreviations:

Ad - Adygeys (*S24; S46*); Al - Albanians (*S47*); Ar - Armenians (*S20*); Az - Azeris (*S20*); Au - Austrians (*S48*); Ba - Bashkirs (*S27*); Bl - Bulgarians (*S20*); Bo - Bosnians (*S22; S49*); Bq - Basques (*S20; S50; S51*); Ch - Chuvash (*S22;S27*); Ci - Canary islanders (*S52*); Cr - Croatians (*S22*); Cz - Czechs (*S22; S53*; *S54*); Eg - Egyptians (*S55*); En - British, Welsh, Cornish (*S20; S29; S53; S56*); Es - Estonians (*S22; S58*); Fi - Finns (*S45; S53; S58; S59*; *S60*; *S61*); Fr - French (*S20; S22; S62*, *S63*, *S64*); Ge - Germans (*S53*; *S65*; *S66*; *S67*; *S68*); Hu - Hungarians (*S22*); Ic - Icelanders (*S53*; *S57; S58*); Iq – Iraqis (*S20*); Ira - Iranians (*S69*); Irl - Irish (*S20*); It - Italians (*S20; S22; S24; S70*; *S71*; *S72*); Ka - Karelians (*S58*); Kb - Kabardinians (*S46*); Ko - Komi (*S22; S27*); Kr - Greeks (*S20*); Ku - Kurds (*S20*); Kz - Kazakhs (*S14*); La - Latvians (*S22, S23*); Lt - Lithuanians (*S21*, *this study*); Ma - Mari (*S27*); Mn - Mongolians (*S73*); Mo - Mordvinians (*S22*; *S27*); Mr - Moroccans (*S22; S74*); Ms - Mansi (*S26*); Ne - Nenets (*S75*); No - Norwegians (*S29; S76*; *S77*; *S78*); Os – Northern Ossetians (*S46*); Pa - Palestinians (*S20*); Pl - Poles (*S82*); Po - Portuguese (*S79*); Ro -Romanians (*S20*); Ru - Russians (*S20; S22; S80*; *S81; S82*); Sa - Saami (*S58*, *S76*); Sc - Scots (*S29*); Se - Swedes (*S22*); Si - Sicilians (*S20*; *S22; S83*); Sp - Spanish (*S24; S51*; *S84*; *S85*); Sr - Sardinians (*S20; S86*); Sv - Slovenians (*S22; S49*; *S87*); Sw - Swiss (*S59*; *S88*); Ta - Tatars (*S27*); Tu - Turks (*S20*; *S22*); Ud - Udmurts (*S22; S27*); Uk - Ukrainians (*S22; S81*); WA - West-Africans (*S74*).

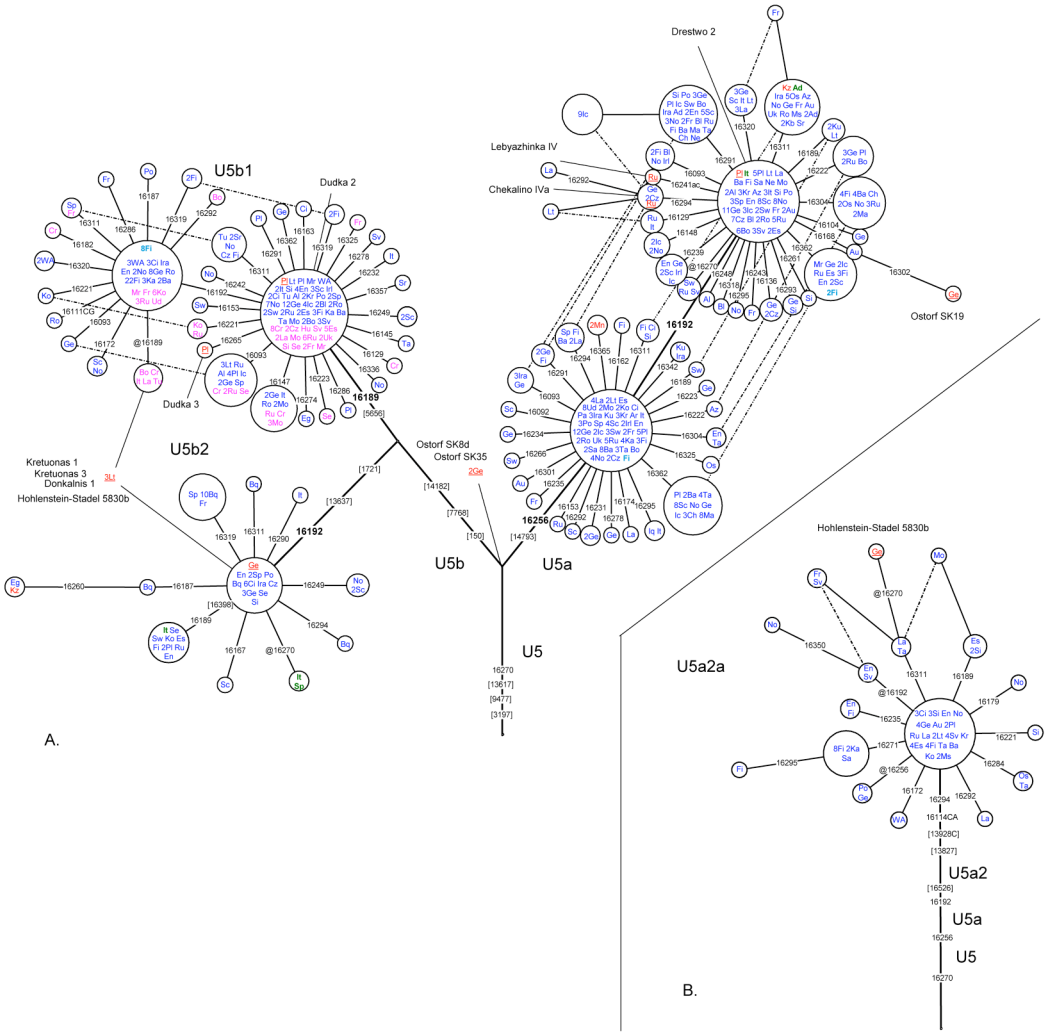
U5b1
Dudka 2
Dudka 3
U5b2
Kretuonas 1
Kretuonas 3
Donkalnis 1
Hohlenstein-Stadel 5830b
Ostorf SK8d
Ostorf SK35
Drestwo 2
Lebyazhinka IV
Chekalino IVa
Ostorf SK19
U5b
U5a
U5
A.
Hohlenstein-Stadel 5830b
U5a2a
U5a2
U5a
U5
B.

Table S1. Archaeological details on the 36 analysed individuals. (*) Radiocarbon dates were calibrated by employing the program CalPal on the basis of Intcal 04 (see tab. 1). Corrections of reservoir effects were applied where identified.

Country	Site, skeleton	Dating *	Archaeological context	Anthropological features	Sampled bones
Lithuania	Donkalnis 1	Kunda Culture? (*S6*)	burial	female individual, ca. 20 years old	Don 1a: Teeth 47, 48 Don 1b: Tooth 18
	Donkalnis 2	Kunda Culture ca. 6300 calBC (CAMS-85221) (*S6*)	burial	female individual, 20-25 years old	Don 2a: Tooth 38 Don 2b: Cranial fragment
	Donkalnis 4	Kunda Culture ca. 5900 calBC (OxA-5924) (*S6*)	burial	male individual, 50-55 years old	Don 4a: Femur Don 4b: Tooth 36
	Kretuonas 1	Narva Culture ca. 4200 calBC (OxA-5935)	burial	female individual, 20-25 years old	Kre 1a: Teeth 37, 38 Kre 1b: Femur Kre 1c: Tibia
	Kretuonas 3	Narva Culture ca. 4450 calBC (OxA-5926)	burial	male individual, 50-55 years old	Kre 2a: Teeth Kre 2b: Tibia
	Plinkaigalis	Corded Ware Culture ca. 2600 calBC (OxA-5928)	burial	female individual, mature	Pli 1a: Teeth 48/38, 32, 33 Pli 1b: Femur Pli 1c: Tibia Pli 1d: Humerus
	Spiginas 3	Kunda Culture ca. 6600 calBC (OxA-5925) (*S6*)	burial	probably female individual, adult	Spi 3a: Femur Spi 3b: Radius
	Spiginas 4	Kunda Culture ca. 6350 calBC (GIN-5571) (*S6*)	burial	female individual, 30-35 years old	Spi 4a: 5th metacarpal Spi 4b: Tooth 18 Spi 4c: Humerus
	Sventoji	Narva Culture	occasional find peat-bog layer		Sve 1a: Tooth 36 Sve 1b: Tooth 38
Poland	Drestwo 1	Kunda Culture ca. 4900 calBC (Ua-13084)	burial	male individual	Dre 1a: Tooth 18 Dre 1b: Bone
	Drestwo 2	ca. 2250 calBC (Ua-13085)	burial	male individual	Dre 2a: Tooth 18 Dre 2b: Bone
	Dudka 1	Globular Amphorae Culture , ca. 3500 calBC (Ki – 5718)	secondary burial with 3 individuals	adult	Dud 1a: Tooth 26 Dud 1b: Tooth 27
	Dudka 2	Zedmar culture?	multiple burial (4 individuals)	adult	Dud 2a: Tooth 38 Dud 2b: Tooth 47
	Dudka 3	Zedmar Culture	burial	adult	Dud 3a: Tooth 27 Dud 3b: Tooth 17 Dud 3c: Tooth 18
Russia	Chekalino IVa	Comb-Marked Pottery or Yelshanskaya Culture	burial	.	Che 4a: Tibia Che 4b: Femur Che 4c: Femur
	Lebyazhinka IV	Yelshanskaya Culture	burial	male	Le1 a: Tibia Le1 b: Tibia Le1 c: Tibia
Germany	Bad Dürrenberg 1	Late Mesolithic ca. 6850 calBC (OxA-3136)	double burial	Infans I	Bdü 1a: Femur Bdü 1b: Femur
	Bad Dürrenberg 2	Late Mesolithic ca. 6850 calBC (OxA-3136)	double burial	female, young adult	Bdü 2a: Tooth 33 Bdü 2b: Femur
	Unseburg	Late Mesolithic ca. 6550 calBC (OxA-2918)	burial	female, matur	Uns 1a: Femur Uns 1b: Tooth 38
	Bottendorf I	Late Mesolithic ca. 4850 calBC (OxA-2919)	burial	male, adult?	Bot 1a: Tooth Bot 1b: Tooth

	Ostorf SK11a	Funnel Beaker Culture ca. 3200 calBC	burial	male, 20-25 years old	SK11a-A: Femur, r SK11a-B: Femur, l
	Ostorf SK28a	Funnel Beaker Culture ca. 3200 calBC	burial	male, 35-45 years old	SK28a-A: Femur, r SK28a-B: Tibia, l
	Ostorf SK8d	Funnel Beaker Culture ca. 3200 calBC	burial	male ?, 15-20 years old	SK8d-A: Tibia, r, SK8d-B: Tibia, l,
	Ostorf SK3	Funnel Beaker Culture ca. 3000 calBC	burial	male, 8-10 years old	SK3-A: Femur, l, proximal SK3-B: Femur, l, distal
	Ostorf SK35	Funnel Beaker Culture ca. 3100 calBC	burial	female, 25-35 years old	SK35a-A: Femur, r SK35a-B: Femur, l
	Ostorf SK12a	Funnel Beaker Culture ca. 3000 calBC	burial	male, 25-35 years old	SK12a-A: Femur SK12a-B : Tibia
	Ostorf SK20a	Funnel Beaker Culture ca. 3000 calBC	burial	female, 25-40 years old	SK20a-A: Femur, l, prox. SK20a-B: Femur, l, dist.
	Ostorf SK45a	Funnel Beaker Culture ca. 3000 calBC	burial	?, 12-16 years old	SK45a-A: Femur, r SK45a-B: Femur, l
	Ostorf SK18	Funnel Beaker Culture ca. 3000 calBC	burial	male, 20-30 years old	SK18-A: Femur, r SK18-B: Femur, l
	Ostorf SK19	Funnel Beaker Culture ca. 2950 calBC	burial	male, 40-60 years old	SK19-A: Femur, r SK19-B: Femur, l
	Hohlenstein-Stadel, Lonetal, HS 5830a	Mesolithic, Beuronien, ca. 6743 ± 139 calBC (ETH-5732)	Skull burial with three skulls from a cave	male, adult	30a: Tooth 38
	Hohlenstein-Stadel, Lonetal, HS 5830b	Mesolithic, Beuronien, ca. 6743 ± 139 calBC (ETH-5732)	Skull burial with three skulls from a cave	female, adult	30b: Tooth 38
	Hohle Fels, Schelklingen, Alb-Donau-Kreis, HF49 Ib1 66	Magdalenian, ca. 13400 calBC	Isolated bones in a cave		49I: Femur, r
	Hohle Fels, Schelklingen, Alb-Donau-Kreis, HF10 Ic 405	Magdalenian ca. 13400 calBC	Isolated bones in a cave		10I: Femur, l
	Falkensteiner Höhle	Mesolithic 7220 ± 120 calBC (ETH-7615)	Isolated bones in a cave	male, 30-40 years old	FH: Post-cranial fragment
The Netherlands	De Bruin G1	Late Mesolithic ca. 5100 calBC	burial	male, 29-44 years old	Deb-A: Tibia Deb-B: Humerus

Table S2. MtDNA types of archaeologists and anthropologists. mtDNA positions were numbered according to the rCRS (*S42*), minus 16000. The sequence range is 15997-16409.

Country	Individual	HVS-I sequence
Lithuania	1	CRS
Lithuania	2	257t
Russia	1	224c 258g 311c
Poland	1	261t 311c
Poland	2	278t
Poland	3	a183c 189c 278t 356c
North Germany (Ostorf)	1	126c 163g 186t 189c 294t
North Germany (Ostorf)	2	192t 311c
North Germany (Ostorf)	3	256t 270t 399g
North Germany (Ostorf)	4	126c 163g 186t 189c 294t
North Germany (Halle)	1	078g 126c 294t 296
North Germany (Halle)	2	129a 172c 223t 311c 391a
North Germany (Halle)	3	235g 261t 291t 293g
North Germany (Halle)	4	291t 390a

Table S3. HVS-I mtDNA primers used for the analyses. All the targets were amplified at an annealing temperature of 58 °C, except the coding region targets (amplified at 56 °C).

Amplicon	Primer pairs (sequence 5'-3')	*Reference*
1	L15996 (CTCCACCATTAGCACCCAAAGC)	*S41*
	H 16142 (ATGTACTACAGGTGGTCAAG)	*S43*
1a	L16055 (GAAGCAGATTTGGGTACCAC)	*S44*
	H 16142 (ATGTACTACAGGTGGTCAAG)	*S43*
2	L16117 (TACATTACTGCCAGCCACCAT)	*S7*
	H16233 (GCTTTGGAGTTGCAGTTGATGTGT)	*S7*
2a (16189c)	L16189 (ATCCACATCAAAACCCCCC)	*This study*
	H16233 (GCTTTGGAGTTGCAGTTGATGTGT)	*S7*
2b (16207t)	L16117 (TACATTACTGCCAGCCACCAT)	*S7*
	H16270 (GTAGGTTTGTTGGTATCCTAA)	*This study*
3	L16209 (CCCCATGCTTACAAGCAAGT)	*S44*
	L16348 (ATGGGGACGAGAAGGGATTTG)	*S7*
3a (16270t)	L16209 (CCCCATGCTTACAAGCAAGT)	*S44*
	H16270 (GTAGGTTTGTTGGTATCCTAA)	*This study*
3b	L16192 (CACATCAAAACCCCCTCCT)	*This study*
	L16348 (ATGGGGACGAGAAGGGATTTG)	*S7*
3c (16192t)	L16192 (CACATCAAAACCCCCTCCT)	*This study*
(16270t)	H16270 (GTAGGTTTGTTGGTATCCTAA)	*This study*
4	H16287 (CACTAGGATACCAACAAACC)	*S44*
	L16410 (GCGGGATATTGATTTCACGG)	*S44*
5	L16209 (CCCCATGCTTACAAGCAAGT)	*S44*
	H16410 (GCGGGATATTGATTTCACGG)	*S44*
5a	L16310 (CCCACCCTTAACAGTACATAA)	*This study*
	H16439 (CCCGGAGCGAGGAGAGTAA)	*This study*
6	L16209 (CCCCATGCTTACAAGCAAGT)	*S44*
	H16303 (TGGCTTTATGTACTATGTAC)	*S44*
7	L16117 (TACATTACTGCCAGCCACCAT)	*S7*
	H16218 (TGTGTGATAGTTGAGGGTTG)	*S44*
8	L15996 (CTCCACCATTAGCACCCAAAGC)	*S41*
	H16410 (GCGGGATATTGATTTCACGG)	*S44*
Hg U: +12308 *Hin*fI	L12216 (CACAAGAACTGCTAACTCATGC)	*S13*
	H12338 (ATTACTTTTATTTGGAGTTGCACCAAGATT)	*S12*

Hg H: -7028 *Alu*I	L6909 (AAGCAATATGAAATGATCTG)	*S12*
	H7131 (CGTAGGTTTGGTCTAGG)	*S12*
Hg T: +13368 BamHI	L13190 (GCTTAGGCGCTATCACCAC)	*S14*
	H13384 (ATATCTTGTTCATTGTTAAG)	*S14*
M1: SNP 1719 G/A	L1678 (CTTAACTTGACCGCTCTGAGCTAAAC)	*This study*
	H1748 (CGCCAGGTTTCAATTTCTATCG)	*This study*
M1: SNP 8251 G/A	L8226 (TTCATGCCCATCGTCCTAGAATTAA)	*This study*
	H8253 (GAGGGGGTGCTATAGGGTAAATACG)	*This study*
M1: SNP 12372 G/A	L12324 (CCCCAAAAATTTTGGTGCAACT)	*This study*
	H12391 (TAGGGTTAACGAGGGTGGTAAGGA)	*This study*
M1: SNP 14766 C/T	L14717 (CAACCACGACCAATGATATGAAAAAC)	*This study*
	H14784 (GGAGGTCGATGAATGAGTGGTTAATT)	*This study*
M2: SNP 7028 C/T	L6989 (ACTGGCATTGTATTAGCAAACTCATC)	*This study*
	H7054 (CCTCCTATGATGGCAAATACAGCT)	*This study*
M2: SNP 8697 G/A	L8673 (GACTAATCACCACCCAACAATGACT)	*This study*
	H8710 (AAGAGATCAGGTTCGTCCTTTAGTGT)	*This study*
M2: SNP 13708 G/A	L13692 (TAACGAAAATAACCCCACCCTACTAAA)	*This study*
	H13740 (GATGCGGGGGAAATGTTGTTAGT)	*This study*
M2: SNP 14798 T/C	L14783 (ATACGCAAAACTAACCCCCTAATAAAA)	*This study*
SNP 14793 A/G	*H14839 (GCCAAGGAGTGAGCCGAAGTT)*	*This study*

Table S4. Results of physical and genetic analyses on the Stone Age samples. Notes: N.a. = not analysed, A = genome of the archaeologists available for comparison, D = diagenetical analysis, M = multiple extractions and number of these, C = clones of HVS-I and number of these, N = positive amplification of nuclear DNA; Rf = RFLP-Analysis; SNP = SNPs from the coding region of mtDNA obtained by means of multiplex amplification. The mtDNA was sequenced from np 15997 to np 16409. mtDNA positions were numbered according to the rCRS (*S42*), minus 16,000.

Country	Site, skeleton	Analyses	Crystal thick-ness (nm)	HVS-I sequence	RFLP / SNPs	Clade	Average DNA damage (‰)
Lithuania	Donkalnis 1	A, D, M4, C79, N, Rf, SNP	n.a.	192t 270t	+12308 *Hinf*I 14793A	U5b	5.6
	Donkalnis 2	A, M2, C105, N	n.a.	Multiple sequences	n.a.	-	n.a
	Donkalnis 4	A, D, M3, C112,	3.6	Multiple sequences	n.a.	-	n.a
	Kretuonas 1	A, M5, C56, N, Rf, SNP	n.a.	192t 270t	+12308 *Hinf*I 14793A	U5b	5.1
	Kretuonas 3	A, M4, C72, N, Rf, SNP	n.a.	192t 270t	+12308 *Hinf*I 14793A	U5b	5.0
	Plinkaigalis	A, M2	n.a.	-	n.a.		
	Spiginas 3	A, D, M3, C93,	3.6	Contamination?	n.a.	-	0.9
	Spiginas 4	A, M3, C109, Rf	n.a.	356c	+12308 *Hinf*I +7028 *Alu*I	U4	6.1
	Sventoji	A, M2	n.a.	-	n.a.		
Poland	Drestwo 1	D, M4, C93, N	3.6	[213a 270t] From a single DNA-extract.	n.a.	[U5]	4.3
	Drestwo 2	D, M4, C102, N, Rf	4.4	192t 256t 270t	+12308 *Hinf*I	U5a	2.2
	Dudka 1	A, M3, C90	n.a.	Multiple sequences	n.a.	-	-
	Dudka 2	A, M3, C80, N, Rf	n.a.	189c 270t	+12308 *Hinf*I	U5b1	5.8
	Dudka 3	A, M3, C127, Rf	n.a.	189c 265g 270t	+12308 *Hinf*I +7028 *Alu*I	U5b1	4.5
Russia	Chekalino IVa	A, D, M2, C83, Rf	3.8	192t 256t 270t 294t	+12308 *Hinf*I	U5a	2.2
	Lebyazhinka IV	A, D, M2, C60, Rf	3.4	192t 241a/c 256t 270t 399g	+12308 *Hinf*I	U5a1	n.a.
Germany	Bad Dürrenberg 1 (infans I)	A,M3, C130	n.a.	Multiple sequences	-	-	n.a.
	Bad Dürrenberg 2	A, D, M2, C 119, Rf	3.4	356c	+12308 *Hinf*I +7028 *Alu*I	U4	n.a.
	Unseburg	A, D, M2, C81	5.5	Multiple sequences	n.a.	-	n.a.
	Bottendorf	A, M2,	n.a.	Multiple sequences	n.a.	-	n.a.
	Ostorf SK11a	A, M3	n.a.	Multiple sequences	n.a.	-	n.a.

	Ostorf SK28a	A, M2, C18	n.a.	224c 311c	n.a.	K	n.a.
	Ostorf SK8d	A, M2, C16	n.a.	270t	+12308 *Hinf*I +7028 *Alu*I	U5	n.a.
	Ostorf SK3	A, M5	n.a.	Multiple sequences	n.a.	-	n.a.
	Ostorf SK35	A, M2	n.a.	270t	+12308 *Hinf*I +7028 *Alu*I	U5	n.a.
	Ostorf SK12a	A, M2	n.a.	093y 126c 153a 294t	n.a.	T2e	n.a.
	Ostorf SK20a	A, M2, C82	n.a.	Multiple sequences	n.a.	-	n.a.
	Ostorf SK45a	A, M2, C16	n.a.	069t 126c	n.a.	J	n.a.
	Ostorf SK18	A, M4	n.a.	093c 126c 153a 294t	n.a.	T2e	n.a.
	Ostorf SK19	A, M3	n.a.	168t 192t 256t 270t 302g	+12308 *Hinf*I +7028 *Alu*I	U5a	n.a.
	Hohlenstein-Stadel, 5830a	M1, SNP	n.a	114a 192t 256t 294t 311c	1721C 7028 T 12372 A 14766 T 14793 G	U5a	n.a.
	Hohlenstein-Stadel, 5830b	M1, SNP	n.a	192t 270t	1721T 7028 T 12372 A 14766 T 14793 A	U5b2	n.a
	Hohle Fels, 49 I b1 66	M2, SNP	n.a	CRS	1721C 7028 Y 12372 A 14766 T 14793 A	U	n.a
	Hohle Fels, 10 Ic 405	M2, SNP	n.a	Not determined	1721C 7028 T 12372 A 14766 T 14793 A	U	n.a
	Falkensteiner Höhle, FH	M2, SNP	n.a	Not determined	1721T 7028 T 12372 A 14766 T 14793 A	U5b2	n.a
The Netherlands	De Bruin G1	M2	n.a.	-	n.a	-	n.a

Table S5. STR-genotyping of three prehistoric individuals. The STR loci are listed in order of length of the amplicon, where D18 is the longest. Only reproducible results from at least two different extracts are considered for consensus construction. Brackets indicate uncertain allele determinations (*S19, S39*). Amel = Amelogenin, the XY locus for sex determination. MK indicates a co-worker sharing HVS-I sequence 192t 270t.

Ind.	Sample	Extr.	PCR	Amel	D3	D8	D5	vWA	D21	D13	FGA	D7	D18
Kretuonas 1	a	1	1	X	16/17	11/15	11/12	15/16			21		
	a	1	2	X	(15/16/17)	11/15	12	15/16	29		21		
	a	1	3	Y?	16/17	11/15	12	15/16	(28)	(14)	(21)		
	a	2	4	X	16/17	11/15	(12)	16	28/29		(20)		
	a	2	5	X	16/17	11/(15)	(12)	16	28/29	(13)	21		
	b	3	6						28/29				
	b	3	7	X		(10)/(14)/16	13	12					
	b	4	8			11/12					24		
	b	4	9										
	Consensus			**X**	**16/17**	**11/15**	**12/12**	**(15)/16**	**28/29**		**21/(21)**		
Kretuonas 3	a	1	1	XY	15/17	(13)/14/(15)	11/12	15/16					
	a	1	2	XY	15/17	9/14	11/12		28	10/14	20		
	a	1	3	XY	15/17	9/14	11/12	16/18	(28)				
	a	2	4	XY	15/17	(14)		16	28				
	a	2	5	XY	15/17	(14)	(11)	16/18	28	(19)/20			
	b	3	6		(14)/15			19					
	b	3	7			9/(13)/14							
	Consensus			**XY**	**15/17**	**9/14**	**(11/12)**	**16/(18)**	**28/28**				
Drestwo 2	a	XVI	1	X/Y	15/17	13/15	11/13	18/(19)		8			
	b	XII	1	X/Y	15/17	13/15	11	18/19	30/31				
	b	XII	2										
	b	XIV	2	X/Y	15/17	13/15	(11)	18/19			(24)		
	b	XIV	3	X/Y	15/17	13/15	11	18/19	30/31	(8/12)	(24)		
	c	XIV	1	X/Y	17	12	12	18/19	30/31				
	c	XIV	2										
	c	XIV	3										
	Consensus			**X/Y**	**15/17**	**13/15**	**11/?**	**18/19**	**30/31**				
MK				**X/X**	**14/18**	**14/14**	**12/12**	**17/18**	**29/31.2**	**12/12**	**20/20**	**11/13**	**14/17**

Table S6. MtDNA types of Neolithic individuals of the LBK/AVK culture (5,500-5,000 BC) (*S7, S16*)

Location/site	Country	ID	HVS-I sequence (np 15997-16409)	12308 Hinf	7025 Alu I	Clade
Asparn Schletz 2	Austria	ASP2	CRS	-12308 Hinf	-7025 Alu	H*
Derenburg/Meerenstieg II	Germany	DEB1	147a 172c 223t 248t 355t			N1a
Derenburg/Meerenstieg II	Germany	DEB2	224c 311c			K
Derenburg/Meerenstieg II	Germany	DEB3	147a 172c 223t 248t 320t 355t			N1a
Derenburg/Meerenstieg II	Germany	DEB4	311c	-12308 Hinf	+7025 Alu I	HV
Derenburg/Meerenstieg II	Germany	DEB5	311c	-12308 Hinf	+7025 Alu I	HV
Eilsleben	Germany	EIL1	CRS	-12308 Hinf	-7025 Alu I	H*
Flomborn	Germany	FLO1	147a 172c 223t 248t 320t 355t			N1a
Flomborn	Germany	FLO2	093c	-12308 Hinf	-7025 Alu I	H*
Flomborn	Germany	FLO4	126c 294t 304c			T2
Flomborn	Germany	FLO5	224c 311c			K
Flomborn	Germany	FLO6	224c 249c 311c			K
Halberstadt	Germany	HAL1	298c	-12308 Hinf	+7025 Alu I	V
Halberstadt	Germany	HAL2	086c 147a 172c 223t 248t 320t 355t			N1a
Halberstadt	Germany	HAL3	093c 126c 294t 296t 304c			T
Schwetzingen	Germany	SCHWE1	126c 294t 296t 304c			T2
Schwetzingen	Germany	SCHWE2	126c 292t 294t 296t			T2
Schwetzingen	Germany	SCHWE4	286t 304c		-7025 Alu I	H*
Schwetzingen	Germany	SCHWE5	126c 294t 296t			T2
Seehausen	Germany	SEE1	069t 126c			J
Unseburg	Germany	UNS2	223t 292t			W
Unterwiederstedt	Germany	UWS2	224c 249c 311c			K
Unterwiederstedt	Germany	UWS5	129a 147a 154c 172c 223t 248t 320t 355t			N1a
Vaihingen an der Enz	Germany	VAI3	319a 343g			U3
Ecsegfalva 1	Hungary	ECS1	147a 172c 189c 223t 248t 274a 355t			N1a

Table S7 Alignments of the HVS-I sequences is furnished in a separate file and can be downloaded at: http://www.uni-mainz.de/FB/Biologie/Anthropologie/MolA/English/Home/Home.html

References

S1. L .P. Louwe Kooijmans, Ed. *Hardinxveld-Giessendam De Bruin. Een jachtkamp uit het Laat-Mesolithicum en de vroege Swifterbant-cultuur, 5500-4450 v. Chr., RAM 88* (NS Railinfrabeheer B.V., Utrecht, 2001).

S2. M. Porr, K. W. Alt, *Int. J. of Osteoarch*., **16**, 395 (2006).

S3. H. Lübke, F. Lüth, T. Terberger, in: L. Larsson, F. Lüth, T. Terberger, Eds, *Innovation and Continuity – Non-Megalthic Mortuary Practises in the Baltic – New Methods and Research into the Development of Stone Age Society. Berichte Römisch-Germanische Kommission 88*, Frankfurt 2008.

S4. M. Zvelebil, M. Lillie, in *Europe's first farmers,* T. D. Price, Ed. (Cambridge Univ. Press, Cambridge, 2000), pp. 57-92.

S5. I. Antanaitis-Jacobs, A. Girininkas, *Archaeologia Baltica* **5**, 9 (2002).

S6. I.Antanaitis. 1999. Concerning the transition to farming in the East Baltic, Documenta Praehistorica XXVI, ed. by Mihael Budja, Ljubljana University, pp. 89-100.

S7. W. Haak, P. Forster, B. Bramanti, S. Matsumura, G. Brandt, *et al*., *Science* **310**, 1016 (2005).

S8. J. C. Hiller, T. J. Thompson, M. P. Evison, A. T. Chamberlain, T. J. Wess, *Biomaterials* **24**, 5091 (2003).

S9. J. Burger, W. Rosendahl, O. Loreille, H. Hemmer, T. Eriksson, *et al*., *Mol. Phylogenet. Evol*. **30**, 841 (2004).

S10. J. C. Hiller, M. J.Collins, A. T. Chamberlain, T. J. Wess, *J. of Arch. Science* **31**, 1349 (2004)

S11. M. Hofreiter, V. Jaenicke, D. Serre, A. A. Haeseler, S. Pääbo, *Nucleic Acids Res* **29**, 4793 (2001)

S12. A. Torroni, K. Huoponen, P. Francalacci, M. Petrozzi, L. Morelli, *et al*., *Genetics* **144**, 1835 (1996).

S13. N. Izagirre, C. de la Rua , *Am. J. Hum. Genet*. **65**, 199 (1999).

S14. C. Lalueza-Fox, M. L. Sampietro, M. T. Gilbert, L. Castri, F. Facchini, *et al*., *Proc. R. Soc. Lond B Biol. Sci.* **271**, 941 (2004).

S15. A. Brandstätter, T. J. Parsons, W. Parson, *Int J Legal Med* **117**, 291 (2003).

S16. W. Haak, Ph-D thesis, Mainz University (2006).

S17. H. Zischler, M. Höss, O. Handt, A. von Haeseler, *Science* **268**, 1192 (1995).

S18. J. Burger, S. Hummel, B. Herrmann, W. Henke, *Electrophoresis* **20**, 1722 (1999).

S19. B. Bramanti, S. Hummel, T. Schultes, B. Herrmann, *Anthrop. Anz*. **58**, 45 (2000).

S20. M. Richards, V. Macaulay, E. Hickey, E. Vega, B. Sykes, *et al*., *Am. J. Hum. Genet.* **67**, 1251 (2000).

S21. D. Kasperaviciute, V. Kucinskas, M. Stoneking, *Ann. Hum. Genet*. **68**, 438 (2004).

S22. K. Tambets, S. Rootsi, T. Kivisild, H. Help, P. Serk, *et al*., *Am. J. Hum. Genet.* **74**, 661 (2004).

S23. L. Pliss, K. Tambets, E. L. Loogväli, N. Pronina, M. Lazdins, *et al*., *Ann. Hum. Genet.* **70,** 439 (2006).

S24. A. Achilli, C. Rengo, V. Battaglia, M. Pala, A. Olivieri, *et al*., *Am. J. Hum. Genet*. **76,** 883 (2005).

S25. M. van Oven, M. Kayser, *Hum. Mutat*.**30**, E386 (2009). http://www.phylotree.org, mtDNA tree Build 3 (01.03.09).

S26. O. A. Derbeneva, E. B. Starikovskaya, D. C. Wallace, R. I. Sukernik, *Am. J. Hum. Genet.* **70,** 1009 (2002).

S27. M. Bermisheva, K. Tambets, R. Villems, E. Khusnutdinova E., *Mol. Biol. (Mosk)* **36**, 990 (2002).

S28. K. Tambets, H. V. Tolk, T. Kivisild, E. Metspalu, J. Parik, *et al*., in *Examining the farming/language dispersal hypothesis,* P. Bellwood, C. Renfrew, Eds. (McDonald Institute for Archaeological Research Monograph Series, Cambridge Univ. Press, Cambridge, 2003), pp 449–458.

S29. A. Helgason, E. Hickey, S. Goodacre, V. Bosnes, K. Stefansson, *et al*., *Am. J. Hum. Genet.* **68**, 723 (2001).

S30. A. Gonzalez, A. Brehm, J. Pérez, N. Maca-Meyer, C. Flores, *et al*., *Am. J. Phys. Anthropol.* **120,** 391 (2003).

S31. L. Excoffier, L. G. Laval, S. Schneider, *Evolutionary Bioinformatics Online* **1**, 47 (2005).

S32. J. Reynolds, B. S. Weir, C. C. Cockerham, *Genetics* **105**, 767 (1983)

S33. L. Excoffier, J. Novembre, S. Schneider*, J. Hered*., **91**, 506 (2000).

S34. S. Y. Ho, P.Endicott, *Am J Hum Genet*., **83**, 142 (2008).

S35. H. J. Bandelt, Q. P.Kong, H. Richards, V. Macaulay, in *Human Mitochondrial DNA and the Evolution of Homo Sapiens*, H. J. Bandelt, V. Macaulay, M. Richards, Eds. (Springer, Berlin-Heidelberg, 2006) pp. 47-90

S36. http://www.phyton.org

S37. http://www.r-project.org

S38. H. J. Bandelt, P. Forster, A. Röhl, *Mol Biol Evol* **16**, 37 (1999). http://www.fluxus-engineering.com

S39. B. Bramanti, S. Hummel, T. Schultes, B. Herrmann, in *Human population genetics in Europe,* C. Susanne, B. Bodzsár Eds. (Biennial Book of EAA, Eötvös Univ. Press, Budapest, 2000), pp. 147-163.

S40. J. Burger, M. Kirchner, B. Bramanti, W. Haak, M. G. Thomas, *Proc. Natl. Acad. Sci.* **104**, 3736 (2007).

S41. L. Vigilant, R. Pennington, H. Harpending, T. D. Kocher, A. C.Wilson, *Proc. Natl. Acad. Sci. U.S.A* **86**, 9350 (1989).

S42. R. M. Andrews, I. Kubacka, P. F. Chinnery, R.N. Lightowlers, D. M. Turnbull, *et al*., *Nat Genet.* **23**, 147 (1999).

S43. A. C. Stone, M. Stoneking, *Am. J. Hum. Genet.* **62**, 1153 (1998).

S44. O. Handt, M. Krings, R. H. Ward, S. Pääbo, *Am. J. Hum. Genet*. **59**, 368 (1996).

S45. S. Finnilä, I. E. Hassinen, L. Ala-Kokko, K, Majamaa, *Am J Hum Genet*. **66**, 1017 (2000).

S46. V. Macaulay, M. Richards, E. Hickey, E. Vega, F. Cruciani, *et al*., *Am. J. Hum. Genet.* **64**, 232 (1999).

S47. M. Belledi, E. S. Poloni , R. Casalotti, F. Conterio, I. Mikerezi, *et al.*, *Eur. J. Hum. Genet.* **8**, 480 (2000).

S48. W. Parson, T. J. Parsons, R. Scheithauer, M. M. Holland, *Int. J. Legal Med.* **111,** 124 (1998).

S49. B. A. Malyarchuk, T. Grzybowski, M. V. Derenko, J. Czarny, K. Drobnic, *et al*., *Ann. Hum. Genet.* **67**, 412 (2003).

S50. J. Bertranpetit, J. Sala, F. Calafell, P. A. Underhill, P. Moral, *et al*., *Ann. Hum. Genet.* **59**, 63 (1995).

S51. H. B. Côrte-Real, V. A. Macaulay, M. B. Richards, G. Hariti, M. S. Issad, *et al*., *Ann. Hum. Genet.* **60**, 331 (1996).

S52. J. C. Rando, V. M. Cabrera, J. M. Larruga, M. Hernández, A. M. González, *et al*., *Ann. Hum. Genet.* **63**, 413 (1999).

S53. M. Richards, H. Corte-Real, P. Forster, V. Macaulay, H. Wilkinson-Herbots, *et al., Am. J. Hum. Genet.* **59**, 185 (1996).

S54. T. Vanecek, F. Vorel, M. Sip, *Int. J. Legal Med.* **118**, 14 (2004).

S55. M. Krings, A. E. Salem, K. Bauer, H. Geisert, A. K. Malek, *et al*., *Am. J. Hum. Genet.* **64**, 1166 (1999).

S56. R. Piercy, K. M. Sullivan, N. Benson, P. Gill, *Int. J. Legal Med.* **106**, 85 (1993).

S57. A. Helgason, S. SigurDardóttir, J. R. Gulcher, R. Ward, K. Stefánsson, *Am. J. Hum. Genet.* **66,** 999 (2000).

S58. A. Sajantila, P. Lahermo, T. Anttinen, M. Lukka, P. Sistonen, *et al*., *Genome Res.* **5**, 42 (1995).

S59. I. Pult, A. Sajantila, J. Simanainen, O. Georgiev, W. Schaffner, *et al. Biol. Chem. Hoppe Seyler* **375**, 837(1994).

S60. R. A. Kittles, A. W. Bergen, M. Urbanek, M. Virkkunen, M. Linnoila, *et al*., *Am. J. Phys. Anthropol.* **108**, 381 (1999).

S61. M. Meinilä, S. Finnilä, K. Majamaa, *Hum. Hered.* **52**, 160 (2001).

S62. F. Rousselet, P. Mangin, *Int. J. Legal Med.* **111**, 292 (1998).

S63. C. Danan, D. Sternberg, A. Van Steirteghem, C. Cazeneuve, P. Duquesnoy, *et al*., *Am. J. Hum. Genet.* **65**, 463 (1999).

S64. F. Cali, M. G. Le Roux, R. D'Anna, A. Flugy, G. De Leo, *et al*., *Int. J. Legal Med.* **114**, 229 (2001).

S65. S. Hofmann, M. Jaksch, R. Bezold, S. Mertens, S. Aholt, *et al*., *Hum. Mol. Genet.* **6**, 1835 (1997).

S66. A. Baasner, C. Schäfer, A. Junge, B. Madea, *Forensic Sci. Int.* **98**, 169 (1998).

S67. S. Lutz, H. J. Weisser, J. Heizmann, S. Pollak, *Int. J. Legal Med.* **111**, 67 (1998).

S68. H. Pfeiffer, B. Brinkmann, J. Hühne, B. Rolf, A. A. Morris, *et al*., *Int. J. Legal Med.***112**, 291 (1999).

S69. M. Metspalu, T. Kivisild, E. Metspalu, J. Parik, G. Hudjashov, *et al*., *BMC Genet.* **5**, 26 (2004).

S70. P. Francalacci, J. Bertranpetit, F. Calafell, P. A. Underhill, *Am. J. Phys. Anthropol.* **100**, 443 (1996).

S71. A. Torroni, H. J. Bandelt, L. D'Urbano, P. Lahermo, P. Moral, *et al*.. *Am. J. Hum. Genet.* **62**, 1137 (1998).

S72. A. Tagliabracci, C. Turchi, L. Buscemi, C. Sassaroli, *Int. J. Legal Med.* **114**, 224 (2001).

S73. C. Keyser-Tracqui, E. Crubézy, B. Ludes, *Am. J. Hum. Genet.* **73**, 247(2003).

S74. J. C. Rando, F. Pinto, A. M. González, M. Hernández, J. M. Larruga, *et al*., *Ann. Hum. Genet.* **62**, 531(1998).

S75. J. Saillard, I. Evseva, L. Tranebjaerg, S. Norby, in *Archaeogenetics: DNA and the population prehistory of Europe,* C. Renfrew, K. Boyle Eds. (McDonald Institute for Archaeological Research Monograph Series, Cambridge Univ. Press, Cambridge, 2000), pp 255–258.

S76. B. M. Dupuy, B. Olaisen, in *Advances in forensic haemogenetics*, A. Carracedo, B. Brinkmann, W. Bär, Eds. (Springer, Berlin, Heidelberg, New York, 1996), vol 6, pp 23–25.

S77. S. H. Opdal, T. O. Rognum, A. Vege, A. K. Stave, B. M. Dupuy, *et al*., *Acta Paediatr*. **87**, 1039 (1998).

S78. G. Passarino, G. L. Cavalleri, A.A. Lin, L. L. Cavalli-Sforza, A. L. Borresen-Dale, *et al*., *Eur. J. Hum. Genet.* **10**, 521(2002).

S79. L. Pereira, M. J. Prata, A. Amorim, *Ann. Hum. Genet.* **64**, 491 (2000).

S80. V. Orekhov, A. Poltoraus, L. A. Zhivotovsky, V. Spitsyn, P. Ivanov, *et al*., *FEBS Letters*, **445**, 197 (1999).

S81. B. A. Malyarchuk, M. V. Derenko, *Ann. Hum. Genet.* **65**, 63 (2001).

S82. B. A. Malyarchuk, T. Grzybowski, M. V. Derenko, J. Czarny, M. Wozniak, *et al*., *Ann. Hum. Genet.* **66**, 261 (2002).

S83. P. Forster, F. Cali, A. Rohl, E. Metspalu, R. D'Anna, *et al*., *Int. J. Legal Med.* **116**, 99 (2002).

S84. M. Crespillo, J. A. Luque, M. Paredes, R. Fernandez, E. Ramirez, *et al*., *Int. J. Legal Med*. **114**, 130 (2000).

S85. J. M. Larruga, F. Díez, F. M. Pinto, C. Flores, A. M. González, *Eur. J. Hum. Genet*. **9**, 708 (2001).

S86. A. Di Rienzo, A. C. Wilson, *Proc. Natl.Acad. Sci USA* **88**, 1597 (1991).

S87. I. Zupanic Pajnic, J. Balazic, R. Komel, *Int. J. Legal Med.* **118**, 1 (2004).

S88. N. Dimo-Simonin, F. Grange, F. Taroni, C. Brandt-Casadevall, P. Mangin, *Int. J. Legal Med*. **113**, 89 (2000).