Absence of the lactase-persistence-associated allele in early Neolithic Europeans

J. Burger, M. Kirchner, B. Bramanti, W. Haak, and M. G. Thomas

PNAS published online Feb 28, 2007;
doi:10.1073/pnas.0607187104

Absence of the lactase-persistence-associated allele in early Neolithic Europeans

J. Burger[†‡], M. Kirchner[†], B. Bramanti[†], W. Haak[†], and M. G. Thomas[§]

[†]Johannes Gutenberg University, Institute of Anthropology, Saarstrasse 21, D-55099 Mainz, Germany; and [§]Department of Biology, University College London, Wolfson House, 4 Stephenson Way, London NW1 2HE, United Kingdom

Edited by Walter Bodmer, Cancer Research UK, Oxford, United Kingdom, and approved December 27, 2006 (received for review September 4, 2006)

Lactase persistence (LP), the dominant Mendelian trait conferring the ability to digest the milk sugar lactose in adults, has risen to high frequency in central and northern Europeans in the last 20,000 years. This trait is likely to have conferred a selective advantage in individuals who consume appreciable amounts of unfermented milk. Some have argued for the ''culture-historical hypothesis,'' whereby LP alleles were rare until the advent of dairying early in the Neolithic but then rose rapidly in frequency under natural selection. Others favor the ''reverse cause hypothesis,'' whereby dairying was adopted in populations with preadaptive high LP allele frequencies. Analysis based on the conservation of lactase gene haplotypes indicates a recent origin and high selection coefficients for LP, although it has not been possible to say whether early Neolithic European populations were lactase persistent at appreciable frequencies. We developed a stepwise strategy for obtaining reliable nuclear ancient DNA from ancient skeletons, based on (*i*) the selection of skeletons from archaeological sites that showed excellent biomolecular preservation, (*ii*) obtaining highly reproducible human mitochondrial DNA sequences, and (*iii*) reliable short tandem repeat (STR) genotypes from the same specimens. By applying this experimental strategy, we have obtained high-confidence LP-associated genotypes from eight Neolithic and one Mesolithic human remains, using a range of strict criteria for ancient DNA work. We did not observe the allele most commonly associated with LP in Europeans, thus providing evidence for the culture-historical hypothesis, and indicating that LP was rare in early European farmers.

ancient DNA | dairying | selection

Most mammals lose the ability to digest the milk sugar lactose after weaning because of an irreversible reduction in expression of the intestinal enzyme lactase (i.e. lactase phlorizin hydrolase). This pattern is also seen in most humans, but some continue expressing lactase throughout adult life [lactase persistence (LP)]. This dominant Mendelian trait is common in populations of northern and central European descent and shows intermediate frequencies in southern and eastern Europe (1). Africa and the Middle East show a more complex distribution, with pastoralists often having high frequencies of LP, whereas in their nonpastoralist neighbors, it is usually much less common (2). The T allele of C/T polymorphism located 13,910 bp upstream of the lactase (LCT) gene (−13.910*T) has been shown to associate strongly with LP in Europeans (3), and recent *in vitro* studies have indicated that it can directly effect LCT gene promoter activity (4). However, different but closely linked polymorphisms associate with LP in most African groups, indicating either that −13.910*T is not causative of LP and/or that the trait has evolved more than once in humans (2, 5, 6).

It has been suggested that the modern frequency of LP in Europe is the result of a relatively recent and strong selection process (7, 8). Although not fully understood, the biological advantages of LP probably include the continuous availability of an energy- and calcium-rich drink that enables a farming community to overcome poor harvests. Because it is unlikely that LP would have provided a selective advantage in the absence of a supply of fresh milk, and because of observed correlations between the frequency of LP and the extent of traditional reliance on animal milk, the culture-historical hypothesis has been proposed (8–12). Under this model, LP was driven from very rare cases of preadaptation to appreciable frequencies only after the cultural practice of dairying arose. However, an opposing view, the reverse cause hypothesis, has also been proposed (8, 13, 14). According to this model, human populations were already differentiated with regard to LP frequency before the development of dairying, and the presence of LP determined the adoption of milk production and consumption practices (15). Based on the decay of long-range haplotypes and variation in closely linked microsatellites, the inferred age of the −13.910*T allele in Europe has been estimated to be between 2,188 and 20,650 years (16) and between 7,450 and 12,300 years (17), respectively. These dates bracket archaeological estimates for the introduction of domestic cattle breeds into Europe (18), and when considered in conjunction with the modern frequency distribution of the −13.910*T allele in Europe, they indicate a strong selective advantage to LP. However, these date estimates do not exclude the possibility of LP being present in Europe at appreciable frequency before the Neolithic.

Analyzing DNA from archaeological human remains is the only direct method to identify the presence of the −13.910*T allele in specific prehistoric populations. Unfortunately, the authenticity of ancient DNA data, particularly when recovered from human remains, cannot be guaranteed because of the problems of modern DNA contamination (19–23) and postmortem damage (24–29). Although generic validation criteria have been proposed for ancient DNA work (30), they are often not tailored to the specific questions being addressed. Gilbert *et al.* (31) have recently proposed a more flexible approach, whereby the validation criteria used are customized to the particulars of the archaeological source material and the aims of the investigation. In the context of understanding the origins and evolution of LP in Europe, we have engaged this approach to analyze 51 bone samples from early Holocene sites in central and eastern Europe for the preservation of mitochondrial and nuclear DNA and subsequently for the presence of the −13.910*T allele. Using a range of authentication criteria, we obtained high confidence genotypes from eight early Neolithic and one Mesolithic skeletons of central, northeast, and southeast Europe (Fig. 1).

Author contributions: J.B. and M.G.T. designed research; M.K., B.B., and W.H. performed research; J.B., M.K., and M.G.T. analyzed data; and J.B., M.K., B.B., and M.G.T. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS direct submission.

Abbreviations: LP, lactase persistence; STR, short tandem repeat.

[‡]To whom correspondence should be addressed. E-mail: jburger@uni-mainz.de.

This article contains supporting information online at www.pnas.org/cgi/content/full/0607187104/DC1.

Fig. 1. Locations of archaeological sites.

Results

We selected bone samples that showed excellent biomolecular preservation. This characteristic was additionally shown by the reproducible amplification of animal DNA from the same and from neighboring archaeological sites (Szarvas, Albertfalva, Szegvár-Tüzköves, and Derenburg; no animal bones were available from the Baltic sites). Further, all samples typed for the −13.910-C/T polymorphism had previously yielded reproducible mitochondrial DNA (J.B., W.H., and B.B. unpublished data and ref. 32) and short tandem repeat (STR) data. The STR profiles of the 10 individuals are shown in supporting information (SI) Table 2. mtDNA haplotypes are shown in Table 1. In accordance with previous studies (33, 34), STR-allelic dropout in heterozygous individuals was observed at a rate of 10%. Locus dropout happened in 24% of PCRs as measured by using the STR Profiler multiplex assay. Because of the high number of repeated typings, allelic and locus dropout did not influence the final consensus genotype determination (33, 35).

−13.910-C/T genotypes were considered reliable when results were reproduced from at least two extractions and two independent PCRs in combination with negative controls. We obtained genotypes from eight Neolithic skeletons of central, northeast, and southeast Europe ranging in age between 5800 and 5000 B.C. (Table 1). One Mesolithic sample typed is absolutely younger [2267 ± 116 calibrated date (cal) B.C.] Additionally, we typed one Medieval skeleton as a control. We identified 9 of 10 individuals as homozygous C at position −13.910; all were Mesolithic and Neolithic samples. The one Medieval individual was heterozygous for the −13.910-C/T polymorphism. This was the only heterozygous individual, and it never showed allelic dropout. In addition, we typed the −22.018-A/G polymorphism at position 18524 in intron 9 of the *MCM6* gene, which is also associated with LP, although the correlation

Table 1. Samples, culture, archaeological or radiocarbon dating, lactase genotype, and mtDNA haplotype

Sample	Culture	Archaeological or radiocarbon dating	*MCM6* genotype		CRS mtDNA 16209–16303 variable positions (−16,000)
			Intron 13	Intron 9	
ELT 2	Merovingian	A.D. 400–600	Y	A	C294T
DEB 1	Neolithic Linear Pottery	5500–5000 B.C.	C	G	C223T, C248T
DEB 3	Neolithic Linear Pottery	5500–5000 B.C.	C	G	C223T, C248T
DEB 4	Neolithic Linear Pottery	5500–5000 B.C.	C	G	CRS
SZA23.1	Neolithic Körös	5840–5630 B.C. (OxA-9375) human rib, grave 1	C	G	C223T, C257A, C261T
SZA23.2	Neolithic Körös	5840–5630 B.C. (OxA-9375) human rib, grave 1	C	G	C223T
SZA23.3	Neolithic Körös	5840–5630 B.C. (OxA-9375) human rib, grave 1	C	G	CRS
KRE 1	Middle Neolithic Narva	5350 ± 130 B.C. (OxA-5935)	C	G	C270T
KRE 2	Middle Neolithic Narva	5580 ± 65 B.C. (OxA-5926)	C	G	C270T
DR 2	Mesolithic Zedmar	2267 ± 116 cal. B.C.	C	G	C256T, C270T

The samples are described in SI Table 3. CRS, Cambridge reference sequence.

between the presence of the A allele and LP is less strong than for −13.910*T (36). Here, all but the medieval sample were homozygous G. The complete results are shown in Table 1. For comparison, we show in SI Table 4 the genotypes of the main archaeologists and anthropologists who handled the samples and of the lab workers who were present in the lab when the samples were processed.

In 7 of 88 cases −13.910 C residues showed an additional minor T peak after minisequencing. In those cases, PCR products were additionally ligated into a pUC18 (T vector, our own production) and used to transform an *Escherichia coli* culture (RR1). Selected transformant colonies were directly PCR-amplified and sequenced by using universal M13 primers. The 41 clones of six PCR products showed the usual pattern of ancient degradation, with type II transitions being the most common postmortem artifact. Clone sequences of 67-bp length showed between zero and four inconsistent substitutions (average 0.77 per clone sequence). The site of the −13.910-C/T polymorphism was affected by inconsistent postmortem alterations in just one clone.

By treating the eight Neolithic individuals as belonging to the same metapopulation, we estimate that the frequency of the −13.910*T allele in this population should be no more than 0.17 with 95% confidence and no more than 0.25 with 99% confidence. Assuming Hardy–Weinberg equilibrium and that the −13.910*T allele is the only cause of LP in ancestral Europeans, these allele frequencies correspond to LP frequencies of no more than 0.31 and 0.44, respectively. These values represent the upper end of our frequency estimates; actual frequencies are likely to have been lower and could have been zero.

Discussion

Archaeological Human Nuclear DNA. Because we report a lack of polymorphism in nuclear loci from archaeological human skeletons dating to up to 8,000 years old, a justification of an endogenous origin of the DNA is required. Despite the severe rules to avoid contamination at the Mainz ancient DNA facilities, on average 10–15% of clones from amplifications of human archaeological mitochondrial DNA contain sequences of contaminant origin (32). For nuclear DNA, both the success and contamination rates are much lower. Only 27% of samples that yielded mitochondrial sequences also yielded nuclear DNA (37). At the same time, the contamination rate of nuclear DNA is close to zero. Sampietro *et al.* (20) report that mitochondrial contaminations by the archaeologists or other workers involved in the sampling process can make up to 17% of sequenced human clones. This is similar to our finds for mitochondrial DNA. However, when analyzing nuclear DNA markers, especially STRs, and in those cases when researchers involved in sampling, washing, etc., of the samples were known to us, we never observed the amplification of those nuclear DNAs as a contamination. Reports of ancient nuclear DNA so far are mainly based on animal skeleton finds (38–42). There are a few studies that report consistent amplification of STR profiles from human skeletal remains (33, 34, 43, 44). We applied the genotyping and reproduction strategies developed in these studies to our sample of 51 bones that previously yielded reproducible human mitochondrial DNA sequences in association with excellently preserved animal bones. This stepwise strategy [(*i*) excellent biomolecular preservation in animal bones, (*ii*) reproducible human mtDNA sequences, and (*iii*) reliable STR genotypes] resulted in highly confident genotyping of less variable nuclear markers, such as the −13.910*-C/T polymorphism. We observe complete absence of the −13.910*T allele in an Early Holocene sample, a result that is difficult to reconcile with modern contamination, because it is very different from the allelic distribution in modern Europeans. The question remains whether allelic dropout may be an explanation for the observed homozygosity in our sample. We observe allelic dropout in our STR multiplex data at a rate of 10%. However, we did not observe preferential amplification when the two alleles that were present were similar in size. Thus, unless the −13.910*C allele is amplified preferentially, which seems unlikely, repeated typing of the same individual from different extractions and different PCRs allows us to exclude allelic dropout as a cause for high homozygosity.

In summary, the following points add to our confidence that we have excluded contamination. Thereby, we also refer to mtDNA sequence data and STR genotypes previously obtained from the skeletons, which show a higher individual variability and help to point out the authentic origin of the less variable lactase genotype (cf. also ref. 45):

- The frequency of the alleles in our prehistoric sample is different from the frequencies in modern populations and the lab crew (SI Table 4). A random modern contaminant population would look different.
- STR genotypes from all 10 individuals differed from each other. This observation excludes a systematic contamination of more than one skeleton by one or a few individuals.
- Reproducible amplification success was restricted to certain archaeological sites. This is inconsistent with background contamination originating in the DNA laboratory.
- Animal bones that were available as comparison for six of the nine prehistoric human samples yielded high quality bovine and cervid DNA sequences, indicating excellent biomolecular preservation conditions.
- Animal bone DNA from the same excavations amplified with the human lactase primers never yielded positive results.
- In those cases where the archaeologists and anthropologists who worked on the skeletons before sampling for genetic analysis were known to us (SI Table 4), we did not detect the DNA profiles (mtDNA, STRs) of these investigators after cleaning the surface of the samples following our protocols.
- Mitochondrial DNA sequences that were obtained from the 10 samples by direct sequencing were always identical to those from a previous study that used extensive cloning from additional PCRs of the same specimens (ref. 32 and B.B., J.B., and W.H., unpublished data). Here, overlapping amplicons from independent PCRs complemented each other and gave unambiguous haplotypes that correspond to attested branches of the human mtDNA phylogenetic tree. Contaminating sequences were restricted to single amplicons and could be ruled out, because they produced nonsense haplotype mosaics (46). In some cases, we were able to amplify the two SNP loci (−13.910-C/T and −22.018-A/G) in a multiplex approach together with an HVR I-fragment. Again, the SNP alleles and the mtDNA haplotype from all previous experiments were reproducible.

Based on these experimental data and theoretic considerations, we are confident that our data are authentic and can be interpreted in the context of the evolution of LP in central and northern Europeans. However, from a theoretical point of view, it is impossible to detect or completely rule out a systematic contamination of a skeleton. Nonetheless, it is probable that our decontamination procedures would remove a systematic contamination in the unlikely event that a whole skeleton became contaminated.

Absence of LP in Meso- and Neolithic European Populations. Estimates for the age of the 13.910*T allele, based on data from modern individuals, are between 2,188 and 20,650 years (16) and between 7,450 and 12,300 years (17). Thus it is, *a priori*, plausible that the −13.910*T allele was present at appreciable frequencies in early Neolithic, and even pre-Neolithic, European populations. However, these relatively recent date estimates are based on the observation of low haplotype diversity and, when con-

sidered in conjunction with modern allele frequencies, indicate a strong role for selection. It is highly unlikely that milk would have been a major component of the adult diet, and thus that selection would have been acting, before the domestication of goats, sheep, or cattle. Bersaglieri *et al.* (16) estimate the coefficient of selection (S) associated with carrying at least one copy of a −13.910*T allele to be between 0.014 and 0.15, using a European dataset. By applying a simple deterministic model (see *Materials and Methods*) we would expect the −13.910*T allele frequency 7,000 years ago to be ≈0.089 (when $S = 0.014$) and ≈0 (when $S = 0.15$). Thus, the results presented here are consistent with expectation under the culture-historical hypothesis for the origins of LP in Europe (8–12). Although there is only one sample, it is also notable that, as would be expected, the −13.910*T allele is absent in our Mesolithic sample.

The present study should contribute to archaeological debate concerning the origin of dairying in Europe. Although some scholars assume that dairying was not practiced during the earliest phases of the Neolithic and that it first spread over Europe in the 3rd millennium B.C. (47, 48), others, based on osteometric sexing of cattle and goats, suggest an onset of dairying practices in southeastern Europe and southern Germany between 7000 and 6500 B.P. (49). Others claim that the technical skill of dairying came to Europe already as part of the fully developed Neolithic package with the first farmers (50, 51). The latter view is supported by the earliest archaeometric evidence for dairying in Europe ≈7900–7500 B.P. in Hungary (52) and ≈6100 B.P. in Britain (53, 54). These dates suggest that dairying practices came to Europe nearly simultaneously with cereal agriculture and domestic animals. However, the absence of the 13.910*T allele in our Neolithic samples indicates that the early farmers in Europe were not yet adapted to the consumption of unprocessed milk. Dairying is unlikely to have spread uniformly over Europe, and the use of milk in the Early Neolithic may have been rare. Although our data are consistent with strong selection for LP beginning with the introduction of cattle to Europe ≈8800 B.P., it is unlikely that fresh milk consumption was widespread in Europe before frequencies of the 13.910*T allele had risen appreciably during the millennia after the onset of farming.

Important questions remain regarding the geographic location of the earliest −13.910*T-allele-carrying populations, the mode and direction of spread of the allele, and the precise nature of the selective advantage(s) conferred by LP. If dairying was a common feature of European Neolithic populations, and the selection pressure is actually so strong, then an LP-causing mutation occurring anywhere in Europe should rise rapidly in frequency. A number of approaches are available for inferring the geographic origin and direction of spread of an allele by using data from modern populations. The modern distribution of the −13.910*T allele might be taken to indicate a northwestern European origin. However, forward computer simulations have shown that the center of distribution of an allele can be far removed from its location of origin when a population expands along a wave front (55). Geographic structuring of haplotype diversity among −13.910*T-carrying lineages may prove to be a more reliable indicator of the region of origin, and in a preliminary report, Enattah *et al.* (56) have used this approach to infer an eastern, possibly steppe, origin for the −13.910*T allele. But, as with inference based on allele distribution alone, common demographic processes, such as demic diffusion, may obscure signals of geographic origin. Although computer simulations that accommodate or explore more realistic demographic scenarios (e.g., ref. 57) and other parameters, such as selection strengths, offer for the future a better understanding of the origin, spread, and coevolution of LP and dairying in Europe, all inference-based methods for investigating these factors carry a high degree of uncertainty. Thus, we believe that an ancient DNA approach, the feasibility of which is demonstrated here, is the most direct way of studying the evolution of LP in Europe. In this study, we provide high-confidence prehistoric lactase genotypes from reliably dated remains. One direction for future research that should yield a greater insight into the origin, evolution, and spread of LP in human populations would be the incorporation of our data into the wider computer simulation models proposed above.

Conclusions

We accept that, although we are highly confident in our data, we cannot be absolutely certain that systematic contamination of single skeletons has been excluded. But crucially, we believe that the information we provide on the spatiotemporal distribution of the −13.910*T allele competes favorably with equivalent information that could be inferred by using molecular data from modern populations. Although the quantity of data presented is not sufficient to completely reject the reverse cause hypothesis for the origins of LP in Europe (9, 13, 14), we have provided (*i*) evidence in support of the culture-historical hypothesis (8–11), (*ii*) a robust example of the flexible approach advocated by Gilbert *et al.* (31) that uses ancient DNA techniques to address questions of human evolution, and (*iii*) a valuable dataset that should be used in conjunction with data from modern populations to build a more complete model of the origins and spread of LP in Europe.

Materials and Methods

Archaeological Bone and Tooth Samples. Bone and tooth material was excavated in eastern Germany, Hungary, northeastern Poland, and northeastern Lithuania (SI Table 3). Three specimens (DEB 1, DEB 3, and DEB 4) are from the Derenburg Meerenstieg II cemetery in the northern Harz region of eastern Germany. The burials belong to the Neolithic Linear Pottery culture (58) and thus to the earliest farmers of central Europe. Three individuals (SZA23.1, SZA23.2, and SZA23.3) are from the Szarvas site, located on the Great Hungarian Plain, District of Békés, southeastern Hungary. The skeletons belong to the Körös Culture, which appeared in eastern Hungary in the early 8th millennium B.P. (59, 60). Two skeletons are from the burial site of Kretuonas 1B in eastern Lithuania. These burials belong to the Middle Neolithic Narva culture and are radiocarbon-dated to the 6th millennium B.P.

Tooth material of a Mesolithic individual stems from Drestwo, a site in the Suwalki region of northeastern Poland. Individual DR 2 belongs to the Mesolithic Zedmar culture and is radiocarbon-dated to the late 5th millennium cal B.P. (J. Siemaszko, personal communication). One skeleton (burial 382) stems from the early medieval Merovingian burial ground Eltville in Hesse, southwestern Germany and is dated to the 4th to 6th century A.D.

Additionally, we analyzed animal samples from Derenburg, Szarvas, and geographically closely situated sites of similar geology in Hungary.

DNA Extraction. DNA extraction was performed in a laboratory that is dedicated to ancient DNA work and free of other molecular work. The bone and tooth material was exposed to UV light for at least 30 min on each side. A part of the bone surface and as much as possible of the tooth surface was abraded by using a rotary drill (Dremel, Breda, The Netherlands), to remove surface contamination. Approximately 1 cm^3 of bone was drilled out and cut into small pieces. For teeth, only the roots were used. After a second exposure to UV light, bone and tooth material was ground to a fine powder by using a mineralogy mill, (MM200; Retsch, Haan, Germany) and DNA was isolated by following the protocol described by Burger *et al.* (61). At least two independent DNA extractions were performed for each individual. To detect possible contamination by exogenous mod-

ern DNA we used extraction blanks as negative controls. During all steps of sample preparation we took all possible precautions to guard against contamination: gloves, face masks, and overalls were always worn. All laboratory equipment (tubes, pipettes, filter tips, etc.) was sterilized by exposure to UV light. All surfaces were cleaned with soap and bleach before and after working. Water was cleaned under constant agitation by irradiation with a water-proof UV-bulb for 10 h.

STR Analysis. Amplification of STRs was performed by using an AmpF*l*S*T*R Profiler Plus Kit (Applied Biosystems, Foster City, CA) according to the manufacturer's instructions, but halving the recommended final reaction volume to 25 μl: 1× PCR Mix, 1× Primerset, 2.5 units of AmpliTaqGold (Applied Biosystems) and 1–5 μl of DNA extract. Cycle conditions consisted of an initial denaturation of 6 min at 95°C followed by 45 cycles of 1 min at 94°C, 1 min at 59°C, 1 min at 72°C, and a final extension of 2 h at 60°C. Amplification products were analyzed by capillary electrophoresis on an ABI Prism 310 Genetic Analyzer (Applied Biosystems). Amplification was repeated six to nine times from at least two extracts. Allelic dropout rates were calculated in the following way: in cases where the final STR consensus profile was heterozygous, homozygous results in individual amplifications were rated as allelic dropout and divided by the number of amplifications. Locus dropout was considered when no allele was amplified, and simultaneously, amplicons of the same or of greater length were amplifiable in the same reaction.

SNP Analysis. We designed two oligonucleotide primer pairs (lac5 and lac6) (see SI Table 5) spanning the C/T variant in intron 13 of the *MCM6* gene and the G/A variant in intron 9 of the *MCM6* gene, using the program Primer Select (Lasergene 99; DNAStar, Madison, WI). PCR amplifications were performed in a final reaction volume of 50 μl, containing 1× PCR Gold Buffer (Applied Biosystems), 0.2 mM dNTP Mix (MBI Fermentas, Hanover, MD), 2.5 mM $MgCl_2$ solution, 3.5 units of AmpliTaqGold (Applied Biosystems), 0.2 μM each primer, 20 μG of BSA, and 3–5 μl of DNA extract.

Further, a 94-bp fragment of mtDNA (L16209, H16303) (SI Table 5) was coamplified with the two SNP loci in a multiplex PCR approach. Multiplex PCR amplifications were similar to the above-mentioned protocol but with an increase of DNA extract to 5–7 μl.

PCR cycling conditions were as follows: an initial 6 min incubation at 94°C, followed by 48 cycles of 94°C for 30 sec, 53–55°C for 40 sec, 72°C for 30 sec, and final elongation at 64°C for 45 min. One extraction blank and at least two PCR negative controls were included in each PCR experiment. Amplification products purified by using the Invisorb Rapid PCR Purification Kit (Invitek, Berlin, Germany) according to the manufacturer's instructions.

LP-associated SNPs (−13.910*-C/T and −22.018-A/G) were assayed by minisequencing, using the ABI Prism SNaPshot Multiplex kit (Applied Biosystems). For each SNP one minisequencing primer (lac7_mini, lac8_mini, SI Table 5) was computer-evaluated to optimize the melting temperature and to exclude the existence of interfering secondary structures by using IDT's SciTools OligoAnalyzer, version 3.0 (http://biotools.idtdna.com/analyzer/). Primer extensions were carried out in a final reaction volume of 8 μl, with 1 μl of purified PCR product, minisequencing primers (0.15 μM each), HPLC-H_2O, and 3 μl of SNaPshot Multiplex Ready Reaction Mix, and by using 25 cycles of the following: 96°C for 10 sec, 55°C for 5 sec, and 60°C for 30 sec. Postextension products were treated with shrimp alkaline phosphatase (MBI Fermentas) and then analyzed by capillary electrophoresis on an ABI Prism 310 Genetic Analyzer (Applied Biosystems), using the GeneScan Analysis software, version 3.1 (Applied Biosystems).

DNA Cloning and Sequencing. In seven cases, C residues showed an additional minor T peak after minisequencing. In those cases, PCR products were additionally ligated into a pUC18 (T vector, our own production) and used to transform an *E. coli* culture (RR1). Selected clone colonies were directly PCR amplified by using universal M13 primers. After subsequent cycle sequencing and capillary electrophoresis, data were analyzed by using the programs Seqman II and MegAlign (DNASTAR).

Statistical Analysis of Data. The upper estimate of allele frequency (P) given zero observations was made by using the formula $P = 1 - \alpha^{(1/n)}$, where α is the probability of no occurrences of an allele in n observations. Early Neolithic (≈7,000 years ago) −13.910*T allele frequencies were estimated for certain values for selection (S) by using a simple deterministic model of allele frequency change assuming (*i*) a modern-day central European −13.910*T allele frequency of ≈0.5, (*ii*) the coefficient of selection to have been constant over the time, (*iii*) a panmictic central European population, and (*iv*) an intergeneration time of 30 years. The expected generation-by-generation allele frequencies were calculated forward through time in Microsoft Excel, starting with very low frequencies, and using equation 3.5 from Maynard Smith (62). The Neolithic allele frequency was found by looking up which frequency would change to 0.5 in 233 generations for a certain value of S.

We thank Jörg Schibler, Magdalena Bogus, Rüdiger Krause, Sabine Möller-Rieker, Valerie Liebs, and Guido Brandt for their help. We thank Ildiko Pap, Zsuzsanna Guba, Veit Dresely, Rimantas Jankauskas, and Jerzy Siemaszko for providing samples. This work was funded by the Bundesministerium für Bildung und Forschung and the Arts and Humanities Research Council Centre for the Evolution of Cultural Diversity.

1. Swallow DM, Hollox EJ (2000) in *The Metabolic and Molecular Basis of Inherited Disease*, eds Scriver CR, Beaudet AL, Sly WS, Valle D (McGraw–Hill, New York), pp 1651–1663.
2. Mulcare CA, Weale ME, Jones AL, Connell B, Zeitlyn D, Tarekegn A, Swallow DM, Bradman N, Thomas MG (2004) *Am J Hum Genet* 74:1102–1110.
3. Enattah NS, Sahi T, Savilahti E, Terwilliger JD, Peltonen L, Jarvela I (2002) *Nat Genet* 30:233–237.
4. Lewinsky RH, Jensen TG, Moller J, Stensballe A, Olsen J, Troelsen JT (2005) *Hum Mol Genet* 14:3945–3953.
5. Ingram CJ, Elamin MF, Mulcare CA, Weale ME, Tarekegn A, Raga TO, Bekele E, Elamin FM, Thomas MG, Swallow DM (2007) *Hum Genet* 120(6):779–788.
6. Tishkoff SA, Reed FA, Ranciaro A, Voight BF, Babbitt CC, Silverman JS, Powell K, Mortensen HM, Hirbo JB, Osman M, *et al.* (2007) *Nat Genet* 39:31–40.
7. Cavalli-Sforza LL (1973) *Am J Hum Genet* 25:82–104.
8. Simoons FJ (1970) *Am J Dig Dis* 15:695–710.
9. McCracken RD (1971) *Curr Anthropol* 12:479–517.
10. Kretchmer N (1972) *Sci Am* 227(4):71–78.
11. Aoki K, Sato K (1986) *Jpn Heart J* 27:693–699.
12. Beja-Pereira A, Luikart G, England PR, Bradley DG, Jann OC, Bertorelle G, Chamberlain AT, Nunes TP, Metodiev S, Ferrand N, *et al.* (2003) *Nat Genet* 35:311–313.
13. Bayless TM, Paige DM, Ferry GD (1971) *Gastroenterology* 60:605–608.
14. Nei M, Saitou N (1986) *Prog Clin Biol Res* 214:21–37.
15. Aoki K (2001) *Theor Popul Biol* 59:253–261.
16. Bersaglieri T, Sabeti PC, Patterson N, Vanderploeg T, Schaffner SF, Drake JA, Rhodes M, Reich DE, Hirschhorn JN (2004) *Am J Hum Genet* 74:1111–1120.
17. Coelho M, Luiselli D, Bertorelle G, Lopes AI, Seixas S, stro-Bisol G, Rocha J (2005) *Hum Genet* 117:329–339.
18. Bollongino R, Edwards CJ, Alt KW, Burger J, Bradley DG (2006) *Biol Lett* 2:155–159.

19. Hofreiter M, Serre D, Poinar HN, Kuch M, Pääbo S (2001) *Nat Rev Genet* 2:353–359.
20. Sampietro ML, Gilbert MT, Lao O, Caramelli D, Lari M, Bertranpetit J, Lalueza-Fox C (2006) *Mol Biol Evol* 23:1801–1807.
21. Handt O, Krings M, Ward RH, Pääbo S (1996) *Am J Hum Genet* 59:368–376.
22. Handt O, Hoss M, Krings M, Pääbo S (1994) *Experientia* 50:524–529.
23. Bouwman AS, Chilvers ER, Brown KA, Brown TA (2006) *Am J Phys Anthropol* 131:428–431.
24. Hansen A, Willerslev E, Wiuf C, Mourier T, Arctander P (2001) *Mol Biol Evol* 18:262–265.
25. Hofreiter M, Jaenicke V, Serre D, Haeseler AA, Pääbo S (2001) *Nucleic Acids Res* 29:4793–4799.
26. Gilbert MT, Hansen AJ, Willerslev E, Rudbeck L, Barnes I, Lynnerup N, Cooper A (2003) *Am J Hum Genet* 72:48–61.
27. Gilbert MT, Binladen J, Miller W, Wiuf C, Willerslev E, Poinar H, Carlson JE, Leebens-Mack JH, Schuster SC (2007) *Nucleic Acids Res* 35:1–10.
28. Binladen J, Wiuf C, Gilbert MT, Bunce M, Barnett R, Larson G, Greenwood AD, Haile J, Ho SY, Hansen AJ, *et al.* (2006) *Genetics* 172:733–741.
29. Stiller M, Green RE, Ronan M, Simons JF, Du L, He W, Egholm M, Rothberg JM, Keates SG, Ovodov ND, *et al.* (2006) *Proc Natl Acad Sci USA* 103:13578–13584.
30. Cooper A, Poinar HN (2000) *Science* 289:1139 (lett).
31. Gilbert MT, Bandelt HJ, Hofreiter M, Barnes I (2005) *Trends Ecol Evol* 20:541–544.
32. Haak W, Forster P, Bramanti B, Matsumura S, Brandt G, Tanzer M, Villems R, Renfrew C, Gronenborn D, Alt KW, *et al.* (2005) *Science* 310:1016–1018.
33. Bramanti B, Hummel S, Schultes T, Herrmann B (2000) *Anthropol Anz* 58:45–49.
34. Burger J, Hummel S, Herrmann B, Henke W (1999) *Electrophoresis* 20:1722–1728.
35. Balogh MK, Burger J, Bender K, Schneider PM, Alt KW (2003) *Forensic Sci Int* 137:188–195.
36. Troelsen JT (2005) *Biochim Biophys Acta* 1723:19–32.
37. Burger J, Gronenborn D, Forster P, Matsumura S, Bramanti B, Haak W (2006) *Science* 312:1875–1876.
38. Tiedemann R, Moll K, Paulus KB, Scheer M, Williot P, Bartel R, Gessner J, Kirschbaum F (November 17, 2006) *Naturwissenschaften*, 10.1007/s00114-006-0175-1.
39. Römpler H, Rohland N, Lalueza-Fox C, Willerslev E, Kuznetsova T, Rabeder G, Bertranpetit J, Schöneberg T, Hofreiter M (2006) *Science* 313:62.
40. Greenwood AD, Capelli C, Possnert G, Pääbo S (1999) *Mol Biol Evol* 16:1466–1473.
41. Edwards CJ, Connellan J, Wallace PF, Park SD, McCormick FM, Olsaker I, Eythorsdottir E, MacHugh DE, Bailey JF, Bradley DG (2003) *Anim Genet* 34:410–416.
42. Burger J, Pfeiffer I, Hummel S, Fuchs R, Brenig B, Herrmann B (2001) *Ancient Biomol* 3:227–238.
43. Ricaut FX, Keyser-Tracqui C, Crubezy E, Ludes B (2005) *Forensic Sci Int* 151:31–35.
44. Gerstenberger J, Hummel S, Schultes T, Häck B, Herrmann B (1999) *Eur J Hum Genet* 7:469–477.
45. Bramanti B, Hummel S, Chiarelli B, Herrmann B (2003) *Hum Biol* 75:105–115.
46. Bandelt HJ (2005) *Eur J Hum Genet* 13:1106–1112.
47. Sherratt AG (1981) in *Pattern of the Past: Studies in Honour of David Clarke*, eds Hodder I, Isaac G, Hammond N (Cambridge Univ Press, Cambridge, UK), pp 261–305.
48. Harrison RJ (1985) *Proc Prehist Soc* 51:75–102.
49. Benecke N (1994) *Der Mensch und seine Haustiere. Die Geschichte einer jahrtausendealten Beziehung* (Konrad Theiss, Stuttgart).
50. Helmer D, Vigne J-D (2004) in *Approches Fonctionnelles en Préhistoire* (Actes XXVè Congrès Préhistorique Français, Nanterre, France, 24–26 Nov 2000), eds Bodu P, Constantin C (Société Préhistoire Française, Paris), pp 397–407.
51. Bogucki PI (1984) *Oxford J Archaeol* 3:15–30.
52. Craig OE, Chapman J, Heron C, Willis LH, Bartosiewicz L, Taylor G, Whittle A, Collins M (2005) *Antiquity* 79:882–894.
53. Copley MS, Berstan R, Dudd SN, Docherty G, Mukherjee AJ, Straker V, Payne S, Evershed RP (2003) *Proc Natl Acad Sci USA* 100:1524–1529.
54. Copley MS, Berstan R, Dudd SN, Aillaud S, Mukherjee AJ, Straker V, Payne S, Evershed RP (2005) *Antiquity* 79:895–908.
55. Edmonds CA, Lillie AS, Cavalli-Sforza LL (2004) *Proc Natl Acad Sci USA* 101:975–979.
56. Enattah N, Valimaki VV, Valimaki MJ, Loyttyniemi E, Sahi T, Jarvela I (2004) *Calcif Tissue Int* 75:488–493.
57. Currat M, Excoffier L (2005) *Proc Biol Sci* 272:679–688.
58. Müller U (2002) *Jahresschrift für mitteldeutsche Vorgeschichte* 85:77–90.
59. Whittle A, Bartosiewicz L, Borić D, Pettitt P, Richards M (2002) *Antaeus* 25:63–117.
60. Sherratt AG (1982) *Oxford J Archaeol* 1:287–316.
61. Burger J, Rosendahl W, Loreille O, Hemmer H, Eriksson T, Gotherstrom A, Hiller J, Collins MJ, Wess T, Alt KW (2004) *Mol Phylogenet Evol* 30:841–849.
62. Maynard Smith J (1998) *Evolutionary Genetics* (Oxford Univ Press, Oxford, UK).