npg

ARTICLE

Low prevalence of lactase persistence in Neolithic South-West Europe

This article has been corrected since Advance Online Publication and a corrigendum is also printed in this issue

Theo S Plantinga[1,2], Santos Alonso[3], Neskuts Izagirre[3], Montserrat Hervella[3], Rosa Fregel[4], Jos WM van der Meer[1,2], Mihai G Netea[1,2] and Concepcion de la Rúa*[,3]

The ability of humans to digest the milk component lactose after weaning requires persistent production of the lactose-converting enzyme lactase. Genetic variation in the promoter of the lactase gene (*LCT*) is known to be associated with lactase production and is therefore a genetic determinant for either lactase deficiency or lactase persistence during adulthood. Large differences in this genetic trait exist between populations in Africa and the Middle-East on the one hand, and European populations on the other; this is thought to be due to evolutionary pressures exerted by consumption of dairy products in Neolithic populations in Europe. In this study, we have investigated lactase persistence of 26 out of 46 individuals from Late Neolithic through analysis of ancient South-West European DNA samples, obtained from two burials in the Basque Country originating from 5000 to 4500 YBP. This investigation revealed that these populations had an average frequency of lactase persistence of 27%, much lower than in the modern Basque population, which is compatible with the concept that Neolithic and post-Neolithic evolutionary pressures by cattle domestication and consumption of dairy products led to high lactase persistence in Southern European populations. Given the heterogeneity in the frequency of the lactase persistence allele in ancient Europe, we suggest that in Southern Europe the selective advantage of lactose assimilation in adulthood most likely took place from standing population variation, after cattle domestication, at a post-Neolithic time when fresh milk consumption was already fully adopted as a consequence of a cultural influence.

INTRODUCTION

Throughout the world, large differences in the ability to digest lactose after weaning, designated as lactase persistence, are observed between populations. Although lactase persistence reaches up to 60–90% in Central-European and Northern European populations, it is much less frequently present in Southern European, Middle-Eastern, African and some South-Asian populations, while being completely absent in the rest of the world population.[1–3] The persistent activity of the lactase enzyme (ie, lactase phlorizin hydrolase, encoded by the lactase gene *LCT*) during adulthood, expressed especially in the intestine, has been demonstrated to be linked with variations in the lactase promoter, such as a C/T transition in the promoter region 13910 basepairs upstream of the *LCT* gene (rs4988235), which is suggested to influence *LCT* gene expression.[4] The T allele, dominant over the C allele, has been described to be the allele associated with lactase persistence in European populations.[5] In other populations, different polymorphisms in the *LCT* promoter have been also shown to be associated with lactase persistence.[6]

Several studies support the perspective that these differences in lactase persistence between populations can be ascribed to processes of positive selection.[1,2] Indeed, it has been suggested that the frequency of lactase persistence has increased in several populations as cattle domestication and dairy consumption was introduced in Europe. From that point a strong selective advantage of the ability to digest milk would have become apparent, which has resulted in a survival benefit of the already existing rare lactase persistent individuals in the population. As a consequence, the frequency of lactase persistence has increased rapidly in the generations to follow. This hypothesis has been described in literature as the 'culture–historical hypothesis'[7,8] and has been supported by recent studies in the Neolithic samples from Central, Mediterranean and Northern Europe[9–11] and medieval samples from Central Europe.[3] Burger *et al*[9] reported the absence of lactase persistence in a total number of nine early Neolithic Central Europeans (7500 YBP), which argues that in this era Europeans were predominantly lactase deficient in adulthood. The absence of lactase persistence is also reported by Lacan *et al*[11] in 26 samples from a late Neolithic burial located in Southern France. However, in Neolithic Scandinavians the prevalence of the −13910T allele was only 5%.[10] The opposing 'reverse cause hypothesis', states that human populations were already differentiated with regard to lactase persistence frequency before the development of dairying, and that the presence of lactase persistence determined the adoption of milk production and consumption practices.[12]

[1]Department of Medicine, Radboud University Nijmegen Medical Centre, Nijmegen, The Netherlands; [2]Nijmegen Institute for Infection, Inflammation and Immunity (N4i), Radboud University Nijmegen Medical Centre, Nijmegen, The Netherlands; [3]Department of Genetics, Physical Anthropology and Animal Physiology, University of the Basque Country, Bizkaia, Spain; [4]Department of Genetics, University of La Laguna, Santa Cruz de Tenerife, Spain
*Correspondence: Dr C de la Rúa, Department of Genetics, Physical Anthropology and Animal Physiology, University of the Basque Country, Barrio Sarriena s/n 48940 Leioa, Bizkaia, Spain. Tel +34 946 012 544, Fax +34-946013145; E-mail: conchi.delarua@ehu.es
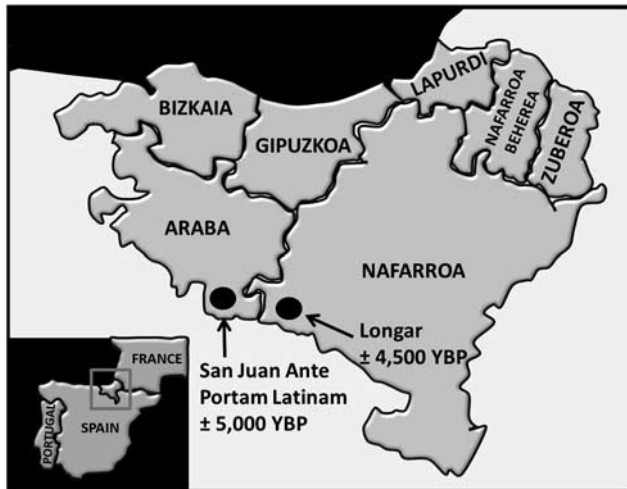
Figure 1 Late Neolithic burial locations in the Basque Country.

In Northern Europe, the advantage of the calcium assimilation hypothesis is likely to be the cause of the high lactase persistence at this latitude,[13] but in Southern Europe the presence of the −13910T allele displays a large variation between present populations.[2] The Iberian Peninsula represents one of the end points of Neolithic migration. The arrival or adoption of the Neolithic way of life in the Northern part of the Iberian Peninsula, here referred to as the Basque Country, is believed to have taken place around 7000 YBP. Several burials dating from this period have been identified and investigated archaeologically in the Basque Country, dating from 7000 to 4000 YBP.[14–16] There are various genetic arguments favouring the Basques as the most homogeneous relict population of the pre-Neolithic inhabitants of Europe. Therefore, the genetic make-up of the Neolithic Basques is likely to mirror the general genetic signature of Neolithic populations in Europe.[17–22] Although the prevalence of the −13910T allele in the modern Basque population reaches 66%,[2] and the selection coefficient for lactase persistence has been shown to be relatively high in the Iberian Peninsula,[23] the processes that shaped this process are controversial.

In order to assess whether lactase persistence was a common feature in Iberian populations immediately after the introduction of cattle domestication, or whether the −13910T high prevalence was a more recent Neolithic event, we performed genetic analysis of the −13910 C/T transition polymorphism on ancient DNA obtained from archaeological remains from two locations in the Basque Country. These include two Late Neolithic collective burials, from the sites of Longar (Navarre) and San Juan Ante Portam Latinam (SJAPL; Araba), dated ∼4500–5000 YBP (Figure 1).

MATERIALS AND METHODS

Temporal and geographical origin of the prehistoric samples
Of the 46 prehistoric samples analysed in this study, lactase profiles were obtained in 26 samples, 19 obtained from the site of SJAPL, which is located in the province of Araba (Basque Country, Spain). [14]C dating of human bone remains from this site dated to 5070 ± 150 YBP, that is, in the Late Neolithic period. Seven additional samples came from the site of Longar, a site located in the South of the province of Navarre (Spain). [14]C dating of human bone remains from this site dated to 4450 ± 70 YBP, that is, Late Neolithic–Early Chalcolithic.

Samples and extraction of DNA
The samples analysed in this study are dental pieces, which compared with bones, are less liable to external contamination. Thus, to prevent the

Table 1 Primer sequences and annealing temperatures used to amplify the promoter region of the *LCT* gene

Primer pair	Forward primer	Reverse primer	Anneal temperature (°C)
1	5′-TACTAGTAGG CCTCTGCGCT-3′	5′-TGCAACCTAA GGAGGAGAGT-3′	54
2	5′-AAATGTACTTAG ACCCTACAATGTA-3′	5′-AACCTAAGGA GGAGAGTTCC-3′	50
3	5′-AAAATGTACTTA GACCCTACAATG-3′	5′-ATACAAATGCA ACCTAAGGAG-3′	50

contamination of the endogenous DNA, we carefully selected those teeth that did not show any signs of caries or deep fissures that might extend into the pulp, our source of DNA.

The samples were processed according to a series of previously detailed criteria.[24–26] Thus, the extraction of DNA and set-up of the PCR were done in a positive pressure, sterile chamber, which was physically separated from the laboratory where the post-PCR processes are usually carried out. All of the working surfaces were regularly cleaned with sodium hypochlorite and were, in addition, regularly irradiated with UV light. If possible, PCR reagents were exposed to intense UV light before use. Suitable disposable clothing was worn (lab coat, mask, gloves and cap).

In order to eliminate surface contamination, the teeth were washed with acids to depurinate possible contaminating external DNA. Besides, the entire tooth surface was irradiated with UV light. Then, after cutting the root of the tooth, it was incubated at 37 °C and with agitation overnight in a lysis buffer (5 ml; 0.5 M EDTA pH 8.0–8.5; 0.5% SDS; 50 mM Tris HCl pH 8.0; 0.01 mg/ml proteinase K). Then, DNA was extracted using the conventional phenol–chloroform procedure. After the extraction, the DNA was concentrated and purified by means of Centricon-30 Amicon spin columns (Millipore, Billerica, MA, USA). Each extraction session involved two contamination controls, which were applied during both the extraction and amplification processes.

Amplification and sequencing of the LCT promoter region
PCR amplification was performed with the incorporation of the extraction controls obtained during the DNA extraction process. Negative controls for the PCR reaction were included. Three different primer pairs were used to amplify the promoter region of the *LCT* gene containing the −13910 position, which all covered a stretch of 80–120 basepairs (Table 1). PCR products were obtained by cycling 96 °C for 1 min for 1 cycle, followed by 45 cycles of 95 °C for 15 s, anneal temperature 30 s, 72 °C for 30 s and a final cycle of 72 °C for 10 min, accompanied by the corresponding annealing temperatures. After the amplification PCR products were purified by ExoSAP-IT (USB Corporation, OH, USA). Both forward and reverse sequences were obtained with the listed primers and Rhodamine or BigDye 1.1 chemistry in an ABI310 (Applied Biosystems, Cleveland, CA, USA). SDs of frequency estimates were calculated with the standard formula: $2 \times SD = \sqrt{(p \times q/2N) \times 2}$.

Mitochondrial DNA analysis
The hypervariable region I of mtDNA was resequenced in all investigated samples as a further control in order to discard possible contamination with the DNA from the investigators. PCR conditions were similar compared with amplification of the lactase amplicons, except the annealing temperatures, which are together with the used primers listed in Table 2.

DNA quantification
We used our standard procedure to quantify the extracted DNA, which consists of measuring the number of molecules of a segment of 113 bp of HVR-I of mtDNA by means of RT-PCR (Step-One, Applied Biosystems). For this, we used oligo 5′-CACCATTAGCACCCAAAGCT-3′ as forward primer and oligo 5′-ACAT AGCGGTTGTTGATGGG-3′ as reverse primer. The sequence of the Taqman probe was as follows: VIC-5′-GAAGCAGATTTGGGTAC-3′ (Applied Biosystems).

Table 2 Primer sequences and annealing temperatures used to amplify the hypervariable region I of mitochondrial (mt)DNA

Primer pair	Forward primer	Reverse primer	Anneal temp. (°C)
1	5'-CACCATTAGC ACCCAAAGCT-3'	5'-ACATAGCGGT TGTTGATGGG-3'	60
2	5'-TGAGCAGATT TGGGTACCAC-3'	5'-GTACTACAGG TGGTCAAGTAT-3'	57
3	5'-CACCATGAAT ATTGTACGGT-3'	5'-TGTGTGATAG TTGAGGGTTG-3'	58
4	5'-CCCCATGCTT ACAAGCAAGT-3'	5'-TGGCTTTATG TACTATGTAC-3'	55
5	5'-CACTAGGATAC CAACAAACCTA-3'	5'-CAAGGGACCC CTATCTGAGG-3'	58
6	5'-CGTACATAGC ACATTACAGT-3'	5'-TGATTTCACG GAGGATGGTG-3'	57

For each sample four replicates were performed, each in 30 μl containing 1X TaqMan Universal PCR Master Mix (Applied Biosystems), 5 μM each primer, 10 μM probe and 10 μl DNA extract (diluted 1/10 with BSA). The cycling conditions were 1 cycle of 50 °C for 2 min, 95 °C for 10 min, followed by 45 cycles of 95 °C for 15 s and 60 °C for 1 min in a StepOne Real-Time PCR System (Applied Biosystems).

For the standard curves, serial dilutions of plasmid pCR2.1-new (3.9 kb), including an insert of 450 bp (Eurofins MWG/Operon, Ebersberg, Germany) containing the HVR-I region of interest, were included in each experiment to generate standard curves. Two different standard curves were performed: one with three points of 1.4×10^4, 1.4×10^5 and 1.4×10^6 molecules/μl for high-concentration samples and a second standard curve for low-concentration samples with three points at 1.12×10^4, 2.24×10^3 and 448 molecules/μl. Four replicates were used for each dilution point. Typical (%efficiency, r^2) values were, respectively, (102%, 0.99) for the high-concentration curve, and (83.2%, 0.95) for the low-concentration samples. Finally, at least three 'no-template-control' samples were included within each experiment. Quantification of results are depicted in Supplementary Table 1.

Reproducibility of the results

Duplication LCT sequencing results. All samples were sequenced a second time with BigDye 3.1 chemistry using the reverse primer only. Both sequences were checked to be coincident at rs4988235.

Cloning of LCT PCR products. To corroborate the results from genomic DNA, the PCR products of all 26 samples were cloned (TOPO TA Cloning Kit, Invitrogen, Carlsbad, CA, USA) and from each cloning reaction 10 colonies picked. In brief, after vector ligation and bacterial expansion, DNA was extracted from selected colonies and was amplified by using universal M13 primers. Subsequently, the original LCT primers were used to perform the sequencing of these products.

Replication in independent laboratories. We assessed the reproducibility of the results obtained in the University of the Basque Country, Bilbao, Spain, by replicating the analysis of six overlapping fragments of mtDNA HVR-I of a second tooth from each of five prehistoric individuals in an additional independent laboratory (Radboud University Nijmegen Medical Centre, Nijmegen, The Netherlands). MtDNA was chosen because its high variability increases the chances to genetically individualise each sample, and therefore to increase the chances to detect (and track) contamination with exogenous DNA. The mtDNA types of most of the personnel that was in contact with the samples were also obtained (see Supplementary Table 2) in order to verify the possibility of contamination originating from the manipulators. Furthermore, lactase genotyping of three ancient DNA samples extracted from a second tooth was performed in a separate laboratory (University of La Laguna, Santa Cruz de Tenerife, Spain).

Table 3 LCT genotype and frequency of lactase persistence for the analysed samples obtained from the burials San Juan Ante Portam Latinam (SJAPL) and Longar

LCT −13910 genotype	SJAPL (N=19)	Longar (N=7)	Total (N=26)
C/C	13	6	19
T/C	2	—	2
T/T	4	1	5
T allele frequency	0.26	0.14	0.23
Frequency lactase persistence	31%	14%	27%

RESULTS

In this study, we have analysed 46 samples that are derived from human teeth from burials originating from the Late Neolithic–Chalcolithic era in the North of the Iberian Peninsula. Previously, these samples yielded reproducible human mitochondrial DNA sequences and were extracted from intact well-preserved teeth.

For the amplification of the promoter region of the LCT gene several primer pairs have been developed and initially tested on modern DNA for their specificity and PCR efficacy. On the basis of these observations, three primer pairs were chosen to use in the PCR amplifications of the ancient DNA (Table 1). In total, 26 ancient DNA samples could be analysed for the LCT −13910 genotype, hence a PCR efficiency of 57%, which revealed an average lactase persistence frequency of 27% (Table 3). Furthermore, mtDNA HVR-I haplotypes of the ancient DNA samples have been generated, as depicted in Supplementary Table 1. Calculation of the SD from frequency estimates was performed after excluding samples that have identical genotypes for both lactase −13910 and for mtDNA HVR-I and originate from the same burial to correct for possible consanguinity. On the basis of these criteria, four samples had to be excluded. With the exclusion of these samples, the calculation revealed $2 \times SD = 0.12$ and lactase persistence frequency is 0.27 with 95% CI 0.15–0.39.

These results have been reproduced by analysis in independent laboratories from five teeth for mitochondrial DNA (Department of Medicine, Radboud University Nijmegen Medical Centre, Nijmegen, The Netherlands) and from three teeth for the lactase genotype (Department of Genetics, University of La Laguna, Santa Cruz de Tenerife, Spain). Owing to the fact that no second tooth was available, we have been unable to replicate the results in the other samples. The replication of these results has been performed to exclude systemic contamination. These analyses confirmed the observations obtained after the initial genotyping.

Furthermore, lactase genotypes have been verified by cloning and resequencing of the PCR products, which revealed identical genotypes as compared with the initial sequencing (data not shown).

Analysis of the HVR-I hypervariable region of the mtDNA and of the LCT genotype in nuclear DNA of the ancient samples and samples from the investigators revealed that only two of the mtDNA genotypes in the investigators, also the most prevalent genotypes in the general population, were found in the ancient DNA samples (Supplementary Table 2).

DISCUSSION

In ancient DNA samples from the Basque Country that are dated ∼5000–4500 YBP we found that the frequency of the genotype −13910T associated with lactase persistence is 27%, while the frequency of LCT persistent allele in the modern Basque Country population is ∼66%.[1,2] These results show a nearly threefold increase

of the frequency of the lactase persistence associated allele (−13910 T) from the Neolithic era to the present time.

Gerbault et al[23] proposed the possibility of a substantial structure in the allele frequencies for LCT in prehistoric Europe. They suggested that the distribution of frequencies of lactase persistence (LCT*P) in Europe is highly compatible with a scenario of high positive selection in Northern Europe for the LCT*P (T) allele, associated to the calcium assimilation hypothesis, while in Southern Europe the substantial variation in the frequency of this allele in populations has not been so far explained satisfactorily. In this regard, the possibility to obtain data from prehistoric populations can provide further insights on this point.

Our data on prehistoric samples, in combination with those of Burger et al,[9] Malmström et al[10] and Lacan et al[11] show variable −13910T allele frequencies in prehistoric Europe. In our sample of 26 Neolithic individuals we find a frequency for the T allele of 0.23 (95% CI 0.11–0.35), whereas in the 8 Neolithic individuals (5000–5840 BC) from a metapopulation of samples from Germany, Hungary and Lithuania, Burger et al[9] report a collective frequency of 0 (95% CI 0–0.14). Similarly, Lacan et al[11] in a late Neolithic settlement from Southern France report a frequency of 0 (95% CI 0–0.056) in 26 samples. On the other hand, Malmström et al[10] report an estimated frequency for the T allele of 0.05 (95% CI 0.001–0.248) in 10 prehistoric individuals from four archaeological sites in the Isle of Gotland (Baltic Sea), which date to the Middle Neolithic, 4800–4200 YBP. Although wider sample sizes are needed to statistically accept that the T allele frequencies are significantly different between these four Neolithic populations, these data point to a scenario of a low allele frequency for the T allele in the prehistoric populations of Northern Europe, which contrasts with a relatively high frequency for this allele in the Iberian Peninsula.

The heritable ability to digest the milk sugar lactose during adulthood (ie, lactase persistence) has evolved to high frequencies in the last 10 000 years of European human history. According to the culture–historical hypothesis, the underlying process responsible for this increase is the evolutionary benefit of carrying this genetic trait at the time that cattle domestication and consumption of dairy products was introduced in Europe. From that point it became advantageous to carry the LCT persistent allele, which has resulted in an increased survival due to better nutrition and also a larger and healthier offspring per generation, or increased calcium assimilation in regions of Northern Europe characterized by low exposure to UV radiation and therefore insufficient vitamin D synthesis in the skin.[23]

However, there is no evidence that the frequency of the 'lactase persistence' phenotype, relatively high in the prehistoric samples from the Iberian Peninsula here analysed, could be attributed to a selective effect of fresh milk consumption for the following reasons. First, even though the prehistoric samples analysed in this work correspond to the Late Neolithic–Chalcolithic period, we still do not know precisely how relevant milk consumption was in the diet of these populations. We nevertheless can assume that it may have not been significant, considering that archaeological data show regional differences in milk consumption, which are associated with environmental conditions.[27] Also, archaeozoological data reveal that in this area wild animal remains from hunting activity represent a similar proportion as those from domestic animals,[28] and the analysis of teeth pathologies like caries patterns indicate, at least for the SJAPL population, a high consumption of carbohydrates from wild fruits and berries, which are rich in fermentable sugars (sucrose particularly) with a high cariogenic power, complemented with proteins of animal origin. It is thought that even from an early age, probably before weaning, there was a consumption of carbohydrates in this population (SJAPL).[16]

A second argument against a major role of selection for explaining the relatively high frequency of the lactase persistence phenotype in the prehistoric samples here analysed, is that genetic drift may have been important at Neolithic times, and consequently, prevailed over any effects of this alleged selection. Although agriculture may have allowed for increased population size by this time, human populations were most likely dispersed in smaller demes. Finally, the first clear evidence for the use of these animals for their milk and other 'secondary' products from living animals is controversial, although the domestication of cattle, sheep and goats had already taken place in the Near East by the eighth millennium BC.[29,30] Organic residues preserved in archaeological pottery have provided the earliest direct evidence for the use of milk in the seventh millennium in the Near East.[27] But, although milking was particularly important in North-Western Anatolia, the data point to regional differences linked with conditions more favourable to cattle husbandry, compared with other regions, where milk usage was less important.[27] Thus, although some researchers suggest that dairy products would have been exploited rapidly after animal domestication,[31] others have suggested that early domestication was predominantly for meat and hides, postulating a 'secondary products revolution' 2000–4000 years after the first domestication of cattle, sheep and goats in the Near East and Europe.[32] In Europe, it is estimated that agriculture has been introduced from 8000 to 6000 YBP,[33,34] and specifically for the Basque Country around 7000 YBP.[35] Taking into account that the full exploitation of secondary products from animal domestication could have happened several millennia after the domestication process itself began, it is therefore unlikely that milk products represented a high proportion of diet in the populations tested here. If the calcium assimilation hypothesis as proposed by Gerbault et al could explain the high values found in the extant populations of Northern Europe, this would not seem to be a proper explanation for those population in the South of Europe, such as the Iberian Peninsula.

Thus, unless selective forces unknown yet were at work, the selective advantage of lactase persistence in the prehistoric populations analysed here was not yet strong enough to increase lactase persistence substantially. Nevertheless, its frequency could have risen in more recent times as a consequence of fresh milk consumption due to cultural pressure. In this case, the rise in the frequency of this phenotype could have started from standing variation, in which the frequency of the T allele was already relatively high due to stochastic reasons. This could also explain the variable frequencies for lactase persistence in present and past Southern Europe (0% in Southern France, Lacan et al, and 27% in the Iberian Peninsula, present study). Therefore, we suggest that in Southern Europe, the selective advantage of lactose assimilation in adulthood most likely took place from standing population variation, after cattle domestication, at a time when fresh milk consumption was already fully adopted as a consequence of a cultural influence.

Recent data have shown that the −13910T variant in the lactase promoter is present on two greatly different haplotypes, one present in European populations, and one in populations from West-Urals and Caucasus.[2] The association of this allele with the phenotype of lactase persistence strongly supports its importance, and is a suggestive example of convergent evolution. The frequency of the −13910T allele in Europe is the greatest in the North-Western parts of the continent, with intermediary and sometimes low frequencies in the Southern parts.[1] Although this distribution would suggest an origin of the mutation in the populations living in the North, a recent modelling of the evolutionary history of the mutation has rather surprisingly suggested its origin in a region located between the Central Balkans and Central Europe, which spread through the dissemination of

Neolithic Linearbandkeramiek culture.[36] Although this may seem counterintuitive on the first sight, other simulations have also shown that the geographic origin of an allele can differ from the region of highest frequency, in particular when it occurs on the wave front of a demographic expansion.[37,38] On the basis of these data, one may speculate regarding the likely arrival of the lactase persistence genotype with Neolithic farmers arriving around 7000 YBP from Central Europe.

Several measures of avoiding contamination have been applied to ensure the validity of our findings. These included reproduction of mtDNA sequencing data previously generated from the same samples, which show a high interindividual variability and help to confirm the authentic origin of the samples.[39] Indeed, only few Neolithic individuals, six in total, were demonstrated to bear the Cambridge Reference Sequence genotype. In the other 20 individuals, 17 different HVR-I haplotypes have been observed, which demonstrates the authenticity of the ancient DNA samples. Furthermore, the data on nuclear DNA reported here have been reproduced for three samples in an independent laboratory by independent workers. The authenticity of the results are reinforced by the observation that the frequency of the alleles in our prehistoric sample collection is different from the frequencies in modern populations. Hence, a random modern contaminant population would look different. Furthermore, mtDNA of archaeologists and lab workers has been analysed, which were demonstrated to be mostly different from the ancient DNA samples investigated, further excluding potential contamination with modern DNA. Finally, nuclear DNA needed to obtain the lactase genotype is generally less likely to be contaminated than mtDNA.[40]

In conclusion, considering the frequencies of lactase persistence in modern populations in Europe that reaches 80%, and the quasi-absence of this genotype in Mesolithic and Early Neolithic Central-European samples, we can conclude that an ancient South-West European population from the Basque Country displays an intermediate prevalence of the −13910T genotype of 27%. This suggests that during, but especially after the Neolithic, a positive selective pressure on lactase persistence may have been exerted after cattle domestication, but this took place from standing population variation, at a time when fresh milk consumption was already fully adopted as a consequence of a cultural influence.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

1 Bersaglieri T, Sabeti PC, Patterson N et al: Genetic signatures of strong recent positive selection at the lactase gene. Am J Hum Genet 2004; 74: 1111–1120.
2 Enattah NS, Trudeau A, Pimenoff V et al: Evidence of still-ongoing convergence evolution of the lactase persistence T-13910 alleles in humans. Am J Hum Genet 2007; 81: 615–625.
3 Nagy D, Tomory G, Csanyi B et al: Comparison of lactase persistence polymorphism in ancient and present-day Hungarian populations. Am J Phys Anthropol 2011.
4 Lewinsky RH, Jensen TG, Moller J, Stensballe A, Olsen J, Troelsen JT: T-13910 DNA variant associated with lactase persistence interacts with Oct-1 and stimulates lactase promoter activity in vitro. Hum Mol Genet 2005; 14: 3945–3953.
5 Enattah NS, Sahi T, Savilahti E, Terwilliger JD, Peltonen L, Jarvela I: Identification of a variant associated with adult-type hypolactasia. Nat Genet 2002; 30: 233–237.
6 Tishkoff SA, Reed FA, Ranciaro A et al: Convergent adaptation of human lactase persistence in Africa and Europe. Nat Genet 2007; 39: 31–40.
7 Simoons FJ: Primary adult lactose intolerance and the milking habit: a problem in biologic and cultural interrelations. II. A culture historical hypothesis. Am J Dig Dis 1970; 15: 695–710.
8 Beja-Pereira A, Luikart G, England PR et al: Gene-culture coevolution between cattle milk protein genes and human lactase genes. Nat Genet 2003; 35: 311–313.
9 Burger J, Kirchner M, Bramanti B, Haak W, Thomas MG: Absence of the lactase-persistence-associated allele in early Neolithic Europeans. Proc Natl Acad Sci USA 2007; 104: 3736–3741.
10 Malmström H, Linderholm A, Liden K et al: High frequency of lactose intolerance in a prehistoric hunter-gatherer population in northern Europe. BMC Evol Biol 2010; 10: 89–96.
11 Lacan M, Keyser C, Ricaut FX et al: Ancient DNA reveals male diffusion through the Neolithic Mediterranean route. Proc Natl Acad Sci USA 2011; 108: 9788–9791.
12 Aoki K: Theoretical and empirical aspects of gene-culture coevolution. Theor Popul Biol 2001; 59: 253–261.
13 Flatz G, Rotthauwe HW: Lactose nutrition and natural selection. Lancet 1973; 2: 76–77.
14 Izagirre N, de la Rua C: An mtDNA analysis in ancient Basque populations: implications for haplogroup V as a marker for a major Paleolithic expansion from southwestern Europe. Am J Hum Genet 1999; 65: 199–207.
15 Márquez B, Gibaja J, González Urquijo J, Ibáñez J, Palomo A: Projectile points as signs of violence in collective burials during the 4th and the 3rd millenium Cal. BC in the NE of the Iberian Peninsula. Digital.CSIC) 2009.
16 de la Rua C, Arriaga H: San Juan ante Portam Latinam and Longar (Final Neolithic - Chalcolithic). Cuadernos de Ciencias Médicas de la Sociedad de Estudios Vascos 2004; 6: 239–301.
17 Chikhi L, Nichols RA, Barbujani G, Beaumont MA: Y genetic data support the Neolithic demic diffusion model. Proc Natl Acad Sci USA 2002; 99: 11008–11013.
18 Lucotte G, Hazout S: Y-chromosome DNA haplotypes in Basques. J Mol Evol 1996; 42: 472–475.
19 Richards M, Macaulay V, Hickey E et al: Tracing European founder lineages in the Near Eastern mtDNA pool. Am J Hum Genet 2000; 67: 1251–1276.
20 Semino O, Passarino G, Oefner PJ et al: The genetic legacy of Paleolithic Homo sapiens sapiens in extant Europeans: a Y chromosome perspective. Science 2000; 290: 1155–1159.
21 Torroni A, Bandelt HJ, D'Urbano L et al: mtDNA analysis reveals a major late Paleolithic population expansion from southwestern to northeastern Europe. Am J Hum Genet 1998; 62: 1137–1152.
22 Torroni A, Bandelt HJ, Macaulay V et al: A signal, from human mtDNA, of postglacial recolonization in Europe. Am J Hum Genet 2001; 69: 844–852.
23 Gerbault P, Moret C, Currat M, Sanchez-Mazas A: Impact of selection and demography on the diffusion of lactase persistence. PLoS One 2009; 4: e6369.
24 Cooper A, Poinar HN: Ancient DNA: do it right or not at all. Science 2000; 289: 1139.
25 Hofreiter M, Serre D, Poinar HN, Kuch M, Paabo S: Ancient DNA. Nat Rev Genet 2001; 2: 353–359.
26 Gilbert MT, Bandelt HJ, Hofreiter M, Barnes I: Assessing ancient DNA studies. Trends Ecol Evol 2005; 20: 541–544, .
27 Evershed RP, Payne S, Sherratt AG et al: Earliest date for milk use in the Near East and southeastern Europe linked to cattle herding. Nature 2008; 455: 528–531.
28 Altuna J, Mariezkurrena K: Types of livestock populations during the Neolithic in the Basque Country and surrounding areas; Tipos de cabañas ganaderas durante el Neolítico del País Vasco y zonas próximas. Archaeofauna 2009; 18: 137–157.
29 Clutton-Brock J: A Natural History of Domesticated Mammals. Cambridge: Cambridge University Press, 1999.
30 Garrard A, Colledge S, Martin L: The emergence of crop cultivation and caprine herding in the 'Marginal Zone' of the Southern Levant. The origin and spread of agriculture and pastoralism in Eurasia. University College London Press, 1996, p. 204–226.
31 Rollefson GO, Kohler-Rollefson I: Early Neolithic exploitation patterns in the Levant: cultural impact of the environment. Popul Environ 1992; 13: 243–254.
32 Sherratt A: The secondary exploitation of animals in the Old World. World Archaeol 1983; 15: 90–104.
33 Craig OE: The development of dairying in Europe: potential evidence from food residues on ceramics. Documenta Praehistorica 2002; 29: 97–107.
34 Spangenberg JE, Matuschik I, Jacomet S, Schibler J: Direct evidence for the existence of dairying farms in prehistoric Central Europe (4th millennium BC). Isotopes Environ Health Stud 2008; 44: 189–200.
35 Zapata L, Pena-Chocarro L, Perez-Jorda G, Stika HP: Early Neolithic Agriculture in the Iberian Peninsula. J World Prehist 2004; 18: 283–325.
36 Itan Y, Powell A, Beaumont MA, Burger J, Thomas MG: The origins of lactase persistence in Europe. PLoS Comput Biol 2009; 5: e1000491.
37 Edmonds CA, Lillie AS, Cavalli-Sforza LL: Mutations arising in the wave front of an expanding population. Proc Natl Acad Sci USA 2004; 101: 975–979.
38 Klopfstein S, Currat M, Excoffier L: The fate of mutations surfing on the wave of a range expansion. Mol Biol Evol 2006; 23: 482–490.
39 Alzualde A, Izagirre N, Alonso S, Alonso A, de la Rua C: Temporal mitochondrial DNA variation in the Basque Country: influence of post-neolithic events. Ann Hum Genet 2005; 69: 665–679.
40 Haak W, Forster P, Bramanti B et al: Ancient DNA from the first European farmers in 7500-year-old Neolithic sites. Science 2005; 310: 1016–1018.

Supplementary Information accompanies the paper on European Journal of Human Genetics website (http://www.nature.com/ejhg)