

Lactase Persistence in Central Asia: Phenotype, Genotype, and Evolution

Author(s): Evelyne Heyer, Lionel Brazier, Laure Ségurel, Tatiana Hegay, Frédéric Austerlitz, Lluís Quintana-Murci, Myriam Georges, Patrick Pasquet, and Michel Veuille

Source: *Human Biology*, 83(3):379-392. 2011.

Published By: Wayne State University Press

DOI: <http://dx.doi.org/10.3378/027.083.0304>

URL: <http://www.bioone.org/doi/full/10.3378/027.083.0304>

BioOne (www.bioone.org) is a nonprofit, online aggregation of core research in the biological, ecological, and environmental sciences. BioOne provides a sustainable online platform for over 170 journals and books published by nonprofit societies, associations, museums, institutions, and presses.

Your use of this PDF, the BioOne Web site, and all posted and associated content indicates your acceptance of BioOne's Terms of Use, available at www.bioone.org/page/terms_of_use.

Usage of BioOne content is strictly limited to personal, educational, and non-commercial use. Commercial inquiries or rights and permissions requests should be directed to the individual publisher as copyright holder.

Lactase Persistence in Central Asia: Phenotype, Genotype, and Evolution

EVELYNE HEYER,¹ LIONEL BRAZIER,² LAURE SÉGUREL,^{1,3} TATIANA HEGAY,⁴
FRÉDÉRIC AUSTERLITZ,^{1,5} LLUIS QUINTANA-MURCI,⁶ MYRIAM GEORGES,¹
PATRICK PASQUET,^{1*} AND MICHEL VEUILLE²

Abstract The aim of the present study is to document the evolution of the lactase persistence trait in Central Asia, a geographical area that is thought to have been a region of long-term pastoralism. Several ethnic groups co-exist in this area: Indo-Iranian speakers who are traditionally agriculturist (Tajik) and Turkic speakers who used to be nomadic herders (Kazakh, Karakalpak, Kyrgyz, Turkmen). It was recently demonstrated that horse milking practice existed in the Botai culture of Kazakhstan as early as 5,500 BP (Outram et al. 2009). However, the frequency of the lactase persistence trait and its genetic basis in Central Asian populations remain largely unknown. We propose here the first genotype-phenotype study of lactase persistence in Central Asia based on 183 individuals, as well as the estimation of the time of expansion of the lactase-persistence associated polymorphism. Our results show a remarkable genetic-phenotypic correlation, with the causal polymorphism being the same than in Europe (-13.910C>T, rs4988235). The lactase persistence trait is at low frequency in these populations: between 25% and 32% in the Kazakh population (traditionally herders), according to phenotype used, and between 11% and 30% in the Tajiko-Uzbek population (agriculturalists). The difference in lactase persistence between populations, even if small, is significant when using individuals concordant for both excretion of breath hydrogen and the lactose tolerance blood glucose test phenotypes ($P = 0.018$, 25% for Kazakh vs. 11% for Tajiko-Uzbeks), and the difference in frequency of the -13.910*T allele is almost significant ($P = 0.06$, 30% for Kazakhs vs. 19% for Tajiko-Uzbeks). Using the surrounding haplotype, we estimate a date of expansion of the T allele around 6,000–12,000 yrs ago, which is consistent with archaeological records for the emergence of agropastoralism and pastoralism in Central Asia.

¹Eco-anthropology and Ethnobiology Unit, CNRS UMR 7206, Muséum National d'Histoire Naturelle, Université Paris Diderot, 75005 Paris, France.

²Origine, Structure et Evolution de la Biodiversité Unit, CNRS UMR 7205, Muséum National d'Histoire Naturelle, 75005 Paris, France.

³Present address: Department of Human Genetics, University of Chicago, Chicago, IL.

⁴Uzbek Academy of Sciences, Institute of Immunology, Tashkent, Uzbekistan.

⁵Écologie, Systématique et Évolution Unit, CNRS UMR 8079, Université Paris Sud, 91405 Orsay, France.

⁶Human Evolutionary Genetics Unit, CNRS URA3012, Institut Pasteur, 7,5015 Paris, France.

*Correspondence to: Patrick Pasquet, Eco-anthropology and Ethnobiology Unit, CNRS UMR 7206, Muséum National d'Histoire Naturelle, Université Paris Diderot, 75005 Paris, France.

Human Biology, June 2011, v. 83, no. 3, pp. 379–392.

Copyright © 2011 Wayne State University Press, Detroit, Michigan 48201-1309

KEY WORDS: LACTASE PERSISTENCE, CENTRAL ASIA, NATURAL SELECTION, LIFESTYLE, PASTORALISM, -13910*T.

The practice of farming during the Neolithic transition, which brought the habit of drinking fresh milk, remains a paradigmatic case of how humans culturally adapted to a natural resource to improve their energetic intake. In mammals, however, the sugar provided by milk, lactose, can only be metabolized during youth, a period in which we synthesize lactase, the lactose-processing enzyme. After weaning, the main source of sugar in the diet of frugivorous mammals becomes sucrose, and lactase is not synthesized any further. Interestingly, in some human populations, a high proportion of individuals maintain the ability to digest fresh milk in adulthood, as a result of a change in gene regulation allowing them to produce lactase throughout their adult life (Ferguson and Maxwell 1967; Ho et al. 1982; Sahi 1974a; Sahi 1974b; Wang et al. 1995). The question of “biological and cultural interrelations” in lactose intolerance was independently put forward by Simoons (1969) and McCracken (1971) based on the observation that the “lactose persistence” phenotype seems to result from the relatively long-term intake of milk in the diet of several human populations. Some authors then formally tested the hypothesis that individuals showing such lactase persistence are found at a higher frequency in populations whose ancestors traditionally consumed fresh milk (Holden and Mace 1997). For example, the frequency of the lactase persistence trait is very high (>80%) in some European populations like the Finnish (Jussila 1969), Czech, or Irish (Holden and Mace 1997) and some African pastoralists like the Beja, Tutsi, Tuareg, or Bedouins (Holden and Mace 1997; Tishkoff et al. 2007), whereas it is extremely low (<5%) in the Pima and Chippewa from South America, the Herero from Africa, or in China (Bolin and Davis 1969; Bryant et al. 1970; Chua and Seah 1973; Holden and Mace 1997; Sung and Shih 1972). Theoretical treatment by Aoki (1986) and Feldman and Cavalli-Sforza (1984) showed that the rise in frequency of lactase persistence since Neolithic times in Europe implies that the advantage conferred to lactose-digesting individuals must have been of the magnitude of some percent of their Darwinian fitness, which is quite considerable. Other evolutionary hypotheses, for example suggesting a selection pressure based on Vitamin D in relation with latitude, remain controversial; they have been excluded using simulation-based approaches in a recent paper (Itan et al. 2009), but not in another (Gerbault et al. 2009).

Even if the precise biological mechanism of lactase persistence remains unknown, several polymorphisms in the vicinity of the LCT gene have proven to be associated with this phenotype. A major polymorphism is the C>T substitution located -13.910 upstream of the LCT gene (rs4988235). It appears as the main allele explaining lactase persistence among Europeans (Enattah et al. 2002). Conversely, the neighboring alleles -13.907*G, -13.915*G, and -14.010*C have been shown to explain the majority of cases of lactase persistence in Africa, an observation that constitutes a clear case of convergent evolution between human populations subject to the same selective pressure (Tishkoff et al. 2007). In Middle East, the allele -13.915*G has also proven to be of main importance

(Enattah et al. 2008; Imtiaz et al. 2007). Using polymorphic markers in the surrounding genomic region, the increase in frequency of these polymorphisms has been dated to 2,000–20,000 yrs ago in Europe (Bersaglieri et al. 2004) and 3,000–7,000 yrs ago in Africa (Tishkoff et al. 2007), consistent with archaeological records on cattle domestication in these areas.

The aim of the present study is to document the evolution of the lactase persistence trait in Central Asia, a geographical area that is thought to have been a region of long-term pastoralism (see for example Frachetti 2008; Jacobson-Tepfer 2008). Furthermore, the vast region of Central Asia is located at the crossroads of cultural transmission routes, between Europe and Asia, as it is located along the Silk Road. Several ethnic groups co-exist in this area: Indo-Iranian-speaking populations that are traditionally agriculturist (Tajik), and Turkic-speaking populations that used to be nomadic herders in the past (Kazakh, Karakalpak, Kyrgyz, Turkmen). Finally, Uzbek seem to be a congregate of tribes with different origins (Martínez-Cruz et al. 2010). It was recently demonstrated that horse milking practice existed in the Botai culture of Kazakhstan as early as 5500 BP (Outram et al. 2009). This indicates that the evolution of strategies for exploiting animals for milk consumption was not contingent on the adoption of the conventional “agricultural package,” from the Fertile Crescent (comprising wheat, barley, lentils, peas, but also cows, goats and sheep), as it appears to have developed independently in the Botai region. However, the frequency of the lactase persistence trait as well as its genetic basis in Central Asian populations remains largely unknown.

In the framework of a wider project aiming to better understand the genetic and cultural diversity of Central Asian populations, we sampled both individuals belonging to a traditional agriculturist population and individuals belonging to a nomadic-herder population in order to (1) measure the frequency of the lactase persistence trait in this poorly known area of the world, (2) perform phenotype-genotype correlations to test which of the known alleles explain this trait in Central Asia, and (3) investigate the time of the expansion of the lactase-persistence associated allele.

Materials and Methods

Population Samples. One hundred eighty-three unrelated adult individuals have been sampled in Uzbekistan: 83 Kazakh living in the city of Gazli situated north of Bukhara city (close to the Kazakh border), belonging to the Turkic language family and being traditionally nomadic herders; 100 Tajiko-Uzbek living in semi-rural area south of Bukhara and speaking both Tajik and Uzbek at home. Importantly, our global genetic diversity study shows that the individuals sampled here clearly cluster with others agriculturalist Tajik populations (Martínez-Cruz et al. 2010).

For all individuals, ethnic affiliation was carefully done with questionnaires assessing self-identified ethnicity, self-identified language spoken, as well

as parents' and grandparents' ethnicity and language spoken, as described in Martínez-Cruz et al. (2010).

Ethics. Approvals for this study were obtained from the Health Ministry of Uzbekistan and the Uzbek Academy of Science. All individuals were fully informed on the goal of our study, and they all signed informed consent. The phenotypic study took place in a medical centre under medical supervision, and preliminary results were given to all participants. Known diabetic subjects and pregnant women were excluded from the test. All individuals with high fasting glucose were recommended for medical follow-up.

Phenotypic Measurements. Donors from the two populations were investigated for their lactase persistence status after ingestion of 50 g of lactose in solution with 250 ml of water after an overnight fast. We used three different tests: increased excretion of breath hydrogen (BH), increased blood glucose (BG), and increased intensity of gastrointestinal symptoms. We measured hydrogen in exhaled air using a portable breath hydrogen analyzer MicroH₂, (Micro Medical Limited, Chapman, UK). First, fasting measurements were done before the lactose ingestion, then measurements were done at 120 min and 150 min after the lactose load. We used the following criterion to classify individuals: an individual with an increased excretion of breath hydrogen ≥ 20 ppm (Caskey et al. 1977) above baseline was classified as lactose maldigester, and one with a rise of less than 20 ppm was diagnosed digester. The capillary blood glucose concentration from a fingertip was measured before lactose ingestion and 20 min and 40 min after ingestion of lactose using a portable glucometer. We used the following criterion: an individual with a rise > 1.1 mmol/l in blood glucose was classified as lactose digester (McMichael et al. 1965). We refer to this phenotype as lactose tolerance blood glucose test. The intensity of four gastrointestinal symptoms (flatulence, abdominal bloating, abdominal pain, and borborygmi) was assessed by the subjects themselves on a five-point scale (1 = no symptoms to 5 = severe symptoms). After a training session, subjects rated possible symptoms before lactose ingestion and a baseline score was established summing the individual scores for each of the four symptoms. After lactose ingestion, perceived intensity for each symptom was rated at 1 hr, 2 hr, 4 hr, 6 hr, 9 hr, and 12 hr. The sum of the maximal peak values of symptom ratings above baseline values during the 12-hr follow-up was calculated, in each subject, for the determination of lactose digester status. A subject with an increase value $> 5/20$ was considered to have clinically significant symptoms and was classified lactose maldigester (adapted from Peuhkuri et al. 1998).

This is the first time to our knowledge that these 3 phenotypes have been measured concomitantly in a field setting. Most of the field studies use only one criterion: Wang et al. (1984) and Ingram et al. (2007) used BH phenotype whereas Tishkoff et al. (2007) used the BG test. Following (Mulcare et al. 2004), we define a "gold standard" phenotype when blood BG and BH are both used.

We will present results for this “gold standard,” but also for BG and BH independently, and then compare both methods. Symptoms will be used in the discussion in relation with the intensity of intolerance.

Polymorphism Typing.

Detection of the -13.910C/T Polymorphism. DNA samples were extracted from blood by standard method (Maniatis et al. 1982). PCR reactions to amplify the region containing the C/T polymorphism were performed in a 10- μ L final volume composed of 1X Eppendorf buffer, 125 μ M of each dNTP, 0.25 U of Eppendorf Taq polymerase, 160 nM of each primer, and 10 ng of DNA. The primers used were Lac-C-L2 (CTGCTTTGGTTGAAGCGAAGAT) and Lac-C-M-U (GCTGGCAATACAGATAAGATAATGGA), which introduces a restriction site by mutagenesis as described earlier (Enattah et al. 2002; Mulcare et al. 2004). The PCR reactions were performed in an Eppendorf Mastercycler with an initial denaturation step at 95°C for 5 min; followed by 30 cycles at 95°C for 1 min, 59°C for 1 min, 72°C for 1 min, and 72°C for 5 min as final extension. The -13.910C>T polymorphism was then detected by RFLP, which allow to discriminate between the C or T alleles. The digestion was performed by adding 15 μ L of digestion mix ($2 \cdot 10^{-3}$ U of HinfI and $2 \times$ digestion buffer) to the 10 μ L of PCR. This reaction led to two types of DNA fragments: nondigested fragments of 201 bp (with allele C) or 177 + 24 bp digested fragments (with allele T). All genotyping were conducted with positive and negative controls and confirmed by two investigators.

Obtention of Haplotypes. Haplotypes were characterized using nine SNP covering a 463-kb fragment overlapping the LCT locus for a total of 0.17 cM. The coordinates and allelic states of the polymorphisms, taking the first ATG codon as a reference, are as follows: -221 kb (G/T rs2011946), -170 kb (C/T rs309137), -120 kb (C/T rs309142), -72 kb (C/T rs309165), -39 kb (C/T rs1057031), +96 kb (C/A rs1438307), +149 kb (A/G rs313528), +196 kb (C/T rs1264936), +240 kb (A/T rs4953953) (see Figure 1). Each SNP was revealed using the tetra-primer ARMS PCR protocol of (Ye et al. 2001).

Sequencing. A 640-bp region were amplified by PCR with ATTATACCT-CAGTCACACTG and TGATTTCGTCTCTCTGAAGCT primers. All reactions were performed under the following conditions: first step at 95°C for 5 min; then 30 cycles at 95°C for 30 sec, 54°C for 30 sec, 72°C for 30 sec; and 72°C for 5 min. Each strand of PCR products was then sequenced commercially with BigDye terminator chemistry (Applied Biosystem) with the two inner sequencing primers AAAGACGACCTTACATCAACC (forward) and TATGGCTCGGAT-GCACTGC (reverse), covering a region from -14.133 to -13.556 bp from the ATG of the LCT gene, which contains the -13.910C>T polymorphism as well as all other causal mutations identified until now (that is, -13.907*G,

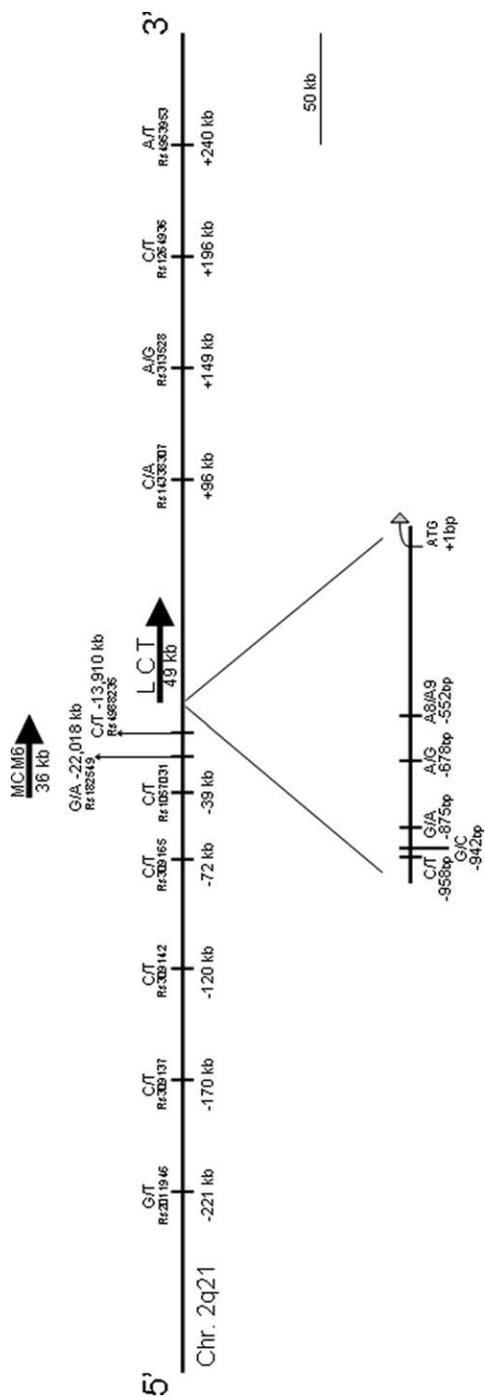


Figure 1. Location of the SNP used for the haplotype study.

Table 1. Frequency of Lactase Persistence in Tajiko-Uzbek and Kazakh Based on the Following Phenotypes: “Gold Standard” (BG + BH), BG Alone, BH Alone and Symptoms

| | <i>Lactase Persistent n (%)</i> | <i>Non-Lactase Persistent n (%)</i> | <i>Fisher Test</i> |
|---|---------------------------------|-------------------------------------|--------------------|
| Gold standard: Breath hydrogen + lactose tolerance blood glucose test | | | |
| Tajiko-Uzbek | 11 (11.0%) | 89 (89.0%) | $p = 0.018$ |
| Kazakh | 21 (25.3%) | 62 (74.7%) | |
| Blood glucose test | | | |
| Tajiko-Uzbek | 19 (19.0%) | 81 (81.0%) | $p = 0.12$ |
| Kazakh | 24 (28.9%) | 59 (71.1%) | |
| Breath hydrogen | | | |
| Tajiko-Uzbek | 30 (30.0%) | 70 (70.0%) | $p = 0.75$ |
| Kazakh | 27 (32.5%) | 56 (67.5%) | |
| Symptoms | | | |
| Tajiko-Uzbek | 40 (40.8%) | 58 (59.2%) | $p = 1.0$ |
| Kazakh | 33 (40.2%) | 49 (59.8%) | |

n: number of individuals.

–13.915*G, and –14.010*C). The sequences were aligned using CodonCode Aligner (CodonCode Corporation).

Statistical Analyses.

Haplotype Inference. Haplotypes were inferred using a Bayesian method (PHASE) (Stephens et al. 2001). We ran five times PHASE and kept the inference with the higher likelihood. However, all five runs were almost 100% congruent.

Estimation of the Age of the Allele and Its Growth Rate. We performed a joint estimation of the age of the allele (i.e., the time since the common ancestor of the carriers of the allele in the population) and the growth rate of the allele (i.e., the increase in number of copies since its appearance in the population) using the method developed by Austerlitz et al. (2003). The growth rate is therefore reflecting the intensity of the selective pressure on the allele. This method uses as inputs the current number of carriers of the allele in the population and the level of linkage disequilibrium with the surrounding haplotypes. The common ancestor can here be either a new mutant or a migrant that brought the gene in the population, or the only individual that left offspring carrying this allele in the present population because of a bottleneck. Recombination rate between SNPs (in cM) were taken from the available estimates based on HapMap linkage disequilibrium data (Frazer et al. 2007).

Results

Frequency of Lactase Persistence. We can see in Table 1 that for Tajiko-Uzbek, the frequency of lactase persistence is 11.0%, 19.0%, 30.0%, and 40.8%, for the phenotype based on BG + BH, BG, BH, and symptoms, respectively. For Kazakhs, these frequencies are 25.3%, 28.9, 32.5%, and 40.2%, respectively. No

Table 2. Genotype–Phenotype Correlation with the $-13.910\text{ C}>\text{T}$ Polymorphism

| Phenotypes/Genotypes | Kazakh | | P Value | Tajiko-Uzbek | | P Value |
|---|--------|-------|-----------------------|--------------|-------|-----------------------|
| | CC | CT/TT | | CC | CT/TT | |
| Gold standard: Breath hydrogen + blood glucose test | | | | | | |
| Non-lactase persistent | 58 | 4 | 5.1×10^{-16} | 82 | 7 | 2.2×10^{-10} |
| Lactase persistent | 0 | 21 | | 0 | 11 | |
| Hydrogen ^a | | | | | | |
| Non-lactase persistent | 54 | 2 | 2.5×10^{-14} | 68 | 2 | 1.2×10^{-8} |
| Lactase persistent | 4 | 23 | | 14 | 16 | |
| Blood glucose test ^a | | | | | | |
| Non-lactase persistent | 56 | 3 | 8.3×10^{-15} | 74 | 7 | 9.1×10^{-6} |
| Lactase persistent | 2 | 22 | | 8 | 11 | |
| Symptoms | | | | | | |
| Non-lactase persistent | 44 | 5 | 5.4×10^{-6} | 51 | 7 | 0.11 |
| Lactase persistent | 14 | 19 | | 30 | 10 | |

a. Results for BH and BG alone are given for comparison because most field studies use only one of this diagnostic.

significant differences were found between the two populations for each phenotype independently, but a significant one was observed for the combined BG + BH phenotype ($P = 0.018$). Furthermore, we observed that the frequency based on BG is lower than the one based on hydrogen, especially for Tajiko-Uzbek. The frequency based on symptoms is also clearly higher than the two others in both populations, suggesting that it does not accurately measure the lactase persistence phenotype.

Genotyping the $-13.910\text{C}>\text{T}$ Polymorphism. Among the Tajiko-Uzbek, we found two individuals with the TT genotype, 16 with the CT, and 82 with the CC. This yields a frequency of 10% of the T allele. Among the Kazakh, we found one individual with the TT genotype, 24 with the CT, and 58 with the CC. This yields a frequency of the T allele of 16.6%. Even if numerically higher in Kazakh, the frequency is not significantly different between the two populations (Fisher test: $P = 0.064$). Based on the allele frequency, the expected carrier frequency in the two populations is 19% and 30.4%, respectively, leading to a F_{st} of 0.9%.

Genotype-Phenotype Correlation with the $-13.910\text{C}>\text{T}$ Polymorphism. As presented in Table 2, when using the “gold standard” as defined by Mulcare et al. (2004), there is a highly significant correlation between phenotype and genotype: $P = 5.1 \times 10^{-16}$ and 2.2×10^{-10} for Kazakh and Tajiko-Uzbek, respectively. Indeed, individuals phenotyped as lactase persistent for both phenotypes show a perfect genotype-phenotype correlation: 11 of the 11 double diagnosed (and 21 of the 21) are carriers of the -13.910*T allele in Tajiko-Uzbek (and in Kazakh). Therefore, the lactase persistent phenotype is in this case entirely explained by this causal allele. Interestingly, we observe 11 individuals carrying a -13.910*T allele and not being lactase persistent, suggesting other factors could play a role in this phenotype.

Considering BH and BG phenotypes separately, we still observe a highly significant association between phenotype and genotype: for BH, $P = 2.5 \times 10^{-14}$ and 1.2×10^{-8} for Kazakh and Tajiko-Uzbek, respectively; and for BG, $P = 8.3 \times 10^{-15}$ and 9.1×10^{-6} , respectively. However, the concordance is here lower, especially in Tajiko-Uzbek: only 58% and 53% of the lactase persistent individuals are carriers of the -13.910^*T allele (according to BG and BH, respectively), as compared with 92% and 85% in Kazakh.

Comparison of the Two Phenotypes. As presented in Table 3, the two phenotypic methods (i.e., BG and BH) are highly correlated in both populations (P value = 5.3×10^{-3} and 1.8×10^{-11} for Tajiko-Uzbek and Kazakh, respectively). The percentage of non-concordant between the two methods is higher in Tajiko-Uzbek (28%) with respect to Kazakh (9%). The rate of false positive (i.e., when a persistent person appears to be a nondigester) is higher for BG than for BH, consistent with Mulcare et al. (2004) findings.

Another Polymorphism? As presented in Table 4, all the lactase persistent individuals for the “gold standard” phenotype carry the -13.910^*T allele. However, we still found 21 discordant individuals (lactase persistent either with BG or BH) that were not carriers of the -13.910^*T allele. In order to ensure there are no other causal polymorphisms associated with the lactase persistent trait in these individuals, we therefore sequenced a 640-bp region covering all other previously known causal mutations (for example the -13.907^*G , -13.915^*G , and -14.010^*C alleles). Among the 21 individuals, we found only one carrier of the -14.010^*C allele, in the Tajiko-Uzbek population. We did not find any new polymorphism.

Dating the Allele. When using the method developed in (Austerlitz et al. 2003), we found a growth rate of the -13.910^*T allele of around 1.02 in Tajiko-Uzbek and 1.03 in Kazakh. The age of the expansion of the allele was found to be between 7,200–12,100 yrs ago in the Tajiko-Uzbek population and between 5,800–9,900 yrs ago in the Kazakh population. Therefore, even though those dates overlap, there is a tendency for a younger date in Kazakh.

Discussion

Our study revealed that the lactase persistence trait reaches a frequency between 25% and 32% (according to the phenotype used) in the Kazakh population, a traditionally herder population of Central Asia. Therefore, the frequency of this trait is lower than in societies where fresh milk is a main source of the diet. This low frequency is consistent with an older study in an eastern area of Central Asia: (Wang et al. 1984) found 23.6% of lactase persistent individuals among Kazakh of North-West China and 12.1% among Mongol of Inner Mongolia, based on hydrogen. Furthermore, and most surprisingly, our results

Table 3. Number of Individuals Diagnosed Lactase Persistent According to the BH and the BG Methods

| | Blood Glucose Test | |
|-----------------|--------------------|-------------|
| | LP | Non-LP |
| Tajiko-Uzbek | | |
| Breath Hydrogen | | |
| LP | 11 (2;9;0) | 19 (0;5;14) |
| Non-LP | 8 (0;0;8) | 62 (0;2;60) |
| Kazakh | | |
| Breath Hydrogen | | |
| LP | 21 (1;20;0) | 6 (0;2;4) |
| Non-LP | 3 (0;1;2) | 53 (0;1;52) |

In parenthesis genotype of individuals (TT;CT;CC).

show small differences of lactase persistence between the traditionally nomadic herder and the agriculturist populations. The differences are significant only when using the “gold standard” phenotype that combines BH and BG ($P = 0.018$): 25% for Kazakh vs. 11% for Tajiko-Uzbeks, and almost significant ($P = 0.06$) for the frequency of the -13.910^*T allele (30% vs. 19% in each population, respectively). Several explanations could account for this small difference: (1) gene flow between the two populations, especially as Central Asia is considered a crossroads of cultural transmission routes; (2) even if all nomadic herders have a high input of dairy products in their diet, some might have a long tradition of not drinking fresh milk but fermented products instead, which would reduce the selective pressure on maintaining lactase into adulthood. According to the last hypothesis, the lactase persistence trait would then not perfectly be correlated with the herder/nonherder lifestyle of these populations. Indeed the F_{st} for the $13.910C>T$ polymorphism (0.9%) is within the range of neutral F_{st} based on our microsatellite study: 0.7%–1% (see Martínez-Cruz et al. 2010), confirming the absence of local differential selection between the two populations. Interestingly, the distribution of this phenotype and its link with pastoralism is indeed not worldwide established. Some populations known to have

Table 4. Age of the Expansion and Growth Rate of the Allele

| Pop | p | N | Growth | | | Age | | | Age | | |
|-----|------|-----------|---------------|--------|--------|--------------|-------|-------|---------------|---------------|----------------|
| | | | Rate | Min | Max | (g) | Min | Max | (yrs) | Min | Max |
| LUZ | 0.1 | 10.000 | 1.0234 | 1.0193 | 1.0353 | 361.1 | 307.4 | 484.4 | 9027.5 | 7685 | 12110 |
| LUZ | 0.1 | 100.000 | 1.0238 | 1.0195 | 1.0357 | 359.5 | 304.2 | 481.5 | 8987.5 | 7605 | 12037.5 |
| LUZ | 0.1 | 1.000.000 | 1.0265 | 1.0209 | 1.0385 | 348.7 | 287.5 | 461.2 | 8717.5 | 7187.5 | 11530 |
| LKZ | 0.15 | 10.000 | 1.0295 | 1.0244 | 1.0441 | 296 | 252.8 | 394.8 | 7400 | 6320 | 9870 |
| LKZ | 0.15 | 100.000 | 1.0301 | 1.0247 | 1.0448 | 294 | 249 | 391.4 | 7350 | 6225 | 9785 |
| LKZ | 0.15 | 1.000.000 | 1.0345 | 1.0274 | 1.0493 | 282.7 | 234 | 370.1 | 7067.5 | 5850 | 9252.5 |

p : frequency of the allele; N : Assumed population size; Age: the values are given in generation and translated in years assuming generation time of 25 years (for consistency with other studies). LUZ: Tajiko-Uzbek; LKZ: Kazakh.

practiced long-term pastoralism exhibit low frequency of the persistent phenotype, for example the Dinka and Nuer in Africa (Bayoumi et al. 1982). Indeed the storage of milk product has been shown recently to be old: in the Middle East, milk was processed even before 6,500 BC, making possible the storage of milk products and providing an explanation why, in spite of lactose intolerance, milk use could have been adopted quickly (Evershed et al. 2008).

Moreover, we have shown that the allele responsible for lactase persistence in Central Asia is the same as the one in Europe (-13.910^*T). Indeed, with the use of both BH and BG (“gold standard” phenotype), we found that all lactase persistent individuals carry this allele in the Kazakh as well as in the Tajiko-Uzbek population. By resequencing a region from -14.133 bp to -13.556 bp from the LCT gene in 21 individuals with a lactase persistent phenotype (either BH or BG) but noncarrier of the -13.910^*T allele, we did not detect any other polymorphism, outside of one individual carrying the 14.010^*C allele. In particular we did not find any carrier of the -13.915^*G allele, which is common in the Middle East. Our results challenged the hypothesis by (Itan et al. 2010) that additional variants could be associated with lactase persistence in Central Asia, but rather show the need for better phenotypic data.

Using a 400-kb-long surrounding haplotype, we then estimated the age of expansion of the allele to be 7,200–12,100 yrs ago in the Tajiko-Uzbek population and 5,800–9,900 yrs ago in the Kazakh population. These estimated dates are consistent with archaeological records in Central Asia. Agropastoralism developed around 6,000 BC with at least two different cultures in the area: the Jeitun in Turkmenistan, which has a Middle Eastern origin, and the Hissar in south Tajikistan, which seems to be an autochthonous development (Brunet 1999; Harris 1996). Even if the origin of pastoralism is not clear in Central Asia, it is clearly attested that mare of milk were used 3,500 BC (5,500 yrs ago) in Kazakhstan (Outram et al. 2009). The expansion time of the -13.910^*T allele in Central Asia is also included in the wide confidence interval of the date estimated in Europe: 2,000–20,000 yrs ago using surrounding haplotypes (Bersaglieri et al. 2004) or 3,400–42,000 yrs ago using microsatellites (Coelho et al. 2005). These large confidence intervals for the time of the European expansion make it difficult to choose between an endogenous expansion in Central Asia vs. recurrent gene flow from Europe. One argument in favor of a low impact of recurrent gene flow is that the higher frequency of lactase persistence is found among Kazakhs who have the lowest proportion of “western” gene pool inferred from admixture analysis from autosomal microsatellite data: the percentage of gene pool from Europe + Middle East is 25% for Kazakh and 55% for Tajiko-Uzbek (Martínez-Cruz et al. 2010). Therefore it could be that a local expansion of the allele arose in Central Asia. This could be an indirect genetic proof of the early domestication of horses for milk products as recently attested from archaeological remains (Outram et al. 2009).

According to Magalon et al. (2008), the estimated growth rate for neutral intergenic sequences is around 1.1% for Uzbeks and 1.5% for Kazaks. With an

estimated expansion coefficient estimated for the allele at 2.4% for Uzbeks and 3% for Kazaks, this led by subtraction to an estimated selection coefficient of 1.3% for Uzbeks and 1.5% for Kazaks. This is in the lower bound of the estimates for Europe: 1 to 15% for (Bersaglieri et al. 2004) and 0.5 to 15% for (Itan et al. 2009).

The intensity of expansion is higher for the nomadic herders Kazakhs than for the agriculturist Tajiko-Uzbek. Even if this difference is small, it is consistent with our phenotypic data: 37% of individuals that are homozygotes CC (genetically nonpersistent) do not describe any bowel syndrome in the Tajiko-Uzbek, while this number is only 20.6% in Kazakh ($P = 0.14$). Similarly, 17.1% of Tajiko-Uzbek do not show any increase in their hydrogen, while it is only 6.9% in Kazakh ($P = 0.12$). Even if not significant, these results suggest that Tajiko-Uzbek could be more tolerant to lactose ingestion than the Kazakhs, by having less deleterious effect of milk consumption for noncarriers of the -13.910^*T allele. Particular food habits could explain such a tendency (Martini and Savaiano 1988). In addition, possible intestinal conditions that may underlie secondary hypolactasia (Bodlaj et al. 2006) could not have been detected by medical history report of the study subjects.

Conclusion. The high prevalence of lactose malabsorption in populations from Central Asia is unexpected, particularly because one of the studied populations is traditionally using milk and has lived as nomadic herders for many generations. This result suggests that this population was not milk dependent as proposed by (Wang et al. 1984) who defined milk dependence as follows: “In contrast to milk use, milk dependence signifies an individual survival value of the ability to consume large amounts of fresh milk without ensuing losses of fluid, energy, minerals and other nutrients by diarrhoea, as is characteristic of lactose absorbers.”

Acknowledgments We thank all individuals who participated to the study, including local nurses and doctors in Uzbekistan. This study was financed by the ESF program Origin of Man, Language, and Languages and the CNRS program “Origines de l’Homme des langues et du langage”; the program ATIP from the CNRS (E. Heyer) and the ANR grant “NUTGENEVOL” (07-BLAN-0064). L.S. was granted by Ministry of Research.

Received 25 June 2010; revision accepted for publication 5 November 2010.

Literature Cited

- Aoki, K. 1986. A stochastic model of gene-culture coevolution suggested by the “culture historical hypothesis” for the evolution of adult lactose absorption in humans. *Proc. Natl. Acad. Sci. USA* 83(9):2929–2933.
- Austerlitz, F., L. Kalaydjieva, and E. Heyer. 2003. Detecting population growth, selection and inherited fertility from haplotypic data in humans. *Genetics* 165(3):1579–1586.
- Bayoumi, R. A., S. D. Flatz, W. Kuhnau et al. 1982. Beja and Nilotes: Nomadic pastoralist groups in the Sudan with opposite distributions of the adult lactase phenotypes. *Am. J. Phys. Anthropol.* 58(2):173–178.
- Bersaglieri, T., P. C. Sabeti, N. Patterson et al. 2004. Genetic signatures of strong recent positive selection at the lactase gene. *Am. J. Hum. Genet.* 74(6):1111–1120.

- Bodlaj, G., M. Stocher, P. Hufnagl et al. 2006. Genotyping of the lactase-phlorizin hydrolase -13910 polymorphism by LightCycler PCR and implications for the diagnosis of lactose intolerance. *Clin. Chem.* 52(1):148–151.
- Bolin, T. D., and A. E. Davis. 1969. Asian lactose intolerance and its relation to intake of lactose. *Nature* 222(5191):382–383.
- Brunet, F. 1999. La Néolithisation en Asie Centrale: Un état de la question. *Paléorient.* 24:27–48.
- Bryant, G. D., Y. K. Chu, and R. Lovitt. 1970. Incidence and aetiology of lactose intolerance. *Med. J. Aust.* 1(26):1285–1288.
- Caskey, D. A., D. Payne-Bose, J. D. Welsh et al. 1977. Effects of age on lactose malabsorption in Oklahoma Native Americans as determined by breath H₂ analysis. *Am. J. Dig. Dis.* 22(2): 113–116.
- Chua, K. L., and C. A. Seah. 1973. Lactose intolerance: Hereditary or acquired? Effect of prolonged milk feeding. *Singapore Med. J.* 14:29–33.
- Coelho, M., D. Luiselli, G. Bertorelle et al. 2005. Microsatellite variation and evolution of human lactase persistence. *Hum. Genet.* 117(4):329–339.
- Enattah, N. S., T. G. Jensen, M. Nielsen et al. 2008. Independent introduction of two lactase-persistence alleles into human populations reflects different history of adaptation to milk culture. *Am. J. Hum. Genet.* 82(1):57–72.
- Enattah, N. S., T. Sahi, E. Savilahti et al. 2002. Identification of a variant associated with adult-type hypolactasia. *Nat. Genet.* 30(2):233–237.
- Evershed, R. P., S. Payne, A. G. Sherratt et al. 2008. Earliest date for milk use in the Near East and southeastern Europe linked to cattle herding. *Nature* 455(7212):528–531.
- Feldman, M. W., and L. L. Cavalli-Sforza. 1984. Cultural and biological evolutionary processes: Gene-culture disequilibrium. *Proc. Natl. Acad. Sci. USA* 81(5):1604–1607.
- Ferguson, A., and J. D. Maxwell. 1967. Genetic aetiology of lactose intolerance. *Lancet* 2(7508):188–190.
- Frachetti, M. 2008. The variability and dynamic landscapes of mobile pastoralism in ethnography and prehistory. In *The Archeology of Mobility*, H. Barnard and W. Wendrich, eds. Los Angeles, CA: Costen Institute of Archeology.
- Frazer, K. A., D. G. Ballinger, D. R. Cox et al. 2007. A second generation human haplotype map of over 3.1 million SNPs. *Nature* 449(7164):851–861.
- Gerbault, P., C. Moret, M. Currat et al. 2009. Impact of selection and demography on the diffusion of lactase persistence. *PLoS ONE* 4(7):e6369.
- Harris, D. R. 1996. The origins and spread of agriculture and pastoralism in Eurasia: An overview. In: *The Origins and Spread of Agriculture and Pastoralism in Eurasia*, D. R. Harris, ed. London, U.K.: University College London Press.
- Ho, M. W., S. Povey, and D. Swallow. 1982. Lactase polymorphism in adult British natives: Estimating allele frequencies by enzyme assays in autopsy samples. *Am. J. Hum. Genet.* 34(4):650–657.
- Holden, C., and R. Mace. 1997. Phylogenetic analysis of the evolution of lactose digestion in adults. *Hum. Biol.* 69(5):605–628.
- Imtiaz, F., F. Savilahti, A. Sarnesto et al. 2007. The T/G 13915 variant upstream of the lactase gene (LCT) is the founder allele of lactase persistence in an urban Saudi population. *J. Med. Genet.* 44(10):e89.
- Ingram, C. J., M. F. Elamin, C. A. Mulcare et al. 2007. A novel polymorphism associated with lactose tolerance in Africa: Multiple causes for lactase persistence? *Hum. Genet.* 120(6):779–788.
- Itan, Y., B. L. Jones, C. J. Ingram et al. 2010. A worldwide correlation of lactase persistence phenotype and genotypes. *BMC Evol. Biol.* 10:36.
- Itan, Y., A. Powell, M. A. Beaumont et al. 2009. The origins of lactase persistence in Europe. *PLoS Comput. Biol.* 5(8):e1000491.
- Jacobson-Tepfer, E. 2008. The emergence of cultures of mobility in the Altai mountains of Mongolia. In *The Archeology of Mobility*, H. Barnard and W. Wendrich, eds. Los Angeles, CA: Costen Institute of Archeology.

- Jussila, J. 1969. Milk intolerance and lactose malabsorption in hospital patients and young servicemen in Finland. *Ann. Clin. Res.* 1(3):199–207.
- Magalon, H., E. Patin, F. Austerlitz et al. 2008. Population genetic diversity of the NAT2 gene supports a role of acetylation in human adaptation to farming in Central Asia. *Eur. J. Hum. Genet.* 16(2):243–251.
- Maniatis, T., E. F. Fritsh, and J. Sambrook. 1982. *Molecular Cloning: A Laboratory Manual*. New York, NY: Cold Spring Laboratory.
- Martínez-Cruz, B., R. Vitalis, L. Ségurel et al. 2011. In the heartland of Eurasia: The multilocus genetic landscape of Central Asian populations. *Eur. J. Hum. Genet.* 19(2):216–223.
- Martini, M. C., and D. A. Savaiano. 1988. Reduced intolerance symptoms from lactose consumed during a meal. *Am. J. Clin. Nutr.* 47(1):57–60.
- McCracken, R. D. 1971. Origins and implications of the distribution of adult lactase deficiency in human populations. *J. Trop. Pediatr. Environ. Child Health* 17(1):7–10.
- McMichael, H. B., J. Webb, and A. M. Dawson. 1965. Lactase deficiency in adults: A cause of “functional” diarrhoea. *Lancet* 1(7388):717–720.
- Mulcare, C. A., M. E. Weale, A. L. Jones et al. 2004. The T allele of a single-nucleotide polymorphism 13.9 kb upstream of the lactase gene (LCT) (C-13.9kbT) does not predict or cause the lactase-persistence phenotype in Africans. *Am. J. Hum. Genet.* 74(6):1102–1110.
- Outram, A. K., N. A. Stear, R. Bendrey et al. 2009. The earliest horse harnessing and milking. *Science* 323(5919):1332–1335.
- Peuhkuri, K., T. Poussa, and R. Korpela. 1998. Comparison of a portable breath hydrogen analyser (Micro H2) with a Quintron MicroLyzer in measuring lactose maldigestion, and the evaluation of a Micro H2 for diagnosing hypolactasia. *Scand. J. Clin. Lab. Invest.* 58(3):217–224.
- Sahi, T. 1974a. The inheritance of selective adult-type lactose malabsorption. *Scand. J. Gastroenterol. Suppl.* 30:1–73.
- Sahi, T. 1974b. Lactose malabsorption in Finnish-speaking and Swedish-speaking populations in Finland. *Scand. J. Gastroenterol.* 9(3):303–308.
- Simoons, F. 1969. Primary adult lactose intolerance and the milking habit: A problem in biologic and cultural interrelations. I. Review of the medical research. *Am. J. Dig. Dis.* 14:819–836.
- Stephens, M., N. J. Smith, and P. Donnelly. 2001. A new statistical method for haplotype reconstruction from population data. *Am. J. Hum. Genet.* 68(4):978–989.
- Sung, J. L., and P. L. Shih. 1972. The jejunal disaccharidase activity and lactose intolerance in Chinese. *Asian J. Med.* 8:149–151.
- Tishkoff, S. A., F. A. Reed, A. Ranciaro et al. 2007. Convergent adaptation of human lactase persistence in Africa and Europe. *Nat. Genet.* 39(1):31–40.
- Wang, Y., C. B. Harvey, W. S. Pratt et al. 1995. The lactase persistence/non-persistence polymorphism is controlled by a cis-acting element. *Hum. Mol. Genet.* 4(4):657–662.
- Wang, Y. G., Y. S. Yan, J. J. Xu et al. 1984. Prevalence of primary adult lactose malabsorption in three populations of northern China. *Hum. Genet.* 67(1):103–106.
- Ye, S., S. Dhillon, X. Ke et al. 2001. An efficient procedure for genotyping single nucleotide polymorphisms. *Nucleic Acids Res.* 29(17):E88–E88.