ANNUAL REVIEWS

Annual Review of Genomics and Human Genetics

On the Evolution of Lactase Persistence in Humans

Laure Ségurel and Céline Bon

Laboratoire Éco-Anthropologie et Ethnobiologie, UMR 7206 CNRS – Muséum national d'Histoire naturelle – Univ Paris Diderot, Sorbonne Paris Cité, F-75016 Paris, France; email: laure.segurel@mnhn.fr, celine.bon@mnhn.fr

Annu. Rev. Genom. Hum. Genet. 2017. 18:297–319

First published as a Review in Advance on April 19, 2017

The *Annual Review of Genomics and Human Genetics* is online at genom.annualreviews.org

https://doi.org/10.1146/annurev-genom-091416-035340

Keywords

lactase, human evolution, Neolithic, milk, natural selection, pastoralism

Abstract

Lactase persistence—the ability of adults to digest the lactose in milk—varies widely in frequency across human populations. This trait represents an adaptation to the domestication of dairying animals and the subsequent consumption of their milk. Five variants are currently known to underlie this phenotype, which is monogenic in Eurasia but mostly polygenic in Africa. Despite being a textbook example of regulatory convergent evolution and gene-culture coevolution, the story of lactase persistence is far from clear: Why are lactase persistence frequencies low in Central Asian herders but high in some African hunter-gatherers? Why was lactase persistence strongly selected for even though milk processing can reduce the amount of lactose? Are there other factors, outside of an advantage of caloric intake, that contributed to the selective pressure for lactase persistence? It is time to revisit what we know and still do not know about lactase persistence in humans.

INTRODUCTION

Mammals are defined by, among other things, the presence of mammary glands, specialized organs that allow females to produce milk to feed their young. Along with this evolutionary novelty arose the need for mammalian newborns to break down lactose, the only carbohydrate found in milk, into smaller, digestible molecules. This is the role of lactase-phlorizin hydrolase, a β-galactosidase also known as the lactase enzyme, which is present in the small intestine and hydrolyzes lactose into galactose and glucose monomers, molecules that are small enough to be absorbed by intestinal cells (113). The expression level of lactase starts declining after weaning and is then very low in all adult mammals (36)—with the notable exception of humans.

Instead, in approximately a third of humans (61), the expression of lactase persists throughout life, a phenotype known as lactase persistence (LP). The frequency of LP varies greatly among populations, ranging from 5% to almost 100%, with the highest frequencies found in people of northern European descent and some populations from West Africa, East Africa, and the Middle East (58, 61, 86, 114, 123, 126) (**Figure 1**). Why is there such a physiological difference among

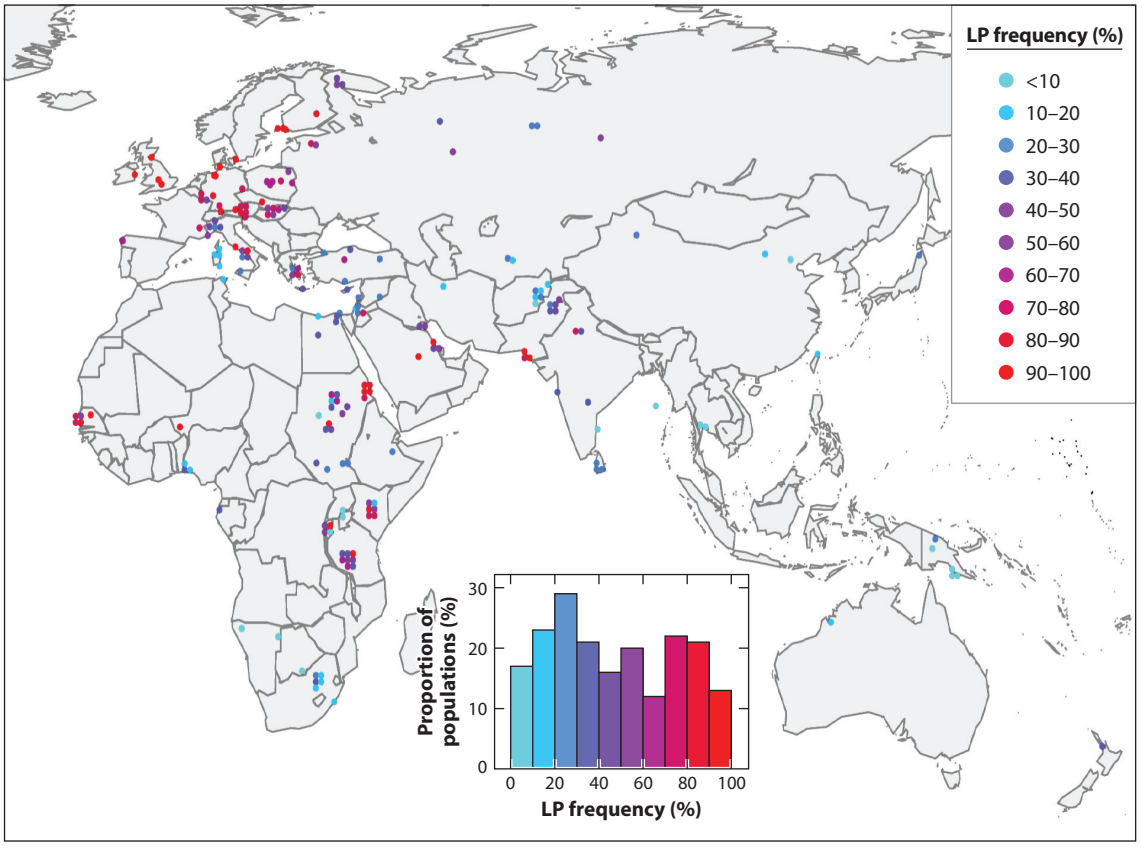

Figure 1

Lactase persistence (LP) phenotypic frequencies in the Old World. The frequencies are from the Global Lactase Persistence Association Database (GLAD; **http://www.ucl.ac.uk/mace-lab/resources/glad**), originally published by Itan et al. (63) and updated in 2013. We kept only populations that included at least ten individuals and removed one recent migrant population (North Africans living in France), resulting in a total of 14,908 individuals in 194 populations. When multiple populations were sampled at the same geographical location, we modified the coordinates so that each point is visible. The histogram represents the density of populations with different LP frequencies.

human populations? Given that the frequency of LP parallels the ancestral milk-drinking habits of populations, with high frequencies found in pastoral or agropastoral populations that traditionally incorporated large amounts of milk into their diets, the increase in LP frequency was suggested to be the result of regular milk consumption in premodern times in populations with domesticated dairying animals, in what is known as the cultural-historical hypothesis (86, 87, 123, 124). After this phenotype was shown to be genetically inherited as an autosomal dominant trait (88, 112), LP became a textbook example of gene-culture coevolution. The story turned out to be that changes in food production practices during the Neolithic revolution (approximately 10,000 years ago, when human populations began domesticating various plants and animals) led to an increase in the frequency of genetic variants that maintain the expression of lactase in adulthood. These variants allowed carrier individuals to broaden their dietary repertoire, which in turn allowed populations to incorporate more milk into their diets. LP also became a prime example of convergent evolution: Different genetic variants that cause the same phenotypic change arose in multiple populations that adopted milking practices in different geographical areas (60, 62, 134).

Moreover, the intensity of natural selection for LP has been estimated to be among the strongest in the human genome, with a selection coefficient of approximately 0.04–0.05 (10, 29, 134). This estimate raises the question of why being able to consume milk as an adult has had such a strong influence on human reproduction and/or survival: Is it only a matter of energy intake, or is there more to the story? Despite nearly half a century of debate around this question (28, 43, 50, 58, 61, 64, 126), the nature of the selective pressures responsible for the increase of LP frequency in humans remains unclear.

Indeed, a number of observations challenge our understanding of this adaptive story. For example, milk consumption does not always correlate with LP frequency: Some pastoral and agropastoral populations have a low frequency even though they consume milk [e.g., traditional herders from Mongolia and Central Asia (55, 147)], and some populations have a high frequency even though they do not consume milk of any kind [e.g., Hadza and Yaaku hunter-gatherers from East Africa (107, 134)]. Also, ancient DNA studies have recently revealed that LP was not common in Europe until the Middle Ages (2, 18, 42, 71, 74, 75, 85, 91, 102), adding some confusion about when LP was first selected for. Given all of these puzzles, it might be time to revisit our understanding of the evolution of LP.

In this review, we discuss only the LP phenotype (also called primary acquired lactase deficiency, adult hypolactasia, or lactose malabsorption), defined as the presence or absence of the lactase enzyme in adults; lactose intolerance, by contrast, refers to the symptoms that can occur when consuming lactose, regardless of whether the cause is lactase nonpersistence (LNP). We do not cover other phenotypes, such as congenital and infant hypolactasia (the underexpression of lactase in newborns or children), secondary hypolactasia (the loss of lactase expression following certain intestinal diseases, surgeries, or drug use), and milk allergies (immune reactions of the body to various components of human or cow milk).

LACTASE PERSISTENCE: A TEXTBOOK EXAMPLE OF ADAPTATION

The Lactase Persistence Phenotype

The most straightforward way to test whether an individual is lactase persistent is to test for the activity of lactase in small-intestinal biopsies (137). However, this approach is highly invasive, so more convenient ways of assessing this phenotype have been developed. First, glucose can be measured in the blood before and after the uptake of a lactose load (typically 50 g of lactose when fasting, which is equivalent to approximately 1 L of cow milk) (93). A significant increase in

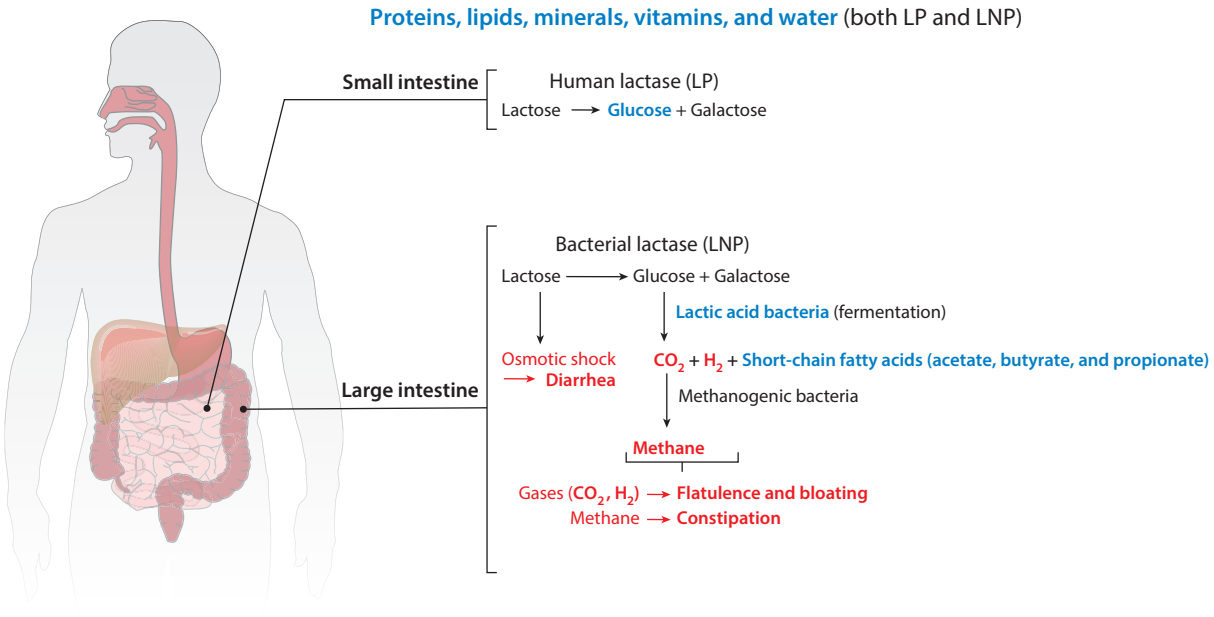

Figure 2

The fate of milk and lactose in the human body. Lactose is first hydrolyzed by human or bacterial lactase enzymes and then fermented in the large intestine (colon) by lactic acid bacteria. Likely benefits of milk are shown in blue; potential harmful effects are shown in red. Abbreviations: LP, lactase persistence; LNP, lactase nonpersistence.

glucose level means that the body is able to break down lactose into glucose in the small intestine. Alternatively, the hydrolysis of lactose can be measured based on the increase of galactose in urine excretion (48).

Lactase activity can also be evaluated by more indirect approaches, notably by looking at the amount of hydrogen in the breath (76) or by investigating the severity of intestinal symptoms (142) after a similar lactose load. Indeed, only LNP individuals will have large amounts of undigested lactose in their colon, which (*a*) will disrupt the osmotic balance and lead to an increase of water in the bowels, resulting in diarrhea (4), and (*b*) will be available to their colonic bacteria for fermentation (**Figure 2**). During this transformation, hydrogen is produced, released into the bloodstream, and then excreted by the lungs. This production of gases during lactose fermentation (not only of hydrogen but also of carbon dioxide and methane) and the acidification of the milieu lead to several symptoms—namely flatulence, bloating, and intestinal cramps (4), although some variability among LNP individuals has been observed (20, 117, 120, 142). Furthermore, up to 8% of lactose can reach the colon in LP individuals, resulting in some additional difficulty in using these indirect tests to distinguish LP from LNP individuals (14).

Despite their limitations, these phenotypic assays, performed individually or in combination, have been used since the 1960s to describe the prevalence of LP across the globe. Notably, worldwide frequencies have been compiled in several reviews (50, 58, 61, 63, 114, 126, 132) and suggest that approximately a third of humans are lactase persistent. The global picture is one of a large area of high frequencies in Europe and patchier distributions of high frequencies in West Africa, East Africa, the Middle East, and South Asia (**Figure 1**).

More specifically, populations (from which at least ten individuals have been sampled) that have an LP frequency higher than 75% are found as follows:

- Uniformly in northern Europe (Ireland, the United Kingdom, Denmark, Estonia, Sweden, and Finland) and central Europe (Austria, the Czech Republic, Germany, Switzerland, the Netherlands, and northern France), and in populations with similar European ancestry (the United States, Canada, and Australia)
- Sparsely in West Africa [Niger (among the Tuareg, camel and goat herders) and Senegal (among the Fula, cattle herders)]
- Sparsely in East Africa [Tanzania (among the Iraqw, agropastoral populations that raise cattle), Kenya (among the Maasai, cattle herders), Rwanda (among the Tutsi, cattle herders), and Sudan (among the Beja, camel and sheep herders)]
- Sparsely in the Middle East [Saudi Arabia and Jordan (among the Bedouin, camel and goat herders)]

In addition, in Pakistan, the results of the only two phenotypic studies are inconsistent with each other: Rab & Baseer (106) found high LP frequencies (80–100%) in five ethnic groups from the south (4–15 individuals per population), whereas Ahmad & Flatz (1) found lower frequencies (30–42%) in four similar ethic groups from the north (8–132 individuals per population). Based on genetic data, Gallego Romero et al. (41) inferred that the LP frequency in India was at most 74% among water buffalo herders from the north of the country.

The distribution of the LP frequency across populations (**Figure 1**) shows that some populations have a frequency of approximately 80%, implying that LP was strongly beneficial in these populations, whereas many others have a frequency of approximately 20%, more likely reflective of admixture with neighboring high-LP populations.

The Genetic Basis of Lactase Persistence

In parallel with the characterization of the LP phenotype, the question arose of whether this trait is genetically encoded or adaptively induced by the environment (e.g., by prolonged milk consumption). After a decade-long controversy, it was finally demonstrated, based on family studies, twin studies, studies of individuals who had divergent genetic ancestries but lived in the same environment, and studies of individuals with admixed genetic ancestry, that LP is an autosomal dominant trait (37, 46, 65, 79, 88, 112, 115). But not until two decades later was the first causal molecular change associated with lactase expression in adulthood finally discovered.

In 2002, a study of Finnish families by Enattah et al. (30) identified the first mutation associated with the LP phenotype: −13.910:C>T (rs4988235). It is surely because this regulatory mutation is located 14 kb upstream of the *lactase* (*LCT*) gene, in intron 13 of the *minichromosome maintenance complex component 6* (*MCM6*) gene, and not within *LCT* or immediately upstream of it, that its discovery took so long. In vitro studies (using the human intestinal cell line Caco-2 with luciferase expression vectors) quickly confirmed that this variant is a *cis*-acting enhancer of the *LCT* promoter (97, 139), and this property was demonstrated in vivo more recently (33). A further study showed that the −13.910:T variant creates a new binding site for octamer-binding protein 1 (Oct-1), a transcription factor that interacts with human hepatocyte nuclear factor 1α (HNF1α) to bind to the *LCT* promoter (77). This allele therefore leads to an alternative path for *LCT* expression that is not downregulated, as the original path is. Interestingly, in vivo studies further revealed that, even though the LP phenotype is considered a binary trait encoded in a dominant manner, lactase activity is instead a codominant quantitative trait, with a clear trimodal distribution (56, 143). This difference in expression between homozygous and heterozygous likely has a minor effect on

the ability to efficiently break down lactose but, as suggested by Swallow (132), it could become important under certain conditions, such as stress or disease.

Enattah et al. (30) also reported another mutation associated with the LP phenotype: −22.018:G>A (rs182549), located in intron 9 of *MCM6*. In Finnish families, −22.018 is in complete linkage disequilibrium with −13.910, and in a sample set of 236 individuals from four populations (Finnish, Italian, German, and Korean), −13.910:T is perfectly associated with LP, whereas all 7 recombinant individuals with genotypes −13.910:C/C and −22.018:G/A are lactase nonpersistent (30), suggesting that −22.018:A is not able to drive LP by itself. Actually, −13.910 and −22.018 might interact epistatically; in vitro studies have shown that the −22.018 region is a weak silencer of the enhancer activity driven by −13.910 (97, 138, 139), and −13.910:T is in very strong linkage disequilibrium with −22.018:A worldwide (31).

As it turns out, −13.910:T is not only a European mutation; it also underlies the LP phenotype all over Asia, including in various populations from Russia, Pakistan, and Iran (31); Central Asia (prevalence of 30% in herders) (55); and India (prevalence of up to 45% in herders from the north) (41). It is also the main LP-associated mutation in Mozabites (27%) (107) and other Berber populations (15–22%) (10, 90) from North Africa, as well as in Fula (37–48%) from both Central Africa (Cameroon and Mali) (60, 80, 107) and East Africa (Sudan) (31). One major haplotype has been found to carry the −13.910:T mutation from Europe to Asia to North Africa, showing that it rapidly spread through gene flow. However, other divergent haplotypes have also been found in a restricted geographical area around the Volga, which could be explained by an independent (convergent) appearance of the −13.910:T allele in this region (31) or by multiple recombination or gene conversion events between haplotypes.

In all of Eurasia, therefore, LP appears to be monogenic. The situation is different in East Africa, where four different mutations have been found to be associated with LP: −13.907:C>G (rs41525747), −13.915:T>G (rs41380347), −14.009:T>G (ss820486563), and −14.010:G>C (rs145946881), all of which cluster in intron 13 of *MCM6*, within a 100-base-pair interval of each other that includes −13.910 (60, 62, 134). All of these mutations result in an increase of lactase activity in vitro (62, 66, 134), and two of them, −13.907:G and −13.915:G, seem to affect binding of Oct-1 (29, 96). While −14.010:G is most prevalent in various Afro-Asian and Nilo-Saharan pastoralists (or agropastoralists) from East Africa (32–46%) (134) and in pastoralists from South Africa (13–20%) (16, 136), and both −13.907:G and −14.009:G are most prevalent among the Beja people of Sudan (21% and 24%, respectively) (62, 107), −13.915:G is the most common variant in camel herders from the Middle East (72–88%) (29, 59, 104).

Other, less common mutations have also been proposed to be associated with LP. Notably, −13.913:T>C was found in one Ethiopian (60), one Jordanian (29), and up to 7.5% of some South African populations (16) but was not confirmed as significantly associated with LP in a larger study of Ethiopian individuals (66). Some very low-frequency variants, such as −14.009:C>G and −14.011:C>T, were further shown to influence lactase activity in vitro (78).

Thus, at least five variants clearly underlie the LP phenotype: −13.910:T (combined with −22.018:A) in most of Eurasia, North Africa, and Central Africa; −13.915:G, mostly in the Middle East; and −13.907:G, −14.009:G, and −14.010:C, mostly in East Africa (see **Table 1**). Interestingly, in Ethiopian populations, all of these mutations coexist, resulting in a higher diversity at this locus in LP as compared with LNP individuals (62, 66). The LP phenotype is therefore a beautiful illustration of regulatory convergent evolution, in which nearby variants underlying the same phenotypic change arose in different populations. Such a parallel increase in the frequency of multiple alleles in different geographical areas clearly suggests the action of natural selection.

Table 1 Known lactase persistence (LP)–associated mutations and their main geographical areas of repartition

LP-associated mutation	Main geographical area of repartition
−13.910:T (combined with −22.018:A)	Eurasia, North Africa, and Central Africa
−13.915:G	Middle East
−13.907:G	East Africa (Ethiopia and Sudan)
−14.009:G	East Africa (Ethiopia and Sudan)
−14.010:C	East Africa (Kenya and Tanzania) and South Africa

Population Genetic Evidence for Natural Selection

Soon after the discovery of the first LP-associated mutation, Bersaglieri et al. (10) showed that haplotypes carrying the −13.910:T variant present typical characteristics of recent and local positive selection. Indeed, in Europe, −13.910:T is located on an unusually long stretch of homozygous markers given its frequency (i.e., a haplotype block of >1 Mb). This observation is unexpected under neutrality, because high-frequency alleles are usually old, and therefore, owing to recombination, they are typically surrounded by short haplotype blocks (94). In addition, the difference in frequency among populations, as measured by the fixation index, is significantly larger than expected under neutrality: The authors calculated it to be 0.53 in 53 worldwide populations, exceeding 99.9% of the values for genome-wide single-nucleotide polymorphisms (10). The authors also estimated the strength and timing of selection from a sample of European Americans and found that the 13.910:T allele arose 2,188–20,650 years before present (BP) and was favored with a selection coefficient of 0.014–0.15. Although these confidence intervals are large, they are consistent with a recent (Neolithic) selection of strong intensity, as confirmed by additional estimates (**Table 2**).

Similar signatures of positive selection have been found in East Africans, with average homozygous tracts of approximately 1.8 Mb, 1.4 Mb, and 1.1 Mb for −14.014:C, −13.907:G, and −13.915:G, respectively (134). The iHS scores of each variant, a statistic that reflects the

Table 2 Estimated selection coefficients of various lactase persistence (LP)–associated mutations in different studies

Study (reference)	Population	Mutation	Selection coefficient (95% CI)	Timing of selection (95% CI) (BP)
Bersaglieri et al. 2004 (10)	European American	−13.910:T	0.014–0.15	2,188–20,650
	Finnish/Swedish	−13.910:T	0.09–0.19	1,625–3,188
Tishkoff et al. 2007 (134)	Kenya-Nilo-Saharan (lowest)	−14.010:C	0.035 (0.008–0.080)	6,925 (2,232–18,496)
	Tanzania-Niger (highest)	−14.010:C	0.077 (0.026–0.142)	2,778 (1,219–6,049)
	European American	−13.910:T	0.039 (0.012–0.107)	9,323 (2,231–19,228)
Enattah et al. 2008 (29)	Saudi Arabian	−13.915:G	0.051 (0.034–0.101)	4,075 (2,050–6,100)
	European American	−13.910:T	0.048 (0.044–0.055)	5,575 (4,950–6,200)
	Western Finnish	−13.910:T	0.043 (0.039–0.049)	5,200 (4,625–5,775)
Itan et al. 2009 (64)	European	−13.910:T	0.095 (0.052–0.159)	7,441 (6,256–8,683)
Gerbault et al. 2009 (44)	European	−13.910:T	0.012 (0.008–0.018)	Not estimated
Peter et al. 2012 (101)	Finnish	−13.910:T	0.025 (0.004–0.20)	11,200 (1,500–64,900)

All coefficients assume a dominant model for genotype-phenotype association. Gerbault et al. (44) did not estimate the timing of selection, because the dates were taken from archeological data as a parameter of the model, and those vary between 7,000 and 8,000 BP. Abbreviations: BP, years before present; CI, confidence interval.

DAIRYING ANIMAL DOMESTICATION IN THE OLD WORLD

The first evidence of dairying animal domestication comes from Anatolia, where goats, sheep, and cattle were domesticated around 10,500 years before present (BP) (82, 140). These species then spread to Europe around 9,000 BP (140) and to Africa around 7,000 BP (8), following the migration of human farmers. In the Indus Valley, the zebu was domesticated around 8,000 BP, followed by the dairy buffalo around 4,500 BP (82, 140). The yak was probably domesticated in Tibet around 4,500 BP (105). Ungulates have also been domesticated for dairying, including the donkey (111) in Arabia or East Africa around 6,000 BP, the camel in Central Asia around 5,000 BP, and the dromedary in Arabia around 3,000 BP (98). Evidence for horse domestication has been more difficult to obtain owing to the high morphological similarity between wild and domesticated horses (148), but remains in Kazakhstan show that horses were harnessed around 5,500 BP (99). The domestication of the reindeer seems to date only from 2,500 BP at the earliest, and its domestication is ongoing (109).

haplotype structure (141), were shown to be highly unusual when compared with an empirical distribution of the rest of the genome. The authors further estimated that −14.010:C has been under selection since between 2,778 and 6,925 BP (95% confidence intervals: 1,219–6,049 and 2,232–18,496 BP, respectively), depending on the population (**Table 2**). They in turn calculated the selection coefficients to be 0.077 and 0.035 (95% confidence intervals: 0.026–0.142 and 0.008–0.080, respectively). A comparison with the numbers obtained in Europeans with the same method suggests that the selective advantage of LP is similar in Europe and Africa (but see 119) and that the timing of selection might be a bit older in Europe, which is consistent with archeological records of pastoralism on both continents (see sidebar titled Dairying Animal Domestication in the Old World). Finally, Enattah et al. (29) also found that the haplotype structure in the Middle East deviates significantly from neutrality, with an estimated date of selection of 4,075 BP (95% confidence interval: 2,050–6,100 BP) and a selection coefficient of 0.051 (95% confidence interval: 0.034–0.101), similar to estimates obtained with the same method in Europeans (**Table 2**). The genetic signatures of selection discussed above are expected under a hard-sweep model, in which only one allele underlies the selected trait and the allele was favored after its introduction into the population (130). However, it might be harder to infer the strength of selection in populations such as Ethiopians or Sudanese, where multiple alleles coexist, potentially resulting in a radically different signature of a soft sweep (62, 66, 107). In any case, it seems that the selection of LP occurred recently and concomitantly in different continents soon after the beginning of cattle and camel domestication (see sidebar titled Dairying Animal Domestication in the Old World).

How strong are these selection coefficients? For comparison, other studies of strongly favored loci in humans have estimated selection coefficients to be (*a*) between 0.04 and 0.09 for genes associated with resistance to malaria (53, 135), (*b*) approximately 0.03 for genes associated with skin pigmentation in Europeans (145), (*c*) between 0.002 and 0.029 for genes associated with hypoxia response to high altitude in Tibetan populations (6), (*d*) 0.036 for the *alcohol dehydrogenase 1B* (*ADH1B*) gene associated with alcohol metabolism in East Asians (101), and (*e*) 0.14 for a signal on the *ectodysplasin A receptor* (*EDAR*) gene, which is involved in the development of hair follicles and associated with an increase in eccrine sweat glands in East Asians (101). Therefore, the signal around *LCT* represents one of the strongest examples of positive selection on the human genome.

THE NATURE OF THE SELECTIVE PRESSURES

Given the evidence of positive selection on the LP phenotype, a natural question is, What drove selection to favor similar regulatory shifts in multiple populations worldwide? As was noticed early on, there is a clear correlation between the milk-drinking habits of populations and their LP frequency, leading to the cultural-historical hypothesis (86, 87, 123, 124). However, the directionality is unclear: Did human populations first start drinking milk in the absence of LP-associated mutations, or were milk-drinking practices favored in populations that already had, in low frequency, LP-associated mutations? If we consider the mutational target of LP-associated mutations to be the Oct-1 binding site, which is 13 base pairs long (60), and use a theta of 0.1% and a generation time of 30 years, the waiting time for a new mutation to arise (let alone reach a substantial frequency) is approximately 7,000 years. This seems too long given that the domestication of dairying animals is no more than 10,000 years old and that two mutations are observed at high frequencies in the Oct-1 binding site: −13.910:T and −13.915:G. One possibility is that the effective population size of humans was much larger in the recent past, decreasing the waiting time for a new mutation. Alternatively, the mutational target might be larger than 13 base pairs; although LP-associated mutations are highly clustered (in a 100-base-pair region), −13.907:G, which is outside the predicted binding site, affects Oct-1 binding (60). Finally, some LP-associated mutations might have already been present at low frequency before animal domestication.

More generally, the validity of the cultural-historical hypothesis has often been questioned (28, 37, 43, 50), and in particular it is unclear whether and (if so) why the ability to consume milk as an adult created such a differential fitness between LP and LNP individuals. The first possibility is that all individuals had an advantage of consuming milk because it is a rich source of proteins, fat, minerals, and vitamins, but only LP individuals could benefit from these without symptoms. There is indeed a cost of drinking milk for LNP individuals, with diarrhea being the most likely cause of a selective disadvantage. However, lactose intolerance symptoms have been investigated by giving 50 g of lactose (equivalent to 1 L of milk) to fasting individuals, and most LNP individuals seem to tolerate moderate amounts of lactose (such as one glass of milk, or up to 15 g of lactose) without any symptoms, especially if they slow down the transit time by consuming other foods (120, 122, 131). Furthermore, the occurrence of diarrhea and the severity of intestinal symptoms depend on the fermentation ability of each individual, which varies widely (51, 52, 149) and depends mostly on the composition of their colonic microbiota (the sum of all microorganisms present in the colon) (4, 81). For example, if colonic bacteria are efficient at fermenting lactose, then the osmotic shock (which leads to diarrhea) will be reduced, but the quantity of gases might increase. In parallel, if methanogenic archaebacteria are prevalent, then carbon dioxide will be transformed mostly into methane, leading to constipation rather than diarrhea, as observed in 30% of individuals (19, 81). In addition, consumption of dairy products by LNP individuals can influence the colonic microbiota composition and lead to a reduction of intestinal symptoms (54, 133). Therefore, it seems that LNP individuals could be able to consume milk in small amounts spread throughout the day, especially if it is taken together with other foods.

LNP individuals can also benefit from proteins and fat if they eat derived dairy products that are low in lactose (35, 81). Notably, milk can be fermented to produce yogurt or various fermented beverages, in which lactose is partially transformed by bacteria and/or yeasts, or it can be processed to obtain cheese, cream, and butter, in which lactose is almost entirely physically eliminated after protein coagulation. Finally, LNP individuals seem to be able to benefit from some amount of glucose when ingesting fermented products such as yogurts, as the lactic acid bacteria present in these products release bacterial lactase, and therefore a certain amount of lactose

can be hydrolyzed during intestinal digestion (70), though it is not clear exactly how much this represents.

What is clear, however, is that LNP individuals cannot derive large amounts of glucose from any dairy products, as lactose, representing about 30% of the calories in human milk (17), is the sole sugar in milk. Flatz & Rotthauwe (37), however, argued that carbohydrates were not scarce in human societies, and they therefore believed that the selective pressure could not be caused by an "unspecific" caloric intake. Although some have proposed that this energetic advantage of milk was most strongly selected for during famines and drought (28, 43), this is not trivially the case, as milk production itself is affected by a scarcity of food for livestock and is therefore minimal in times of paucity (9, 13, 34). Fresh milk therefore might not constitute a realistic alternative in times of food shortages.

Because of these arguments, and because of the imperfect correlation between the degree of pastoralism and LP frequency (with LP being disproportionally present in Europeans), some researchers have proposed alternative explanations for the selective advantage of LP. Notably, Flatz & Rotthauwe (38) proposed the calcium assimilation hypothesis, in which milk would be an important source of calcium, vitamin D, and lactose, the latter two of which facilitate calcium adsorption in humans. This property would be especially advantageous in farmers from high latitudes, such as Europeans, who have a low dietary supply of vitamin D (owing to their shift to cereals) and experience low levels of UV irradiation (which stimulates the production of vitamin D). However, recent studies have challenged the view that lactose facilitates calcium adsorption in humans, proposing instead that LNP individuals could benefit from more calcium when eating derived dairy products than LP individuals do from consuming fresh milk (73).

Other authors proposed that milk might have represented a precious source of uncontaminated water and electrolytes, especially in populations inhabiting dry arid environments and facing, for example, cholera, such as some African and Middle Eastern pastoralists (21, 22). This hypothesis and the calcium assimilation one explain LP only in specific parts of the world, so a combination of them is required to explain the overall pattern. To disentangle these hypotheses and gain some understanding of the selective pressures responsible for the physiological differences among human populations, Holden & Mace (58) performed a joint analysis of LP phenotypic data in 62 populations together with data on dependence on pastoralism, levels of UV irradiation, and intensity of drought, controlling for the phylogenetic relationship between populations. Their conclusion was that the data are concordant only with the cultural-historical hypothesis, because LP frequencies significantly correlate with pastoralism but not with the two other factors.

A more recent study by Itan et al. (64) simulated selection in combination with underlying demographic processes in Europe and reached the same conclusion, finding no evidence that selection is stronger in high latitudes than in lower ones. The authors further estimated the origin of selection to be between central Europe and the northern Balkans 7,500 years ago. By contrast, Gerbault et al. (44) concluded that the most likely scenario in Europe is one of demic diffusion of farming together with an advantage of calcium assimilation in high latitudes, while favoring the cultural-historical hypothesis in Africa. In summary, it appears that LP frequencies in Africa are generally consistent with the cultural-historical hypothesis, whereas in Europe, results on the respective influences of pastoralism, latitude, and demography remain conflicting. Beyond studies of human genetic data, the geographical concordance between human LP frequencies and the genetic diversity of six milk genes in European cattle breeds (7) strongly supports the idea that pastoralism played a primary role in the increase in LP frequency in Europe.

And what about Asia, which is home to the largest populations of herders, the steppe populations? To address this question, we reanalyzed the data from Holden & Mace (58), analyzing

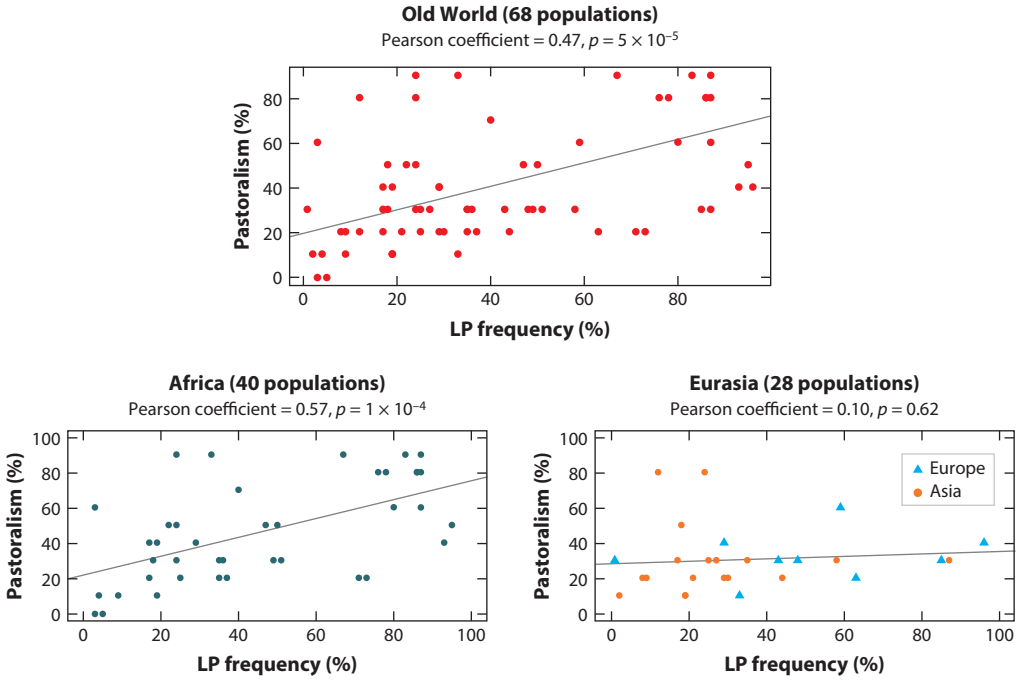

Figure 3

Correlation between lactase persistence (LP) phenotypic frequency and the proportion of pastoralism in populations from the Old World (*top*), Africa (*bottom left*), and Eurasia (*bottom right*). Most of the data are from Holden & Mace (58); to their data, we added 14 populations for which LP frequencies were available in the Global Lactase Persistence Association Database (GLAD; **http://www.ucl.ac.uk/mace-lab/resources/glad**), originally published by Itan et al. (63) and updated in 2013, and for which we could find information on the proportions of pastoralism in Murdock's "Ethnographic Atlas: A Summary" (89) (Somali, !Kung, Wolof, Diola, Serere, Bisharin, Sandawe, Iraqw, Maasai, Kazakh, Burmese, Kashmiri, and Finnish populations as well as Arabs from Israel). We removed Holden & Mace's data from America and Oceania (8 populations) to focus on Africa, Europe, and Asia.

separately populations from Africa, Europe, and Asia (a subset of 54 out of 62 total populations) and adding 14 populations for which LP frequencies were available in the Global Lactase Persistence Association Database (GLAD; **http://www.ucl.ac.uk/mace-lab/resources/glad**) (63) and for which we were able to find information on the proportions of pastoralism in Murdock's "Ethnographic Atlas: A Summary" (89). As expected, we found a highly significant correlation between LP and pastoralism in the worldwide data set (68 populations, Pearson coefficient $=$ $0.47, p = 5 \times 10^{-5}$) and within Africa (40 populations, Pearson coefficient $= 0.57, p = 1 \times 10^{-4}$). However, the correlation was not significant within Europe, Asia, or Eurasia as a whole (10, 18, and 28 populations, respectively; Pearson coefficient $=$ 0.18, 0.01, and 0.10, respectively; $p >$ 0.61) (see **Figure 3**). Even though we did not take into account the phylogenetic relationship between populations, as was done in the original paper (58), our findings suggest that the observed worldwide correlation could be driven mostly by African populations. Furthermore, we can see that Europeans have a high LP frequency despite moderate ancestral levels of pastoralism (because they are traditionally agropastoral populations that derive a considerable amount of energy from domesticated plants), whereas Asians all have low LP frequencies even though some populations have relied heavily on pastoralism for millennia. Why, then, are these Asian herders lactase nonpersistent?

FIRST EVIDENCE OF MILK PROCESSING

The slaughtering-age profiles of early Neolithic cattle (140) and the presence of residual milk proteins and lipids on ceramics as old as 9,000 years before present (BP) (32) revealed that dairying was practiced early on by the first farmers in Anatolia. This suggests that milk was used for human benefit soon after domestication. In Europe, a sieve from around 8,000 BP carries evidence of cheese processing (116), and β-lactoglobulin has been found in calculus, showing that milk or lactose-rich whey was drunk around 5,000 BP (144). In Asia, use of mare milk has been shown in Kazakhstan around 5,500 BP (99). In Africa, analyses of fatty acids in potsherds in Libya revealed the use of milk as early as around 7,000 BP (27). Therefore, populations knew how to make derived dairy products from milk soon after animal domestication.

PUZZLING OBSERVATIONS

Non-Lactase-Persistent Herders

Under the cultural-historical hypothesis, populations with higher levels of milk dependence (notably nomadic herders) should have higher LP frequencies, provided that they did not shift to pastoralism recently and did not substantially admix with nonpastoralist populations. Yet in long-term herders from Asia, such as Mongols and Kazakhs, the observed frequency is quite low. Indeed, LP frequencies are estimated to be 12% in Mongols (147) and 24–30% in Kazakhs (55, 147). In other herders from Central Asia, where the LP phenotype is well correlated with the presence of −13.910:T (55), LP frequency can be estimated from genetic data. We collected such data for ten additional Central Asian herder populations (two Karakalpak, five Kyrgyz, one Kazakh, one Turkmen, and one Uzbek, totaling 301 individuals) and found LP frequencies to lie between 3% and 26%, with an average of 14% (L. Ségurel & E. Heyer, unpublished data). Similarly, in Tibetan populations, all five previously LP-associated mutations are absent, suggesting either a very low frequency of LP or an independent genetic basis for it (100). Therefore, in Asia (outside of northern India and Pakistan), the emerging picture is one of very low LP frequencies, whether populations engaged in a pastoralist lifestyle or not.

Central Asian herders are not the only exception to the expected pattern. Indeed, the Sami reindeer herders in Scandinavia have a lower LP frequency than the rest of the Swedish population (40–75% versus 91%) (110) despite a higher dependence on pastoralism (60% versus 30%). Similarly, some African pastoral ethnic groups who consume milk (50–90% pastoralism) have low LP frequencies, as in the case of the Dinka (LP frequency of 22%) and Nuer (25%) in Sudan (5), the Somali in Ethiopia (24%) (62), and the Herero in South Africa (3%) (58). Finally, in the area where animals were first domesticated (notably in Turkey) and more generally around the Mediterranean, populations that have used milk for millennia (see sidebar titled First Evidence of Milk Processing) have moderate LP frequencies.

How can we explain these discrepancies? First, there could be important differences between pastoral groups in the time since domestication; the first dairying domesticated animals appeared 10,500 BP in the Middle East, whereas the latest ones seem to be the reindeer, 2,500 BP (see sidebar titled Dairying Animal Domestication in the Old World). Milk from different species also contains different amounts of lactose: 100 g of milk from mares, donkeys, and humans contains 6.4–6.9 g of lactose (the highest concentration); from cows, buffalo, yaks, goats, sheep, and camels contains 4.2–5.1 g of lactose; and from reindeer and moose contains 2.6–2.9 g of lactose (the lowest concentration) (35). The low lactose content in reindeer milk, combined with their more

recent domestication, could explain the moderate LP frequency in the Sami people. However, this explanation is not likely for Central Asian and Mongolian herders, who have been consuming high-lactose mare milk for at least 5,500 years (99).

Ingram et al. (62) observed that part of a Somali population from East Africa did not present a baseline release of hydrogen and proposed that these individuals might have a colonic adaptation that reduces the symptoms associated with milk consumption. However, it is unclear whether these individuals are lactase persistent or nonpersistent, as neither the increase of glucose nor the presence of symptoms was investigated. In any case, although these individuals might have reduced symptoms, they are still not able to derive glucose from milk.

Importantly, there could have been major differences in dietary availability among populations during the Neolithic, with European agropastoral populations facing harsh times of food shortage during their transition to agriculture, whereas steppe populations would have undergone a smoother transition from hunter-gathering to herding. It has indeed been proposed that early farmers had much less balanced diets than hunter-gatherers, with many food deficiencies and a shortage of proteins (103, 108). Alternatively, it could be that some pastoral populations are admixing too often or too strongly with neighboring nonpastoral populations, and therefore are unable to maintain a strong signal of local adaptation. This could be the case for Central Asia, which is a migratory crossroads and lies in the middle of nonpastoral groups.

Another important factor is the transformation of milk into derived dairy products. Durham (28) proposed that, although domestication of dairying animals is a prerequisite for LP to evolve, it is expected to be high not in all pastoral populations, but rather only in those where milk is not entirely processed into low-lactose products and/or those under high dietary stress. He further showed that LP indeed correlates better with the amount of fresh milk consumed than with the proportion of pastoralism (28), even though such data are not reliably available for all populations. Notably, this explanation might account for the intermediate LP frequency in populations around the Mediterranean and north of the Middle East (39%), where dairying animals were first domesticated, given that these populations consume moderate amounts of fresh milk (102 L per person per year) and transform a large proportion of milk into cheese (38% on average) (28). Northern and central Europeans, by contrast, consume much more fresh milk (489 L per person per year), transform a lower proportion into cheese (18% on average), and have a very high LP frequency (91%). If so, given that all populations had access early on to milk-processing techniques (see sidebar titled First Evidence of Milk Processing), the question becomes, Why have they not avoided the selective pressure for LP by a more rapid cultural adaptation, i.e., milk transformation into secondary products? Indeed, why would a population continue to consume fresh milk despite having intestinal symptoms for thousands of years (the time it would take for the LP-associated mutations to increase substantially enough that an appreciable fraction of the population could enjoy the nutritional benefit of lactose) when the cultural adaptation of milk processing would allow them to benefit from proteins, lipids, vitamins, and minerals at a lower cost?

Cultural differences in terms of preferences or beliefs that influence the practices around milk could have been important, as was nicely described for South Asian populations by Simoons (125). Climatic differences could also influence the need or ability to preserve dairy products outside of fresh milk. (The primary function of fermentation, for example, is to increase the conservation time of dairy products.) A strong seasonality of milk availability would also increase the need to process milk in order to store dairy products for the winter. Furthermore, differences in mobility could influence consumption practices, given that a highly nomadic lifestyle likely favors the transportation of processed dairy products over a large quantity of fresh milk. Finally, differences in carbohydrate availability in the rest of the diet might matter, with some populations experiencing a stronger pressure to increase their protein and fat intake, whereas others place

more importance on carbohydrates. Therefore, differences in cultural, ecological, nutritional, and environmental factors likely contribute to the quantities and relative amounts of fresh and processed milk consumed in different human populations (12).

Ancient DNA

Another puzzling observation that added some confusion about when and where LP started to increase in frequency is the late appearance of the −13.910:T LP-associated allele in ancient DNA from Europe (2, 18, 42, 71, 74, 75, 85, 91, 102). Ancient DNA studies are especially helpful because they can provide a direct temporal snapshot of the evolution of the LP phenotype. Furthermore, they provide information about where the selective constraints started, before further migration and admixture events occurred. Because of degradation, access to ancient DNA from before 10,000 BP and DNA conserved in warm areas is challenging, but fortunately, data from Europe from the Neolithic and more recent times are expected to yield substantial information, and indeed such data have started to accumulate about the −13.910:T mutation.

As expected, analysis of hunter-gatherers from the Paleolithic and Mesolithic revealed that no individual carried the LP allele before 5,000 BP (39, 40, 42, 67, 121) (**Figure 4**), except one out of eight sequencing reads covering the −13.910 polymorphism in a Spanish individual from the Mesolithic (95), which has been interpreted as caused by a cytosine deamination, a typical damage in ancient DNA. The LP allele was therefore not present as a standing variation (at least at appreciable frequency) in pre-Neolithic European hunter-gatherer populations.

More surprisingly, no LP allele was found in 69 ancient Europeans dating back to the early or middle Neolithic (approximately 8,500–5,000 BP), whether from the Epicardial culture in

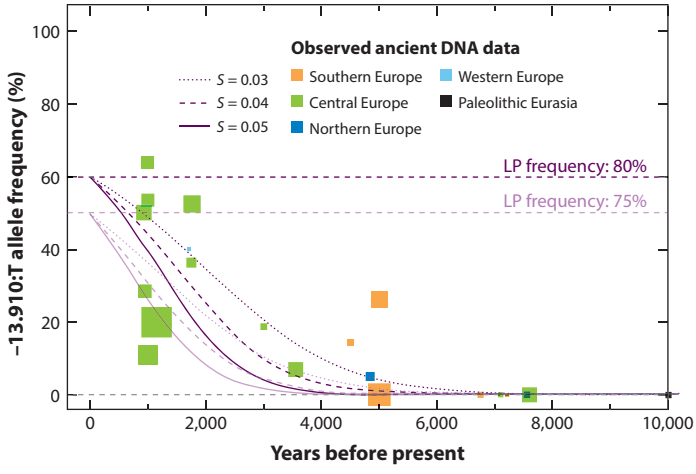

Figure 4

Evolution of lactase persistence (LP) in Europe over the last 10,000 years. The figure shows the theoretical expectations of the trajectory of an allele under selection for various selection coefficients (*S*) with a final allele frequency of 50% (*light purple lines*) or 60% (*dark purple lines*) superposed on −13.910:T allele frequencies observed in ancient DNA data sets (*colored squares*). The allele frequencies of 50% and 60% correspond to LP frequencies of 75% and 80%, respectively, as observed in modern populations from central Europe. The sizes of the colored squares are proportional to the number of samples (from 1 to 35), and the colors indicate the area of origin of the human remains. We used a generation time of 30 years to obtain dates in years. The frequencies are taken from References 2, 18, 39, 40, 42, 57, 67, 71, 74, 75, 83, 84, 91, 102, and 146. The value for Reference 85 was not numerically available and is therefore not included here.

southern Europe (75) or from the Linearbandkeramic (LBK)–associated cultures in central Europe (42, 146) (**Figure 4**). Both of these cultures are thought to have emerged from demographic migrations of the agropastoralist populations from Anatolia (42, 127, 128), who were the first to consume milk (see sidebar titled First Evidence of Milk Processing). Given that the LBKs emerged in the same area (the Danube basin) and the same time (approximately 6,000 BP) (8) that were inferred for selection of LP (**Table 1**), and that these populations practiced dairying, the appearance of LP in Europe is often associated with the LBK complex (64).

LP was actually found for the first time in Europe in the late Neolithic (approximately 5,000–4,300 BP) (**Figure 4**). It was found (*a*) surprisingly, in one out of ten hunter-gatherers (−13.910:T frequency of 5%) from the Pitted Ware culture in Sweden (83); (*b*) in two burials sites from the Corded Ware culture in northern Spain (−13.910:T frequency of 14% and 26%) (102); and (*c*) in central Europeans (−13.910:T frequency of approximately 5%), with the oldest occurrence of LP being in a 4,350-year-old individual from the Bell-Beaker culture, which is strongly associated with the Corded Ware culture (85). Allentoft et al. (2) further estimated that the LP allele frequency was 7% during the Bronze Age in central Europe (i.e., in descendants from the Corded Ware culture). The first observation of LP in Europe is therefore concomitant with an important demographic and cultural event: the massive migration into Europe of eastern steppe populations related to the Yamnaya culture (2, 49), a nomad steppe culture heavily reliant on cattle herding and sporadic agriculture (69), which subsequently admixed with European populations, resulting in the development of the Corded Ware culture in central Europe. Interestingly, this association between LP and steppe ancestry was also supported by a data set of 13 ancient Hungarians from the Iron Age (42): The individual with the LP allele was the only one with a high proportion of genetic ancestry from steppe populations. The absence of LP before the late Neolithic and the correlation between its appearance and migration from the steppes have therefore led to the alternative hypothesis of LP first arising in a pastoralist steppe population and then being brought to western Europe at the beginning of the Corded Ware culture (2).

However, these steppe populations have been estimated to display a very small amount of LP (−13.910:T frequency of 0% and approximately 6% in two studies from the Bronze Age) (2, 85), challenging the idea that they are the source populations for LP. When imputing the −13.910:T frequency based on surrounding markers, Allentoft et al. (2) reported a high LP allele frequency (approximately 20%) in ancient steppe populations. However, the major haplotype currently carrying the −13.910:T allele is also found carrying the −13.910:C (LNP) allele both in high frequency in modern Eurasian populations (31) and in one early Neolithic individual (the "Stuttgart" individual) (68). Therefore, the −13.910 genotype cannot be reliably imputed from surrounding sequences, whether from modern or ancient data.

In summary, the LP allele has not been found anywhere before 5,000 BP; its frequency then reached 5–26% (depending on the study) in late-Neolithic Europeans associated with the Corded Ware culture, 7% during the Bronze Age, and 19% during the Iron Age (in the Hallstatt culture in Poland) (146) (**Figure 4**). The allele frequency then clearly increased, reaching 36% and 53% at two Roman sites in Poland (11 and 20 individuals, respectively) (146).

In the Middle Ages, 80 individuals from four contemporaneous archaeological sites from the same area in Poland had heterogeneous LP allele frequencies (20–64%) (146). This heterogeneity was also found in Germany and Hungary, where Middle Ages populations had an LP allele frequency of 50% (71) and 11% (91), respectively. In the latter case, two different cultural backgrounds were differentiated: the "commoners," a third of whom displayed the LP phenotype, and the "conquerors," who belonged to pastoralist nomad tribes who invaded Hungary in 895 AD, in whom no LP allele was found. The high variability of LP allele frequency may be due either to a

strong stratification of protohistoric and historic populations or to noise, given the small number of individuals who have been analyzed from each population.

To compare the observed ancient DNA data with theoretical expectations, we calculated the expected trajectory of a selected allele under a dominant model with a selection coefficient of 0.03, 0.04, and 0.05, reaching a modern allele frequency of 50% or 60% (corresponding to an LP frequency of 75% and 80%, respectively, as observed in central Europe), and assuming a constant population size and strength of selective pressure. As shown in **Figure 4**, except for the case of the 0.03 selection coefficient (which is lower than most estimated values based on modern data; **Table 2**), we do not expect to observe a substantial LP allele frequency before 3,000 BP. The theoretical allele trajectory is therefore broadly compatible with the observed ancient DNA data, as was also concluded by a recent study (92). The main puzzle is actually in the reverse direction than that previously highlighted: The estimated LP frequency for late-Neolithic European populations (14% and 26% in 7 and 19 individuals from northern Spain, respectively) is higher than expected, especially for a south European population. This discrepancy could result from the sampling error, because of particular demographic events in these populations, or it could indicate that the selection coefficient has not been constant in time, with a higher selective pressure in the distant past than at present.

In conclusion, although ancient DNA data are valuable in this context and have allowed investigators to confirm a progressive increase of LP in Europe since the late Neolithic, the questions around the timing of selection, the strength of the selective pressure, and the geographical origin of the LP allele (LBK or steppe population) are still open. The accumulation of more population data from the late Neolithic, including samples from steppe populations that predate the Yamnaya, should allow a better evaluation of the frequency of the LP allele and the strength of selection in Europe.

CONCLUSION AND PERSPECTIVES

Although LP has been investigated in an impressive number of individuals since the 1960s, some geographical areas remain phenotypically understudied, notably in North Africa and West Africa and around the Caucasus. The situation should also be clarified in western Asia, notably in Pakistan, where contrasting results have been obtained, and in Tibetans, for whom we lack phenotypic data entirely. Despite the identification of five LP-associated alleles, some additional molecular basis of the trait remains to be uncovered. Indeed, two populations with substantial LP completely lack known LP-associated mutations: the Hadza from Tanzania, who have an LP frequency of 47% (estimated in 19 individuals with the hydrogen test) (107, 134), and the Wolof from Senegal, who have an LP frequency of 51% (although the phenotypic and genetic data for the latter did not come from the same study) (3, 62). The case of the Hadza is particularly puzzling, because no associated mutation has been found despite a large resequencing effort, including intron 9 of *MCM6* (1.3 kb), intron 13 of *MCM6* (3.2 kb), and the *LCT* promoter (2 kb) (107, 134). More generally, the current known alleles have been calculated to account for at most 45% of the phenotypic variation in African populations (107). Future studies assessing genotype-phenotype associations would benefit from testing LP with both the glucose and hydrogen tests, as indirect approaches are influenced by several confounding factors and likely have a high error rate. An interesting research direction would be to study the gut microbiome composition in LNP individuals, particularly to understand which factors are associated with intestinal symptoms. Additional molecular work (both in vitro and in vivo), notably testing the effect of multiple mutations on lactase expression thanks to mutagenesis, would also help to understand the size of the mutational target and the pathways involved in the regulation of *LCT* expression.

Concerning the nature of the selective pressures responsible for the increase in LP frequency in multiple populations, it seems clear that broadening the dietary repertoire and being able to derive glucose from milk were strongly favored in some populations. Why this was not the case in all pastoralists and was especially the case in Europeans is not entirely clear but could have resulted from a combination of cultural, nutritional, and environmental factors, such as the preference or need to ferment and transform milk into low-lactose dairy products, the stability and availability of other food, the constraints resulting from seasonality and mobility, and the type of livestock. There are a spectrum of dietary practices, ranging from nonmilking pastoralists (although these are rare) to mostly fermenting pastoralists to heavily milk-drinking pastoralists, and this variety seems to be better correlated with LP frequency than the levels of pastoralism are (28). In any case, more nutritional anthropological studies investigating the amount, type, and seasonality of milk and dairy products consumed and the perception of these foods in traditional populations would be helpful in understanding why some populations took up drinking fresh milk, whereas others mostly transformed it. Additionally, patterns of admixture between pastoral and nonpastoral populations might have played an important role in limiting the efficacy of natural selection.

Interestingly, although most of the focus has been on the cost of drinking milk for LNP individuals, arguments can also be made that consuming milk or dairy also has advantages for these individuals. Indeed, recent studies have shown that the gut microbiota of LNP individuals differ from those of LP individuals from the same population in that they have a higher prevalence of *Bifidobacterium* (11, 15, 47), which is explained by the larger amounts of lactose available for bacterial fermentation. In fact, this represents one of the strongest associations to date between genetic variants and variation in gut microbiome composition (11, 15, 47). A consequence is that LNP individuals who consume fermented dairy products have higher levels of short-chain fatty acids, the end products of fermentation (**Figure 2**), which represent an important source of energy for hosts (between 10% and 30% of their basal metabolic needs) (23). However, most of these fatty acids are usually derived from carbohydrates other than lactose, notably from starch and nonstarchy polysaccharides (4, 81). It has been estimated that the ingestion of lactose provides 4 kcal/g when digested in the small intestine (as in LP), whereas it yields approximately 2 kcal/g if fermented in the colon (as in LNP) (118). The difference in energetic uptake between LP and LNP is therefore not that high (especially if adding to that the benefits from proteins and fat). Outside of the energetic value of lactose, some have proposed that lactose should be considered a prebiotic for LNP individuals (133) because it stimulates the growth of lactic acid bacteria that are thought to be beneficial for human health, notably owing to their production of antibacterial peptides and their stimulation of the host immune system (26, 72). This feature could therefore represent an alternative selective advantage of fermenting milk in populations with low amounts of vegetal carbohydrates or high pathogenic loads.

Finally, the lactase enzyme—more accurately known as the lactase-phlorizin hydrolase—has the ability to hydrolyze not only lactose but also other β-galactosides and β-glucosides, such as phlorizin and flavonoid glucosides in plants (24, 129). Because LP frequencies are high in some nonpastoral populations with no milk in their diets [namely the Khoisan-speaking Hadza hunter-gatherers (47%) and the Afro-Asiatic Yaaku hunter-gatherers (78%)], some have proposed that LP could also be selected for its broader role in the hydrolysis of vegetal molecules (107, 134). However, as pointed out by Gerbault et al. (45), this hypothesis may not be realistic, given that another intestinal enzyme, cytosolic β-glucosidase, is also able to hydrolyze the same compounds (25). In the end, more research will be needed to explain why these hunter-gatherers are lactase persistent.

DISCLOSURE STATEMENT

The authors are not aware of any affiliations, memberships, funding, or financial holdings that might be perceived as affecting the objectivity of this review.

ACKNOWLEDGMENTS

We wish to thank Molly Przeworski for her helpful comments on an earlier version of this review.

LITERATURE CITED

1. Ahmad M, Flatz G. 1984. Prevalence of primary adult lactose malabsorption in Pakistan. *Hum. Hered.* 34:69–75
2. Allentoft ME, Sikora M, Sjögren K-G, Rasmussen S, Rasmussen M, et al. 2015. Population genomics of Bronze Age Eurasia. *Nature* 522:167–72
3. Arnold J, Diop M, Kodjovi M, Rozier J. 1980. Lactose intolerance in adults in Senegal. *C. R. Seances Soc. Biol. Fil.* 174:983–92
4. Arola H, Tamm A. 1994. Metabolism of lactose in the human body. *Scand. J. Gastroenterol. Suppl.* 202:21–25
5. Bayoumi RA, Flatz SD, Kuhnau W, Flatz G. 1982. Beja and Nilotes: nomadic pastoralist groups in the Sudan with opposite distributions of the adult lactase phenotypes. *Am. J. Phys. Anthropol.* 58:173–78
6. Beall CM. 2014. Adaptation to high altitude: phenotypes and genotypes. *Annu. Rev. Anthropol.* 43:251–72
7. Beja-Pereira A, Luikart G, England PR, Bradley DG, Jann OC, et al. 2003. Gene-culture coevolution between cattle milk protein genes and human lactase genes. *Nat. Genet.* 35:311–13
8. Bellwood PS. 2005. *First Farmers: The Origins of Agricultural Societies*. Malden, MA: Blackwell
9. Bernus E. 1988. Seasonality, climatic fluctuations, and food supplies (Sahelian nomadic pastoral societies). In *Coping with Uncertainty in Food Supply*, ed. I de Garine, GA Harrison, pp. 318–36. Oxford, UK: Clarendon
10. Bersaglieri T, Sabeti PC, Patterson N, Vanderploeg T, Schaffner SF, et al. 2004. Genetic signatures of strong recent positive selection at the lactase gene. *Am. J. Hum. Genet.* 74:1111–20
11. Blekhman R, Goodrich JK, Huang K, Sun Q, Bukowski R, et al. 2015. Host genetic variation impacts microbiome composition across human body sites. *Genome Biol.* 16:191
12. Bloom G, Sherman P. 2005. Dairying barriers affect the distribution of lactose malabsorption. *Evol. Hum. Behav.* 26:301–12
13. Bodley JH. 2011. African cattle peoples: tribal pastoralists. In *Cultural Anthropology: Tribes, States, and the Global System*, pp. 99–128. Lanham, MD: AltaMira
14. Bond JH, Levitt MD. 1976. Quantitative measurement of lactose absorption. *Gastroenterology* 70:1058–62
15. Bonder MJ, Kurilshikov A, Tigchelaar EF, Mujagic Z, Imhann F, et al. 2016. The effect of host genetics on the gut microbiome. *Nat. Genet.* 48:1407–12
16. Breton G, Schlebusch CM, Lombard M, Sjodin P, Soodyall H, Jakobsson M. 2014. Lactase persistence alleles reveal partial East African ancestry of southern African Khoe pastoralists. *Curr. Biol.* 24:852–58
17. Brussow H. 2013. Nutrition, population growth and disease: a short history of lactose. *Environ. Microbiol.* 15:2154–61
18. Burger J, Kirchner M, Bramanti B, Haak W, Thomas MG. 2007. Absence of the lactase-persistence-associated allele in early Neolithic Europeans. *PNAS* 104:3736–41
19. Campbell AK, Waud JP, Matthews SB. 2005. The molecular basis of lactose intolerance. *Sci. Prog.* 88:157–202
20. Cavalli-Sforza LT, Strata A. 1987. Double-blind study on the tolerance of four types of milk in lactose malabsorbers and absorbers. *Hum. Nutr. Clin. Nutr.* 41:19–30
21. Cook GC. 1978. Did persistence of intestinal lactase into adult life originate on the Arabian peninsula? *Man* 13:418–27
22. Cook GC, al-Torki MT. 1975. High intestinal lactase concentrations in adult Arabs in Saudi Arabia. *Br. Med. J.* 3:135–36

23. Cummings JH, Macfarlane GT. 1991. The control and consequences of bacterial fermentation in the human colon. *J. Appl. Bacteriol.* 70:443–59
24. Day AJ, Canada FJ, Diaz JC, Kroon PA, McLauchlan R, et al. 2000. Dietary flavonoid and isoflavone glycosides are hydrolysed by the lactase site of lactase phlorizin hydrolase. *FEBS Lett*. 468:166–70
25. Day AJ, DuPont MS, Ridley S, Rhodes M, Rhodes MJ, et al. 1998. Deglycosylation of flavonoid and isoflavonoid glycosides by human small intestine and liver beta-glucosidase activity. *FEBS Lett*. 436:71–75
26. Di Cerbo A, Palmieri B, Aponte M, Morales-Medina JC, Iannitti T. 2016. Mechanisms and therapeutic effectiveness of lactobacilli. *J. Clin. Pathol.* 69:187–203
27. Dunne J, Evershed RP, Salque M, Cramp L, Bruni S, et al. 2012. First dairying in green Saharan Africa in the fifth millennium BC. *Nature* 486:390–94
28. Durham WH. 1991. Cultural mediation: the evolution of adult lactose absorption. In *Coevolution: Genes, Culture and Human Diversity*, pp. 226–85. Stanford, CA: Stanford Univ. Press
29. Enattah NS, Jensen TG, Nielsen M, Lewinski R, Kuokkanen M, 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.* 82:57–72
30. Enattah NS, Sahi T, Savilahti E, Terwilliger JD, Peltonen L, Jarvela I. 2002. Identification of a variant associated with adult-type hypolactasia. *Nat. Genet.* 30:233–37
31. Enattah NS, Trudeau A, Pimenoff V, Maiuri L, Auricchio S, et al. 2007. Evidence of still-ongoing convergence evolution of the lactase persistence T-13910 alleles in humans. *Am. J. Hum. Genet.* 81:615–25
32. Evershed RP, Payne S, Sherratt AG, Copley MS, Coolidge J, et al. 2008. Earliest date for milk use in the Near East and southeastern Europe linked to cattle herding. *Nature* 455:528–31
33. Fang L, Ahn JK, Wodziak D, Sibley E. 2012. The human lactase persistence-associated SNP −13910*T enables in vivo functional persistence of lactase promoter-reporter transgene expression. *Hum. Genet.* 131:1153–59
34. FAO (Food Agric. Organ. UN). 2003. *Transhumant Grazing Systems in Temperate Asia*. Plant Prod. Prot. Ser. No. 31. Rome: FAO
35. FAO (Food Agric. Organ. UN). 2013. *Milk and Dairy Products in Human Nutrition*. Rome: FAO
36. Flatz G. 1987. Genetics of lactose digestion in humans. *Adv. Hum. Genet.* 16:1–77
37. Flatz G, Rotthauwe HW. 1971. Evidence against nutritional adaption of tolerance to lactose. *Humangenetik* 13:118–25
38. Flatz G, Rotthauwe HW. 1973. Lactose nutrition and natural selection. *Lancet* 302:76–77
39. Fu Q, Li H, Moorjani P, Jay F, Slepchenko SM, et al. 2014. Genome sequence of a 45,000-year-old modern human from western Siberia. *Nature* 514:445–49
40. Fu Q, Posth C, Hajdinjak M, Petr M, Mallick S, et al. 2016. The genetic history of Ice Age Europe. *Nature* 534:200–5
41. Gallego Romero I, Basu Mallick C, Liebert A, Crivellaro F, Chaubey G, et al. 2012. Herders of Indian and European cattle share their predominant allele for lactase persistence. *Mol. Biol. Evol.* 29:249–60
42. Gamba C, Jones ER, Teasdale MD, McLaughlin RL, Gonzalez-Fortes G, et al. 2014. Genome flux and stasis in a five millennium transect of European prehistory. *Nat. Commun.* 5:5257
43. Gerbault P, Liebert A, Itan Y, Powell A, Currat M, et al. 2011. Evolution of lactase persistence: an example of human niche construction. *Philos. Trans. R. Soc. Lond. B* 366:863–77
44. Gerbault P, Moret C, Currat M, Sanchez-Mazas A. 2009. Impact of selection and demography on the diffusion of lactase persistence. *PLOS ONE* 4:e6369
45. Gerbault P, Roffet-Salque M, Evershed RP, Thomas MG. 2013. How long have adult humans been consuming milk? *IUBMB Life* 65:983–90
46. Gilat T, Benaroya Y, Gelman-Malachi E, Adam A. 1973. Genetics of primary adult lactase deficiency. *Gastroenterology* 64:562–68
47. Goodrich JK, Davenport ER, Beaumont M, Jackson MA, Knight R, et al. 2016. Genetic determinants of the gut microbiome in UK twins. *Cell Host Microbe* 19:731–43
48. Grant JD, Bezerra JA, Thompson SH, Lemen RJ, Koldovsky O, Udall JN Jr. 1989. Assessment of lactose absorption by measurement of urinary galactose. *Gastroenterology* 97:895–99

49. Haak W, Lazaridis I, Patterson N, Rohland N, Mallick S, et al. 2015. Massive migration from the steppe was a source for Indo-European languages in Europe. *Nature* 522:207–11
50. Harrison GG. 1975. Primary adult lactase deficiency: a problem in anthropological genetics. *Am. Anthropol.* 77:812–35
51. He T, Priebe MG, Harmsen HJ, Stellaard F, Sun X, et al. 2006. Colonic fermentation may play a role in lactose intolerance in humans. *J. Nutr.* 136:58–63
52. He T, Priebe MG, Vonk RJ, Welling GW. 2005. Identification of bacteria with beta-galactosidase activity in faeces from lactase non-persistent subjects. *FEMS Microbiol. Ecol.* 54:463–69
53. Hedrick PW. 2011. Population genetics of malaria resistance in humans. *Heredity* 107:283–304
54. Hertzler SR, Savaiano DA. 1996. Colonic adaptation to daily lactose feeding in lactose maldigesters reduces lactose intolerance. *Am. J. Clin. Nutr.* 64:232–36
55. Heyer E, Brazier L, Ségurel L, Hegay T, Austerlitz F, et al. 2011. Lactase persistence in central Asia: phenotype, genotype, and evolution. *Hum. Biol.* 83:379–92
56. Ho MW, Povey S, Swallow D. 1982. Lactase polymorphism in adult British natives: estimating allele frequencies by enzyme assays in autopsy samples. *Am. J. Hum. Genet.* 34:650–57
57. Hofmanova Z, Kreutzer S, Hellenthal G, Sell C, Diekmann Y, et al. 2016. Early farmers from across Europe directly descended from Neolithic Aegeans. *PNAS* 113:6886–91
58. Holden C, Mace R. 1997. Phylogenetic analysis of the evolution of lactose digestion in adults. *Hum. Biol.* 69:605–28
59. Imtiaz F, Savilahti E, Sarnesto A, Trabzuni D, Al-Kahtani K, 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:e89
60. Ingram CJ, Elamin MF, Mulcare CA, Weale ME, Tarekegn A, et al. 2007. A novel polymorphism associated with lactose tolerance in Africa: multiple causes for lactase persistence? *Hum. Genet.* 120:779–88
61. Ingram CJ, Mulcare CA, Itan Y, Thomas MG, Swallow DM. 2009. Lactose digestion and the evolutionary genetics of lactase persistence. *Hum. Genet.* 124:579–91
62. Ingram CJ, Raga TO, Tarekegn A, Browning SL, Elamin MF, et al. 2009. Multiple rare variants as a cause of a common phenotype: several different lactase persistence associated alleles in a single ethnic group. *J. Mol. Evol.* 69:579–88
63. Itan Y, Jones BL, Ingram CJ, Swallow DM, Thomas MG. 2010. A worldwide correlation of lactase persistence phenotype and genotypes. *BMC Evol. Biol.* 10:36
64. Itan Y, Powell A, Beaumont MA, Burger J, Thomas MG. 2009. The origins of lactase persistence in Europe. *PLOS Comput. Biol.* 5:e1000491
65. Johnson JD, Simoons FJ, Hurwitz R, Grange A, Mitchell CH, et al. 1977. Lactose malabsorption among the Pima Indians of Arizona. *Gastroenterology* 73:1299–304
66. Jones BL, Raga TO, Liebert A, Zmarz P, Bekele E, et al. 2013. Diversity of lactase persistence alleles in Ethiopia: signature of a soft selective sweep. *Am. J. Hum. Genet.* 93:538–44
67. Jones ER, Gonzalez-Fortes G, Connell S, Siska V, Eriksson A, et al. 2015. Upper Palaeolithic genomes reveal deep roots of modern Eurasians. *Nat. Commun.* 6:8912
68. Key FM, Fu Q, Romagné F, Lachmann M, Andrés AM. 2016. Human adaptation and population differentiation in the light of ancient genomes. *Nat. Commun.* 7:10775
69. Kohl PL. 2007. *The Making of Bronze Age Eurasia*. Cambridge, UK: Cambridge Univ. Press
70. Kolars JC, Levitt MD, Aouji M, Savaiano DA. 1984. Yogurt—an autodigesting source of lactose. *N. Engl. J. Med.* 310:1–3
71. Krüttli A, Bouwman A, Akgül G, Della Casa P, Rühli F, Warinner C. 2014. Ancient DNA analysis reveals high frequency of European lactase persistence allele (T-13910) in medieval central Europe. *PLOS ONE* 9:e86251
72. Kumari A, Catanzaro R, Marotta F. 2011. Clinical importance of lactic acid bacteria: a short review. *Acta Biomed.* 82:177–80
73. Kwak H, Lee W, Lee M. 2012. Revisiting lactose as an enhancer of calcium absorption. *Int. Dairy J.* 22:147–51

74. Lacan M, Keyser C, Ricaut F-X, Brucato N, Duranthon F, Guilaine J. 2011. Ancient DNA reveals male diffusion through the Neolithic Mediterranean route. *PNAS* 108:9788–91
75. Lacan M, Keyser C, Ricaut F-X, Brucato N, Tarrús J, et al. 2011. Ancient DNA suggests the leading role played by men in the Neolithic dissemination. *PNAS* 108:18255–59
76. Levitt MD, Donaldson RM. 1970. Use of respiratory hydrogen (H_2) excretion to detect carbohydrate malabsorption. *J. Lab. Clin. Med.* 75:937–45
77. Lewinsky RH, Jensen TG, Møller J, Stensballe A, Olsen J, Troelsen JT. 2005. T-13910 DNA variant associated with lactase persistence interacts with Oct-1 and stimulates lactase promoter activity in vitro. *Hum. Mol. Genet.* 14:3945–53
78. Liebert A, Jones BL, Danielsen ET, Olsen AK, Swallow DM, Troelsen JT. 2016. In vitro functional analyses of infrequent nucleotide variants in the lactase enhancer reveal different molecular routes to increased lactase promoter activity and lactase persistence. *Ann. Hum. Genet.* 80:307–18
79. Lisker R, Gonzalez B, Daltabuit M. 1975. Recessive inheritance of the adult type of intestinal lactase deficiency. *Am. J. Hum. Genet.* 27:662–64
80. Lokki AI, Jarvela I, Israelsson E, Maiga B, Troye-Blomberg M, et al. 2011. Lactase persistence genotypes and malaria susceptibility in Fulani of Mali. *Malaria J.* 10:9
81. Lomer MC, Parkes GC, Sanderson JD. 2008. Review article: lactose intolerance in clinical practice—myths and realities. *Aliment. Pharmacol. Ther.* 27:93–103
82. MacHugh DE, Larson G, Orlando L. 2017. Taming the past: ancient DNA and the study of animal domestication. *Annu. Rev. Anim. Biosci.* 5:329–51
83. Malmström H, Linderholm A, Lidén K, Storå J, Molnar P, et al. 2010. High frequency of lactose intolerance in a prehistoric hunter-gatherer population in northern Europe. *BMC Evol. Biol.* 10:89
84. Martiniano R, Caffell A, Holst M, Hunter-Mann K, Montgomery J, et al. 2016. Genomic signals of migration and continuity in Britain before the Anglo-Saxons. *Nat. Commun.* 7:10326
85. Mathieson I, Lazaridis I, Rohland N, Mallick S, Patterson N, et al. 2015. Genome-wide patterns of selection in 230 ancient Eurasians. *Nature* 528:499–503
86. McCracken RD. 1971. Lactase deficiency: an example of dietary evolution. *Curr. Anthropol.* 12:479–517
87. McCracken RD. 1971. Origins and implications of the distribution of adult lactase deficiency in human populations. *J. Trop. Pediatr. Environ. Child Health* 17:7–10
88. Métneki J, Czeizel A, Flatz SD, Flatz G. 1984. A study of lactose absorption capacity in twins. *Hum. Genet.* 67:296–300
89. Murdock GP. 1967. Ethnographic Atlas: a summary. *Ethnology* 6:109–236
90. Myles S, Bouzekri N, Haverfield E, Cherkaoui M, Dugoujon JM, Ward R. 2005. Genetic evidence in support of a shared Eurasian-North African dairying origin. *Hum. Genet.* 117:34–42
91. Nagy D, Tömöry G, Csányi B, Bogácsi-Szabó E, Czibula Á, et al. 2011. Comparison of lactase persistence polymorphism in ancient and present-day Hungarian populations. *Am. J. Phys. Anthropol.* 145:262–69
92. Nakagome S, Alkorta-Aranburu G, Amato R, Howie B, Peter BM, et al. 2016. Estimating the ages of selection signals from different epochs in human history. *Mol. Biol. Evol.* 33:657–69
93. Newcomer AD, McGill DB. 1966. Lactose tolerance tests in adults with normal lactase activity. *Gastroenterology* 50:340–46
94. Nordborg M, Tavare S. 2002. Linkage disequilibrium: what history has to tell us. *Trends Genet.* 18:83–90
95. Olalde I, Allentoft ME, Sánchez-Quinto F, Santpere G, Chiang CWK, et al. 2014. Derived immune and ancestral pigmentation alleles in a 7,000-year-old Mesolithic European. *Nature* 507:225–28
96. Olds LC, Ahn JK, Sibley E. 2011. −13915*G DNA polymorphism associated with lactase persistence in Africa interacts with Oct-1. *Hum. Genet.* 129:111–13
97. Olds LC, Sibley E. 2003. Lactase persistence DNA variant enhances lactase promoter activity in vitro: functional role as a *cis* regulatory element. *Hum. Mol. Genet.* 12:2333–40
98. Orlando L. 2016. Back to the roots and routes of dromedary domestication. *PNAS* 113:201606340
99. Outram AK, Stear NA, Bendrey R, Olsen S, Kasparov A, et al. 2009. The earliest horse harnessing and milking. *Science* 323:1332–35
100. Peng MS, He JD, Zhu CL, Wu SF, Jin JQ, Zhang YP. 2012. Lactase persistence may have an independent origin in Tibetan populations from Tibet, China. *J. Hum. Genet.* 57:394–97

101. Peter BM, Huerta-Sanchez E, Nielsen R. 2012. Distinguishing between selective sweeps from standing variation and from a de novo mutation. *PLOS Genet.* 8:e1003011
102. Plantinga TS, Alonso S, Izagirre N, Hervella M, Fregel R, et al. 2012. Low prevalence of lactase persistence in Neolithic South-West Europe. *Eur. J. Hum. Genet.* 20:778–82
103. Prentice AM. 2005. Starvation in humans: evolutionary background and contemporary implications. *Mech. Ageing Dev.* 126:976–81
104. Priehodova E, Abdelsawy A, Heyer E, Cerny V. 2014. Lactase persistence variants in Arabia and in the African Arabs. *Hum. Biol.* 86:7–18
105. Qiu Q, Wang L, Wang K, Yang Y, Ma T, et al. 2015. Yak whole-genome resequencing reveals domestication signatures and prehistoric population expansions. *Nat. Commun.* 6:10283
106. Rab SM, Baseer A. 1976. High intestinal lactase concentration in adult Pakistanis. *Br. Med. J.* 1:436
107. Ranciaro A, Campbell MC, Hirbo JB, Ko WY, Froment A, et al. 2014. Genetic origins of lactase persistence and the spread of pastoralism in Africa. *Am. J. Hum. Genet.* 94:496–510
108. Rocha J. 2012. The evolution of lactase persistence. *Anthropol. Porteguesa* 29:121–37
109. Røed KH, Flagstad O, Nieminen M, Holand Ø, Dwyer MJ, et al. 2008. Genetic analyses reveal independent domestication origins of Eurasian reindeer. *Proc. R. Soc. B* 275:1849–55
110. Ross AB, Johansson A, Ingman M, Gyllensten U. 2006. Lifestyle, genetics, and disease in Sami. *Croat. Med. J.* 47:553–65
111. Rossel S, Marshall F, Peters J, Pilgram T, Adams MD, O'Connor D. 2008. Domestication of the donkey: timing, processes, and indicators. *PNAS* 105:3715–20
112. Sahi T. 1974. The inheritance of selective adult-type lactose malabsorption. *Scand. J. Gastroenterol. Suppl.* 30:1–73
113. Sahi T. 1978. Dietary lactose and the aetiology of human small-intestinal hypolactasia. *Gut* 19:1074–86
114. Sahi T. 1994. Genetics and epidemiology of adult-type hypolactasia. *Scand. J. Gastroenterol. Suppl.* 202:7–20
115. Sahi T, Isokoski M, Jussila J, Launiala K, Pyorala K. 1973. Recessive inheritance of adult-type lactose malabsorption. *Lancet* 302:823–26
116. Salque M, Bogucki PI, Pyzel J, Sobkowiak-Tabaka I, Grygiel R, et al. 2013. Earliest evidence for cheese making in the sixth millennium BC in northern Europe. *Nature* 493:522–25
117. Savaiano DA, Levitt MD. 1987. Milk intolerance and microbe-containing dairy foods. *J. Dairy Sci.* 70:397–406
118. Schaafsma G. 2008. Lactose and lactose derivatives as bioactive ingredients in human nutrition. *Int. Dairy J.* 18:458–65
119. Schlebusch CM, Sjodin P, Skoglund P, Jakobsson M. 2013. Stronger signal of recent selection for lactase persistence in Maasai than in Europeans. *Eur. J. Hum. Genet.* 21:550–53
120. Scrimshaw NS, Murray EB. 1988. The acceptability of milk and milk products in populations with a high prevalence of lactose intolerance. *Am. J. Clin. Nutr.* 48:1079–159
121. Seguin-Orlando A, Korneliussen TS, Sikora M, Malaspinas A-S, Manica A, et al. 2014. Genomic structure in Europeans dating back at least 36,200 years. *Science* 346:1113–18
122. Shaukat A, Levitt MD, Taylor BC, MacDonald R, Shamliyan TA, et al. 2010. Systematic review: effective management strategies for lactose intolerance. *Ann. Intern. Med.* 152:797–803
123. Simoons FJ. 1969. Primary adult lactose intolerance and the milking habit: a problem in biological and cultural interrelations. I. Review of the medical research. *Am. J. Dig. Dis.* 14:819–36
124. Simoons FJ. 1970. Primary adult lactose intolerance and the milking habit: a problem in biologic and cultural interrelations. II. A culture historical hypothesis. *Am. J. Dig. Dis.* 15:695–710
125. Simoons FJ. 1970. The traditional limits of milking and milk use in southern Asia. *Anthropos* 65:557–93
126. Simoons FJ. 1978. The geographic hypothesis and lactose malabsorption. A weighing of the evidence. *Am. J. Dig. Dis.* 23:963–80
127. Skoglund P, Malmström H, Omrak A, Raghavan M, Valdiosera C, et al. 2014. Genomic diversity and admixture differs for stone-age Scandinavian foragers and farmers. *Science* 344:747–50
128. Skoglund P, Malmström H, Raghavan M, Storå J, Hall P, et al. 2012. Origins and genetic legacy of Neolithic farmers and hunter-gatherers in Europe. *Science* 336:466–69

129. Skovbjerg H, Sjöström H, Norén O. 1981. Purification and characterisation of amphiphilic lactase/phlorizin hydrolase from human small intestine. *Eur. J. Biochem.* 114:653–61
130. Smith JM, Haigh J. 1974. The hitch-hiking effect of a favourable gene. *Genet. Res.* 23:23–35
131. Suarez FL, Savaiano D, Arbisi P, Levitt MD. 1997. Tolerance to the daily ingestion of two cups of milk by individuals claiming lactose intolerance. *Am. J. Clin. Nutr.* 65:1502–6
132. Swallow DM. 2003. Genetics of lactase persistence and lactose intolerance. *Annu. Rev. Genet.* 37:197–219
133. Szilagyi A. 2015. Adaptation to lactose in lactase non persistent people: effects on intolerance and the relationship between dairy food consumption and evolution of diseases. *Nutrients* 7:6751–79
134. Tishkoff SA, Reed FA, Ranciaro A, Voight BF, Babbitt CC, et al. 2007. Convergent adaptation of human lactase persistence in Africa and Europe. *Nat. Genet.* 39:31–40
135. Tishkoff SA, Varkonyi R, Cahinhinan N, Abbes S, Argyropoulos G, et al. 2001. Haplotype diversity and linkage disequilibrium at human G6PD: recent origin of alleles that confer malarial resistance. *Science* 293:455–62
136. Torniainen S, Parker MI, Holmberg V, Lahtela E, Dandara C, Jarvela I. 2009. Screening of variants for lactase persistence/non-persistence in populations from South Africa and Ghana. *BMC Genet.* 10:31
137. Townley RR, Khaw KT, Shwachman H. 1965. Quantitative assay of disaccharidase activities of small intestinal mucosal biopsy specimens in infancy and childhood. *Pediatrics* 36:911–21
138. Troelsen JT. 2005. Adult-type hypolactasia and regulation of lactase expression. *Biochim. Biophys. Acta* 1723:19–32
139. Troelsen JT, Olsen J, Møller J, Sjöström H. 2003. An upstream polymorphism associated with lactase persistence has increased enhancer activity. *Gastroenterology* 125:1686–94
140. Vigne J-D. 2012. Elevage laitier au Néolithique. In *Dictionnaire des cultures alimentaires*, ed. J-P Poulain, pp. 442–51. Paris: Press. Univ. France
141. Voight BF, Kudaravalli S, Wen X, Pritchard JK. 2006. A map of recent positive selection in the human genome. *PLOS Biol.* 4:e72
142. Vonk RJ, Priebe MG, Koetse HA, Stellaard F, Lenoir-Wijnkoop I, et al. 2003. Lactose intolerance: analysis of underlying factors. *Eur. J. Clin. Investig.* 33:70–75
143. Wang Y, Harvey CB, Pratt WS, Sams VR, Sarner M, et al. 1995. The lactase persistence/non-persistence polymorphism is controlled by a *cis*-acting element. *Hum. Mol. Genet.* 4:657–62
144. Warinner C, Hendy J, Speller C, Cappellini E, Fischer R, et al. 2014. Direct evidence of milk consumption from ancient human dental calculus. *Sci. Rep.* 4:7104
145. Wilde S, Timpson A, Kirsanow K, Kaiser E, Kayser M, et al. 2014. Direct evidence for positive selection of skin, hair, and eye pigmentation in Europeans during the last 5,000 y. *PNAS* 111:4832–37
146. Witas HW, Płoszaj T, Jêdrychowska-Dańska K, Witas PJ, Masłowska A, et al. 2015. Hunting for the LCT-13910*T allele between the Middle Neolithic and the Middle Ages suggests its absence in dairying LBK people entering the Kuyavia region in the 8th millennium BP. *PLOS ONE* 10:1–24
147. Yongfa W, Yongshan Y, Jinjin X, Ruofu D, Flatz SD, et al. 1984. Prevalence of primary adult lactose malabsorption in three populations of northern China. *Hum. Genet.* 67:103–6
148. Zeder MA. 2006. *Documenting Domestication: New Genetic and Archaeological Paradigms*. Berkeley: Univ. Calif. Press
149. Zhong Y, Priebe MG, Vonk RJ, Huang CY, Antoine JM, et al. 2004. The role of colonic microbiota in lactose intolerance. *Dig. Dis. Sci.* 49:78–83