Lactase persistence and milk consumption in Europe: an interdisciplinary approach involving genetics and archaeology

Michela Leonardi

AG Palaeogenetik, Institute of Anthropology, Johannes Gutenberg University, Mainz, D
Human Evolutionary Ecology Group, Department of Anthropology, University College London, London, UK
m.leonardi@ucl.ac.uk

ABSTRACT - The ability to digest milk during adulthood (lactase persistence) is a genetically determined trait present only in humans. Its origin and diffusion are correlated with the development of pastoralism and the consumption of fresh milk. This work will present the genetic and archaeological data that allow the reconstruction of the co-evolutionary process between dairying culture and lactase persistence, as well as a discussion of the chronology and the way lactase persistence spread in Europe.

IZVLEČEK – Sposobnost presnavljanja mleka v odrasli dobi (laktazna persistenca) je genetsko pogoje- na značilnost, navzoča le pri ljudem. Njen izvor in razširitev sta povezana z razvojem pastirstva in uživanjem svežega mleka. V članku predstavljamo genetske in arheološke podatke, ki nam omogočajo rekonstrukcijo koevolucijskih procesov med mlečno-gospodarskimi kulturami in laktazno persis- stenco, poleg tega pa komentiramo kronologijo in način, kako se je laktazna persistenca širila v Evropo.

KEY WORDS – lactase persistence; gene-culture coevolution; European Neolithic; genetics; dairying

Introduction

Mammals have the ability to digest lactose – the main sugar contained in milk – only until weaning is over. After that, there is a significant reduction in the production of lactase, the enzyme that allows the absorption of lactose. In humans, however, at least four genetic mutations have been associated with the production of lactase into adulthood (lactase persistence). The frequency of lactase persistent individuals varies significantly between and within continents, and in some cases even between neighbouring regions. In Europe the distribution of lactase persistent individuals follows a cline showing lower frequencies in the south, and higher frequencies, reaching as much as 98%, in the north.

The study of ancient DNA shows that the lactase persistence associated variant was absent or present in very low frequencies in most regions before and during the Neolithic (Burger et al. 2007; Lacan et al. 2011; Malmström et al. 2009; Plantinga et al. 2012). This result suggests that, before the beginning of animal husbandry, the European population was not able to drink milk during adulthood without suffering from very unpleasant and sometimes deleterious symptoms. After the domestication of cattle, sheep and goats, milk became available as a source of nutrition for adults, and the frequency of the lactase persistence-associated mutation increased rapidly in the population, reaching the present-day rates.

The origin and spread of lactase persistence is a very complex process that, to be understood well, must be considered within its archaeological, genetic and social context (Gerbault et al. 2011; Leonardi et al. 2012).
The digestion of milk

Lactose is a disaccharide sugar that can be found in different percentages in the milk of almost all mammals apart from the platypus and some marine mammals (Reich, Arnould 2007). Lactose cannot be digested in its disaccharide form; to be assimilated, it has to be hydrolysed and separated into two monosaccharide sugars (glucose and galactose). This is performed in the mammal intestine with the enzyme lactase (lactase phlorizin hydrolase or LCT). The production of lactase in the small intestine is usually already high at birth (e.g., in humans; Wang et al. 1994), or peaks a few days later (e.g., in rodents; Troelsen 2005), and remains at more or less the same level until weaning is over, when it decreases significantly.

When this happens, the undigested lactose reaches the colon, where it is fermented by colonic bacteria. The glucose is then fermented involving the production of short chain fatty acids and gases, entailing osmotic effects. The result of this process is the appearance of very unpleasant symptoms such as bloating, chronic diarrhoea and abdominal cramps (Ingram et al. 2009a). The severity of symptoms can vary between individuals showing the same levels of lactase production (Vonk et al. 2003), since colonic adaptation can lead to a better capacity to ferment lactose and a reduction in the production of hydrogen by the colonic microbiota (Szilagyi et al. 2002).

In about 35% of adults worldwide, lactase is produced throughout their life, allowing them to drink milk at any age without experiencing any of the above-mentioned ill effects (Ingram et al. 2009a).

The genetics of lactase persistence

Lactase persistence is inherited as a dominant Mendelian trait (Ferguson, Maxwell 1967; Sahi et al. 1973; Sahi, Launiala 1977). A single gene located in chromosome 2, called LTC, codes for lactase. With the neighbouring MCM6 gene, a region (intron 13) seems to have an enhancer action on the LCT gene (Fang et al. 2012; Jensen et al. 2011; Letinsky et al. 2005; Olds et al. 2011; Olds, Sibley 2003; Troelsen et al. 2003). Several different single nucleotide polymorphisms (SNPs) in this region show a strong association with lactase persistence, and their distribution is geographically structured. The first mutation found to be linked to this trait is a cytosine to thymine transition 13910 C/T (Enattah et al. 2002). The derived variant at this locus (–13910*T) is associated with lactase persistence throughout Europe (Anagnostou et al. 2009; Ingram et al. 2009a; Itan et al. 2010; Manco et al. 2013; Nagy et al. 2009; Sun et al. 2007; Torniainen et al. 2009), Central Asia (Heyer et al. 2011) and India (Gallego Romero et al. 2012).

However, in Africa –13910 C/T alone does not explain the observed pattern of lactase persistence distribution, and several other SNPs within the same genetic region appear to be associated with the ability to digest large quantities of raw milk by adults. The same is true of the Middle East and in Tibet (Al-Abri et al. 2012; Enattah et al. 2008; Imtiaz et al. 2007; Ingram et al. 2007; Ingram et al. 2009b; Peng et al. 2012; Tishkoff et al. 2007). For this reason, several independent origins for lactase persistence have been proposed (Enattah et al. 2008; Ingram et al. 2007; Peng et al. 2012; Tishkoff et al. 2007).

Lactase persistence in modern worldwide populations

The frequency of lactase persistent individuals varies significantly between geographical regions (Itan et al. 2010). In Eastern Asia and in Native American populations a small percentage of the population is persistent (Itan et al. 2010). In Africa, the distribution is irregular, with frequencies that can reach as much as 92% in pastoralist populations, but can fall as low as 0.02% in non-pastoralist human groups (for a table listing the frequencies in a great number of global populations see http://www.ucl.ac.uk/mace-lab/resources/glad, partially published in Itan et al. 2010). A similar pattern can be observed in the Arabian Peninsula and in Jordan, with relevant differences between Bedouin and non-Bedouin neighbouring groups (Al-Abri et al. 2012; Ingram et al. 2009a). In the Indian subcontinent, however, a clinal pattern can be detected, declining from northwest to south-east (Gallego Romero et al. 2012; Ingram et al. 2009a; Itan et al. 2010).

Also in Europe, a similar cline with frequencies increasing from south to north can be observed. In the northern part of the continent almost all the sampled individuals are lactase persistent, with frequencies ranging between 96% and 83% in Finland, Denmark, Ireland and the United Kingdom; in the Mediterranean area, the opposite trend can be observed (Itan et al. 2010; Manco et al. 2013).
Lactase persistence and milk consumption in Europe: an interdisciplinary approach involving genetics and archaeology

Archaeology

Before exploring the origin and spread of lactase persistence in Europe in more detail, it is necessary to consider the cultural and archaeological context related to the beginning of milk consumption in Europe; this will be summarised in the next two sections.

The Mesolithic-Neolithic transition in Europe

The so-called Neolithic revolution marks the transition between a lifestyle based on hunting and gathering to one based on food production through cultivation and animal management and exploitation. The transition from the Palaeolithic-Mesolithic to the Neolithic entailed many different changes, not only related to material culture (e.g., the acquisition of pottery) and knowledge of new skills (e.g., those related to animal and plant domestication) but also to the social structure of the human groups (e.g., sedentism and a different redistribution of wealth).

The first Neolithic cultures seem to have developed around the Fertile Crescent some 12,000 years ago, and then to have spread to neighbouring regions, including Europe, during the following few thousand years (Flannery 1973). Two opposite models have been proposed to reconstruct the way in which Neolithic cultures developed and spread in Europe.

The demic diffusion model suggests that the Neolithic culture and lifestyle spread from the Near East into Europe through the migration of farmers/agriculturalists (and possibly domesticates) with no substantial admixture with local hunter-gatherer populations (Ammerman, Cavalli-Sforza 1984). On the other hand, the cultural diffusion model theorises a step-by-step process whereby local Mesolithic groups learnt new skills from neighbouring Neolithic populations, without a replacement of people (Zvelebil, Zvelebil 1988, 574–583). The more recent studies suggest that a more complex intermediate model involving a succession of migration phases interleaved by local admixture could be a better representation of what actually happened (Whittle, Cummings 2007). Genetic data have been used in order to differentiate between the two hypotheses, but the subject is still debated (Barbujani et al. 2012; Thomas et al. 2013).

The role of lactase persistence in the Neolithic revolution

Before the Neolithic, milk was available only during the first years of life (i.e. breastfeeding); only after the beginning of animal domestication did it become a possible source of nourishment for adults. Lactase persistence, even if already present in some individuals, would have been of no utility until the beginning of the human management of cattle, sheep and...
The origins of lactase persistence in Europe

The mutation associated with lactase persistence in Europe (-13 910*T) is relatively recent. Estimates based on the method of long-range haplotype conservation suggest its origin between 2188 and 20 650 years ago (Bersaglieri et al. 2004) while an analysis of variation in closely linked micro-satellites has dated it to a period between 7450 and 12 300 years ago (Coelho et al. 2005). It is interesting to consider that the dates estimated for the origin of one of the African variants (14 010°C) are similar, even if slightly more recent (Tishkoff et al. 2007).

Given the scenario presented, one of the most interesting questions that has been addressed during the last decade is whether the -13 910*T variant was already present in Europe in significant frequencies at the beginning of the Neolithic (e.g., due to random mutation and the effects of genetic drift) or if the spread of the mutation associated with lactase persistence was purely the result of selection acting after the beginning of animal domestication, when milk became available as a nutritional source throughout the entire life of the individuals. The advent and further development of technologies allowing the sequencing of DNA extracted from archaeological specimens has helped answer this question.

The first study to address this issue was performed on a set of samples from Central Europe (Germany, Hungary, Lithuania and Poland) dating from the Mesolithic (one individual) to the Early Neolithic period (eight individuals) (Burger et al. 2007). All of them were homozygotes for the ancestral allele, and therefore unable to digest fresh milk as adults. The authors performed statistical analyses that showed that the frequencies of the derived alleles must have been very low, if not nil, in order to have such a pattern in the sample.

More recently, several studies have been carried out in different regions of Europe clarifying this assertion (Burger, Thomas 2011; Linderholm 2011). In France, the mentioned SNP has been typed in 26 skeletons dated to the end of the Neolithic, before the beginning of the Bell-Beaker Culture, and all of these were homozygotes for the C variant (Lacan et al. 2011). In southern Scandinavia only one out of ten Middle Neolithic hunter-gatherers carried the derived allele, and was heterozygote (Malmström et al. 2009).

In contrast, in northern Spain, out of 26 Middle Neolithic individuals, seven (five of them homozygotes) carried the T allele (Plantinga et al. 2012). The authors tend to explain this difference in frequencies, compared to the other European samples, as the effect of genetic drift rather than the result of natural selection.

Additionally, some more recent samples, from the Middle Ages, have been analysed. A single individual
from Germany was lactase persistent, being heterozygous (Burger et al. 2007), and out of 23 samples from Hungary only three carried the derived haplotype (Nagy et al. 2011). It has nevertheless to be born in mind that for periods such as the Middle Ages high levels of gene flow and important migrations occurred in Europe, and this must be considered when interpreting such data (Nagy et al. 2011; Reich, Arnauld 2007).

Selection on lactase persistence

The above-mentioned data support a massive increase of LP frequency from virtually 0 to close to 100% in Northern Europe in a few thousand years, which is a very short period when considering the evolution of our species. This evidence suggests strong natural selection acting on this locus: the coefficient of selection has been estimated as reaching almost 0.2 in Northern European populations (Bersaglieri et al. 2004). Such a strong indication of selection poses the problem of understanding why drinking fresh milk had such a significant selective advantage.

It is important to consider that the intake of small quantities of fresh milk does not always cause unpleasant symptoms in non-persistent individuals, and that when milk is processed or fermented, such as in yogurt or cheese, the amount of lactose decreases and the consumption of these products does not cause symptoms to non-persistent individuals (Hammer et al. 1998). Moreover, as already discussed, gut flora adaptation can in some cases allow individuals to avoid the symptoms of lactose malabsorption (Szilagy et al. 2002).

The evidence of cheese production during the Early Neolithic mentioned above has a special meaning. It has already been discussed that lactase persistence was very rare or absent during the Neolithic, and this would not allow human groups to drink fresh milk. Processed milk, on the other hand, can be digested by non-persistent individuals without unpleasant symptoms. The production of cheese at the beginning of the Neolithic suggests that at that time human groups may have already recognised the high nutritional benefits available from milk, and, in order to digest it, they processed it. It can be presumed, therefore, that the reason lactase persistence spread and became so common in Europe is not to be found in the simple nutritional benefits of consuming fresh milk.

The observations above raise a question: why is lactase persistence so strongly selected if non-persistent individuals can consume milk by processing it, thereby gaining many of milk’s nutritional benefits? Several possible explanations have been proposed, but it is still under debate as whether any is convincing enough to justify levels of positive selection as high as those observed. As previously discussed, high frequencies of lactase persistence tend to correlate with a traditionally pastoralist lifestyle or with high consumption of fresh milk, even if this is not always the case (Ingram et al. 2009b). This evidence could be the result of two quite different scenarios.

The reverse-cause argument (McCracken 1971): after its first appearance, a mutation associated to lactase persistence could have grown in frequency within one or several small human groups only because of genetic drift. Milk drinking could then have been adopted since the group would have been able to tolerate it.

Gene culture co-evolution (McCracken 1971. 497–517, Simoons 1970.695–710): lactase persistence could have been positively selected in dairying populations, which had access to fresh milk throughout life.

As already noted in the previous section, ancient DNA data support the second hypothesis. During the Neolithic, when dairying practices were already established, lactase persistence appears to be virtually absent, or present in very low frequencies in Europe. It is also important to consider that the random independent origin and then the increase of the frequencies of different lactase persistence-associated alleles due to the random fluctuation (genetic drift) in different regions would be a very complex and unlikely scenario, for which no possible explanation is available at the moment.

The calcium assimilation hypothesis (Flatz, Rotthauwe 1973): exposure to sunlight allows human skin to produce vitamin D, which is an element of great importance for the absorption of calcium in the bones. When vitamin D, taken in through a diet rich in fish or produced thanks to exposure to the sun, is not sufficient, bone development can be seriously compromised. In addition to other nutrients, milk contains small amounts of vitamin D and large amounts of calcium. The frequency of lactase persistence in Europe correlates with latitude and sunlight, and this evidence suggests a cause-effect relationship between the two. In Northern Europe, it is possible to observe a transition from a diet rich in fish and marine organisms (rich in vitamin D) during
the Mesolithic to one based mainly on cereals (poor in vitamin D) in the Late Neolithic (Eriksson et al. 2008; Lidén, Eriksson 2007; Lidén et al. 2004). The Meso-Neolithic transition has been shown by many scholars to be associated with a deterioration in health (Cohen 2008; Eriksson et al. 2008; Eshed et al. 2010; Hershkovitz, Gopher 2008; Larsen 1995). Drinking milk would have helped populations living where sunlight is low to avoid rickets and similar problems related to such a nutritional change.

**Adaptation to arid environments** (Cook, al-Forki 1975.135–136): in a normal situation, raw milk and cheese could provide the same nutritional benefits, but in arid environments milk could be an uncontaminated source of fluid, while lactose intolerance related symptoms (mainly diarrhoea) could lead to dehydration and, eventually, death. Nevertheless, this hypothesis is not easily applicable to Europe, where the climate is temperate.

A statistical test was performed to see whether lactase persistence was more likely to be correlated with dairying practice, calcium assimilation or adaptation to arid environments (Holden, Mace 1997. 605–628). The best correlation was obtained between the ability to digest milk and pastoralism. This study suggests, in accordance with the palaeogenetic data, that pastoralism was more probably adopted before lactase persistence arose or became frequent.

**The spread of lactase persistence through Europe**

The pattern of genetic diversity that can be observed in a population is the result of many different biological and social processes that contribute to shaping it during its history: random mutation, natural selection, genetic drift (and, consequently, demography), social factors influencing mating choices (e.g., social structuring) etc.

Computer simulations are very powerful instruments for testing various hypotheses through models with different degrees of complexity. In this way, the most important factors playing a role in shaping the diversity observed can be explored both in a computationally effective way, removing the confounding effect of minor elements, and within a realistic framework integrating all the information that appears to be relevant to the problem. Computer simulations have been applied to investigate the evolution of LP since the late 1980s (Aoki 1986). As already mentioned, lactase persistence is not entirely correlated with pastoralism and milk drinking, since some individuals (or human groups) are able to drink milk without being lactase persistent. Aoki tried to test if this could be the result of a process of gene-culture co-evolution and, by using computer simulations, he confirmed that this is the case. Moreover, his study showed that the incomplete correlation could be easily linked to the stochastic nature of the process.

Two recent studies have reached different results when simulating the spread of lactase persistence in Europe. As already discussed, it is still debated whether the beginning of the Neolithic in Europe was linked to the migration of people from the Middle East or to simple cultural transmission of skills and techniques from agriculturalist populations to neighbouring hunter-gatherers. The colonisation and expansion in Europe by farmer populations could have had an important impact on the diffusion of the lactase persistence-associated allele, since some demographic processes could mimic the effects of selection (Klopfstein et al. 2006). The impact of demographic effects and differential selection based on latitude have been taken into account in a recent study simulating the spread of lactase persistence in Europe (Gerbaulet et al. 2009). The authors tested two different scenarios for the spread of the Neolithic in Europe (the demic and cultural diffusion models). On the basis of the dates for the beginning of the Neolithic in each region, they simulated the evolution of the frequencies of the lactase persistence-associated allele through time. Selection was incorporated in three different ways: constant throughout the continent, increasing towards the north (to test the calcium assimilation hypothesis) or higher in central European Early Neolithic (Linearband Ceramic) populations.

The results show that the present-day frequencies of milk digesters in southern Europe could be due to genetic drift linked to the arrival of Neolithic farmers from the Near East, but selection is required to reproduce the modern frequencies observable in the northern part of the continent. The authors then support the demic diffusion model associated with the calcium assimilation hypothesis.

Spatially explicit simulations were also applied to the same subject (Itan et al. 2009). After creating a geographical background as close to Europe as possible, they modelled the evolution of the lactase persistence associated variant in three humans groups: hunter-gatherers (already present in the continent prior to the beginning of the simulations, 9000 years
ago), dairying farmers, and non-dairying farmers. Food producers were allowed to reach a higher population density, and they spread from the Middle East towards Europe. Gene flow between groups, long-distance migrations and the cultural diffusion of subsistence practice were also included in the model. Positive selection acted only on dairying farmers. The best simulations were chosen on the basis of the fit with the arrival time of agriculture and modern frequencies of lactase persistence in 12 different locations. The method used to analyse the results allowed the estimation of migration rates, selection coefficients and time and geographical coordinates for the beginning of selection on lactase persistence; the best fit was reached when selection started in a region between the Balkans and Central Europe, from 6256 to 8683 years BP, and differential selection in Northern Europe was not necessary to reach modern frequencies. The times and regions mentioned are in great agreement with the development of the Linearband Ceramic culture (Pavňák 2005).

An analysis of the differences between the two mentioned studies can be found in Leonardi et al. (2012). The simulation model used in Gerbault et al. (2009) is less complex than that used in Itan et al. (2009) in several ways. The model from the former that can be better compared with the latter is the so-called LBK scenario, where selection is higher only for the descendants of LBK groups. In this model, the frequencies of lactase persistence in Northern European populations do not reach present-day values, showing that higher levels of selection only in LBK-related populations is not enough to reconstruct in a satisfactory way the modern distribution of lactase persistence in Europe. Since gene flow and long-distance migrations between populations are not explicitly modelled, this result does not contradicts the findings of Itan et al. (2009) but suggests that selection was not constant through time and space, a hypothesis that has not been rejected by the latter study, where the selection coefficient was constant.

Conclusion

Lactase persistence is an amazing example of gene-culture co-evolution. The modern-day pattern of diversity at this locus in Europe is the result of the interaction of a large number of factors such as physiology, genetics, demography, migrations of people, social structuring, and cultural contact. The study of such a complex subject must start from a very careful analysis of the archaeological, historical and biological contexts and should be performed using methods that allow as far as possible integrations of the different types of information available. During the last decade, the increase of computational power coupled with more effective sequencing techniques has led to the possibility to simulate in silico more models, and to reconstruct with better precision the history of our and other species. A better understanding of the origin and spread of lactase persistence in Europe will definitely pass along this path.

ACKNOWLEDGEMENTS

This work was supported by the LeCHE EU Marie Curie FP7 Framework Programme grant (grant ref. 215362–2). The author thanks Joachim Burger and Mark Thomas for scientific support; Pascale Gerbault and Jonathan Cole for their useful discussion of the manuscript.

References


Arenas M., Francois O., Currat M., Ray N. and Excoffier L. 2013. Influence of admixture and paleolithic range con-


