What makes us human?

Lactase persistence: a case of evolution in modern humans

Dallas M. Swallow
(University College London, UK)

We modern humans may be able to control our environment and protect ourselves from the adverse consequences of disease, but our evolution never stops. The selective pressures that affect the numbers of viable offspring carrying novel DNA changes alter in time and space. Changes in lifestyle in the last 10,000 years have led to major changes in diet, making new dietary components potential selective agents. One of the best examples of this comes from the genetically determined variation in our ability to digest the lactose in milk.

Baby mammals suckle milk from their mothers as their first source of food. The major carbohydrate component in most milk is lactose, a disaccharide that has to be digested into its constituent monosaccharides, glucose and galactose, by the small intestinal enzyme lactase. These monosaccharides can then be transported across the epithelial brush border membranes (by a glucose and galactose transporter) and on into the bloodstream. Lactase activity is therefore essential for the development of young mammals that depend on lactose-containing milk.

In such mammals, including most humans, lactase expression decreases after the weaning period is over, when it is no longer needed, since lactose is not found elsewhere in the diet. In approximately 45% of humans, however, lactase activity persists into adult life (so-called lactase persistence or LP). As a result, these people can digest lactose as adults, and can thus readily use fresh milk from other mammalian species as a food source. Humans milk cows, sheep, goats, horses, camels and other animals (but not pigs), which all have similar concentrations of lactose in their milk. In northern Europe, LP is the most prevalent phenotype, leading mid-20th Century investigators to think it was the normal human situation even though adult milk intolerance had been recognized by early doctors such as Galen in the Second Century AD.
When undigested lactose reaches the colon, it is fermented by colonic bacteria that release gases including hydrogen. As a result, the passage of lactose into the colon in non-lactase producers can lead to gastrointestinal symptoms, such as bloating, flatulence and abdominal pain, as well as diarrhoea from the osmotic effect of the undigested lactose. These symptoms are described as ‘lactose intolerance’. Lactose intolerance should be distinguished from milk allergy which involves an immune response, usually to milk protein.

The age at which lactase is down-regulated in lactase non-persistent people varies from one population to another; it has been reported that most Chinese and Japanese people lose their lactase at between 1 and 5 years of age, whereas in non-persistent Finns, this does not occur until somewhat later (usually at 5–10 years of age). Even though the mechanisms of developmental lactase down-regulation are not well understood, it is clear that it is not reversible. This contrasts with secondary lactose intolerance, due to intestinal damage or inflammation, or diarrhoea, which is usually reversed after the causal problem is removed.

In this article, I summarize what is known about the molecular and population genetics and evolution of the LP trait, but further details and references can be found elsewhere.

Genetically determined lactase persistence and non-persistence

Lactase persistence (LP) is inherited in an autosomal dominant manner, meaning heterozygous carriers can digest large quantities of lactose. The frequency of the trait differs worldwide; LP is particularly common in northern Europe, with frequencies of ~89–96% in the British Isles and southern Scandinavia, and a declining gradient towards the south and east in the rest of Europe. Outside Europe, there are general north-east to south-west gradients across the African subcontinent and a north to south gradient across India. However, LP is not evenly distributed within particular geographic regions. Indeed, in Africa and the Middle East, LP is often found at very different frequencies in neighbouring populations, such as 64% in the Beni Amir (pastoralists) and 23% in the Dounglawi (non-pastoralists) in Sudan. LP frequency has been shown to correlate strongly with a tradition of milk drinking and pastoralism.

A single gene (LCT) codes for lactase, and the lactase enzyme itself is functionally indistinguishable in most people, and between children and adults.

However, several single nucleotide changes have been found in a cis-acting enhancer located 13–14 kb upstream of the start codon, in an intron of the adjacent gene MCM6. This enhancer function has been demonstrated in vitro (Figure 1) and studies show that several different nucleotide changes can increase transcription and alter transcription factor binding in vitro. From this, it is inferred that in adult life these nucleotide changes affect RNA transcription in vivo and thus how much enzyme is made. The first of these variants to be found, a single nucleotide transition at −13910 (−13910C>T), occurs at high frequency in Europe, whereas others, which occur within 100 nucleotides of each other (Figure 2), were first found in Africa and on the Arabian peninsula (Figure 3).
Natural selection and evolution of lactase persistence in humans

There is strong molecular and population genetic evidence to suggest that the lactase gene region is unusual. The age estimates for the regulatory sequence variants range between 2000 and 21 000 years ago for the −13910*T allele and between 1000 and 23 000 years ago for the −14010*C allele, one of the major alleles in Africa. Interestingly, these date estimates bracket those for the domestication of milkable animals and the spread of agriculture and herding obtained from archaeological data (cave paintings, distributions of animal bones and dairy fat residues in pots). A low frequency or absence of −13910*T variation in early Neolithic central European farmers and early Neolithic farmers from north-east Iberia, middle Neolithic Scandinavian hunter-gatherers and late Neolithic farmers from southern France suggests that dairying was practised before LP arose or became common, confirming the young age of the variant alleles (see Gerbault et al.5 for references).

All studies so far show that the chromosomal region carrying the European −13910*T and other LP alleles is notably lacking variability in terms of single nucleotide and simple repeat (microsatellite) differences across different individuals over a long genomic distance. The existence of this extended stretch of DNA, or haplotype, together with the frequency reached over a very short time period is good evidence for rather strong positive selection.

Thus several lines of evidence (genetics, anthropology and archaeology) suggest that, once there was a supply of lactose-containing dairy products, LP provided a selective advantage. This implies that these traits evolved as the result of a co-evolutionary process involving both genes and culture. The strength of natural selection estimated for the LP-associated alleles is very high at 5–16% for −13910*T, and 1–15% for −14010*C, among the highest estimates for any human genes in the last 30 000 years.

Computer simulation studies of this gene–culture co-evolutionary process in Europe together with archaeological (pottery) data indicate that the expansion of LP and dairying began between 6000 and 9000 years ago, and that the European allele might have originated in a region around modern-day Hungary6. The other alleles are likely to have arisen in Africa or the Middle East, but as yet this is based only on geographic distribution rather than simulation studies.
In the east of Africa, several different persistence alleles have been selected in parallel, resulting in a very different molecular genetic pattern of diversity, namely more diversity in the LP people, but providing further evidence of the advantageous effect of the trait\textsuperscript{1,2}. Our results also indicate that multiple subtly different mechanisms involving differential transcription factor binding can lead to the same result: that is continued expression of lactase into adult life\textsuperscript{1,2}. The nature of the selective advantage is, however, less clear, but suggestions include the benefit of easy consumption of ready calories in times of shortage of other foodstuffs, clean water in times of drought, and a source of calcium and vitamin D in areas of low sunlight. It is likely that the relative importance of each of these was different in different geographic regions, being influenced by the occurrence of famine, drought and diarrhoeal disease, and that the spread has also been strongly influenced by demographic effects such as expanding population size and patterns of migration. There are suggestions in some studies that lactase-persistent people are of larger stature than non-lactose-persistent people, but it is very difficult to exclude the confounding effects of other genetic heterogeneity in such studies, just as it is in studies which have attempted to examine LP and milk drinking in relation to bone fractures, or even disadvantageous effects such as cataracts or even heart disease. If lactase-persistent people are indeed bigger, presumably as a result of drinking more milk, how does this relate to the numbers of viable offspring?

The other difficult feature to come to terms with is that, in parallel with this genetic adaptation, there has also been cultural adaptation. Fermentation of milk, which occurs naturally as milk sours, is accompanied by a reduction in lactose concentration, and humans have exploited the souring process to make products such as cheese, some of which contain only traces of lactose. This cultural practice allows consumption of milk products with the retention of most of the nutrients and calories. So why should the selection pressure for LP have been so strong? Perhaps the key is in the whey, either for its water content or some other component?

To add further complexity, the microbial community within the human gastrointestinal tract can be altered by diet, and both a change of the composition of the microbiota and an increase in faecal β-galactosidase activity have been observed after daily milk feeding in association with a reduction in lactose intolerance symptoms. There are also inter-individual differences and changes in the production of gases (such as hydrogen), which cause much of the discomfort in lactose intolerance. Indeed milk and lactose-containing products can often be consumed without provoking symptoms of intolerance. This all shows that dietary adaptation unrelated to lactase expression can allow significant milk consumption in some cases, as observed by us in some milk-drinking Somali camel herders.

Despite these complexities, the evidence for recent genetic adaptation is overwhelming, thus providing a clear example of how multiple regulatory mutations can arise: regulatory sequences being much more evolutionarily malleable than coding sequences.

Dallas Swallow is an Emeritus Professor at University College London who has had a long interest in functional genetic variation in the context of both health and disease, starting by studying the biochemical genetics of enzyme and other protein variation. Exploiting, first, monoclonal antibodies, then gene cloning, and other molecular techniques as they developed, she has contributed over many years (and is still doing so) to the understanding of lactase persistence, at both the molecular and population levels.

email: d.swallow@ucl.ac.uk.