# Escape from epigenetic silencing of lactase gene expression triggered by a single nucleotide change

Dallas M. Swallow and Jesper T. Troelsen

The importance of subtle gene regulation and epigenetics in determining complex human traits is increasingly being recognized. However, bridging the gap between environmental, epigenetic and genetic influences and unravelling causal relationships remains a big challenge. Labrie et al. now report an example of epigenetic changes influenced by genetic factors that are involved in the regulation of lactase gene expression.

Intestinal lactase is required for the digestion of lactose, the major carbohydrate in mammalian milk, but lactase gene expression is usually down-regulated after weaning. In many humans, however, lactase persists into adult life allowing digestion of milk, an evolutionary adaptation to dietary change. Reduced intestinal lactase levels, whether due to developmental down-regulation or secondary loss, due for example to enteritis, results in incomplete digestion of lactose, which can lead to symptoms of diarrhea and flatulence when fresh milk is consumed. However, people who are lactase non-persistent (LNP) do not necessarily suffer from symptomatic lactose intolerance, as there are large inter-individual differences due to variables such as the gut microbiome composition and gut transit time, as well as to a person's sensitivity to the symptoms.

While lactase persistence (LP) is clearly not a disease, a study in this issue by Labrie *et al.*[1] that reports a genetic and epigenetic interplay in the regulation of lactase expression provides us with a curiously good example of what we may hope to find in the future for complex human diseases, such as diabetes, heart disease, multiple sclerosis and schizophrenia, for which genome-wide association studies are not yet able to explain all of the inherited component[2].

There is good evidence that the persistence of high intestinal lactase activity into adult life is attributable to regulation of transcription of the lactase gene (*LCT*). This is a primarily heritable trait that varies dramatically in prevalence in different parts of the world (Box 1). In Europeans, a single nucleotide substitution (-13910 C>T) within intron 13 of *MCM6*, the gene adjacent to *LCT*, is strongly associated with lactase persistence. The C>T substitution has been shown experimentally to reside within an enhancer for *LCT* and induces increased promoter activity[3-6]. In Middle Eastern and African populations four other clearly functional alleles within the same enhancer have so far been identified (see for example [7]) (Box 1). However, a key remaining question is how these SNPs drive allele specific expression in adults, while all haplotypes are expressed at high levels in babies.

Labrie *et al.* use human intestinal samples from Europeans, and also conduct experiments in mice. In humans, having confirmed the genetically controlled inter-individual transcriptional differences in *LCT* expression, they find seven regulatory regions, in which changes in DNA methylation are correlated with *LCT* mRNA levels. DNAase hypersensitivity and histone modification sites (H3K27ac, H3K4me1, H3K4me3) derived from Epigenome Roadmap data overlap these same seven sites. In contrast, in white

blood cells, which do not express lactase, these regions are fully methylated.

The authors used CRISPR-Cas9-induced genetic deletions in human tissue culture and in the mouse to validate the regulatory function of these epigenetically-modified elements. They also mapped the promoter of a long non-coding RNA (lncRNA), LOC100507600, to the intron 5 region and show that it regulates lactase expression by using RNA interference (RNAi).

The seven regions identified also include the *LCT* enhancer located within intron 13 of *MCM6*, which houses the functional DNA variants. Indeed, one of the most powerful observations in the current study relates to inter-individual differences in DNA methylation at this enhancer region. These show the greatest correlation with genotype for the known functional European LP SNP as well as SNPs that form part of the LP haplotype. The strongly stratified difference in modification of this enhancer region provides, in our view, additional evidence for the functional role of this enhancer in LP.

The second very important observation is that even in adults (the only samples available) a significant age-related change in DNA methylation can be seen, providing for the first time a mechanism for age-related difference in lactase gene expression. Apart from the lncRNA region, there tends to be an increase in DNA modification with age for the ancestral haplotype and a non-significant decrease in the case of LP haplotypes. Overall, modifications at *LCT* and *MCM6* regulatory elements in enterocytes remain relatively low except in the *MCM6* regions of the ancestral haplotype, where in intron 13 DNA methylation is already at about 80% even in young adults.

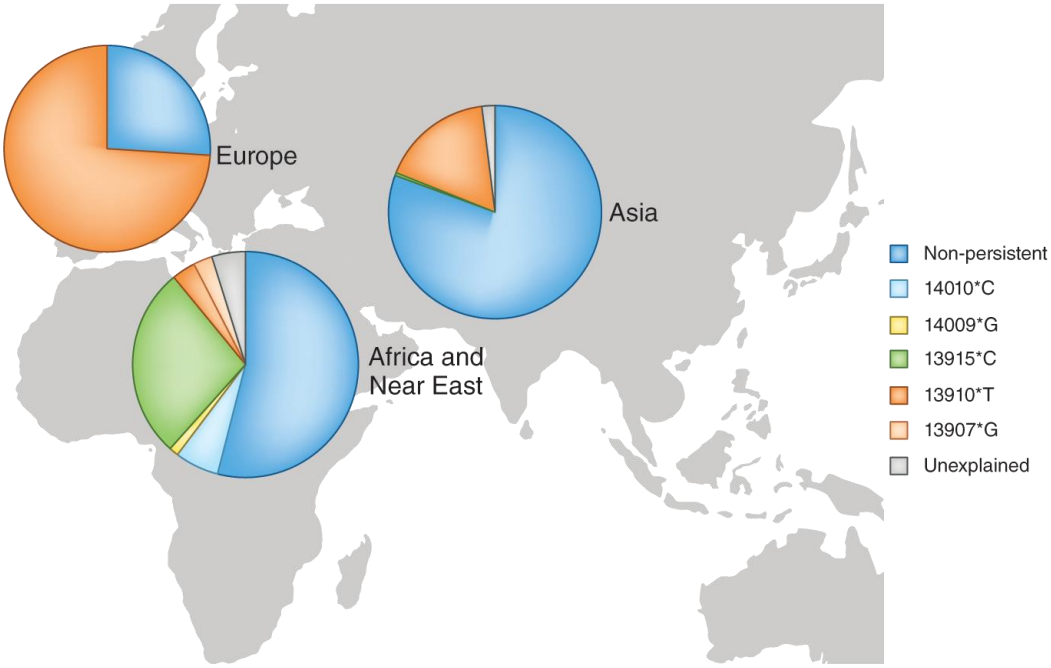
In summary, the study provides functional evidence for several regulatory sites for *LCT* expression and reveals that lactase down-regulation is correlated with accumulation of transcriptionally suppressive epigenetic changes on haplotypes carrying the *C-13910* allele, while *T-13910* containing haplotypes escape from epigenetic inactivation, facilitating lactase persistence (Fig. 1). In the future, it will be intriguing to examine DNA methylation profiles of fetuses, where there is low lactase expression from both LP and non-LP haplotypes, and of children, where expression from both LP and non LP haplotypes is high[8]. Despite the current lack of these data we now begin to get insight into how lactase down-regulation is effected and how non-genetic factors might occasionally tip the pattern of expression, such that more or less lactase may be expressed than predicted from genotype.

The current study provides a strong rationale for further investigating how an escape from inactivation is brought about. It should in this context be noted that in Europeans the only nucleotide difference from the ancestral sequence within the *LCT* enhancer is 13910 C>T. However, in other populations, individuals who carry different functional LP alleles show high expression of lactase and good lactose digestion, [9-11] despite being homozygous CC at -13910, suggesting that the other enhancer SNPs presumably also allow escape from epigenetic down-regulation in a *cis*-acting manner. *In vitro* effects observed for the various functional enhancer SNPs, which include differential binding of transcription factors, suggest that differences in the relevant binding sites in the enhancer might in turn lead to histone and DNA modification. Differential epigenetic changes along the crypt-villus axis associated with transition to cellular differentiation[12] and altered coordination of changes in transcription factor expression may play a part in this. As in all genetic/epigenetic studies it is difficult to determine the causal direction of the effects

observed, but this report describes a very important step forward and emphasizes how age-related epigenetic changes may interact with common DNA variants that are benign in childhood, but associated with disease in the adult.
The authors declare no conflict of interest

*Box 1 Lactase persistence distribution assessed by the frequency of known functional alleles and assumed dominance of the trait.*

The pale blue segment shows the proportion of lactase non-persistence, while the other colours show LP accounted for by the different functional alleles described so far (see for example[10,11,13,14] and references therein). The dark orange segments show lactase persistence attributable to the allele/haplotype studied by Labrie *et al.* The pale grey segment is a very rough estimate of the 'causal alleles' that may be missing, assessed by genotype and phenotype frequency comparisons[15]. Note that in most studies lactase persistence is inferred indirectly from lactose tolerance tests—the physiological response to a lactose load after an overnight fast. While lactose digestion tests are not 100% accurate, there is good indirect evidence that a few carefully tested people express high lactase levels as adults although no candidate functional DNA variant has been found.



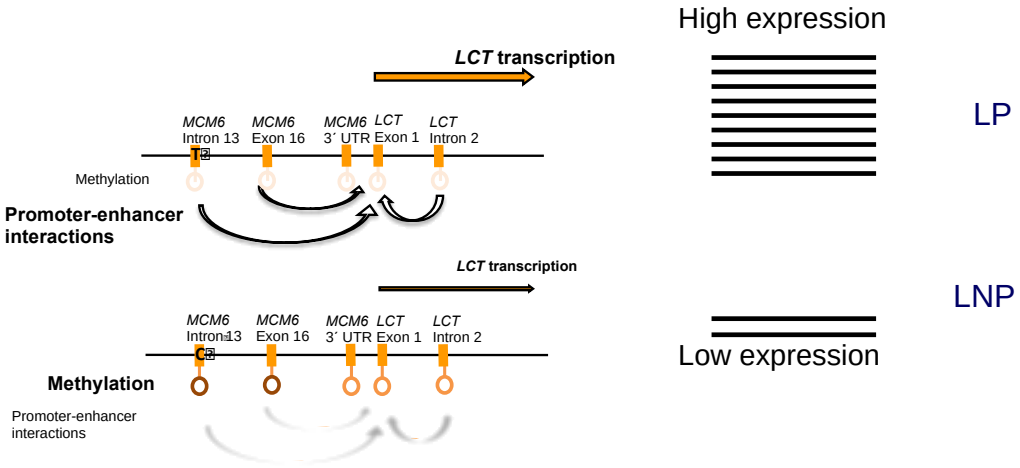



*Figure 1 Inverse relationship between MCM6/LCT enhancer methylation and lactase gene expression.* Putative interactions between the enhancers and promoter are shown as arrows. Shaded lollipops represent levels of DNA methylation of regulatory regions identified by Labrie *et al* . LP denotes lactase persistence haplotype and LNP denotes lactase non-persistence haplotype.

## References


1. Labrie, V.e.a. *Nat Struct Mol Biol* **23**(2016).
2. Tak, Y.G. & Farnham, P.J. Making sense of GWAS: using epigenomics and genome engineering to understand the functional relevance of SNPs in non-coding regions of the human genome. *Epigenetics Chromatin* **8**, 57 (2015).
3. Wang, Z., Maravelias, C. & Sibley, E. Lactase gene promoter fragments mediate differential spatial and temporal expression patterns in transgenic mice. *DNA Cell Biol* **25**, 215-22 (2006).
4. Troelsen, J.T., Olsen, J., Moller, J. & Sjostrom, H. An upstream polymorphism associated with lactase persistence has increased enhancer activity. *Gastroenterology* **125**, 1686-94 (2003).
5. Olds, L.C. & Sibley, E. Lactase persistence DNA variant enhances lactase promoter activity in vitro: functional role as a cis regulatory element. *Hum Mol Genet* **12**, 2333-40 (2003).
6. Lewinsky, R.H. et al. 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 (2005).

7. Jensen, T.G. et al. The -14010*C variant associated with lactase persistence is located between an Oct-1 and HNF1alpha binding site and increases lactase promoter activity. *Hum Genet* **130**, 483-93 (2011).
8. Wang, Y. et al. The genetically programmed down-regulation of lactase in children. *Gastroenterology* **114**, 1230-6 (1998).
9. Imtiaz, F. et al. 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 (2007).
10. Jones, B.L. et al. Diversity of lactase persistence in African milk drinkers. *Hum Genet* **134**, 917-25 (2015).
11. Tishkoff, S.A. et al. Convergent adaptation of human lactase persistence in Africa and Europe. *Nat Genet* **39**, 31-40 (2007).
12. Sheaffer, K.L. et al. DNA methylation is required for the control of stem cell differentiation in the small intestine. *Genes Dev* **28**, 652-64 (2014).
13. Ranciaro, A. et al. Genetic origins of lactase persistence and the spread of pastoralism in Africa. *Am J Hum Genet* **94**, 496-510 (2014).
14. Liebert, A. Evolutionary and molecular genetics of regulatory alleles responsible for lactase persistence. *Ph.D Thesis* **University College London**(2014).
15. Itan, Y., Jones, B.L., Ingram, C.J., Swallow, D.M. & Thomas, M.G. A worldwide correlation of lactase persistence phenotype and genotypes. *BMC Evol Biol* **10**, 36 (2010).