

Evolution of lactase persistence: an example of human niche construction

Pascale Gerbault, Anke Liebert, Yuval Itan, Adam Powell, Mathias Currat, Joachim Burger, Dallas M. Swallow and Mark G. Thomas

Phil. Trans. R. Soc. B 2011 **366**, doi: 10.1098/rstb.2010.0268, published 14 February 2011

References

[This article cites 92 articles, 34 of which can be accessed free](#)

<http://rstb.royalsocietypublishing.org/content/366/1566/863.full.html#ref-list-1>

[Article cited in:](#)

<http://rstb.royalsocietypublishing.org/content/366/1566/863.full.html#related-urls>

Subject collections

Articles on similar topics can be found in the following collections

[ecology](#) (543 articles)

[evolution](#) (715 articles)

[molecular biology](#) (189 articles)

Email alerting service

Receive free email alerts when new articles cite this article - sign up in the box at the top right-hand corner of the article or click [here](#)

Research

Evolution of lactase persistence: an example of human niche construction

Pascale Gerbault^{1,*}, Anke Liebert¹, Yuval Itan^{1,3}, Adam Powell^{1,4},
Mathias Currat⁵, Joachim Burger², Dallas M. Swallow¹
and Mark G. Thomas^{1,3,4,6}

¹Research Department of Genetics, Evolution and Environment, University College London, Wolfson House, 4 Stephenson Way, London NW1 2HE, UK

²Institute of Anthropology, Johannes Gutenberg University, AG Palaeogenetik, SBII-2 Stock-Raum 02-333, Colonel Kleinmann-Weg 2, 55128 Mainz, Germany

³CoMPLEX (Centre for Mathematics and Physics in the Life Sciences and Experimental Biology), University College London, Physics Building, Gower Street, London WC1E 6BT, UK

⁴AHRC Centre for the Evolution of Cultural Diversity, Institute of Archaeology, University College London, 31–34 Gordon Square, London WC1H 0PY, UK

⁵Laboratory of Anthropology, Genetics and Peopling History (AGP), Department of Anthropology and Ecology, University of Geneva, 12 rue Gustave-Revilliod, 1227 Geneva, Switzerland

⁶Evolutionary Biology Centre, Department of Evolutionary Biology, Uppsala University, Norbyvägen 18D, 752 36 Uppsala, Sweden

Niche construction is the process by which organisms construct important components of their local environment in ways that introduce novel selection pressures. Lactase persistence is one of the clearest examples of niche construction in humans. Lactase is the enzyme responsible for the digestion of the milk sugar lactose and its production decreases after the weaning phase in most mammals, including most humans. Some humans, however, continue to produce lactase throughout adulthood, a trait known as lactase persistence. In European populations, a single mutation (-13910^*T) explains the distribution of the phenotype, whereas several mutations are associated with it in Africa and the Middle East. Current estimates for the age of lactase persistence-associated alleles bracket those for the origins of animal domestication and the culturally transmitted practice of dairying. We report new data on the distribution of -13910^*T and summarize genetic studies on the diversity of lactase persistence worldwide. We review relevant archaeological data and describe three simulation studies that have shed light on the evolution of this trait in Europe. These studies illustrate how genetic and archaeological information can be integrated to bring new insights to the origins and spread of lactase persistence. Finally, we discuss possible improvements to these models.

Keywords: lactase persistence; niche construction; Neolithic; domestic animals; dairying; natural selection

1. THE BIOLOGY OF LACTASE PERSISTENCE

(a) *Niche construction, lactase persistence phenotype and genotypes*

In biological evolution, natural selection acts on traits that are heritable. The vast majority of these traits are inherited by the transmission of DNA sequences from parent to offspring, i.e. genetic inheritance. But other aspects of the biology of an organism can be inherited extra-genetically, such as certain culturally transmitted behaviours [1,2] and features of the environment that have been shaped by ancestral populations [3]. Since this kind of extra-genetic inheritance can play a key role in the survival of an organism and the

evolutionary trajectory of a species, it has been assigned the term ‘niche construction’ [4,5]. A deeper understanding of the relationship between genetic evolution and niche construction can come from evolutionary theory, notably by recognizing that humans are far from unique in their ability to change their own selective environments [6]. However, because human culture has strongly modified our environments with such remarkable ecological and evolutionary consequences [2], human gene–culture coevolution provides some of the clearest and most spectacular examples of niche construction.

Gene–culture coevolutionary theory integrates cultural variation in the analysis of differential transmission of genes from one generation to the next [7]. This theory can be explicitly used to explore the evolutionary consequences of niche construction when a stable transmission of learned information is conveyed

* Author for correspondence (p.gerbault@ucl.ac.uk).

One contribution of 13 to a Theme Issue ‘Human niche construction’.

between successive generations [8,9]. Indeed, cultural processes can change the human selective environment and thereby affect which genotypes survive and reproduce [1,2]. If the cultural inheritance of an environment-modifying human activity persists for long enough to generate a stable selection pressure, it will be able to co-direct human evolution. There are many examples of this in human evolution [1–3,10] but none are so well studied, clear-cut, widespread and well supported as the coevolution of lactase persistence (LP) and dairying [4,11–13].

Lactose is the main carbohydrate in milk and is a major energy source for most young mammals. The enzyme responsible for hydrolysis of lactose into glucose and galactose is lactase (or lactase-phlorizin-hydrolase, LPH). Without this enzyme, mammals are unable to break down and thus use lactose, and since milk is the essential component of young mammals' diet, lactase activity is fundamental to the early development of most mammals. After the weaning period is over, lactase production usually declines, although the mechanisms and evolutionary reasons for this downregulation are not fully understood. However, some humans continue to express lactase throughout adult life, and are thus able to digest the lactose found in fresh milk. This trait is called LP.

The LP trait frequency is found in around 35 per cent of adults living in the world today [14,15], but varies widely among human populations, both between and within continents (figure 1*a*). High frequencies of LP are generally observed in northern European populations. Indeed, LP frequency can vary from 15–54% in eastern and southern Europe to 62–86% in central and western Europe, and to as high as 89–96% in the British Isles and Scandinavia [15,16]. In India, LP frequency is higher in the north (63%) than further south (23%) or east [17]. There are relatively little data on East Asians but it seems that the trait is rare there. In Africa, the distribution of LP is very patchy, with high frequencies being observed mainly in traditionally pastoralist populations [14,18–20]. For example, LP reaches 64 per cent in Beni Amir pastoralists (Sudan), whereas in Dounglawi (Sudan), a neighbouring non-pastoralist population, LP frequency is around 20 per cent [11,21,22].

In recent years, a number of single nucleotide polymorphisms (SNPs) have been found in association with the LP trait in different populations. The first to be identified, -13910^*T , is found not in the *LCT* gene (the lactase gene) but within an intron of a neighbouring gene, *MCM6* [23]. *In vitro* studies indicate that this nucleotide change affects lactase promoter activity [24,25] and thus is highly likely to cause LP, although it is currently not possible to exclude tight linkage disequilibrium with another, as yet unobserved, functional variant. An interpolated map showing the global distribution of -13910^*T can be seen in figure 1*b*, using published data [14,15,26–30] supplemented with recently collected data. A cursory comparison of figure 1*a,b* shows that while the -13910^*T allele may explain the distribution of LP in Europe, it cannot explain the distribution of LP in Africa or the Middle East. Indeed, Mulcare *et al.* [31]

showed this formally using a robust statistical framework, and other studies have identified additional LP-associated alleles that explain much of the distribution of LP in Africa [16,20,32]. Interestingly, all these variants are located within 100 nucleotides of -13910^*T in the same intron of the *MCM6* gene, a region that is functionally important for the expression of lactase *in vitro* [24,33]. Because these various LP-associated alleles are found on several different haplotypic backgrounds [14,20,34], it is now clear that LP has evolved multiple times and is thus an example of convergent evolution [35].

Using genetic variation in regions surrounding *LCT*, it is possible to obtain estimates of the age of specific LP-associated alleles. Dates of origin for -13910^*T ranging between 2188 and 20 650 years ago [36], and between 7450 and 12 300 years ago [37] have been obtained using extended haplotype homozygosity (EHH) and variation at closely linked microsatellites, respectively. Similar dates (1200–23 200 years old) were also obtained for one of the major African variants (-14010^*C) using EHH [20]. These date estimates are remarkably recent for alleles that are found at such high frequencies in multiple populations. It is easy to envisage recent alleles being rare since they change in frequency slowly, and in a directionless way, by genetic drift. However, a recent allele that has reached such high population frequencies requires more than genetic drift alone; it requires the extra 'kick' of natural selection. Indeed, the estimated selection strengths required to explain the age/frequency distributions of -13910^*T [36]—and of -14010^*C [20]—are enormous (1.4–19 and 1–15%, respectively), which are among the highest estimated for any human genes in the last approximately 30 000 years [36,38].

(b) *Selection hypotheses on lactase persistence*

The reasons why LP should provide such a selective advantage are still open to debate (see discussion below). However, since this trait has been mainly identified in dairying-practising or pastoralist populations [11] and since fresh milk and some milk products are the only known naturally occurring sources of lactose, it is unlikely that LP would be selected without a supply of fresh milk. Interestingly, the date estimates for the emergence of -13910^*T and -14010^*C bracket archaeological dates for the spread of domestic animals and dairying into Europe and the spread of pastoralism in Africa (south of the Sahara, into Kenya and northern Tanzania), respectively [11,20,39–42]. This supports the idea that LP coevolved with the cultural adaptation of dairying as a gene–culture coevolution process.

Nonetheless, the correlation between LP and milk consumption is not complete [11,16,31,32]. In lactase non-persistent individuals, the fermentation by colonic bacteria and osmotic effects of undigested lactose often cause symptoms such as abdominal pain, bloating, flatulence and diarrhoea. However, it has been shown that some lactase non-persistent individuals can consume lactose-containing products without any obvious ill effects. For example, the low LP

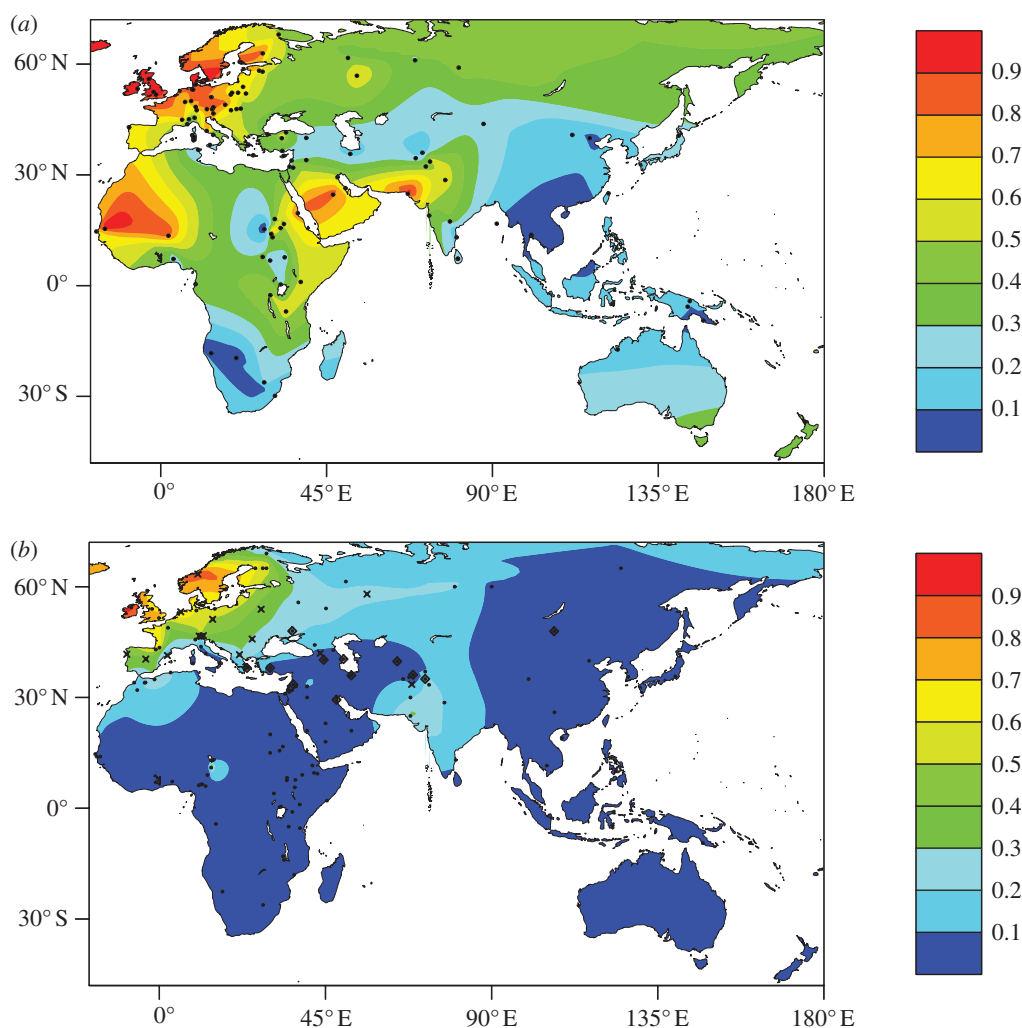


Figure 1. Interpolated maps of the distribution of LP and the -13910^*T allele in the 'old world'. (a) LP phenotype distribution. Data points (dots) were taken from the literature (see text and [14] for details). (b) Distribution of the allele -13910^*T , associated to LP. Dots represent sample data taken from a previous review [14,26–30]; crosses represent data for new locations not previously tested and diamonds correspond to locations where additional data have been added. Regularly updated frequency data are available at <http://www.ucl.ac.uk/mace-lab/GLAD/> website.

frequency of Somali people living in Ethiopia does not prevent them from drinking more than 500 ml of milk per day without any obvious discomfort [16]. This inter-individual variation of the amount of lactose tolerated by lactase non-persistent people may be a result of variation in the composition of the gut flora (particularly the presence of lactic acid bacteria) [16,43]. Also, fermented dairy products (i.e. yoghurt or cheese) contain less lactose, allowing consumption by non-persistent individuals without any of the expected symptoms [43].

Two contrasting theories have been proposed to explain the co-distribution of LP and dairying practices. The *culture-historical* hypothesis [11,12,44] argues LP developed, and was consequently selected, after milk production and dairy consumption spread. The opposing hypothesis (mentioned in McCracken [12] as the *reverse-cause* hypothesis) proposes that only populations whose frequency of LP was high enough adopted dairying. In other words, human groups were differentiated with regards to LP frequency by a process unrelated to milk consumption—through genetic drift [45]—before the invention of dairying. This view specifically assumes that drinking milk would not

necessarily have conferred any selective advantage. The *culture-historical* hypothesis is better supported by recent findings from archaeological research (see §1c for details). In particular, organic residues preserved in archaeological pottery provided evidence for the use of milk after 8500 BP in the western part of Turkey [40], a region where LP frequency is low today (figure 1a). It suggests that domestic animals were milked before LP arose or was present at appreciable frequencies.

(c) *The advantage of being lactase persistent*

Several explanations have been proposed for how and why LP may have been selected. For example, it may simply be that milk is a good source of calories, or specifically an important source of protein and fat. The milk production of a prehistoric cow has been estimated to range between 400 and 600 kg per weaning period. Even when the milk necessary for the raising of the calves is subtracted, some 150–250 kg remains [46]. This is almost equivalent to the calorie gain from the meat of a whole cow. Hence, over the years, milking may have resulted in a greater energy

yield than the use of cattle for meat. But it is likely that the benefits LP gave to early dairying populations extend beyond simply increasing the food supply or making more economic use of livestock. Furthermore, the specific benefits of dairying may have varied in space and time.

Strong selective pressures on LP may have been episodic and occurred only under certain extreme circumstances, such as drought, epidemic or famine. For example, milk would have represented an alternative food resource in between periods of crop cultivation. When no cereal food was available, for example between harvesting seasons or in periods of crop failure, LP individuals would have had an advantage. This is especially true for children after the period when lactase production is normally downregulated, a phase of life that shows an increased mortality according to osteological investigation of prehistoric skeletal collections [47,48]. In addition, Cook & Al-Torki [49] hypothesized that in regions where water was scarce, milk would have been used by pastoralist groups as relatively pathogen-free fluid. If by drinking fresh milk, lactase non-persistent individuals were at risk from the potentially dehydrating effects of diarrhoea under such conditions, selection may have been strong in lactase-persistent individuals.

However, this arid climate hypothesis is less likely to have been relevant in Europe, where LP frequencies are at their highest. The observed correlation between latitude and LP frequency in Europe led Flatz & Rotthauwe [50] to propose the *calcium assimilation* hypothesis. Calcium is essential for bone health and its absorption in the gut is dependent on the presence of vitamin D. While some food sources such as fish are rich in this vital nutrient, most people in the world produce the majority of their vitamin D photochemically in the skin through the action of UVB on 7-dehydrocholesterol. However, UVB exposure is insufficient to produce the required quantities of vitamin D in people living at high latitudes for much of the year. This is unlikely to have been a problem for pre-Neolithic European hunter-gatherers who would have had a vitamin D-rich diet through their consumption of marine foods [51], but may have been a problem for early agriculturalists. Additionally, it has been proposed that a fibre-rich diet, owing to high consumption of cereal grains, can lead to a reduction in the amount of plasma 25-hydroxyvitamin D₃ [52]. Thus, the consumption of milk has been suggested to have conferred an advantage to early LP farmers in regions such as the circum Baltic/North Sea area, where climatic conditions allowed high crop cultivation and consumption. Because milk contains small quantities of vitamin D and plenty of calcium, it can provide a valuable supplement at low-sunlight latitudes.

Bloom & Sherman [53] formulated the ecological dairying barrier hypothesis to complement the culture-historical model. They suggested that dairying, and therefore the evolution of LP, required environments that are favourable for raising milk-producing ungulates. Others proposed that the LP distribution is related to malaria [54]. They implied that LP is the ancestral phenotype and that lactase non-

persistence would have been selected in regions where the disease was frequent. However, this hypothesis was not supported by studies of lactase non-persistence prevalence in glucose-6-phosphate dehydrogenase deficient subjects from Sardinia [55,56]. In a contrasting hypothesis, it has been suggested that a milk diet would provide protection against malaria by impairing a part of plasmodia metabolism (folate synthesis) [52] and thus lead to selection for LP. Pointing to the tight relation between dairying and LP, it has also been hypothesized that milk drinking was a privilege restricted to some individuals in highly hierarchical societies and that it spread as prestige class behaviour [13,44,57]. (For further discussion on the spread of a prestige-associated culture, see [58].)

A common feature of most populations with high frequencies of LP is a history of dairying activity [11]. The availability of fresh milk to some human groups has challenged their niche, thereby creating a potential genetic feedback for a need of continuous lactase expression throughout adult life. As they appear to be interdependent feedback processes, analyses of LP genetic and cultural diversity can hardly be conducted separately [13]. Hence, it is likely that the study of the process of when, how and why some populations kept and exploited ungulates will shed new light on our understanding of the distribution of LP. Indeed, archaeological data can be used to provide evidence for the presence of this cultural behaviour in past populations. For example, the analysis of milk residues [59] and the determination of kill-off patterns of animals on archaeological sites [60] constitute two archaeological methods that can inform on the emergence of dairying. Consequently, to fully understand the origins and evolution of LP, it is necessary to consider archaeological and archaeozoological research on the origins of domestication and dairying. In Europe, the spread of domestic animals is tightly associated with the diffusion of the Neolithic from the Near East.

2. LACTASE PERSISTENCE, THE NEOLITHIC TRANSITION AND THE HISTORY OF DAIRYING

(a) *The spread of the Neolithic*

Before 8400 BP, hunter-gathering was the only subsistence strategy in Europe, but by 6000 BP, when farming had spread over most of the continent (figure 2), it had become rare. The spread of farming into Europe was dependent on earlier developments in the Neolithic core zone of the Near East and Anatolia. Archaeologically, this process defined a new period referred to as the 'Neolithic revolution'. It is characterized by the presence of polished stone tools and pottery, a more sedentary lifestyle and the management and subsequent domestication of certain animal and plant species. These features are often characterized as the 'Neolithic package', and this term has been widely used to define Neolithic sites, even though all these features are not always present together at these sites.

Chronologically and geographically, the Neolithic culture spread from its original core zone to other parts of western Eurasia and simultaneously to north

Africa. There are two opposing models for the spread of farming into Europe [62]. One, the *cultural diffusion*, or *acculturation* model, favours the idea that local hunter–gatherers adopted Neolithic practices once they gained access to them from farming neighbours. The alternative *demic diffusion* model holds that the Neolithic was carried throughout Europe by the movement of farming peoples (and consequently their genes) spreading into the territory of foragers. It is highly unlikely that the Neolithic transition could be wholly explained by either of these models alone. The reality is likely to have been more complex and involved local heterogeneity of farming adoption processes [63].

Analyses of European genetic diversity have attempted to address the relative role of acculturation and migrations on the European gene pool during the spread of farming and yielded contrasting results, depending on the methodology employed, the loci studied and the proxies for ancestral source populations used [64–70]. Recently, Bramanti *et al.* [71] presented ancient DNA evidence for a genetic discontinuity between late hunter–gatherers and early farming populations, at the beginning of the central European Neolithic 7500 BP, thus providing support for a migrating farmers model in this region. A full Neolithic lifestyle was established within a few generations in this area and cultural contact between farmers and local foragers persisted for some time. But the way in which Mesolithic hunter–gatherers and the Early Neolithic farmers interacted and how this led to an environment peopled exclusively by farmers is still debated [10,63,72].

(b) Domestic animals associated with the Neolithic transition

Archaeozoological data on domestic animals have been particularly valuable in establishing a more precise picture of the spread of the Neolithic and the new subsistence strategies associated with it. This involves bone analysis (animal species, sex, age at death, morphology, presence of cutting and cooking signs [14], C direct dating and ancient DNA analyses) and placing this within their archaeological context. Morphological changes [73] and culling strategies [74] can be used to identify where and when wild animals started to become herded livestock. During the pre-pottery Neolithic B phase (PPNB) in the Neolithic core zone, several human groups started to manage wild animals at approximately the same time, in different places. After a period of hundreds or even thousands of years, this process finally resulted in phenotypes characteristic of domesticates. These domestication-associated phenotypes illustrate one of the consequences of human cultural niche construction on the evolution of other species [3,8,10,75].

In the context of the Neolithic, the oldest evidence for domestication is for goat and sheep (11 000 BP) followed by pig and cattle (10 500 BP) [73,74,76,77]. These domesticates soon spread from the Neolithic core region to a large part of the Near East, including Cyprus [77,78] and central Anatolia. However, true animal husbandry as a major economic activity in the

Near East only began during the 10th millennium BP. Subsequently, but not earlier than 9000 BP, these domesticates spread into western and southern Anatolia [72,79]. From there, the spread of domesticates together with husbandry techniques followed two main routes (figure 2): (i) the Mediterranean route via the Aegean and Adriatic Seas, then moving further west to southern Italy, the Thyrenic Islands, southern France and the Iberian Peninsula; and (ii) a Danubian route via the Balkan Peninsula, further to the south-western part of central Europe and finally into central and northern Europe. The two routes might have met in Greece, in the Rhine valley and in the northwest of the continent before crossing to the British Isles [80].

Domestic goat (*Capra aegagrus*) and sheep (*Ovis aries*) must have been introduced to these areas by humans, as there is no evidence of indigenous wild progenitors in Europe. Ancient DNA studies indicate that cattle (*Bos taurus*) were also imported from Anatolia and, once in Europe, did not mix substantially with European wild cattle, the aurochs (*Bos primigenius*) [81–83]. This is probably because humans managed to keep these two forms of cattle reproductively separate, although substantial size differences and other consequences of domestication may also have limited interbreeding. In contrast, while domestic pigs seem to have been introduced into Europe from the Neolithic core zone, ancient DNA evidence indicates that they mixed substantially with local wild boar [84].

In terms of dairying, the further development of herd structure after the arrival of early domesticates in Europe is crucial. Indeed, neither the Early Neolithic in the Balkans (8200–7500 BP) nor the earliest Neolithic in central Europe (Linearbandkeramik, 7500–7000 BP) yield clear archaeozoological signs of highly specialized dairying. However, in the following periods, especially in central Europe, there is evidence of a progressive increase in herding associated with dairying, i.e. prevalence of female over male animals. This particularly holds true for cattle and goat [46]. Similarly, in southeast Europe, a transition to a more specialized dairying economy becomes increasingly apparent as early as the 7th millennium BP. It should be noted that it is very likely that dairying was practised long before this period [40,42] and may have represented a means of managing wild animals as early as the beginning of the domestication process. However, it did not become a substantial economic factor before the 7th millennium BP in most parts of Europe.

(c) Evidence of the consumption of dairy products

In addition to faunal remains, the detection of dairy fats associated with archaeological pottery is a powerful line of evidence for dairying activities in pre-history [59,85]. Investigations of organic residues in archaeological pottery have revealed a wide range of compound types, including dairy fats [85]. Using this method, the processing of milk has been identified in the western part of the present day Turkey as early as 8500 BP [40]. Similarly, evidence for the use of dairy products has been shown in Neolithic sites of

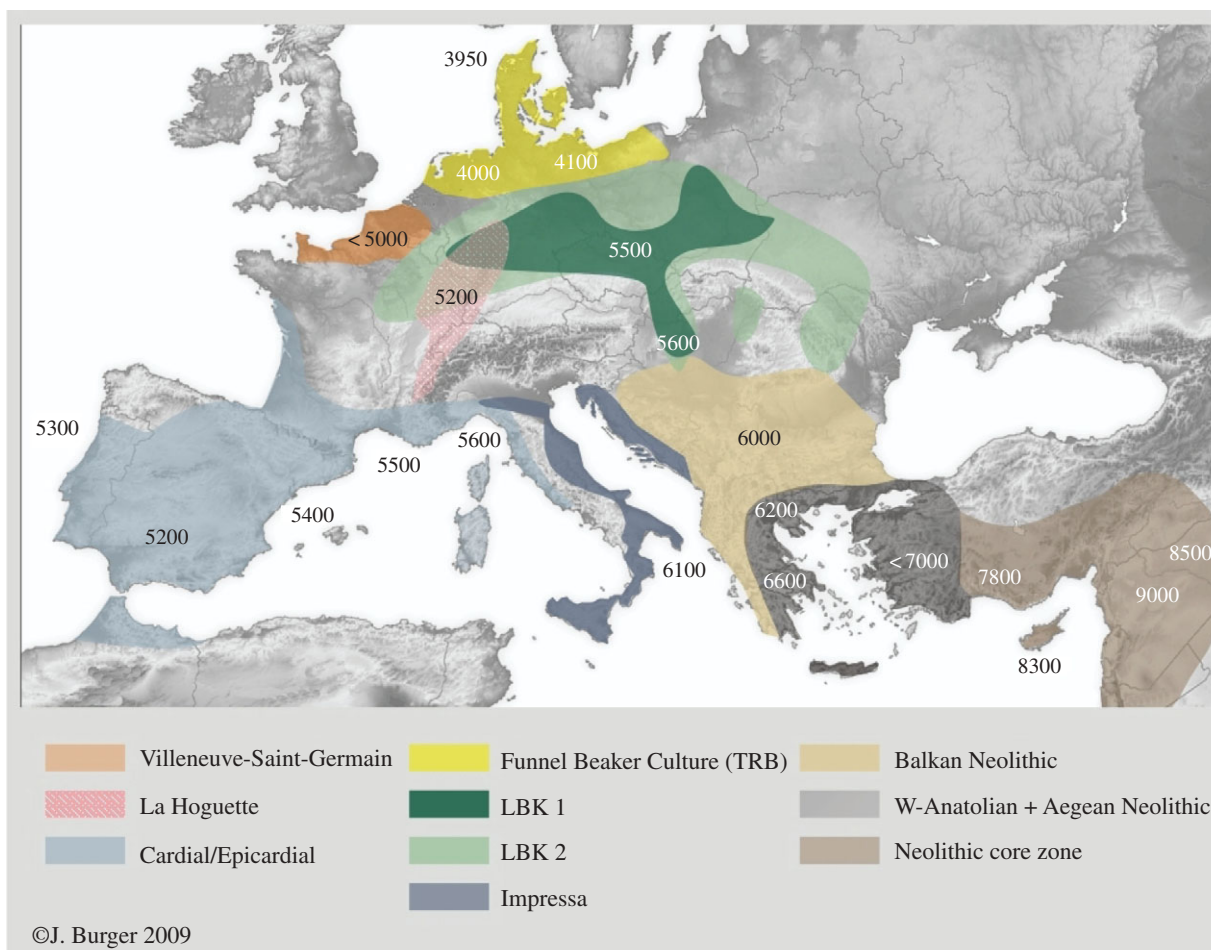


Figure 2. Arrival dates and approximate geographical expansions of defined Early Neolithic cultures from [61].

Romania and Hungary around 7900–7450 BP [39], Britain around 6100 BP [86] and Scotland around 3000 BP [87]. Thus, the archaeozoological and residual lipid data clearly indicate that dairying was practised early in Neolithic Europe. However, ancient DNA data from central and northern Europe suggest that LP frequency was low during this period [88]. Thus, we may hypothesize that Early Neolithic people, among whom LP was rare or absent, initially practised dairying in south-eastern Europe and later migrated towards central and northern Europe, an area inhabited by foragers that occupied a different niche [10].

Indeed, stable isotope studies of bone collagen show that while hunter–gatherers largely relied on marine food, the farmers’ diet, even on coastal sites, was mainly terrestrial [51]. However, other authors found no evidence of this switch in diet [89–91]. This may be because the localization of hunter–gatherers compared with farmer sites is biased towards coastal regions [90,92], or there may be other methodological issues [91,92]. An additional possibility is that in regions with abundant aquatic resources, the process of subsistence change offered by Neolithic farmers would have been slower. For example, in Scandinavia, where marine resources are abundant, there is no direct evidence of a sudden dietary transition [90]. In fact, instead of a replacement of hunter–gatherers by farmers, hunters from the Pitted Ware Culture (PWC) [89] coexisted for nearly 1000 years with the Neolithic Funnel Beaker Culture (TRB). However, it

should be noted that recent ancient DNA evidence indicates little genetic continuity between PWC hunter–gatherers and modern Scandinavians [93].

In summary, LP and the main LP-associated allele in Europe, -13910^*T , are found at highest frequencies in northwest Europe where dairying arrived latest (figures 1*a,b* and 2). A simple—phylogeographic—interpretation of the latter would be that LP first evolved in northwest Europe. But computer simulation studies have shown that when a population expands, the centre of distribution of an allele can be far removed from its location of origin [94,95]. This process is called ‘allele surfing’ and is thought to have occurred with the spread of farmers in Europe [65]. Furthermore, selection has—to an extent—shaped the distribution of LP [14,36,37], although it is unclear whether this selection was continuous or episodic, and whether it varied by latitude [50] or ecological zones [53]. It is therefore clear that to obtain a more complete picture of the coevolution of LP and dairying in Europe it is necessary to integrate cultural, demographic and selective processes. Computer simulation represents one of the most promising approaches to understand these processes because it can combine multiple sources of information. However, as with all modelling approaches, it is necessary to identify the key evolutionary parameters that have shaped the observed data. This is a non-trivial task when one considers the range of proposed hypotheses to explain the current distribution of LP in Europe

today. In the following section, we review three of the simulation studies that have deepened our understanding of the coevolution of LP and dairying in Europe.

3. REVIEW OF SIMULATION STUDIES

(a) *Aoki's model of lactase persistence and niche construction*

The coevolution for LP and dairying simulated by Aoki [96] captures many of the features of niche construction, as later proposed by Laland *et al.* [5]. The evolution of populations, made up of four compound phenotypes (two genetic: LP versus non-LP individuals and two cultural: milk users versus non-users), was simulated. The compound phenotype of being an LP milk user was selectively advantageous (fitness of 1), whereas the other three phenotypes were assigned a selective value of $1 - s$ (where s is the coefficient of selection). The effective population size (N_e) assumed was 100 individuals. At each generation, the model considered: (i) random mating between individuals (the random combinations of alleles and cultural behaviour from a parental generation to the next), (ii) the transmission of the cultural trait occurred with distinct probabilities of becoming a milk user, i.e. $f(y)$ for an LP individual and $g(y)$ for a lactase non-persistent individual, and (iii) random sampling of the individuals after the action of selection. At the end of the process, a correlation between the fixation of both the LP trait and the cultural behaviour was measured, and an average time until fixation of LP was calculated. This study addressed three questions: (i) if the evolution of both LP and milk consumption is mutually dependent, what would the correlation between their frequencies be? (ii) Does the type of selection, either strictly genetic or culturally induced, produce differences in the rate at which an allele reaches fixation? (iii) Do the frequency estimates of European LP (falling in the interval (0.05–0.70)) fit the hypothesis of a gene–culture coevolution process that started 6000 years ago?

Although this model did not take into account cultural diffusion from neighbouring populations, an interesting feature is the use of distinct probabilities for transmission of the culture of milk consumption according to the LP phenotype of an individual. A key result was that an incomplete correlation between LP and milk consumption frequencies is expected. This actually represents inter-individual variability in the non-persistence phenotype, and corresponds either to LP individuals that do not drink milk or to lactase non-persistent individuals that do drink milk. It explicitly shows that the gene–culture coevolution hypothesis is still credible even if the correlation between the genetic and the cultural traits is incomplete. This is in accordance with previous analytical gene–culture coevolution studies [7,97,98] which have notably shown that genotype–phenotype correlation can be reduced by a factor depending on environmental and selection variances [97].

Another observation drawn from this model is that the change in LP frequency is slower when the selection is conditioned by culture. This means that in order to explain such high frequency of LP in

north-western Europe, we must consider that either every LP individual actually did drink milk, or that other processes have been involved to generate this distribution. The author consequently suggested that to be able to detect such a change in allele frequency since the start of dairying (then considered to have occurred 6000 years ago), either the effective population (N_e) was relatively small (100 individuals as simulated in the study), or the selection coefficient in favour of the phenotype was very high (more than 5% if $N_e = 500$). Such a high value of 5 per cent falls well within the confidence intervals ((1.4–19%) and (1–15%)) of recent studies ([20,36], respectively) inferred directly from genetic data. However, lower values of the selection coefficient may be expected to fit the observed LP pattern, as it is now known that the start of dairying was much earlier than 6000 years ago [40], leaving more time for the genotype to be selected. In addition, both genetic and archaeological studies point to the importance of the demographic process when studying the Neolithic diffusion in Europe [99,100], whereas Aoki's simulations were performed with a constant population size of 100 individuals. Moreover, demographic expansion has been shown to have a potentially dramatic effect on the diffusion of new mutations [94,95,101]. Hence, we may expect that the assumption of a constant population size could lead to an overestimation of the rate of frequency change and of the selection coefficients.

(b) *Spatial variation in selection intensity*

While the selection coefficients explored in Aoki's model were assumed to be constant, in a recent study, Gerbault *et al.* [18] modelled a geographical structuring of selection pressure by latitude, thereby testing explicitly the calcium assimilation hypothesis [50]. The evolution of a dominant allele associated with LP was simulated in four Near-Eastern and 22 European populations since the Neolithic transition in their respective regions, according to two demographic models: cultural diffusion and demic diffusion [63]. Each of these models was tested using either a constant selection coefficient or a selection coefficient varying with latitude. Thus, a total of four scenarios were tested combining the two selection models and the two demographic models. The program used (called SELECTOR) models four parametrized processes: (i) random genetic drift, (ii) logistically regulated population growth, (iii) positive selection varying between 0 and 3 per cent, and (iv) time elapsed since the onset of the Neolithic. To simulate the cultural diffusion model, the initial LP allele frequency was set at 1 per cent in all 26 populations and evolved according to the above processes. One important assumption is thus that the LP-associated allele was already present in Europe before the Neolithic but at a low frequency (1%). Under the demic diffusion model, three populations were used as sources in the Near East (Lebanon, Syria and Iran). The remaining populations, including the fourth Near Eastern population (Cyprus), were populated by sampling from neighbouring populations along the presumed route of the spread of farming, and at

the times indicated from archaeological data [41]. A maximum-likelihood test was performed to evaluate independently in each sample what selection coefficient best fitted the observed data (the LP frequency for each of the 26 populations). Moreover, the four simulated scenarios were formally compared using an approximate Bayesian computation (ABC) approach [102] based on LP frequencies within samples [103].

In this study [18], the scenario that gave the highest relative probability of obtaining the observed data was the demic diffusion model combined with a selection coefficient varying with latitude. Indeed, according to this model, genetic drift alone can explain frequencies as low as those observed in southern Europe, as previously suggested by Nei & Saitou [45]. However, in northern Europe, positive selection coefficients were required to drive LP frequencies to their observed values (figure 3). This result supports the calcium assimilation *hypothesis*. A complementary explanation for increased selection strength at higher latitudes is that the lower temperatures would allow milk to remain fresh for longer [44,57]. If this applies, LP would illustrate a specific case of niche construction where localized access to a resource is variable [5], i.e. fresh milk. Indeed, niche construction modelling [104] has shown that in such a situation a strong association between a cultural and a genetic trait is expected, as is observed for LP and dairying in northern Europe.

Even though the model of Gerbault *et al.* [18] does not simulate a wave of advance, it illustrates the importance of demography, since the demic diffusion model with variation in the selection coefficient was far more likely (99.1%) than the cultural diffusion model (0.9%). The selection coefficients estimated range between 0.8 and 1.8 per cent, and the authors suggested that if the simulations were closer to a wave of advance model, this range might have been lower. It should be noted that these estimates are not constant over all the populations considered, as under the best-fitted models the selection coefficient varied latitudinally and approached zero in southern Europe.

(c) *Demic and cultural diffusion of farming*

In a more complex spatially explicit model, Itan *et al.* [105] simulate the spread of farmers from the Near East to Europe over the last 9000 years, taking into account potential interactions with hunter-gatherers. The geographical unit of this simulation was a deme, and in total the simulated world was made of 2375 land demes, each containing three interacting populations: hunter-gatherers, dairying farmers and non-dairying farmers. At each generation, each population underwent a succession of seven processes: (i) logistically regulated population growth, in which each deme had a fixed carrying capacity determined by climatic and elevation factors. The carrying capacities of dairying and non-dairying farmers were equal, whereas those of hunter-gatherers were 50 times smaller [106,107]. (ii) Unidirectional migration process modelled as a stochastic Gaussian random walk from one deme to another (a process equivalent

to demic diffusion). (iii) Cultural diffusion, where a proportion of individuals from one culture 'converted' to one of the two other cultural groups, based on the relative 'dominance' of the other groups in the focal and eight surrounding demes. (iv) Intra-demic gene flow between different cultural groups within a deme. (v) Inter-demic gene flow between neighbouring demes belonging to the same cultural group. (vi) Selection acting on an LP allele only in the dairying farmers cultural group. (vii) Drift of the LP allele in all cultural groups in all demes, modelled as a binomial sampling process. The LP-associated allele was set to appear in a randomly chosen location when the population size of dairying farmers in this deme reached a critical value (20 individuals). This ensured that the LP-associated allele appeared on, or near, the wave-front of dairying farmer diffusion. Its frequency was updated according to the evolution of a dominant allele with parametrized selective advantage [108], which remained constant in each simulation. This selection in turn drove additional increases (over and above the logistically regulated population growth) in the population sizes of dairying farmers [108].

After the simulations, the authors estimated parameters of the model using ABC [102] with a particular focus on two related questions: (i) when and where in Europe did LP-dairying coevolution begin, and (ii) what factors drove the spread of LP to get the observed European pattern of LP frequencies? Simulations were fitted to both observed $-13910 \times T$ frequencies at 12 European locations and the corresponding dates of arrival of farmers [41] at 11 of these locations. Both the dates estimated (between 6256 and 8683 years BP) and the geographical area identified (the region between central Europe and northern Balkans—figure 4) for the origin of LP-dairying coevolution correlate well with the time and location of the emergence of the Linearbandkeramik (LBK) culture (figure 2). The LBK is recognized to have been the direct forerunner of a Neolithic cattle-based economy [109], whereas on the Mediterranean Neolithic sites, faunal assemblages are more variable in composition and in some places dominated by sheep and goat [42,46,109]. Indeed, cows provide quantitatively more milk than sheep or goats, allowing cattle-based farmers to have a better supply, thereby supporting a stronger selection for LP. Even though these origin time and location estimates were not independently derived, as simulations were conditioned on known farming arrival dates [41], recent results from spatially explicit niche-construction modelling [9] corroborate this hypothesis. In this latter model [9], local mating of niche constructors (i.e. LBK-economy distribution and diffusion) has the effect of increasing the local concentration of the resource (i.e. fresh milk), thereby generating stronger selection in favour of the resource-dependent trait (i.e. LP).

Indeed, from the study of Itan *et al.* [105], the selection coefficients necessary to explain the $-13910 \times T$ allele frequency distribution in Europe were estimated to range from 5.2 to 15.9 per cent. Although very high, this range falls within the range estimated (1.4–19%) from molecular data [36], but shows less overlap with the range estimated by Gerbault *et al.* [18].

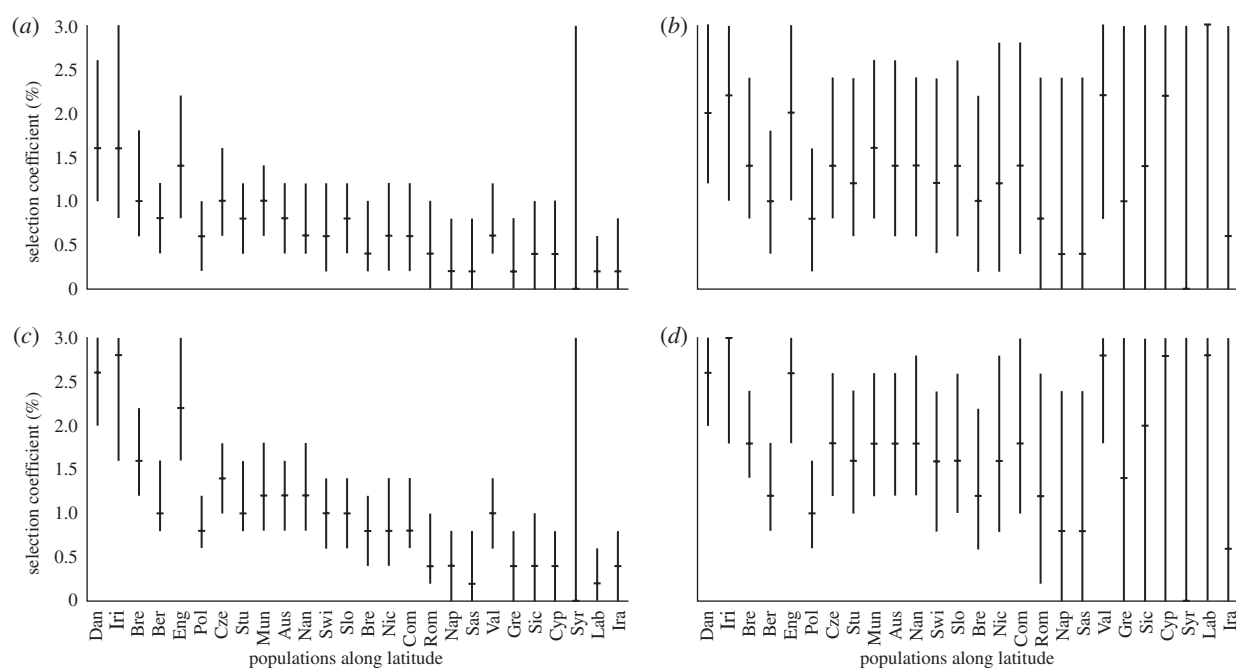


Figure 3. Scenarios simulated in [18] and selection coefficients required to fit the observed estimates of LP frequencies (taken from [18]). Bars represent the 95% confidence interval of the selection coefficient estimated for the population and the central point is the MLE (maximum-likelihood estimate). Populations are ordered from the highest to the lowest latitude: Danish (Dan), Irish (Iri), German from Bremen (Bre), German from Berlin (Ber), English (Eng), Polish (Pol), Czech (Cze), German from Stuttgart (Stu), German from Munchen (Mun), Austrian (Aus), French from Nantes (Nan), Swiss (Swi), Slovenian (Slo), Italian from Brescia (Bre), French from Nice (Nic), Spanish from Santiago de Compostela (Com), Italian from Roma (Rom), Italian from Napoli (Nap), Sardinian (Sas), Spanish from Valencia (Val), Greek (Gre), Sicilian (Sic), Cypriot (Cyp), Lebanese (Leb) and Iranian (Ira) (doi:10.1371/journal.pone.0006369.g002). (a) Demic diffusion, gene-culture coevolution; (b) demic diffusion, calcium assimilation; (c) cultural diffusion, gene-culture coevolution; and (d) cultural diffusion, calcium assimilation.

However, as the selection strength inferred by Itan *et al.* [105] applies to less than half of the overall population (dairying farmers), the population-wide selection coefficient is likely to be around half of the value reported. A key difference between the two studies is that Gerbault *et al.* [18] explicitly modelled, and found support for, a latitudinal effect on LP selection coefficients [50] while Itan *et al.* [105] did not. Even though it is difficult to separate the effects of selection and demography in their model (as demographic parameters and the selection coefficient were estimated simultaneously), Itan *et al.* [105] argued that such a latitudinal effect is not *necessary* to explain the higher observed frequencies of LP in northern Europe. Indeed, they found that without any assumption of a latitudinal effect on selection, their model under the estimated parameters predicted higher LP frequencies in northern Europe than observed, and typically lower frequencies in southern Europe than observed. Thus, the addition of latitudinal effect on selection should drive an even poorer fit to the LP data and so could be thought of as having ‘negative explanatory power’. However, it should be noted that Itan *et al.* [105] did not explicitly model a latitudinal effect on selection, and so have not formally rejected it.

Another key difference between both studies is the values estimated for the selection coefficient, which are larger in Itan *et al.* [105] than in Gerbault *et al.*

[18]. In the study of Itan *et al.* [105], the best fit was obtained when the allele was first selected in dairy farmers, in central Europe/the northern Balkans, and subsequently spread into neighbouring regions. High selection coefficients were thus required to obtain the high frequencies currently observed in north and central Europe in a relatively short period of time. In the model of Gerbault *et al.* [18], the allele was assumed to have reached non-negligible frequencies before it had spread into central Europe as it was already present in the Near East at the onset of the Neolithic transition. Hence, while selection was required, this was to a lesser extent than if it had started in central Europe. There is a potential ‘trade-off’ between selection intensity and the timing and location of origin of the gene-culture coevolutionary process, and in future work, combinations of values for these parameters will bring insights into the means of diffusion not only of LP but of other genetic traits as well.

4. DISCUSSION AND PERSPECTIVES

(a) Simulation overview

Although simulation studies are unlikely to recover the precise evolutionary history of LP and dairying, they have two main advantages over purely genetic or purely archaeological studies. First, they allow the integration of information from multiple sources

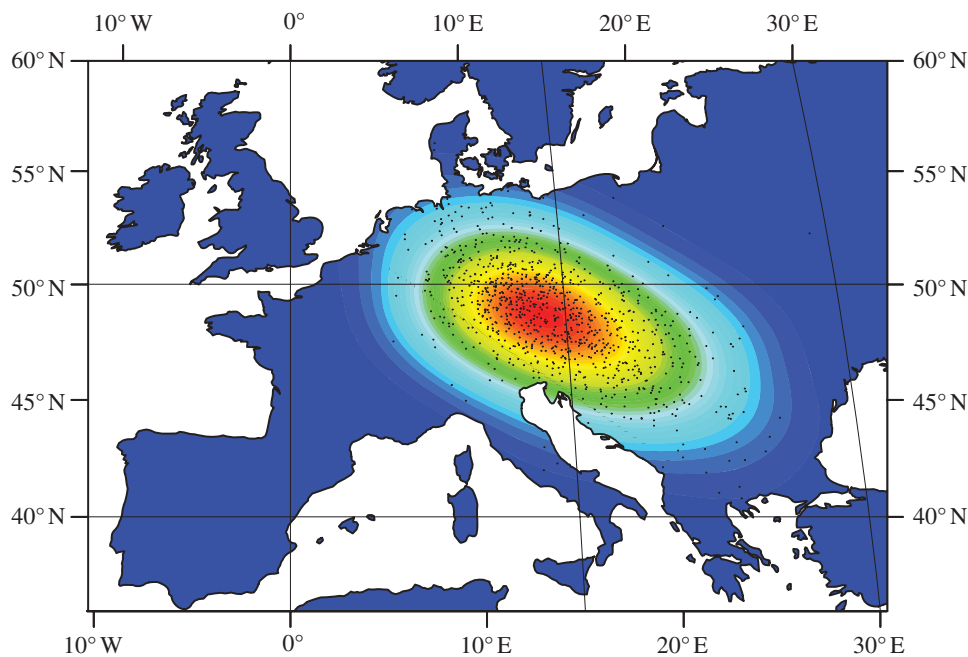


Figure 4. Approximate posterior density of region of origin for LP—dairying coevolution (taken from [105]). Points represent regression-adjusted latitude and longitude coordinates from simulations accepted at the 0.5% tolerance level. Shading was added using two-dimensional kernel density estimation (doi:10.1371/journal.pcbi.1000491.g003).

(genetics, archaeology, ecology). Second, they provide a formal comparison of alternative scenarios, according to the parameters taken into account. Many combinations of parameters can be tested, and extreme scenarios (such as cultural or demic diffusion) can be evaluated statistically, and potentially excluded. Furthermore, ABC methods [102,103] have proved to be very useful when evaluating the fit of observed data on complex scenarios tested, adding power to simulation studies [18,105]. Clearly, simulations performed at the continental scale help to understand general features and processes of the European Neolithic transition, and to explore alternative possibilities that may have led to observed patterns of genetic diversity.

(b) *Demography and niche construction*

The above simulation studies of LP/dairying coevolution, as well as ecological and archaeological information, suggest that selection on LP is unlikely to have been constant over time and space during its spread throughout Europe, and show that the role of demography in its diffusion cannot be ignored. The extent to which combinations of these two phenomena have shaped the LP distribution as it is observed in Europe remains unanswered. While Itan and colleagues [105] used a complex demographic model, the population growth was logistically regulated for each group, with no direct consequence of the growth of one group on the growth of the others (i.e. competition between cultural groups for land resources). But as almost no hunter–gatherers survive in Europe today, it is logical to assume that niche modification by farmers may have had an effect on the distribution of hunter–gatherers that led ultimately to their disappearance (because of the distinct abilities

of both populations to niche construct [10]). As an adaptation of the model of Itan *et al.* [105], an alternative dynamic could be proposed to investigate how demographic constraints may have affected the diffusion of the trait. Inclusion of density-dependent competition, as implemented by Currat & Excoffier [110], should bring new insights into the effect of these demographic processes. This improvement would allow regional variation in the simulation of the Neolithic spread, with a higher carrying capacity of hunter–gatherer populations in areas where marine resources are abundant, thereby potentially changing the competition dynamics between hunter–gatherers and farmers. This would also enable account to be taken both of localized resource availability and spatially structured populations for investigating the evolution of the association of LP (i.e. the resource-dependent trait) and dairying (i.e. the niche-constructing practice) [9,104]. This may be relevant for the co-habitation of hunters from the PWC [89] with the TRB farmers, who coexisted for nearly 1000 years in Scandinavia. Further analyses assessing this co-habitation would be appropriate to explicitly evaluate how the co-influences of genes, environment and pre-existing culture affect differential transmission of Neolithic cultural forms [111].

The spread of LP is a good example of niche construction for two reasons. First, because human groups who started to drink milk modified their pre-existing selection pressure, thereby generating an evolutionary feedback that may have advantaged some individuals against others in certain environmental conditions. Second, by perpetuating this behaviour from the Neolithic to modern days, the coevolution of LP and dairying illustrates how culture can affect the genetic diversity of human populations. That is why niche construction factors should be considered

when studying human adaptation. Gene–culture co-evolution studies have shown that both genes and culture exert distinct, but interacting, influences in the evolution of human phenotypes [98,111]. More precisely, cultural processes have been shown to alter the outcome expected under purely genetic transmission [7,97,98]. Even though dairying culture has been taken into account, it should be noted that the effect of differential cultural transmission has not been explicitly modelled in the three simulation studies described above. It is thus expected that the combination of demographic information and differential cultural and genetic transmission influenced by selection would enlarge the possible evolutionary conditions that may have shaped the worldwide LP distribution observed today.

(c) *Migration and environment*

Other environmental parameters are likely to have affected the distribution of the culturally distinct populations, and the speed of the Neolithic transition. Indeed, niche-construction studies have shown that environmental variables should not be seen as static but as dynamic features that change according to climate and human activities [5], thereby modifying the selection pressure to which humans are subjected [35]. For example, Itan *et al.* [105] took into account environmental factors to determine different carrying capacities according to the economy type of the populations. In addition to this, other environmental pressures, such as vegetation cover and natural water access, may have conditioned the movement of farming populations with their domesticated animals [57]. These environmental variables can allow the modelling of potential regional variation in the spread of the Neolithic, as suggested by archaeological studies [112]. This will have an initial advantage of bringing a better understanding of the effect of the environment on the spread of agriculture and farming. Then, the relative importance of environmental variables that influence migrations can be evaluated. However, it should be noted that increasing the number of parameters in a model is generally to be avoided unless necessary. In particular, because the more complex a model is (the closer to the reality it is), the more difficult it becomes to interpret, as isolating the effects of different processes, and parameters, can be extremely complicated. Furthermore, the parameter space of a complex model can be very difficult to explore by simulation.

(d) *The importance of an interdisciplinary approach*

While selection on LP has been inferred, it has not been shown directly [113]. Other pleiotropic factors may be involved that affect the ability to digest milk on the one hand or that affect selection on *LCT* on the other. What is clear from the study of LP-associated genetic variants is that it has emerged independently in different regions of the world [14,20,23,31,32,114]. While some alleles show evidence of strong directional selection, soft selective sweeps [115]—whereby ancestral genetic variation is associated with different adaptive substitutions

[116]—have been invoked [16] to explain how several alleles can be maintained in a population through mutation and migration processes. It would be of interest to assess whether such a process applies more widely than in Africa and the Middle East. Ongoing research on the *LCT* gene and the regulation of its expression, and further sampling, with associated sequencing and phenotyping will be necessary to give a more complete picture of the structure of LP-associated genetic diversity [15]. Additionally, archaeological assemblage studies will provide useful information on the time and location of the development of milk-drinking behaviour. Finally, simulation studies can assess different possible scenarios by taking into account information obtained from different fields of study, particularly by employing ABC methods [102,103]. LP evolution will be better understood by combining data from these different fields, and by generating expectations of different data types in simulations. More broadly, the study of LP evolution is likely to shed light on the interactions and relative contributions of distinct processes involved in shaping human variation.

5. CONCLUSION

The coevolution of LP and dairying appears to be not as straightforward a process as it may have first seemed. While several alleles seem to be associated with LP, the pattern of LP in Europe appears peculiar, as a single variant (-13910^*T) has been identified as explaining most of the modern distribution of LP [15]. In this continent, the spread of dairying-associated practices goes back to Early Neolithic periods, and it is likely that archaeological studies will soon give a more precise picture of the early development of dairying in Europe. Archaeology notably underlines how complex the migration processes may have been. It may not be surprising then, even though the models were different, that simulation studies on the spread of LP in Europe highlight the importance of demography on the distribution of the -13910^*T allele. Some of these simulation studies also hint that selection may not have been constant over time and space. It thus appears that, in order to disentangle the effects of selection from those of demography on the distribution of both LP phenotype and genotypes, there is a need to better understand spatial patterns of genetic variation under neutrality [66,101,117,118]. In turn, the extent to which LP-associated genetic variation can be explained by these patterns will determine what can be said about the adaptation process [119]. Inference on the coevolution of LP (and its associated alleles) and dairying is likely to provide a useful framework for future studies on the evolution of other adaptive traits in which niche construction is invoked.

We thank two anonymous referees for their comments. P.G. and A.L. are funded by an EU Marie Curie FP7 Framework Programme grant (LeCHE, grant ref: 215362-2). Y.I. was funded by the B'nai B'rith/Leo Baeck London Lodge and Annals of Human Genetics scholarships. We also thank the AHRC Center for the Evolution of Cultural Diversity (CECD) and the Center for Mathematics and Physics in

the Life Sciences and Experimental Biology (CoMPLEX), UCL, for supporting this research.

REFERENCES

- Shennan, S. 2011 Property and wealth inequality as cultural niche construction. *Phil. Trans. R. Soc. B* **366**, 918–926. (doi:10.1098/rstb.2010.0309)
- Gintis, H. 2011 Gene–culture coevolution and the nature of human sociality. *Phil. Trans. R. Soc. B* **366**, 878–888. (doi:10.1098/rstb.2010.0310)
- Smith, B. D. 2011 General patterns of niche construction and the management of ‘wild’ plant and animal resources by small-scale pre-industrial societies. *Phil. Trans. R. Soc. B* **366**, 836–848. (doi:10.1098/rstb.2010.0253)
- Laland, K. N., Odling-Smee, J. & Feldman, M. W. 2000 Niche construction, biological evolution, and cultural change. *Behav. Brain Sci.* **23**, 131–146. (Discussion 46–75). (doi:10.1017/S0140525X00002417)
- Laland, K. N., Olding-Smee, F. J. & Feldman, M. W. 1996 On the evolutionary consequences of niche construction. *J. Evol. Biol.* **9**, 293–316. (doi:10.1046/j.1420-9101.1996.9030293.x)
- Laland, K. N. & Sterelny, K. 2006 Perspective: seven reasons (not) to neglect niche construction. *Evolution* **60**, 1751–1762. (doi:10.1111/j.0014-3820.2006.tb00520.x)
- Feldman, M. W. & Cavalli-Sforza, L. L. 1976 Cultural and biological evolutionary processes, selection for a trait under complex transmission. *Theor. Popul. Biol.* **9**, 238–259. (doi:10.1016/0040-5809(76)90047-2)
- Odling-Smee, F. J., Laland, K. N. & Feldman, M. 2003 *Niche construction: the neglected process in evolution*. Princeton, NJ: Princeton University Press.
- Rendell, L., Fogarty, L. & Laland, K. N. 2011 Runaway cultural niche construction. *Phil. Trans. R. Soc. B* **366**, 823–835. (doi:10.1098/rstb.2010.0256)
- Rowley-Conwy, P. & Layton, R. 2011 Foraging and farming as niche construction: stable and unstable adaptations. *Phil. Trans. R. Soc. B* **366**, 849–862. (doi:10.1098/rstb.2010.0307)
- Holden, C. & Mace, R. 1997 Phylogenetic analysis of the evolution of lactose digestion in adults. *Hum. Biol.* **69**, 605–628.
- McCracken, R. D. 1971 Lactase deficiency: an example of dietary evolution. *Curr. Anthropol.* **12**, 497–517. (doi:10.1086/201234)
- Durham, W. H. 1991 *Cultural mediation: the evolution of adult lactose absorption. Coevolution: genes, culture and human diversity*, pp. 226–285. Stanford, CA: Stanford University Press.
- Ingram, C. J., Mulcare, C. A., Itan, Y., Thomas, M. G. & Swallow, D. M. 2009 Lactose digestion and the evolutionary genetics of lactase persistence. *Hum. Genet.* **124**, 579–591. (doi:10.1007/s00439-008-0593-6)
- Itan, Y., Jones, B. L., Ingram, C. J., Swallow, D. M. & Thomas, M. G. 2010 A worldwide correlation of lactase persistence phenotype and genotypes. *BMC Evol. Biol.* **10**, 36. (doi:10.1186/1471-2148-10-36)
- Ingram, C. J. *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**, 579–588. (doi:10.1007/s00239-009-9301-y)
- Swallow, D. M. & Hollox, E. J. 2000 The genetic polymorphism of intestinal lactase activity in adult humans. In *The metabolic and molecular basis of inherited disease* (eds C. R. Scriver, A. L. Beaudet, W. S. Sly & D. Valle), pp. 1651–1662. New York, NY: McGraw-Hill.
- 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. (doi:10.1371/journal.pone.0006369)
- Swallow, D. M. 2003 Genetics of lactase persistence and lactose intolerance. *Annu. Rev. Genet.* **37**, 197–219. (doi:10.1146/annurev.genet.37.110801.143820)
- Tishkoff, S. A. *et al.* 2007 Convergent adaptation of human lactase persistence in Africa and Europe. *Nat. Genet.* **39**, 31–40. (doi:10.1038/ng1946)
- Bayoumi, R. A., Flatz, S. D., Kuhnu, 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**, 173–178. (doi:10.1002/ajpa.1330580208)
- Bayoumi, R. A., Saha, N., Salih, A. S., Bakkar, A. E. & Flatz, G. 1981 Distribution of the lactase phenotypes in the population of the Democratic Republic of the Sudan. *Hum. Genet.* **57**, 279–281. (doi:10.1007/BF00278944)
- Enattah, N. S., Sahi, T., Savilahti, E., Terwilliger, J. D., Peltonen, L. & Jarvela, I. 2002 Identification of a variant associated with adult-type hypolactasia. *Nat. Genet.* **30**, 233–237. (doi:10.1038/ng826)
- Lewinsky, R. H., Jensen, T. G., Moller, J., Stensballe, A., Olsen, J. & Troelsen, J. T. 2005 T-13910 DNA variant associated with lactase persistence interacts with Oct-1 and stimulates lactase promoter activity *in vitro*. *Hum. Mol. Genet.* **14**, 3945–3953. (doi:10.1093/hmg/ddi418)
- Olds, L. C. & Sibley, E. 2003 Lactase persistence DNA variant enhances lactase promoter activity *in vitro*: functional role as a *cis* regulatory element. *Hum. Mol. Genet.* **12**, 2333–2340. (doi:10.1093/hmg/ddg244)
- Anagnostou, P., Battaglia, C., Coia, V., Capelli, C., Fabbri, C., Pettener, D., Destro-Bisol, G. & Luiselli, D. 2009 Tracing the distribution and evolution of lactase persistence in Southern Europe through the study of the T(-13910) variant. *Am. J. Hum. Biol.* **21**, 217–219. (doi:10.1002/ajhb.20851)
- Coelho, M., Sequeira, F., Luiselli, D., Beleza, S. & Rocha, J. 2009 On the edge of Bantu expansions: mtDNA, Y chromosome and lactase persistence genetic variation in southwestern Angola. *BMC Evol. Biol.* **9**, 80. (doi:10.1186/1471-2148-9-80)
- Nagy, D., Bogacsi-Szabo, E., Varkonyi, A., Csanyi, B., Czibula, A., Bede, O., Tari, B. & Rasko, I. 2009 Prevalence of adult-type hypolactasia as diagnosed with genetic and lactose hydrogen breath tests in Hungarians. *Eur. J. Clin. Nutr.* **63**, 909–912. (doi:10.1038/ejcn.2008.74)
- Sun, H. M., Qiao, Y. D., Chen, F., Xu, L. D., Bai, J. & Fu, S. B. 2007 The lactase gene –13910T allele cannot predict the lactase-persistence phenotype in north China. *Asia Pac. J. Clin. Nutr.* **16**, 598–601.
- Torniaainen, S., Parker, M. I., 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. (doi:10.1186/1471-2156-10-31)
- Mulcare, C. A., Weale, M. E., Jones, A. L., Connell, B., Zeitlyn, D., Tarekegn, A., Swallow, D., Bradman, N. & Thomas, M. 2004 The T allele of a single-nucleotide polymorphism 13.9 kb upstream of the lactase gene (LCT) (C-13.9kbT) does not predict or cause the lactase-persistence phenotype in Africans. *Am. J. Hum. Genet.* **74**, 1102–1110. (doi:10.1086/421050)
- Ingram, C. J. *et al.* 2007 A novel polymorphism associated with lactose tolerance in Africa: multiple causes for lactase persistence? *Hum. Genet.* **120**, 779–788. (doi:10.1007/s00439-006-0291-1)
- Troelsen, J. T. 2005 Adult-type hypolactasia and regulation of lactase expression. *Biochim. Biophys. Acta* **1723**, 19–32. (doi:10.1016/j.bbagen.2005.02.003)

- 34 Enattah, N. S. *et al.* 2007 Evidence of still-ongoing convergence evolution of the lactase persistence T-13910 alleles in humans. *Am. J. Hum. Genet.* **81**, 615–625. (doi:10.1086/520705)
- 35 Post, D. M. & Palkovacs, E. P. 2009 Eco-evolutionary feedbacks in community and ecosystem ecology: interactions between the ecological theatre and the evolutionary play. *Phil. Trans. R. Soc. B* **364**, 1629–1640. (doi:10.1098/rstb.2009.0012)
- 36 Bersaglieri, T., Sabeti, P. C., Patterson, N., Vanderploeg, T., Schaffner, S. F., Drake, J. A., Rhodes, M., Reich, D. & Hirschhorn, J. 2004 Genetic signatures of strong recent positive selection at the lactase gene. *Am. J. Hum. Genet.* **74**, 1111–1120. (doi:10.1086/421051)
- 37 Coelho, M., Luiselli, D., Bertorelle, G., Lopes, A. I., Seixas, S., Destro-Bisol, G. & Rocha, J. 2005 Microsatellite variation and evolution of human lactase persistence. *Hum. Genet.* **117**, 329–339. (doi:10.1007/s00439-005-1322-z)
- 38 Sabeti, P. C. *et al.* 2006 Positive natural selection in the human lineage. *Science* **312**, 1614–1620. (doi:10.1126/science.1124309)
- 39 Craig, O. E., Chapman, J., Heron, C., Willis, L. H., Bartosiewicz, L., Taylor, G., Whittle, A. & Collins, M. J. 2005 Did the first farmers of central and eastern Europe produce dairy foods? *Antiquity* **79**, 882–894.
- 40 Evershed, R. P. *et al.* 2008 Earliest date for milk use in the Near East and southeastern Europe linked to cattle herding. *Nature* **455**, 528–531. (doi:10.1038/nature07180)
- 41 Pinhasi, R., Fort, J. & Ammerman, A. J. 2005 Tracing the origin and spread of agriculture in Europe. *PLoS Biol.* **3**, e410. (doi:10.1371/journal.pbio.0030410)
- 42 Vigne, J.-D. 2006 Maîtrise et usages de l'élevage et des animaux domestiques au Néolithique: quelques illustrations au Proche-Orient et en Europe. In *Populations néolithiques et environnements* (ed. J. Guilaine), pp. 87–114. Paris, France: Errance éd.
- 43 Lomer, M. C., Parkes, G. C. & Sanderson, J. D. 2008 Review article: lactose intolerance in clinical practice—myths and realities. *Aliment. Pharmacol. Ther.* **27**, 93–103. (doi:10.1111/j.1365-2036.2007.03557.x)
- 44 Simoons, F. J. 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. (doi:10.1007/BF02235991)
- 45 Nei, M. & Saitou, N. 1986 Genetic relationship of human populations and ethnic differences in reaction to drugs and food. *Prog. Clin. Biol. Res.* **214**, 21–37.
- 46 Benecke, N. 1994 *Archäozoologische Studien zur Entwicklung der Haustierhaltung in Mitteleuropa und Südkandinavien von den Anfängen bis zum ausgehenden Mittelalter*. Schriften zur Ur- und Frühgeschichte 46. Berlin, Germany: Akademie Verlag.
- 47 Herring, D. A., Saunders, S. R. & Katzenberg, M. A. 1998 Investigating the weaning process in past populations. *Am. J. Phys. Anthropol.* **105**, 425–439. (doi:10.1002/(SICI)1096-8644(199804)105:4<425::AID-AJPA3>3.0.CO;2-N)
- 48 Lohrke, B. 2004 *Kinder in der Merowingerzeit: Gräber von Mädchen und Jungen in der Alemannia*. Rahden/Westf: Leidorf.
- 49 Cook, G. C. & Al-Torki, M. T. 1975 High intestinal lactase concentrations in adult Arabs in Saudi Arabia. *Br. Med. J.* **3**, 135–136. (doi:10.1136/bmj.3.5976.135)
- 50 Flatz, G. & Rotthauwe, H. W. 1973 Lactose nutrition and natural selection. *Lancet* **2**, 76–77. (doi:10.1016/S0140-6736(73)93267-4)
- 51 Richards, M. P., Schulting, R. J. & Hedges, R. E. 2003 Archaeology: sharp shift in diet at onset of Neolithic. *Nature* **425**, 366. (doi:10.1038/425366a)
- 52 Cordain, L., Hickey, M. S. & Kim, K. In press. Malaria and rickets represent selective forces for the convergent evolution of adult lactase persistence. In *Biodiversity in agriculture: domestication, evolution and sustainability* (eds P. Gepts, R. Bettinger, S. B. Brush, T. Famula, P. E. McGuire & C. O. Qualset). Cambridge, UK: Cambridge University Press.
- 53 Bloom, G. & Sherman, P. W. 2005 Dairying barriers affect the distribution of lactose malabsorption. *Evol. Hum. Behav.* **26**, 301–312. (doi:10.1016/j.evolhumbehav.2004.10.002)
- 54 Anderson, B. & Vullo, C. 1994 Did malaria select for primary adult lactase deficiency? *Gut* **35**, 1487–1489. (doi:10.1136/gut.35.10.1487)
- 55 Meloni, T., Colombo, C., Ogana, A., Mannazzu, M. C. & Meloni, G. F. 1996 Lactose absorption in patients with glucose 6-phosphate dehydrogenase deficiency with and without favism. *Gut* **39**, 210–213. (doi:10.1136/gut.39.2.210)
- 56 Meloni, T., Colombo, C., Ruggiu, G., Dessena, M. & Meloni, G. F. 1998 Primary lactase deficiency and past malarial endemicity in Sardinia. *Ital. J. Gastroenterol. Hepatol.* **30**, 490–493.
- 57 Gillis, R. E. 2003 A study of the development of lactose tolerance, as an example of nutritional genetic adaptation to domesticated animal products in the Old World populations, through the analysis of archaeological and biomolecular evidence. M.Sc. thesis, Sheffield University, Sheffield, UK.
- 58 Ihara, Y. 2011 Evolution of culture-dependent discriminate sociality: a gene–culture coevolutionary model. *Phil. Trans. R. Soc. B* **366**, 889–900. (doi:10.1098/rstb.2010.0247)
- 59 Dudd, S. N. & Evershed, R. P. 1998 Direct demonstration of milk as an element of archaeological economies. *Science* **282**, 1478–1481. (doi:10.1126/science.282.5393.1478)
- 60 Vigne, J. D. 2008 Zooarchaeological aspects of the Neolithic diet transition in the Near East and Europe, and their putative relationships with the Neolithic demographic transition. In *The Neolithic demographic transition and its consequences* (eds J. P. Bocquet-Appel & O. Bar-Yosef), pp. 179–205. Berlin, Germany: Springer.
- 61 Burger, J. & Thomas, M. G. 2011 The palaeopopulation genetics of humans, cattle and dairying in Neolithic Europe. In *The bioarchaeology of the transition to agriculture* (eds R. Pinhasi & J. Stock), pp. 371–384. Chichester, UK: Wiley Blackwell.
- 62 Ammerman, A. J. & Cavalli-Sforza, L. L. 1984 *The Neolithic transition and the genetics of populations in Europe*. Princeton, NJ: Princeton University Press.
- 63 Whittle, A. & Cummings, V. 2007 *Going over: the Mesolithic–Neolithic transition in North-West Europe*. New York, NY: Oxford University Press.
- 64 Guimaraes, S. *et al.* 2009 Genealogical discontinuities among Etruscan, Medieval, and contemporary Tuscans. *Mol. Biol. Evol.* **26**, 2157–2166. (doi:10.1093/molbev/msp126)
- 65 Semino, O. *et al.* 2000 The genetic legacy of Paleolithic *Homo sapiens sapiens* in extant Europeans: a Y chromosome perspective. *Science* **290**, 1155–1159. (doi:10.1126/science.290.5494.1155)
- 66 Currat, M. & Excoffier, L. 2005 The effect of the Neolithic expansion on European molecular diversity. *Proc. R. Soc. B* **272**, 679–688. (doi:10.1098/rspb.2004.2999)

- 67 Chikhi, L., Nichols, R. A., Barbujani, G. & Beaumont, M. A. 2002 Y genetic data support the Neolithic demic diffusion model. *Proc. Natl Acad. Sci. USA* **99**, 11 008–11 013. (doi:10.1073/pnas.162158799)
- 68 Dupanloup, I., Bertorelle, G., Chikhi, L. & Barbujani, G. 2004 Estimating the impact of prehistoric admixture on the genome of Europeans. *Mol. Biol. Evol.* **21**, 1361–1372. (doi:10.1093/molbev/msh135)
- 69 Belle, E. M., Landry, P. A. & Barbujani, G. 2006 Origins and evolution of the Europeans' genome: evidence from multiple microsatellite loci. *Proc. R. Soc. B* **273**, 1595–1602. (doi:10.1098/rspb.2006.3494)
- 70 Richards, M. *et al.* 2000 Tracing European founder lineages in the Near Eastern mtDNA pool. *Am. J. Hum. Genet.* **67**, 1251–1276. (doi:10.1016/S0002-9297(07)62954-1)
- 71 Bramanti, B. *et al.* 2009 Genetic discontinuity between local hunter-gatherers and central Europe's first farmers. *Science* **326**, 137–140. (doi:10.1126/science.1176869)
- 72 Guilaine, J. 2003 *De la Vague a la Tombe: la conquete neolithique de la Mediterranee (8000-2000 avant J.C.)*. Paris, France: Vle Editions du Seuil.
- 73 Helmer, D., Gourichon, L., Monchot, H., Peters, J. & Sana Segui, M. 2005 Identifying early domestic cattle from pre-pottery Neolithic sites on the Middle Euphrates using sexual dimorphism. In *First steps of animal domestication new archaeozoological approach* (eds J. D. Vigne, D. Helmer & J. Peters), pp. 86–95. Oxford, UK: Oxbow.
- 74 Peters, J., Von Den Driech, A. & Helmer, D. 2005 The Upper Euphrates–Tigris Basin: cradle of agro-pastoralism? In *First steps of animal domestication new archaeozoological approach* (eds J. D. Vigne, D. Helmer & J. Peters), pp. 96–124. Oxford, UK: Oxbow.
- 75 Riede, F. 2011 Adaptation and niche construction in human prehistory: a case study from the southern Scandinavian Late Glacial. *Phil. Trans. R. Soc. B* **366**, 793–808. (doi:10.1098/rstb.2010.0266)
- 76 Peters, J., Helmer, D., Von Den Driech, A. & Sana Segui, M. 1999 Early animal husbandry in the Northern Levant. *Paleorient* **25**, 27–47. (doi:10.3406/paleo.1999.4685)
- 77 Vigne, J. D., Carrere, I. & Guilaine, J. 2001 Unstable status or early domestic ungulates in the Near East: the example of Shillourokambos (Cyprus, IX–VIIIth millennia cal. B.C.). In *Actes du Colloque International Organise par le Departement des Antiquites de Chypre et l'Ecole Francaise d'Athenes*, 17–19 May 2003, Nicosia (Cyprus).
- 78 Vigne, J. D., Zazzo, A., Saliege, J. F., Poplin, F., Guilaine, J. & Simmons, A. 2009 Pre-Neolithic wild boar management and introduction to Cyprus more than 11,400 years ago. *Proc. Natl Acad. Sci. USA* **106**, 16 135–16 138.
- 79 Özdoğan, M. & Basgelen, N. 1999 *Neolithic in Turkey. The cradle of civilization*. Istanbul: Ancient Anatolians Civilizations Series 3.
- 80 Tresset, A. & Vigne, J. D. 2007 Substitution of species, techniques and symbols at the Mesolithic–Neolithic transition in Western Europe. In *Going over: the Mesolithic–Neolithic transition in the North-West Europe* (eds A. Whittle & V. Cummings), pp. 189–210. Oxford, UK: Oxford University Press.
- 81 Bollongino, R., Elsner, J., Vigne, J. D. & Burger, J. 2008 Y-SNPs do not indicate hybridisation between European aurochs and domestic cattle. *PLoS ONE* **3**, e3418. (doi:10.1371/journal.pone.0003418)
- 82 Edwards, C. J. *et al.* 2007 Mitochondrial DNA analysis shows a Near Eastern Neolithic origin for domestic cattle and no indication of domestication of European aurochs. *Proc. R. Soc. B* **274**, 1377–1385. (doi:10.1098/rspb.2007.0020)
- 83 Troy, C. S., MacHugh, D. E., Bailey, J. F., Magee, D. A., Loftus, R. T., Cunningham, P., Chamberlain, A. T., Sykes, B. C. & Bradley, D. G. 2001 Genetic evidence for Near-Eastern origins of European cattle. *Nature* **410**, 1088–1091. (doi:10.1038/35074088)
- 84 Larson, G. *et al.* 2007 Ancient DNA, pig domestication, and the spread of the Neolithic into Europe. *Proc. Natl Acad. Sci. USA* **104**, 15 276–15 281. (doi:10.1073/pnas.0703411104)
- 85 Copley, M. S., Berstan, R., Dudd, S. N., Docherty, G., Mukherjee, A. J., Straker, V., Payne, S. & Evershed, R. P. 2003 Direct chemical evidence for widespread dairying in prehistoric Britain. *Proc. Natl Acad. Sci. USA* **100**, 1524–1529. (doi:10.1073/pnas.0335955100)
- 86 Copley, M. S., Berstan, R., Mukherjee, A. J., Dudd, S. N., Straker, V., Payne, S. & Evershed, R. P. 2005 Dairying in antiquity III. Evidence from absorbed lipid residues dating to the British Neolithic. *J. Archaeol. Sci.* **32**, 523–546. (doi:10.1016/j.jas.2004.08.006)
- 87 Craig, O. E., Taylor, G., Mulville, J., Collins, M. & Parker Pearson, M. 2005 The identification of prehistoric dairying activities in the western Isles of Scotland: an integrated biomolecular approach. *J. Archaeol. Sci.* **32**, 91–103. (doi:10.1016/j.jas.2004.06.009)
- 88 Burger, J., Kirchner, M., Bramanti, B., Haak, W. & Thomas, M. G. 2007 Absence of the lactase-persistence-associated allele in Early Neolithic Europeans. *Proc. Natl Acad. Sci. USA* **104**, 3736–3741. (doi:10.1073/pnas.0607187104)
- 89 Fornander, E., Eriksson, G. & Liden, K. 2008 Wild at heart: approaching pitted ware identity, economy and cosmology through stable isotopes in skeletal material from the Neolithic site Korsnas in Eastern Central Sweden. *J. Anthropol. Archaeol.* **27**, 281–297. (doi:10.1016/j.jaa.2008.03.004)
- 90 Milner, N., Craig, O. E., Bailey, G. N., Pedersen, K. & Andersen, S. H. 2004 Something fishy in the Neolithic? A re-evaluation of stable isotope analysis of Mesolithic and Neolithic coastal populations. *Antiquity* **78**, 9–22.
- 91 Evershed, R. P. 2007 Exploiting molecular and isotopic signals at the Mesolithic–Neolithic transition. In *Going over: the Mesolithic–Neolithic transition in North-West Europe* (eds A. Whittle & V. Cummings), pp. 141–164. New York, NY: Oxford University Press.
- 92 Boric, D., Grupe, G., Peters, J. & Zivko, M. 2004 Is the Mesolithic–Neolithic subsistence dichotomy real? New stable isotope evidence from the Danube Gorges. *Eur. J. Archaeol.* **7**, 221–248. (doi:10.1177/1461957104056500)
- 93 Malmström, H. *et al.* 2009 Ancient DNA reveals lack of continuity between Neolithic hunter-gatherers and contemporary Scandinavians. *Curr. Biol.* **19**, 1758–1762. (doi:10.1016/j.cub.2009.09.017)
- 94 Edmonds, C. A., Lillie, A. S. & Cavalli-Sforza, L. L. 2004 Mutations arising in the wave front of an expanding population. *Proc. Natl Acad. Sci. USA* **101**, 975–979. (doi:10.1073/pnas.0308064100)
- 95 Klopstein, S., Currat, M. & Excoffier, L. 2006 The fate of mutations surfing on the wave of a range expansion. *Mol. Biol. Evol.* **23**, 482–490. (doi:10.1093/molbev/msj057)
- 96 Aoki, K. A. 1986 Stochastic model of gene–culture coevolution suggested by the ‘culture historical hypothesis’ for the evolution of adult lactose absorption in humans. *Proc. Natl Acad. Sci. USA* **83**, 2929–2933. (doi:10.1073/pnas.83.9.2929)

- 97 Cavalli-Sforza, L. L. & Feldman, M. W. 1976 Evolution of continuous variation: direct approach through joint distribution of genotypes and phenotypes. *Proc. Natl Acad. Sci. USA* **73**, 1689–1692. (doi:10.1073/pnas.73.5.1689)
- 98 Feldman, M. & Cavalli-Sforza, L. 1989 On the theory of evolution under genetic and cultural transmission with application to the lactase absorption problem. In *Mathematical evolutionary theory* (ed. M. Feldman), pp. 145–173. Princeton, NJ: Princeton University Press.
- 99 Bocquet-Appel, J. P. & Naji, S. 2006 Testing the hypothesis of a worldwide Neolithic demographic transition. Corroboration from American Cemeteries (with comments). *Curr. Anthropol.* **47**, 341–365. (doi:10.1086/498948)
- 100 Guilaine, J. & Manen, C. 2007 From Mesolithic to Early Neolithic in the western Mediterranean. In *Going over: the Mesolithic–Neolithic transition in the North–West Europe* (eds A. Whittle & V. Cummings), pp. 21–52. New York, NY: Oxford University Press.
- 101 Hofer, T., Ray, N., Wegmann, D. & Excoffier, L. 2009 Large allele frequency differences between human continental groups are more likely to have occurred by drift during range expansions than by selection. *Am. Hum. Genet.* **73**, 95–108. (doi:10.1111/j.1469-1809.2008.00489.x)
- 102 Beaumont, M. A., Zhang, W. & Balding, D. J. 2002 Approximate Bayesian computation in population genetics. *Genetics* **162**, 2025–2035.
- 103 Sousa, V. C., Fritz, M., Beaumont, M. A. & Chikhi, L. 2009 Approximate Bayesian computation without summary statistics: the case of admixture. *Genetics* **181**, 1507–1519. (doi:10.1534/genetics.108.098129)
- 104 Silver, M. & Di Paolo, E. 2006 Spatial effects favour the evolution of niche construction. *Theor. Popul. Biol.* **70**, 387–400. (doi:10.1016/j.tpb.2006.08.003)
- 105 Itan, Y., Powell, A., Beaumont, M. A., Burger, J. & Thomas, M. G. 2009 The origins of lactase persistence in Europe. *PLoS Comput. Biol.* **5**, e1000491. (doi:10.1371/journal.pcbi.1000491)
- 106 Bellwood, P. S. 2005 *The first farmers: the origins of agricultural societies*. Malden, MA: Blackwell Publishing.
- 107 Hassan, F. A. 1981 *Demographic archaeology*. New York, NY: Academic Press.
- 108 Maynard Smith, J. 1998 *Evolutionary genetics*, 2nd edn. Oxford, UK: Oxford University Press.
- 109 Benecke, N. 1994 *Der Mensch und seine Haustiere. Die Geschichte einer jahrtausendealten Beziehung*. Stuttgart, Germany: Theiss Verlag.
- 110 Currat, M. & Excoffier, L. 2004 Modern humans did not admix with Neanderthals during their range expansion into Europe. *PLoS Biol.* **2**, e421. (doi:10.1371/journal.pbio.0020421)
- 111 Durham, W. H. 1991 *The relationship of genes and culture. Coevolution: genes, culture and human diversity*, pp. 154–225. Stanford, CA: Stanford University Press.
- 112 Whittle, A. 2007 Going over: people and their times. In *Going over: the Mesolithic–Neolithic transition in the North–West Europe* (eds A. Whittle & V. Cummings), pp. 617–628. Oxford, UK: Oxford University Press.
- 113 Nielsen, R. 2009 Adaptionism: 30 years after Gould and Lewontin. *Evolution* **63**, 2487–2490. (doi:10.1111/j.1558-5646.2009.00799.x)
- 114 Enattah, N. S. *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. (doi:10.1016/j.ajhg.2007.09.012)
- 115 Hermisson, J. & Pennings, P. S. 2005 Soft sweeps: molecular population genetics of adaptation from standing genetic variation. *Genetics* **169**, 2335–2352. (doi:10.1534/genetics.104.036947)
- 116 Pennings, P. S. & Hermisson, J. 2006 Soft sweeps II: molecular population genetics of adaptation from recurrent mutation or migration. *Mol. Biol. Evol.* **23**, 1076–1084. (doi:10.1093/molbev/msj117)
- 117 Excoffier, L. & Ray, N. 2008 Surfing during population expansions promotes genetic revolutions and structuration. *Trends Ecol. Evol.* **23**, 347–351. (doi:10.1016/j.tree.2008.04.004)
- 118 Francois, O., Currat, M., Ray, N., Han, E., Excoffier, L. & Novembre, J. 2010 Principal component analysis under population genetic models of range expansion and admixture. *Mol. Biol. Evol.* **27**, 1257–1268. (doi:10.1093/molbev/msq010)
- 119 Novembre, J. & Di Rienzo, A. 2009 Spatial patterns of variation due to natural selection in humans. *Nat. Rev. Genet.* **10**, 745–755. (doi:10.1038/nrg2632)