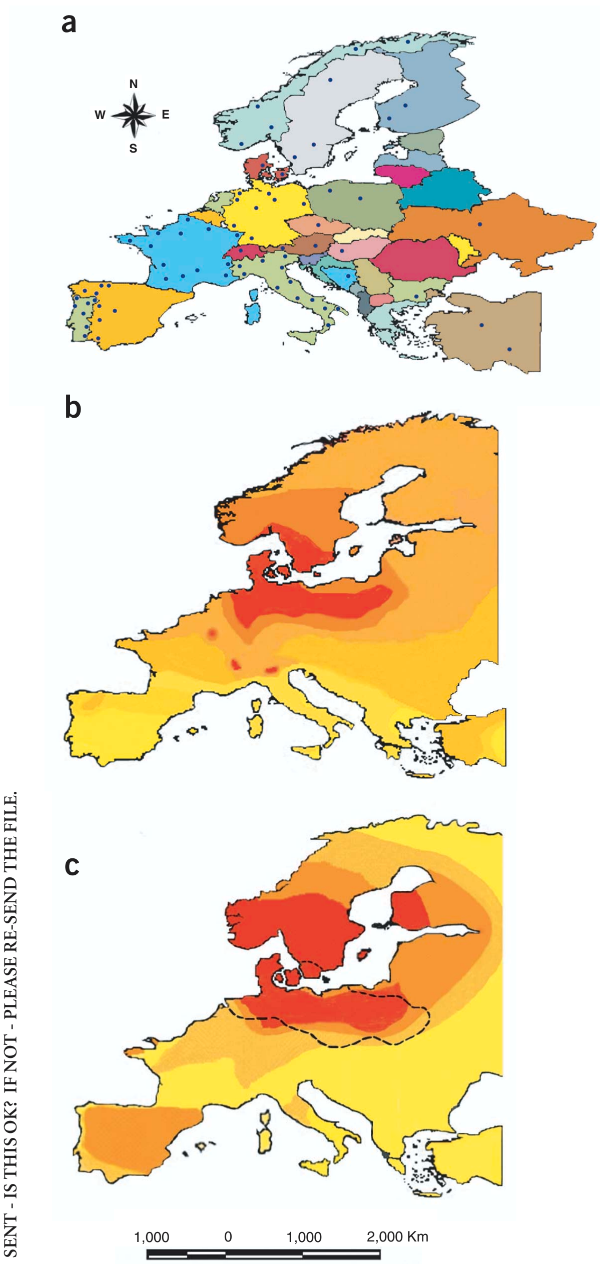
Gene-culture coevolution between cattle milk protein genes and human lactase genes

Albano Beja-Pereira[1,2], Gordon Luikart[2], Phillip R England[2], Daniel G Bradley[3], Oliver C Jann[4], Giorgio Bertorelle[5], Andrew T Chamberlain[6], Telmo P Nunes[7], Stoitcho Metodiev[8], Nuno Ferrand[2,9] & Georg Erhardt[4]

Milk from domestic cows has been a valuable food source for over 8,000 years, *[AUTHOR: REF. CITATIONS ARE NOT PERMITTED IN ABSTRACT.]* especially in lactose-tolerant human societies that exploit dairy breeds. We studied geographic patterns of variation in genes encoding the six most important milk protein in 70 native European cattle breeds. We found substantial geographic coincidence between high diversity in cattle milk genes, locations of the European Neolithic cattle farming sites (>5,000 years ago) and present-day lactose tolerance in Europeans. This suggests a gene-culture coevolution between cattle and humans.

Some, but not all, human populations have the genetically determined ability to digest milk lactose in adulthood, thereby benefiting from the rich food resources in cow's milk[1]. These societies (*e.g.*, Northern Europe) are lactose-tolerant and highly dependent on milk products. Lactose tolerance is an example of selection-based evolutionary change in humans from milk-drinking cultures[2]. Has there also been a detectable evolutionary change in the gene pool of domestic cattle from these cultures?

Our study of nonsynonymous mutations in [AUTHOR: CORRECT?] six milk protein genes in ~20,000 cattle from 70 breeds across



AUTHOR: SCALE IS MISSING FROM THE POWERPOINT FILE YOU SENT - IS THIS OK? IF NOT - PLEASE RE-SEND THE FILE.

Figure 1 ***[AUTHOR: PLEASE PROVIDE BRIEF TITLE FOR WHOLE OF FIG. 1.]*** (**a**) Geographic distribution of the 70 cattle breeds (blue dots) across Europe and Turkey. (**b**) Synthetic map showing the first principal component resulting from the allele frequencies at the cattle genes. (**c**) Geographic distribution of the lactase persistence allele in contemporary Europeans (including thousands of individuals from nearly all countries). The dashed black line indicates the limits of the geographic distribution of early Neolithic cattle pastoralist (Funnel Beaker Culture) inferred from archaeological data[15]. **[AUTHOR: EXPLAIN THE LEGEND FOR B & C].**

[1]LECA - Génomique des Populations et Biodiversité, CNRS UMR 5553, Université Joseph Fourier, B.P. 53, 38041 Grenoble, Cedex 9, France. [2]Centro de Investigação em Biodiversidade e Recursos Genéticos (CIBIO-UP) and Secção Autónoma de Ciências Agrárias, Faculdade de Ciências, Universidade do Porto, Campus Agrário de Vairão, Rua Padre Armando Quintas, 4485-661 Vairão (VCD), Portugal. [3]Department of Genetics, Smurfit Institute, Trinity College, Dublin 2, Ireland. [4]Institut für Tierzucht und Haustiergenetik, Justus-Liebig-Universität Gieβen, Ludwigstr. 21b, 35390 Gieβen, Germany. [5]Department of Biology, University of Ferrara, 44100 Ferrara, Italy. [6]Department of Archaeology and Prehistory, University of Sheffield, Sheffield S1 4ET, UK. [7]CIISA/UISEE/DETSA, Faculdade de Medicina Veterinária, Polo Universitário da Ajuda, Rua Prof. Cid dos Santos, 1300-477 Lisboa, Portugal. [8]Department of Genetics and Animal Breeding, Thracian University, Agricultural Faculty, 6000 Stara Zagora, Bulgaria. [9]Departamento de Zoologia/Antropologia da Faculdade de Ciências Praça Gomes Teixeira, 4099-002 Porto, Portugal. Correspondence should be addressed to A.B.-P. (albano.beja-pereira@ujf-grenoble.fr).

Published online ′ November 2003; doi:10.1038/ng′

Table 1 Spearman correlation coefficient values between the first principal component from the milk protein gene frequencies, the lactase persistence allele frequency and the presence or absence of archeological evidence for Neolithic cattle pastoralists

	Spearman correlation	Degrees of freedom	*P*
First principal component versus neolithic cattle pastoralists	–0.750	21.2–27.3	<0.0005
First principal component versus lactase persistence allele frequency	–0.593	17.7–24.6	<0.01
Neolithic cattle pastoralists versus lactase persistence allele frequency	0.730	19.3–24.8	<0.0005

Europe (**Fig. 1a**) found high allelic richness and genetic distinctiveness in the native cattle from North Central Europe (NCE), as illustrated by the synthetic map of cattle milk protein genes (**Fig. 1b** and **Supplementary Tables 1** and **2**, **Supplementary Fig. 1** and **Supplementary Methods** online). Notably, this synthetic map (inner contour) closely matches the European distribution of the allele for human lactase persistence that is most frequent in NCE ($P < 0.0005$; **Table 1**). This is in stark contrast to the lower levels of lactose tolerance found in people of Southern Europe and the Near East. There was also strong concordance ($P < 0.001$) of the geographic distribution of cattle milk gene diversity with the early Neolithic distribution of a European cattle pastoralist society[3] (**Fig. 1c**).

How can we explain the strong geographic concordance between cattle milk gene diversity, human lactose tolerance and the distribution of the earliest European cattle pastoralists? We propose that since Neolithic times, there has been gene-culture coevolution between the domestic cattle and human culture driven by the advantages conferred by milk consumption. This led to the maintenance of larger herds and selection for increased milk yield and altered milk protein composition. This coevolution seemingly influenced the frequencies of the important milk protein genes in cattle and the gene encoding lactase in humans. In fact, a recent study suggested that the relatively old variant for lactose tolerance was only recently driven to high frequencies in North Central Europeans after the introduction of dairy culture in this region[4].

This scenario is also supported by evidence for selection at milk protein loci in bovids[5]. For example, directional selection can explain high intraspecific divergence and low intraspecific polymorphism in *k*-casein sequences across bovids[5]. Our data also show patterns consistent with selection: 19 NCE breeds deviated significantly from neutrality (in Ewens-Watterson test, 32% of 114 tests with $P < 0.01$ versus 2% of 306 tests with $P < 0.01$ in the 51 non-NCE breeds; in Fu test and Tajima test, all NCE breeds showed $P < 0.05$ versus 4 of 51 non-NCE cattle).

Our genetic data corroborate recent archaeological evidence suggesting that the early European cattle pastoralists in NCE were dependent on milk[6,7], as early Neolithic sites in NCE are rich in cattle remains[4]. Based on the analysis of intratooth change in nitrogen isotope ratios from archaeological cattle teeth, it seems that cattle herds were managed for early weaning of calves, making cow's milk more available for human consumption. Meat production, practiced outside NCE, necessitates later weaning to optimize weight gain[7].

Among several phenomena that might have shaped our data, selection seems the most probable explanation. [**AUTHOR: OK AS EDITED?**] Recent studies have shown that high diversity in human genes can evolve rapidly due to selection[8]. In addition, analysis of bovine myostatin alleles showed signals of balancing selection in a number of independently occurring mutations that cause double-muscling in beef breeds[9].

Given that population surveys of mtDNA sequence, microsatellite markers and classical polymorphisms in European cattle breeds show no evidence of elevated diversity in NCE[12-14], it is likely that selection pressure imposed by early pastoralists and their successors in different regions of NCE has left the legacy of high allelic diversity at these specific milk genes. [**AU: OK AS EDITED? PLEASE CHECK REFERENCES ARE CORRECT**] It is also possible that some of the diversity represents relatively recent mutation (<10,000 years), although, under a neutral model, mutation rates are too low (10^{-6}–10^{-9}; ref. 5) for this to be a primary factor. Selection may have maintained many favorable new mutations by protecting them from the normal process of attrition due to drift.

Another possible source of the unique diversity found in cattle in NCE is historical gene flow from an as yet unidentified origin. Two candidates for this source are local wild aurochs (*Bos primigenius*), which persisted in NCE until the sixteenth century, and domestic cattle other than those that gave rise to present day European cattle (outside NCE). Extensive wild ox introgression seems unlikely, and no mtDNA sequences have been detected in European cattle which match aurochs sequences identified using ancient DNA sequencing[14]. [**AU: OK AS EDITED? PLEASE CHECK REFERENCE IS CORRECT**].

Notably, our findings contradict the results of previous surveys of genetic variation in European cattle[11–13], which suggested that diversity declines with distance from the Fertile Crescent region. This discrepancy could be explained by selection on the milk genes, and it may also reflect different sampling strategies. Our analysis is based on a sample set that is unprecedented in size, geographic coverage and breed diversity. Furthermore, unlike previous studies, we analyzed only nonsynonymous polymorphisms in strong candidate genes most likely to yield unusual geographic patterns in milk gene diversity.

Our study provides evolutionary insights and identifies high diversity in cattle genes that are economically important, suggesting that cattle in NCE are a potentially precious genetic resource for future agricultural productivity. Farming practices since the Neolithic seem to have left reciprocal genetic signatures in cattle and human populations from NCE. This may represent a rare example of cultural-genetic coevolution between humans and another species. Other examples of coevolution have been documented for human genes and genes of parasites, such as *Plasmodium*[14]. But our study represents the first non-disease-related example of genetic coevolution between humans and domestic animals, reflecting the extent to which domestication has shaped human societies and the genomes of both humans and cattle.

Note: Supplementary information is available on the Nature Genetics website.

ACKNOWLEDGMENTS
We thank M. Zvelebil for ideas and discussion. A.B.-P. is supported by a grant from Fundação para a Ciência e Tecnologia through the Graduate Programme in Areas of Basic and Applied Biology, and the work was partially supported by a Praxis project grant. G.L. and P.R.E. were funded the European Union (Econogene). D.G.B. is a Science Foundation Ireland Investigator.

COMPETING INTERESTS STATEMENT
The authors declare that they have no competing financial interests.

Received 19 June; accepted 29 October 2003
Published online at http://www.nature.com/naturegenetics/

1. Schrimshaw, N.S. & Murray, E.B. *Am. J. Clin. Nutr.* **48**, 1059–1179 (1988).
2. Feldman, M.W. & Cavalli-Sforza, L.L. in *Mathematical evolutionary theory* (ed. Feldman, M.W.) 145–173 (Princeton University Press, Princeton, New Jersey, 1989).
3. Midgley, M.S. *TRB Culture: The First Farmers of the North European Plain* (Edinburgh University Press, Edinburgh, 1992).
4. Enattah, N.S. *et al. Nat. Genet.* **30**, 233–237 (2002).
5. Ward, R., Honeycutt, L. & Derr, J.N. *Genetics* **147**, 1863–1872 (1997).
6. Dudd, S. & Evershed, P. *Science* **282**, 1478–1480 (1998).
7. Balasse, M. & Tresset, A. *J. Archaeol. Sci.* **29**, 853–859 (2002).
8. Tishkoff, S.A. *et al. Science* **293**, 455–462 (2001).
9. Dunner, S. *et al. Genet. Sel. Evol.* **35**, 103–118 (2003).
10. Troy, C. *et al. Nature* **410**, 1088–1091 (2001).
11. Loftus, R.T. *et al. Mol. Ecol.* **8**, 2015–2022 (1999).
12. MacHugh, D.E., Loftus, R.T., Cunningham, P. & Bradley, D.G. *Anim. Genet.* **29**, 333–340 (1998).
13. Medjugorac, I., Kustermann, W., Lazar, P., Russ, I. & Pirchner, F. *Anim. Genet.* **25**, 19–27 (1994).
14. Hill, A.V., Jepson, A., Plebanski, M. & Gilbert, S.C. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **352**, 1317–1325 (1997).
15. Zvelebil, M. in *Archaeogenetics: DNA and the Population History of Europe* (ed. Boyle, K.) 57–79 (MacDonald Institute Cambridge, Cambridge, 2000).