

Livestock genetic origins: Goats buck the trend

David E. MacHugh*[†] and Daniel G. Bradley[‡]

*Department of Animal Science and Production and Conway Institute of Biomolecular and Biomedical Research, Faculty of Agriculture, University College, Dublin 4, Ireland; and [‡]Department of Genetics, Smurfit Institute, Trinity College, Dublin 2, Ireland

The domestic goat (*Capra hircus*) often is dismissed as the “poor man’s cow” for its ability to thrive on meager fodder and cope with harsh environments. However, this belies the economic and archaeological importance of the species. From an agricultural standpoint, the world’s 700 million goats provide reliable access to meat, milk, skins, and fiber for small farmers—particularly in developing countries. In addition, accumulating archaeological evidence indicates that goats, in the form of their wild progenitor—the bezoar (*Capra aegagrus*), were the first wild herbivores to be domesticated (Fig. 1). These studies suggest that this happened ≈10,000 years ago at the dawn of the Neolithic in the region known as the Fertile Crescent (1, 2). Therefore these sturdy animals may have been the first “walking larders” and by example, could have triggered subsequent domestications of the full repertoire of Euroasian livestock species that have provided the bulk of the animal protein consumed by ever-expanding human populations.

In this issue of PNAS, Luikart and colleagues (3) add *C. hircus* to the growing list of domestic animals that have been widely surveyed for mtDNA sequence variation (4–9). In their survey, Luikart *et al.* (3) demonstrate that the structure and distribution of mtDNA variation in domestic goats are qualitatively different from the patterns observed in other large Eurasian herbivores domesticated for food, skins and fiber (cattle, buffalo, pigs, and sheep; see Fig. 2).

Diversity, Capture, and Genetic Inertia

There has been a long tradition of interest in animal domestic origins among breeders, geneticists, and archaeologists. With the maturation of molecular population genetics during the last decade, the tools are now available to systematically investigate the problem at the phylogeographic level. In this regard, mtDNA has represented the most informative genomic element for teasing out the what, where and (admittedly with less confidence), the when of livestock domestication. Mammalian mitochondrial chromosomes display a maternal mode of genetic transmission and an absence of recombination (10) and



Fig. 1. A third or fourth century BC Mesopotamian stone carving of a man carrying either a domestic or a wild goat. This piece is displayed in the Louvre Museum in Paris. (Figure courtesy of Mike Schwartz.)

are subject to a relatively rapid mutation rate (particularly in control region sequences). Because of these genetic features, mtDNA studies of livestock provide valuable information about the domestication process.

First, clonal transmission of intact mtDNA haplotypes sans recombinational noise means that it is possible to discern discrete maternal lineages within domestic populations that may have complex genetic histories. Consequently, sequences that descend from different captures from a diverse wild species maintain a phylogenetic distinction even after millennia of domestic interbreeding. Second, the rates of substitution accumulation within sequences of moderate length are of a similar order to the time depth of domestication. This allows the resolution of predomestic and postdomestic patterns of sequence diversity—even allowing for the calibration difficulties that are inherent in control region variation. Third, there

seems to be significant temporal inertia in the geography of domestic mtDNA. The high disparity between male and female reproductive variance under managed breeding implies that maternal lineages are likely to show some geographical inertia, especially with respect to secondary introgression. This would suggest that genetic change has been predominantly male-mediated. For example, all African cattle sampled to date display uniformly *Bos taurus* mtDNA sequences despite a widespread and substantial *Bos indicus* introgression from the East over several millennia (11).

Finally, a widely perceived weakness of mtDNA phylogeography is that it deals with only a single segregating locus and one not always representative of the ancestry of a whole genome. However, in the context of animal domestication this may actually be an advantage. Unlike other genomic components, a domestic mtDNA lineage must at some point in its history have entered the domestic pool through the physical capture of a wild female animal. Nuclear gene histories may have been complicated through more ephemeral encounters between wild males and tame females (as still observed in species such as Asian mithan where the progenitor remains accessible).

An East-West Duality in Animal Domestication?

Fig. 2 shows neighbor-joining trees that summarize mtDNA sequence diversity in five domestic ungulate data collections. The most striking feature (with the exception of the horse) is that sequences invariably cluster into one of two groups. These distinct clades are separated by a dominant internal branch—a topology akin to a double-headed broomstick. A second important feature of the diversity is that, in each case, the two clades have a tendency to be geographically distributed—primarily along an East-West division. The phylogenetic divisions within cattle and water buffalo follow (with some qual-

See companion article on page 5927.

[†]To whom reprint requests should be addressed. E-mail: david.machugh@ucd.ie.

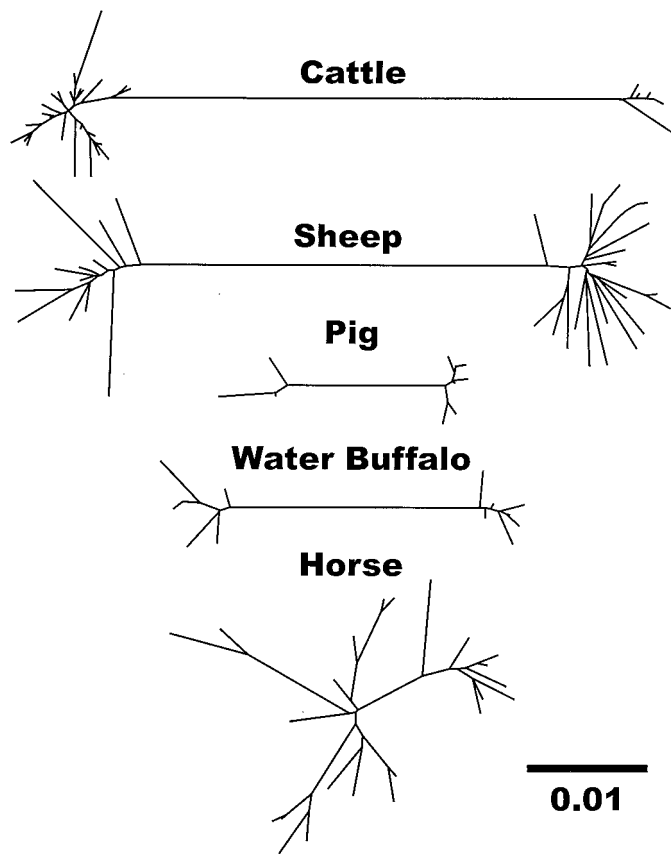


Fig. 2. Unrooted neighbor-joining phylogenies constructed by using uncorrected mtDNA sequence divergences. These data are as follows: 43 complete cattle control region sequences sampled in disparate locations on three continents (4, 14); 42 complete sheep control region sequences from a mixed breed New Zealand sample (16); 11 partial pig control region sequences that include both East Asian and European varieties (15); 15 cytochrome *b* sequences from water buffalo (12, 13); and 18 complete horse control region sequences (25).

ification) established taxonomic categorizations based on morphology, i.e., that between humped and humpless cattle and that between river and swamp buffalo (4, 12–14). The division in pigs is also between European samples and those from Asia (8, 15). The pattern in sheep is less clear, and the sequences used here were sampled from New Zealand herds of necessarily exotic origin (16). However with the corresponding restriction fragment length polymorphism haplotypes, one was found in a European sheep sample only and the second was detected in animals of both European and Central Asian provenance (7).

A third piece of information inherent in these phylogenies is the quantitative divergence between the dual clades. Under the assumption of a molecular clock, it is possible to estimate the coalescence time between the two clusters in each species. In each case, the time to the most recent common ancestor is estimated in hundreds of thousands of years (4, 7–9, 12, 13, 15). Notwithstanding the difficulties in calibrating mtDNA sequence divergence,

particularly within the control region, these estimates comfortably predate the time depth of domestic history, which stretches to only *ca.* 10,000 years ago (17). Finally, in the first four species, the diversity within each sister clade is of similar magnitude, although the numbers sampled in each sometimes differ. Therefore taking these factors into account, the diversity within each of the broom-head clusters probably is derived from the attenuated sampling of a subset of wild variation that a domestication event (or series of localized events) would entail. The extant sequence diversity also would comprise the limited variation that has accrued through mutation within domestic history. The quantitative divergence between each pair of clusters combined with their East-West geographical separation provides strong support for at least two domestication centers for cattle, sheep, pig, and water buffalo. The last species, horse, yields a more complex pattern that suggests a domestication process on the Eurasian steppes that was not so constrained within time and space (9).

Domestic Goat Phylogeography

The paper in this issue by Luikart *et al.* (3) describes the first comprehensive description of goat mtDNA variation and brings closure to the first phase of molecular study of the origins of the major Eurasian domestic animals. It highlights similarities, but also interesting differences with the other Fertile Crescent species. In the initial part of the study, Luikart and his colleagues (3) sequenced the first hypervariable segment of the mtDNA control region from 406 goats, broadly sampled across the Old World. They also analyzed the same mtDNA segment from 14 wild *Capra* species, including four wild bezoar (*C. aegagrus*). When a phylogenetic tree is constructed from these data it suggests that goats, unlike cattle, sheep, or pigs (see Fig. 2) seem to have three, not two, deep matrilineal roots (termed lineages A, B, and C, respectively). Surprisingly, none of the four *C. aegagrus* sequences emerge within any of these three domestic haplogroups. However, this may be an artifact of the small bezoar sample because a previous, albeit less comprehensive study from the same laboratory has found *C. aegagrus* museum samples that do cluster within a *C. hircus* mtDNA haplogroup (18).

To gain a better understanding of the relationships among the domestic goat haplogroups, Luikart *et al.* (3) then turned their attention to a region of the mtDNA molecule that evolves less rapidly—the cytochrome *b* gene. They selected six individual goat samples, two from each of the three lineages and obtained complete 1,140-bp gene sequences. Using published sheep cytochrome *b* sequences and established dates for the evolutionary split between *Capra* and *Ovis*, Luikart *et al.* (3) estimate a coalescence time for domestic goat mtDNA of 200–280 thousand years ago. Armed with this information they infer that there were at least three geographically and temporally separate captures of founder female bezoar goats during the formation of early domestic populations. As Luikart *et al.* (3) demonstrate it is unlikely that a single local domestication could explain the observed pattern, because the number of reproductively active females necessary to maintain the three ancestral haplotypes in the source population would have been improbably large.

Mismatch distribution analysis is a useful statistical method to infer the ancient genetic demography of a population (19). Luikart *et al.* (3) use this technique to reveal patterns suggestive of population expansion in all three domestic goat lineages. Assuming that lineage A (the most diverse haplogroup) started expanding with domestication 10,000 years ago, they then estimate corresponding expansion

start points for lineages B and C of about 2,000 and 6,000 years ago, respectively. This finding suggests to Luikart *et al.* (3) that the minor haplogroups derive from secondary and tertiary expansions that trail the initial expansion of lineage A. However, it is important to emphasize that additional samples may alter this preliminary snapshot of goat genetic history. In addition, the wild diversity involved in the original captures may have been quite different. Further work on the phylogeography within the major haplogroups should help to clarify these points.

Another intriguing aspect of Luikart *et al.*'s (3) study is the biogeography of the three domestic goat mtDNA haplogroups. Again, the pattern is strikingly different from that observed for cattle, sheep, and pigs. Although we need to keep in mind the inherent and generally unavoidable bias in sampling, it is clear that lineage A predominates across the globe (>90% of samples). Lineage B, on the other hand, is present in about 6% of the animals sampled and seems essentially confined to breeds from southern Asia. This finding leads Luikart *et al.* (3) to speculate that this lineage may derive from a local domestication in the region, perhaps within the Neolithic culture of Baluchistan in western Pakistan. Based on the total sample, lineage C is even rarer, confined to a small number of European breeds and a single animal from Mongolia. The absence of lineage C from Near Eastern populations is puzzling. However, further sampling, particularly in Central Asia may shed some light on the origins of this haplogroup. From a wider perspective, these rare but distinct lineages raise the possibility that cryptic legacies of additional domestication may persist within

other domestic species that are less comprehensively sampled. It therefore may be too early to presume duality as a paradigm of domestication.

The global distribution of goat mtDNA variation has an additional surprise—a remarkably low level of phylogeographic structure (particularly when compared with domestic cattle). In other words, geographical location has little relevance to the mtDNA type a particular animal possesses. Based on the antiquity of goat domestication and the documented presence of goats in all corners of the Old World stretching deep into prehistory (20), we might expect that goats should display relatively high levels of geographic structuring. It therefore suggests to Luikart and his colleagues (reasonably in our view), that goats have been a highly mobile species, probably as small and portable units of human trade throughout history (3).

Future Prospects: The Y Imperative

Additional sampling, particularly in regions likely to harbor the minor haplogroups will improve our understanding of the where of goat origins. Clearly this is also necessary in the other mammalian domesticates and it is apparent that there is a long way to go before we gain a detailed overview of the biological history of these important species.

Studies of mtDNA variation, although highly informative, are only one part of the puzzle. The next phase should be to develop suites of information-rich Y chromosome DNA markers for each species. Surveys of variation in the nonrecombining portion of this chromosome have been immensely valuable in complementing and adding to the picture of recent human evolution that emerged from mtDNA surveys (21). Fortu-

nately in many cases, animal geneticists should be able to short-circuit Y chromosome marker development. Based on comparisons with autosomal markers, it is clear that the rapidly evolving microsatellite-type markers developed for one artiodactyl species often will amplify in other taxa (22, 23).

Once an adequate number of Y-specific single-nucleotide polymorphisms and microsatellites become available for each domestic species, they should prove invaluable for studies of genetic diversity. These tools will be particularly useful for goats, where males from any number of wild populations or subspecies could have contributed genes to domestic herds. If Y haplotype systems of sufficient resolution can be developed they also should be able to provide detailed information concerning recent geographical movements and perhaps display the breed-specific characteristics that mtDNA stubbornly refuses to. Obviously, the autosomal genome should not be neglected either and as studies of microsatellites in cattle have shown, highly polymorphic diploid markers also can shed light on recent population movements and reveal the fine grain of admixture between divergent populations (11, 24).

The phylogeographic surveys required to augment our current knowledge of livestock diversity will be challenging and expensive because certain key regions also may be the most inaccessible and politically unstable. However, we feel the effort will be worth it; the genetic origins of domestic animals, unlike our own species, are genuinely deep-rooted and multifaceted and there should be plenty more surprises in store.

We thank all our colleagues for helpful comments and insights. This work was supported by the Wellcome Trust.

- Legge, T. (1996) in *The Origins and Spread of Agriculture and Pastoralism in Eurasia*, ed. Harris, D. R. (UCL Press, London), pp. 238–262.
- Zeder, M. A. & Hesse, B. (2000) *Science* **287**, 2254–2257.
- Luikart, G., Gielly, L., Excoffier, L., Vigne, J.-D., Bouvet, J. & Taberlet, P. (2001) *Proc. Natl. Acad. Sci. USA* **98**, 5927–5932.
- Loftus, R. T., MacHugh, D. E., Bradley, D. G., Sharp, P. M. & Cunningham, P. (1994) *Proc. Natl. Acad. Sci. USA* **91**, 2757–2761.
- Bradley, D. G., MacHugh, D. E., Cunningham, P. & Loftus, R. T. (1996) *Proc. Natl. Acad. Sci. USA* **93**, 5131–5135.
- Vilà, C., Savolainen, P., Maldonado, J. E., Amorim, I. R., Rice, J. E., Honeycutt, R. L., Crandall, K. A., Lundeberg, J. & Wayne, R. K. (1997) *Science* **276**, 1687–1689.
- Hiendleder, S., Mainz, K., Plante, Y. & Lewalski, H. (1998) *J. Hered.* **89**, 113–120.
- Giuffra, E., Kijas, J. M., Amarger, V., Carlborg, Ö., Jeon, J.-T. & Andersson, L. (2000) *Genetics* **154**, 1785–1791.
- Vilà, C., Leonard, J. A., Götherström, A., Marklund, S., Sandberg, K., Lidén, K., Wayne, R. K. & Ellegren, H. (2001) *Science* **291**, 474–477.
- Ingman, M., Kaessmann, H., Pääbo, S. & Gyllenstein, U. (2000) *Nature (London)* **408**, 708–713.
- MacHugh, D. E., Shriver, M. D., Loftus, R. T., Cunningham, P. & Bradley, D. G. (1997) *Genetics* **146**, 1071–1086.
- Tanaka, K., Solis, C. D., Masangkay, J. S., Maeda, K., Kawamoto, Y. & Namikawa, T. (1996) *Biochem. Genet.* **34**, 443–452.
- Lau, C. H., Drinkwater, R. D., Yusoff, K., Tan, S. G., Hetzel, D. J. & Barker, J. S. (1998) *Anim. Genet.* **29**, 253–264.
- Mannen, H., Tsuji, S., Loftus, R. T. & Bradley, D. G. (1998) *Genetics* **150**, 1169–1175.
- Watanabe, T., Okumura, N., Ishiguro, N., Nakano, M., Matsui, A., Sahara, M. & Komatsu, M. (1999) *Mol. Ecol.* **8**, 1509–1512.
- Wood, N. J. & Phua, S. H. (1996) *Anim. Genet.* **27**, 25–33.
- Smith, B. D. (1998) *The Emergence of Agriculture* (Scientific American Library, New York).
- Manceau, V., Després, L., Bouvet, J. & Taberlet, P. (1999) *Mol. Phylogenet. Evol.* **13**, 504–510.
- Rogers, A. R. & Harpending, H. (1992) *Mol. Biol. Evol.* **9**, 552–569.
- Porter, V. (1996) *Goats of the World* (Farming Press, Ipswich, U.K.).
- de Knijff, P. (2000) *Am. J. Hum. Genet.* **67**, 1055–1061.
- Moore, S. S., Sargeant, L. L., King, T. J., Mattick, J. S., Georges, M. & Hetzel, D. J. (1991) *Genomics* **10**, 654–660.
- Kemp, S. J., Hishida, O., Wambugu, J., Rink, A., Longeri, M. L., Ma, R. Z., Da, Y., Lewin, H. A., Barendse, W. & Teale, A. J. (1995) *Anim. Genet.* **26**, 299–306.
- Loftus, R. T., Ertugrul, O., Harba, A. H., El-Barody, M. A., MacHugh, D. E., Park, S. D. & Bradley, D. G. (1999) *Mol. Ecol.* **8**, 2015–2022.
- Kim, K. I., Yang, Y. H., Lee, S. S., Park, C., Ma, R., Bouzat, J. L. & Lewin, H. A. (1999) *Anim. Genet.* **30**, 102–108.