

LONDON
SCHOOL of
HYGIENE
& TROPICAL
MEDICINE



Chessa, B; Pereira, F; Arnaud, F; Amorim, A; Goyache, F; Mainland, I; Kao, RR; Pemberton, JM; Beraldi, D; Stear, MJ; Alberti, A; Pittau, M; Iannuzzi, L; Banabazi, MH; Kazwala, RR; Zhang, YP; Arranz, JJ; Ali, BA; Wang, ZL; Uzun, M; Dione, MM; Olsaker, I; Holm, LE; Saarma, U; Ahmad, S; Marzanov, N; Eythorsdottir, E; Holland, MJ; Ajmone-Marsan, P; Bruford, MW; Kantanen, J; Spencer, TE; Palmarini, M (2009) Revealing the History of Sheep Domestication Using Retrovirus Integrations. *Science (New York, NY)*, 324 (5926). pp. 532-536. ISSN 0036-8075 DOI: <https://doi.org/10.1126/science.1170587>

Downloaded from: <http://researchonline.lshtm.ac.uk/5322/>

DOI: [10.1126/science.1170587](https://doi.org/10.1126/science.1170587)

Usage Guidelines

Please refer to usage guidelines at <http://researchonline.lshtm.ac.uk/policies.html> or alternatively contact researchonline@lshtm.ac.uk.

Available under license: <http://creativecommons.org/licenses/by-nc-nd/2.5/>

Published in final edited form as:

Science. 2009 April 24; 324(5926): 532–536. doi:10.1126/science.1170587.

REVEALING THE HISTORY OF SHEEP DOMESTICATION USING RETROVIRUS INTEGRATIONS

B. Chessa^{1,2}, F. Pereira³, F. Arnaud¹, A. Amorim³, F. Goyache⁴, I. Mainland⁵, R.R. Kao¹, J. M. Pemberton⁶, D. Beraldi⁶, M. Stear¹, A. Alberti², M. Pittau², L. Iannuzzi⁷, M.H. Banabazi⁸, R. Kazwala⁹, Y.-P. Zhang¹⁰, J.J. Arranz¹¹, B.A. Ali¹², Z. Wang¹³, M. Uzun¹⁴, M. Dione¹⁵, I. Olsaker¹⁶, L.-E. Holm¹⁷, U. Saarma¹⁸, S. Ahmad¹⁹, N. Marzanov²⁰, E. Eythorsdottir²¹, M.J. Holland^{22,23}, P. Ajmone-Marsan²⁴, M.W. Bruford²⁵, J. Kantanen²⁶, T.E. Spencer²⁷, and M. Palmarini^{1,*}

¹Institute of Comparative Medicine, Faculty of Veterinary Medicine, University of Glasgow, UK
²Dipartimento di Patologia e Clinica Veterinaria, Università degli Studi di Sassari, Sassari, Italy
³Instituto de Patologia e Imunologia Molecular da Universidade do Porto (IPATIMUP); Faculdade de Ciências da Universidade do Porto, Porto, Portugal
⁴Área de Genética y Reproducción Animal, SERIDA-Somió, Gijón, Spain
⁵Division of Archaeological, Geographical and Environmental Sciences, University of Bradford, Bradford, UK
⁶Institute of Evolutionary Biology, School of Biological Sciences, University of Edinburgh, Edinburgh, UK
⁷National Research Council (CNR), ISPAAM, Naples, Italy
⁸Dept. of Biotechnology, Animal Science Research Institute of Iran (ASRI), Karaj, Iran
⁹Dept. of Veterinary Medicine and Public Health, Sokoine University of Agriculture, Morogor, Tanzania
¹⁰State key Laboratory of Genetic Resources and Evolution, Kunming Institute of Zoology, Chinese Academy of Sciences, Kunming, China
¹¹Departamento de Producción Animal, Facultad de Veterinaria, Universidad de León, Spain
¹²Genetic Engineering and Biotechnology Research Institute, Mubarak City for Scientific Research and Technology Applications, Alexandria, Egypt
¹³National Diagnostic Center for Exotic Animal Diseases, China Animal Health and Epidemiology Centers, Qingdao, China
¹⁴School of Health Science, Canakkale Onsekiz Mart University, Canakkale, Turkey
¹⁵International Trypanotolerance Centre, Banjul, The Gambia
¹⁶Dept. of Basic Sciences and Aquatic Medicine, The Norwegian School of Veterinary Science, Oslo, Norway
¹⁷Department of Genetics and Biotechnology, Faculty of Agricultural Sciences, University of Aarhus, Tjele, Denmark
¹⁸Department of Zoology, Institute of Ecology and Earth Sciences, University of Tartu, Tartu, Estonia
¹⁹NWFP Agricultural University, Peshwar, Pakistan
²⁰All-Russian Research Institute of Animal Husbandry, Russian Academy of Agricultural Sciences, Dubrovitsy, Russia
²¹Agricultural University of Iceland, Hvanneyri, Iceland
²²Medical Research Council Laboratories, Fajara, Banjul, Gambia
²³London School of Hygiene and Tropical Medicine, London, UK
²⁴Istituto di Zootecnica, Università Cattolica del Sacro Cuore, Piacenza, Italy
²⁵School of Biosciences, Cardiff University, Cardiff, UK
²⁶Biotechnology and Food Research, MTT Agrifood Research Finland, Jokioinen, Finland
²⁷Center for Animal Biotechnology and Genomics, Texas A&M University, College Station, USA

Abstract

*To whom correspondence should be addressed. m.palmarini@vet.gla.ac.uk.

Supporting Online Material

www.sciencemag.org

Materials and Methods

Fig. S1, S2, S3

Table S1, S2

The domestication of livestock represented a crucial step in human history. By using endogenous retroviruses as genetic markers, we found that sheep differentiated on the basis of their “retrotype” and morphological traits, dispersed across Eurasia and Africa via separate migratory episodes. Relicts of the first migrations include the Mouflon, as well as breeds previously recognized as “primitive” on the basis of their morphology, such as the Orkney, Soay and the Nordic short-tailed sheep now confined to the periphery of NW Europe. A later migratory episode, involving sheep with improved production traits, shaped the vast majority of present-day breeds. The ability to differentiate genetically primitive sheep from more modern breeds provides valuable insights into the history of sheep domestication.

The first agricultural systems, based on the cultivation of cereals, legumes and the rearing of domesticated livestock developed within South-West Asia approximately 11,000 years before present (YBP) (1, 2). By 6,000 YBP, agro-pastoralism introduced by the Neolithic agricultural revolution became the main system of food production throughout prehistoric Europe, from the Mediterranean north to Britain, Ireland and Scandinavia (3), south into North Africa (4) and east into West and central Asia (5).

Sheep and goats were the first livestock species to be domesticated (6). Multiple domestication events, as inferred by multiple mitochondrial lineages, gave rise to domestic sheep and similarly other domestic species (7-10). Initially, sheep were reared mainly for meat but, during the fifth millennium BP in South-West Asia and the fourth millennium BP in Europe, specialization for ‘secondary’ products such as wool became apparent. Sheep selected for secondary products appear to have replaced more primitive domestic populations. The issue of whether specialization for secondary products occurred first in South-West Asia or occurred throughout Europe is not known with certainty, owing to the lack of definitive archaeological evidence for the beginning of wool production (6, 11, 12).

In this study we used a family of endogenous retroviruses (ERVs) as genetic markers to study the history of the domestic sheep. ERVs result from the stable integration of the retrovirus genome (‘provirus’) into the germline of the host (13) and are transmitted vertically from generation to generation in a Mendelian fashion. The sheep genome contains at least 27 copies of endogenous retroviruses related to the exogenous and pathogenic Jaagsiekte sheep retrovirus (enJSRVs) (14-16). Most enJSRVs loci are fixed in domestic sheep but some are differentially distributed between breeds and individuals (i.e., they are insertionally polymorphic) (14). enJSRVs can be used as highly informative genetic markers because the presence of each endogenous retrovirus in the host genome is the result of a single integration event in a single animal and is irreversible, so populations sharing the same provirus in the same genomic location are *de facto* phylogenetically related.

We analyzed genomic DNA samples collected from 1362 animals belonging to 133 breeds of the domestic sheep (*Ovis orientalis aries* usually referred to as *Ovis aries*) and closest wild relatives (see below) divided into 65 groups formed by one or more breeds sharing a common geographical location and/or breeding links (Table S1) (17). Samples tested also included the Urial sheep (*Ovis vignei*), and the Mediterranean and Asiatic Mouflon (*Ovis orientalis musimon*, *Ovis orientalis ophion* and *Ovis orientalis orientalis*). The overwhelming majority of breeds that we studied are local, historically related to specific geographical areas and not subjected to the intensive breeding programs of commercial flocks.

Samples were tested for the presence or absence of six independently inherited insertionally polymorphic enJSRVs (enJSRV-18, enJSRV-7, enJSRV-8, enJSRV-15, enJSRV-16, and enJS5F16) by polymerase chain reaction (PCR) employing two sets of primers that amplify respectively the 5’ and 3’ long terminal repeats (LTR) of each provirus (including the

flanking genomic DNA sequences of the host) as previously described (14, 17). Provirus enJSRV-18 had by far the highest frequency in our dataset (85%), enJSRV-7 and enJS5F16 were detected in 27% and 30% of the samples, respectively, while enJSRV-15, enJSRV-16 and enJSRV-8 were present in only 3-5% of the samples (Fig. 1A).

We inferred the distribution of the insertionally polymorphic enJSRV loci in the earliest domesticated sheep by determining their occurrence in the Urial sheep and in the Mediterranean/Asiatic Mouflon, and then by verifying the molecular signatures indicative of the age of a provirus. The estimated divergence between the Urial (one of the closest living relatives of the domestic sheep) and the domestic sheep is approximately 800,000 YBP (18). Consequently, any provirus that is shared between these two species will predate the process of domestication. The same is true for the Asiatic Mouflon which is believed to be the direct ancestor of the domestic sheep (19-21), while the closely related Mediterranean Mouflon is thought to be the remnant of the first domesticated sheep readapted to feral life (19, 22, 23). Despite its widespread distribution in the samples tested, enJSRV-18 was absent from the Urial sheep (n=5), the Mediterranean Mouflon (n=17) and the Asiatic Mouflon (n=15). By contrast, the relatively rarer enJSRV-7 was detected in three of five Urial sheep, in most (86%) Asiatic Mouflons and in all Mediterranean Mouflons. These data suggest that the integration of enJSRV-7 in the germline of the host predates the integration of enJSRV-18. Differences between the proximal (5') and distal (3') long terminal repeats (LTRs) of enJSRV-7 confirm its antecedence. The divergence between the 5' and 3' LTR of an endogenous retrovirus gives an estimate of the "age" of an endogenous provirus because upon infection, retroviruses reverse transcribe their genome from RNA into DNA and during this process they duplicate the genomic ends giving rise to two identical LTRs. Proximal and distal LTRs of an endogenous retrovirus must be identical upon integration, but can diverge over time at the same rate as non-coding sequences ($\sim 2.3\text{-}5 \times 10^{-9}$ substitutions per site per year). enJSRV-7 appears to be the oldest provirus in our samples since it displays five nucleotide substitutions between 5' and 3' LTRs (445nt long), while all the other insertionally polymorphic proviruses (including enJSRV-18) possess identical LTRs. These data suggest that the populations originating from the earliest domesticated sheep did not carry any of the insertionally polymorphic enJSRVs used in this study or carried enJSRVs-7.

The geographical variation of all enJSRV loci was visualized by constructing interpolation maps from their insertion frequency values (Fig. 1B-F). The highest frequency of enJSRV-7 was found in the Mediterranean Mouflon and in Soay sheep now inhabiting the island of St. Kilda off NW Scotland (Fig. 1C). enJSRV-18 was uniformly distributed at very high frequencies throughout the Old World. Low frequencies of enJSRV-18 were observed in the islands inhabited by the Mediterranean Mouflon and in peripheral regions of NW Europe. Two enJSRV proviruses, enJS5F16 and enJSRV-8, showed a similar geographical pattern with a high frequency in the British Isles and Scandinavia (Fig. 1E and F). The less common enJSRV-15 and enJSRV-16 had less obvious geographical patterns (Fig. S1).

We then analyzed the combination of insertionally polymorphic enJSRVs (which we call 'retrotype') in each of the populations analysed (Fig. 2). The R2 retrotype (representing the presence of enJSRV-18 only) was the predominant retrotype in most of the populations tested. Interestingly the R4 retrotype, indicating presence of enJSRV-18 and enJSRV-7 together (Fig. 2), was another common retrotype in the area corresponding to the historical Phoenicia and in Southern Europe, suggesting that maritime trade and colonization had a major influence on sheep movement in the Mediterranean, confirmed by studies using sheep mitochondrial DNA variation (24, 25). Additional enJSRV insertions accounted for more complex retrotypes of populations in Northern Europe (see also supporting online text). Sheep populations in Africa, Pakistan and China displayed a similarly homogenous R2 retrotype pattern common to the populations in South-West Asia, suggesting direct

migratory links of domestic sheep between these areas. Most of the populations from Scandinavia displayed similar retrotypes to Icelandic and the Faeroe Island populations, supporting the historically registered movements of the Norse settlers during the later first millennium AD (26). To visualise the genetic relationship of the tested populations we analyzed the data using two different approaches: a multidimensional scaling (MDS) plot obtained from the interpopulation matrix of Nei's unbiased genetic distances and principal component analysis (PCA) computed from the correlation matrix among enJSRV insertion frequencies.

The MDS analysis revealed a marked separation (particularly evident in the first dimension) between the great majority of domestic breeds and an outer group formed by the Mouflon, Soay sheep, Hebrideans, Orkney sheep, Icelandic and Nordic breeds (Figure 3A). Similar results were obtained by PCA (Figure 3B). Collectively, the data we obtained indicate that relicts of the first migrations are still present in the Mouflon of Sardinia, Corsica and Cyprus and in breeds in peripheral North European areas. On the basis of their retrotypes these primitive populations are characterized by the absence of enJSRV-18 (fixed in most of the modern breeds) and either the presence of enJSRV-7 in high frequency or the lack of insertionally polymorphic enJSRVs (including enJSRV-7). By contrast, the retrotypes of the great majority of sheep breeds cluster together and are characterized by the high frequency or fixation of enJSRV-18.

The homogenous retrotypes (R2, R2-R4) that we observed in the sheep of modern day Turkey, Iran, Saudi Arabia, Syria, Israel and Egypt, combined with available archaeological evidence, suggest that selection of domestic sheep with the desired secondary characteristics common to the modern breeds occurred first in South-West Asia and then spread successfully into Europe and Africa, and the rest of Asia. This may provide genetic support to the theory that specialized wool production arose in South-West Asia and then spread throughout Europe (11). The primitive breeds survived the second migrations of improved breeds from South-West Asia by returning to a feral or semi-feral state in islands without predators or by occupying inaccessible areas less prone to commercial exchanges and associated introgression. Most, if not all, of the breeds we identified as of ancient origin were already considered primitive on the basis of morphological traits such as a darker and coarser hair (instead of a whiter woolly fleece), a moulting coat and the frequent presence of horns in females as well as males (Fig. 4).

Our study also provides genetic evidence supporting the anecdotal origin of some less common sheep breeds. For example, one of the 10 populations analysed from the British Isles, the Jacob sheep, displayed a homogenous R2 retrotypic very different from the other British populations and more similar to the South-Western Asiatic and African breeds. The origins of the Jacob are unknown. This breed owes its name to the Biblical story of Jacob who took "every speckled and spotted sheep" as a wage from his father in law Laban (Genesis 30:25-43; probably the first recorded use of selective breeding in livestock). Our retrotypic analysis supports a direct link between the Jacob sheep and breeds in South-West Asia or Africa rather than other British breeds. Our study also firmly links the Soay sheep with the Mediterranean and Asiatic Mouflon rather than the Nordic breeds.

In conclusion, the polymorphic nature of enJSRVs revealed a remarkable secondary population expansion of improved domestic sheep, most likely out of South-West Asia, providing valuable insights into the history of pastoralist societies whose economy included sheep husbandry. By differentiating for the first time genetically primitive breeds from modern ones, our study offers a rationale for identifying and preserving rare gene pools. Finally, we demonstrate the utility of endogenous retroviruses as a new class of genetic markers used to unravel the history of a domesticated species.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We would like to thank the ECONOGENE, North-SheD and NordGen consortia, Caroline Leroux, Jim DeMartini, James Trafford, Susan Haywood, Valgerdur Andresdottir, Michael Reichert, Yolanda Bayon Gonzalez, Katja Voigt, Vladimir Feinstein, Fermin San Primitivo and Paul Halstead for help in obtaining some of the samples used in this study and for useful comments. We would like to thank Pablo Murcia for coining the term ‘retrotypes’. We are grateful to G.P. Di Meo and A. Perucatti for FISH analysis. We also thank Brent Huffman (www.ultimateungulate.com), NordGen, Sven Jeppson and Kenneth Headspeath for pictures reproduced in Fig. 4. This study was supported by the BBSRC, the Wellcome Trust, NIH Grant HD05274, the Scottish Funding Council through a Strategic Research Developmental Grant and in part from Fundação para a Ciência e a Tecnologia, “Misura P5 Biodiversita’ animale” from the Regione autonoma della Sardegna and the Chinese Academy of Sciences (KSCX2-YW-N-018). The Soay sheep project in St. Kilda is supported by NERC. IPATIMUP is partially supported by “Programa Operacional Ciência e Inovação 2010” (POCI 2010), VI Programa Quadro (2002–2006). MP is a Wolfson-Royal Society Research Merit awardee.

References

- Zeder MA. Proc Natl Acad Sci U S A. 2008; 105:11597. [PubMed: 18697943]
- College S, Conolly J, Shennan S. European Journal of Archaeology. 2005; 8:137.
- Price, TD. Europe’s First farmers. Cambridge University Press; Cambridge: 2000.
- Barker, G. Examining the Farming/Language Dispersal Hypothesis. Bellwood, P.; Renfrew, C., editors. McDonald Institute Monographs; Cambridge: 2002. p. 151-162.
- Harris, DR.; Gosden, C. The Origins and Spread in Agriculture and Pastoralism in Eurasia. Harris, D., editor. UCL press; London: 1996. p. 370-389.
- Ryder, ML. Sheep & Man. Gerald Duckworth & Co.; London: 1983.
- Pedrosa S, et al. Proc Biol Sci. 2005; 272:2211. [PubMed: 16191632]
- Tapio M, et al. Mol Biol Evol. 2006; 23:1776. [PubMed: 16782761]
- Meadows JR, Cemal I, Karaca O, Gootwine E, Kijas JW. Genetics. 2007; 175:1371. [PubMed: 17194773]
- Naderi S, et al. Proc Natl Acad Sci U S A. 2008; 105:17659. [PubMed: 19004765]
- Sherratt, A. Patterns of the Past: Studies in Honour of David Clark. Hodder, I.; Isaacs, G.; Hammond, N., editors. Cambridge University Press; Cambridge: 1981.
- Helmer D, Gourichon L, Vila E. Anthropozoologica. 2007; 42:41.
- Boeke, JD.; Stoye, JP. Retroviruses. Coffin, JM.; Hughes, SH.; Varmus, HE., editors. Cold Spring Harbor Laboratory Press; Plainview, NY: 1997. p. 343-436.
- Arnaud F, et al. PLoS Pathogens. 2007; 3:e170. [PubMed: 17997604]
- Palmarini M, Sharp JM, De las Heras M, Fan H. Journal of Virology. 1999; 73:6964. [PubMed: 10400795]
- Arnaud F, Varela M, Spencer TE, Palmarini M. Cell Mol Life Sci. 2008
- Materials and Methods are available as supporting material on *Science* Online.
- Hernandez Fernandez M, Vrba ES. Biol Rev Camb Philos Soc. 2005; 80:269. [PubMed: 15921052]
- Hiendleder S, Kaup B, Wassmuth R, Janke A. Proc Biol Sci. 2002; 269:893. [PubMed: 12028771]
- Clutton-Brock, J. A Natural History of Domesticated Mammals. ed. 2nd. Cambridge University Press; Cambridge: 1999.
- Zohary D, Tchernov E, Horwitz L. Journal of Zoology. 1998; 245:129.
- Poplin F. Annales de Genetique et Selection Animale. 1979; 11:133.
- Vigne, J-D. The Holocene history of the European vertebrate fauna. Modern aspects of research. Benecke, N., editor. Verlag Marie Leidorf GmbH, Rahden/Westf; Germany: 1999. p. 295-322.

24. Pedrosa S, et al. *Genet Sel Evol.* 2007; 39:91. [PubMed: 17212950]
25. Pereira F, et al. *Mol Biol Evol.* 2006; 23:1420. [PubMed: 16672283]
26. McGovern TH. *Annual Review of Anthropology.* 1990; 19:331.
27. Excoffier L, Laval G, Schneider S. *Evolutionary Bioinformatics Online.* 2005; 1:47. [PubMed: 19325852]
28. Miller, MP. Department of Biological Sciences, Northern Arizona University; Flagstaff, AZ: 1997.
29. Cavalli-Sforza, LL.; Menozzi, P.; Piazza, A. *The history and geography of human genes.* Princeton University Press; Princeton, USA: 1994.

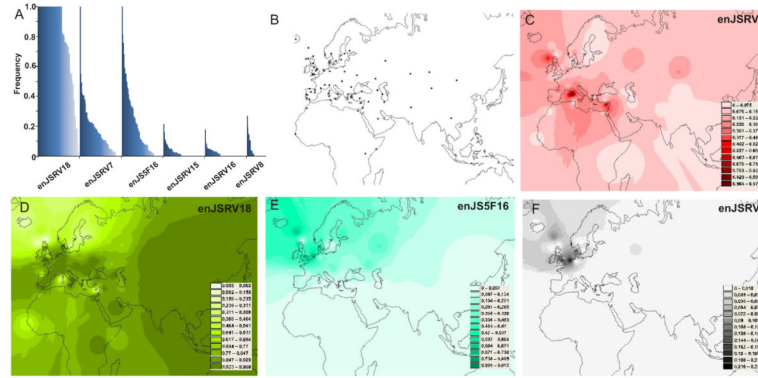


Fig. 1. Worldwide distribution of insertionally polymorphic enJSRVs. Distribution of the insertionally polymorphic enJSRV loci analysed in this study in 65 sheep populations representing local breeds from the old world. **(A)** Frequencies of each enJSRV locus in each population are represented by a vertical bar and arranged in a descending order. Insertion frequencies were obtained using the software Arlequin 3.11 (27) treating the absence of a specific enJSRV provirus as a recessive allele. **(B)** Locations of sheep populations sampled. **(C-F)**. Interpolation maps displaying the spatial distribution of estimated enJSRVs frequencies. The geographical variation was visualized using the ‘Spatial Analyst Extension’ of ArcView GIS 3.2 software (ESRI, Redlands, CA, USA; <http://www.esri.com>). Interpolated map values were calculated employing the inverse distance–weighted with 12 nearest neighbours and a power of two, and interpolation surfaces were divided into 13 classes with higher insertion frequencies indicated by darkest shading. The central point of the sampling area was used as geographic coordinates for each population (Table S1).

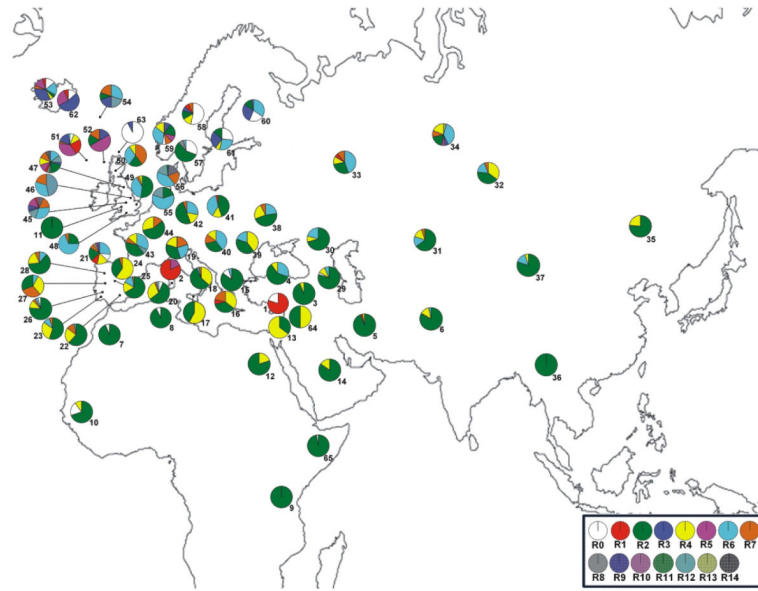


Fig. 2.

Combination of enJSRV proviruses (retrotypes) in the domestic sheep. Pie charts in the figure represent the frequency of each retrotype in the 65 populations tested. Each sheep tested was assigned a retrotype on the basis of the combination of insertionally polymorphic enJSRV proviruses present in their genome. Retrotypes were defined R0 to R14 as follows: **R0** = no insertionally polymorphic enJSRVs; **R1** = enJSRV-7; **R2** = enJSRV-18; **R3** = enJS5F16; **R4** = enJSRV-7 + enJSRV-18; **R5** = enJSRV-7 + enJS5F16; **R6** = enJSRV-18 + enJS5F16; **R7** = enJSRV-7 + enJSRV-18 + enJS5F16; **R8** = enJSRV-8; **R9** = enJS5F16 + enJSRV-8; **R10** = enJSRV-7 + enJS5F16 + enJSRV-8; **R11** = enJSRV-18 + enJSRV-8; **R12** = enJSRV-18 + enJS5F16 + enJSRV-8; **R13** = enJSRV-7 + enJSRV-18 + enJSRV-8; **R14** = enJSRV-7 + enJSRV-18 + enJS5F16 + enJSRV-8. Each retrotype is represented with a different colour (and pattern) as indicated in the figure. Numbers beside each pie chart indicate each of the 65 populations tested as indicated in Table S1. Note that most of the populations in South-West Asia, Central Asia, Southern Europe and Africa possess R2 (i.e. presence of enJSRV-18 only, shown in green) as the predominant retrotype. Around the Mediterranean basin there is also a high proportion of R4 given by the contemporary presence of enJSRV-7 and enJSRV-18 (shown in yellow). The primitive breeds are characterized by a high proportion of animals with R0 (no insertionally polymorphic proviruses, shown in white) or R1 (presence of enJSRV-7 only, shown in red). A ‘Nordic’ retrotype R3 (shown in blue) was characterized by a low frequency of enJSRV-18 and a high frequency of enJS5F16; Nordic populations also had a relatively high frequency of sheep with none of the insertionally polymorphic proviruses tested.

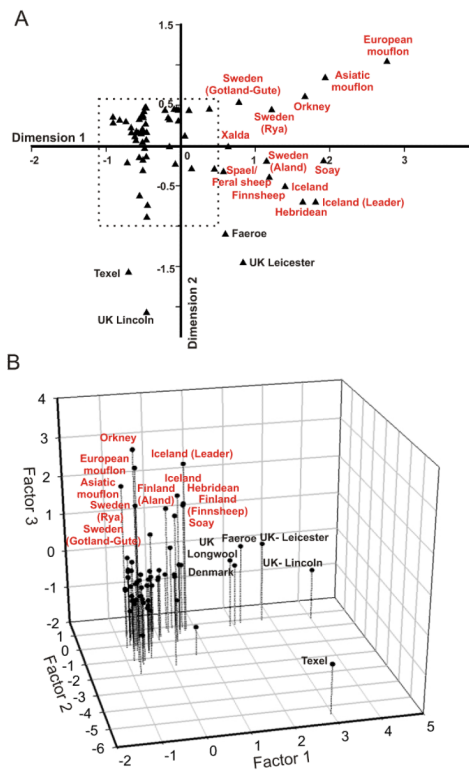


Fig. 3.

Genetic distances between sheep populations on the basis of enJSRVs insertion frequencies. **(A)** Multidimensional (MDS) scaling plot computed from the matrix of Nei's unbiased genetic distances (TFPGA 1.3 software) (28). The dominant nature of the enJSRVs as genetic markers was considered in all analyses. The matrix of interpopulation distances was summarized in two dimensions by use of MDS analysis as implemented by the STATISTICA '99 software package (StatSoft Inc., Tulsa, OK, USA). Each triangle represents one of the 65 populations tested. Note that in the graph only those populations outside the main cluster (enclosed within the square with the broken line and including the great majority of breeds from Africa, Asia and Europe) have been named. **(B)** Tri-dimensional plot summarising data obtained by principal component analysis (PCA) of the insertionally polymorphic enJSRV proviruses in the 65 sheep populations tested using the Proc Factor of the statistical package SAS/STAT® (SAS Institute Inc, Cary NC) according to the recommendations by Cavalli-Sforza *et al.* (29). Four factors, accounting for 86.66% of variation, with eigenvalue ≥ 1 were identified. Factor 1 (on the X-axis), explained 30.09% of variation and can be interpreted as the 'Northern Sea factor', distinguishing between a group of populations formed from some UK and continental European (including Denmark and Texel) sheep populations and the others. Factor 2 (on the Y-axis), explaining 23.58% of variation separating the Texel population from the rest. Factor 3 (on the Z-axis), explained 22.92% of variation and can be interpreted as the 'primitive breed factor', distinguishing between the group of populations formed by the mouflon and Scandinavian populations (including the Hebridean, Orkney and Soay populations) from the rest. Note that to increase the clarity of the figure, the populations that form the main cluster have not been named.



Fig. 4. Morphological characteristics of primitive breeds. Breeds identified in this study as remnants of the first sheep migrations possess morphological characteristics (such as darker coarser fleece, moulting coat, frequent presence of horns in females), similar to wilder sheep and the Mouflon. (A) Urial sheep; (B) Cyprus Mouflon; (C) Mediterranean Mouflon; (D) Orkney sheep; (E) Soay sheep; (F) Gute sheep; (G) Åland sheep; (H) Icelandic sheep; (I) Hebridean sheep.