

*This copy is for your personal, non-commercial use only.*

**If you wish to distribute this article to others**, you can order high-quality copies for your colleagues, clients, or customers by [clicking here](#).

**Permission to republish or repurpose articles or portions of articles** can be obtained by following the guidelines [here](#).

**The following resources related to this article are available online at [www.sciencemag.org](http://www.sciencemag.org) (this information is current as of April 6, 2010 ):**

**Updated information and services**, including high-resolution figures, can be found in the online version of this article at:

<http://www.sciencemag.org/cgi/content/full/324/5926/532>

**Supporting Online Material** can be found at:

<http://www.sciencemag.org/cgi/content/full/324/5926/532/DC1>

A list of selected additional articles on the Science Web sites **related to this article** can be found at:

<http://www.sciencemag.org/cgi/content/full/324/5926/532#related-content>

This article **cites 18 articles**, 9 of which can be accessed for free:

<http://www.sciencemag.org/cgi/content/full/324/5926/532#otherarticles>

This article has been **cited by** 9 article(s) on the ISI Web of Science.

This article has been **cited by** 4 articles hosted by HighWire Press; see:

<http://www.sciencemag.org/cgi/content/full/324/5926/532#otherarticles>

This article appears in the following **subject collections**:

Evolution

<http://www.sciencemag.org/cgi/collection/evolution>

Biomedical Sciences, Texas A&M University, College Station, TX 77843–4461, USA. <sup>21</sup>The Roslin Institute, Royal (Dick) School of Veterinary Studies, The University of Edinburgh, Roslin, Midlothian, E25 9PS, UK. <sup>22</sup>Direction Nationale de l'Élevage, Post Office Box 559, Conakry, Guinea. <sup>23</sup>Ethiopian Institute of Agricultural Research, Post Office Box 2003, Addis Ababa, Ethiopia. <sup>24</sup>Embrapa Dairy Cattle Center, Rua Eugênio do Nascimento, 610, Juiz de Fora, MG 36038–330, Brazil. <sup>25</sup>Animal Genomics, AgResearch, Ruakura, Post Box 3123, Hamilton 3240, New Zealand. <sup>26</sup>Embrapa Beef Cattle Center, Rod. BR 262, km 4, Campo Grande, MS 79002–970, Brazil. <sup>27</sup>Research and Development, LIC, Post Box 3016, Hamilton 3240, New Zealand. <sup>28</sup>Animal Genetics Resources Characterization, International Livestock Research Institute, Post Office Box 5689, Addis Ababa, Ethiopia. <sup>29</sup>Department of Biological Statistics and Computational Biology, Cornell University, 101 Biotechnology Building, Ithaca, NY 14853, USA. <sup>30</sup>Veterinary Biomedical Sciences, Royal (Dick) School of Veterinary Studies, The University of Edinburgh Summerhall, Edinburgh,

EH9 1QH Scotland. <sup>31</sup>World Guernsey Cattle Federation, The Hollyhocks, 10 Clos des Goddards, Rue des Goddards, Castel, Guernsey, GY5 7JD, Channel Islands, UK. <sup>32</sup>Agricultural Food and Nutritional Science, University of Alberta, 410 AgFor Centre, Edmonton, AB, T6G 2P5, Canada. <sup>33</sup>Molecular Genetics Research Unit, USDA–ARS, U.S. Meat Animal Research Center, Post Office Box 166, Clay Center, NE 68933, USA. <sup>34</sup>Dipartimento di Produzioni Animali, Università della Tuscia, via de Lellis, Viterbo 01100, Italy. <sup>35</sup>Department of Veterinary Pathobiology, College of Veterinary Medicine and Biomedical Sciences, Texas A&M University, College Station, TX 77843–4461, USA. <sup>36</sup>Genome Dynamics and Evolution, Wellcome Trust Sanger Institute, Hinxton, CB10 1SA, UK. <sup>37</sup>Animal Genetics and Genomics, Department of Primary Industries, 475 Mickelham Road Attwood, VIC 3031, Australia. <sup>38</sup>Faculty of Land and Food Resources, University of Melbourne, Royal Parade, Parkville, VIC 3010, Australia. <sup>39</sup>Smurfit Institute of Genetics, Trinity College, Dublin 2, Ireland. <sup>40</sup>Department of Molecular Biology and Biochemistry,

Simon Fraser University, 8888 University Drive, Burnaby, BC, V5A 1S6, Canada.

\*To whom correspondence and requests for materials should be addressed. E-mail: curt.vantassell@ars.usda.gov (C.P.V.T.), taylorjerr@missouri.edu (J.F.T.), and agibbs@bcm.tmc.edu (R.A.G.)

†In memoriam.

## Supporting Online Material

www.sciencemag.org/cgi/content/full/324/5926/528/DC1

Materials and Methods

Figs. S1 to S27

Tables S1 to S9

References

31 October 2008; accepted 16 March 2009

10.1126/science.1167936

# Revealing the History of Sheep Domestication Using Retrovirus Integrations

Bernardo Chessa,<sup>1,2</sup> Filipe Pereira,<sup>3</sup> Frederick Arnaud,<sup>1</sup> Antonio Amorim,<sup>3</sup> Félix Goyache,<sup>4</sup> Ingrid Mainland,<sup>5</sup> Rowland R. Kao,<sup>1</sup> Josephine M. Pemberton,<sup>6</sup> Dario Beraldi,<sup>6</sup> Michael J. Stear,<sup>1</sup> Alberto Alberti,<sup>2</sup> Marco Pittau,<sup>2</sup> Leopoldo Iannuzzi,<sup>7</sup> Mohammad H. Banabazi,<sup>8</sup> Rudovick R. Kazwala,<sup>9</sup> Ya-ping Zhang,<sup>10</sup> Juan J. Arranz,<sup>11</sup> Bahy A. Ali,<sup>12</sup> Zhiliang Wang,<sup>13</sup> Metehan Uzun,<sup>14</sup> Michel M. Dione,<sup>15</sup> Ingrid Olsaker,<sup>16</sup> Lars-Erik Holm,<sup>17</sup> Urmas Saarma,<sup>18</sup> Sohail Ahmad,<sup>19</sup> Nurbay Marzanov,<sup>20</sup> Emma Eythorsdottir,<sup>21</sup> Martin J. Holland,<sup>22,23</sup> Paolo Ajmone-Marsan,<sup>24</sup> Michael W. Bruford,<sup>25</sup> Juha Kantanen,<sup>26</sup> Thomas E. Spencer,<sup>27</sup> Massimo Palmarini<sup>1\*</sup>

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 northwest Europe. A later migratory episode, involving sheep with improved production traits, shaped the great 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 Southwest Asia ~11,000 years before present (yr B.P.) (1, 2). By 6000 yr B.P., 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 B.P. in Southwest Asia and the fourth millennium B.P. in Europe, specialization

for “secondary” products such as wool became apparent. Sheep selected for secondary products appear to have replaced more primitive domestic populations. Whether specialization for secondary products occurred first in Southwest 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).

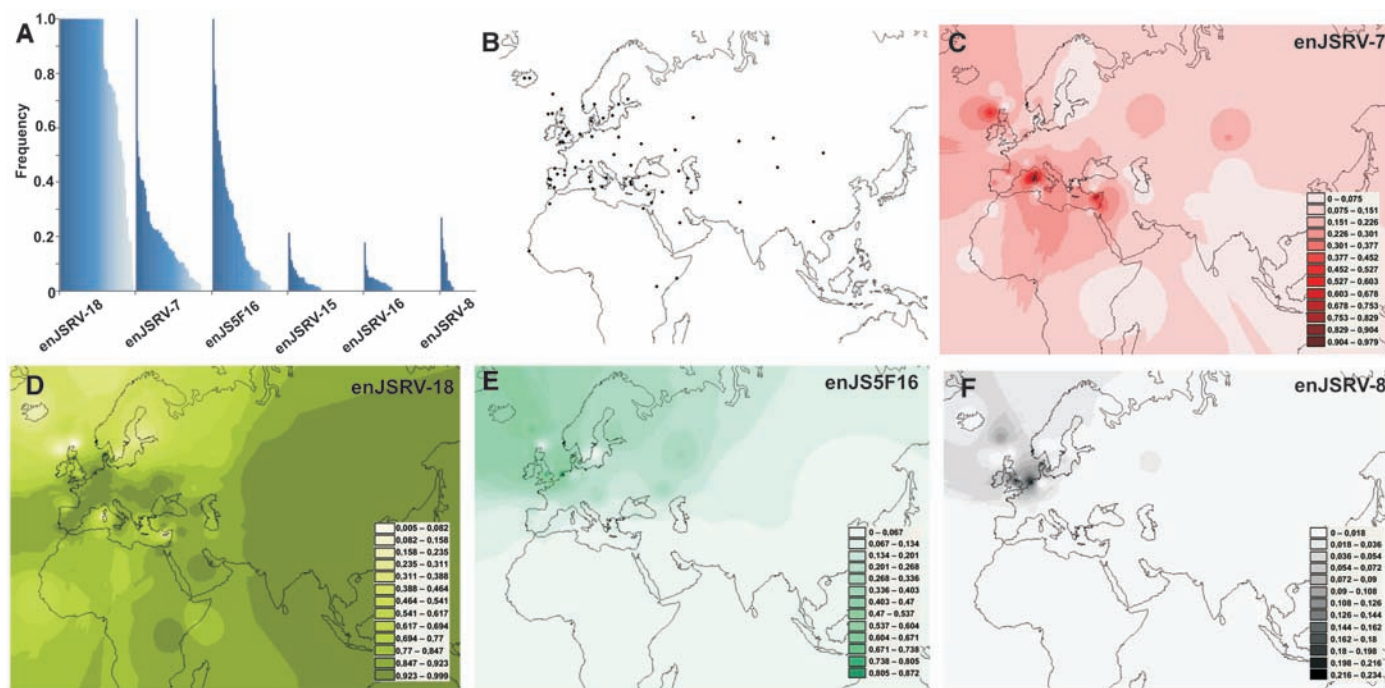
For this study, we used a family of endogenous retroviruses (ERVs) as genetic markers to examine the history of the domestic sheep. ERVs result from the stable integration of the retrovirus genome (“provirus”) into the germline of the host (13) and are transmitted vertically from generation to generation in a Mendelian fashion. The sheep genome contains at least 27 copies of ERVs 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

<sup>1</sup>Institute of Comparative Medicine, Faculty of Veterinary Medicine, University of Glasgow, Glasgow G61 1QH, UK.

<sup>2</sup>Dipartimento di Patologia e Clinica Veterinaria, Università degli Studi di Sassari, 07100 Sassari, Italy. <sup>3</sup>Instituto de Patologia e Imunologia Molecular da Universidade do Porto (IPATIMUP), Faculdade de Ciências da Universidade do Porto, 4200-465 Porto, Portugal. <sup>4</sup>Área de Genética y Reproducción Animal, SERIDA-Somío, E-33203 Gijón, Spain. <sup>5</sup>Division of Archaeological, Geographical and Environmental Sciences, University of Bradford, Bradford BD7 1DP, UK. <sup>6</sup>Institute of Evolutionary Biology, School of Biological Sciences, University of Edinburgh, Edinburgh EH9 3JT, UK. <sup>7</sup>National Research Council (CNR), ISPAAM, 80147 Naples, Italy. <sup>8</sup>Department of Biotechnology, Animal Science Research Institute of Iran (ASRI), 3146618361 Karaj, Iran. <sup>9</sup>Department of Veterinary Medicine and Public Health, Sokoine University of Agriculture, Morogoro, Tanzania. <sup>10</sup>State Key Laboratory of Genetic Resources and Evolution, Kunming Institute of Zoology, Chinese Academy of Sciences, Kunming 650223, China. <sup>11</sup>Departamento de Producción Animal, Facultad de Veterinaria, Universidad de León, 24071 León, Spain. <sup>12</sup>Genetic Engineering and Biotechnology Research Institute, Mubarak City for Scientific Research and Technology Applications, New Borg El-Arab City, 21934, Alexandria, Egypt. <sup>13</sup>National Diagnostic Center for Exotic Animal Diseases, China Animal Health and Epidemiology Centers, Qingdao 266032, China. <sup>14</sup>School of Health Science, Canakkale Onsekiz Mart University, 17100 Canakkale, Turkey. <sup>15</sup>International Trypanotolerance Centre, Banjul, Gambia. <sup>16</sup>Department of Basic Sciences and Aquatic Medicine, Norwegian School of Veterinary Science, 0033 Oslo, Norway. <sup>17</sup>Department of Genetics and Biotechnology, Faculty of Agricultural Sciences, University of Aarhus, 8830 Tjele, Denmark. <sup>18</sup>Department of Zoology, Institute of Ecology and Earth Sciences, University of Tartu, 51014 Tartu, Estonia. <sup>19</sup>NWFP Agricultural University, Peshwar, Pakistan. <sup>20</sup>All-Russian Research Institute of Animal Husbandry, Russian Academy of Agricultural Sciences, Dubrovitsy 142132, Russia. <sup>21</sup>Agricultural University of Iceland, Hvanneyri, IS-311 Borgarnes, Iceland. <sup>22</sup>Medical Research Council Laboratories, Fajara, Banjul, Gambia. <sup>23</sup>London School of Hygiene and Tropical Medicine, London WC1E 7HT, UK. <sup>24</sup>Istituto di Zootecnica, Università Cattolica del Sacro Cuore, 29100 Piacenza, Italy. <sup>25</sup>School of Biosciences, Cardiff University, Cardiff CF10 3AX, UK. <sup>26</sup>Biotechnology and Food Research, MTT Agrifood Research Finland, 31600 Jokioinen, Finland. <sup>27</sup>Center for Animal Biotechnology and Genomics, Texas A&M University, College Station, TX 77843, USA.

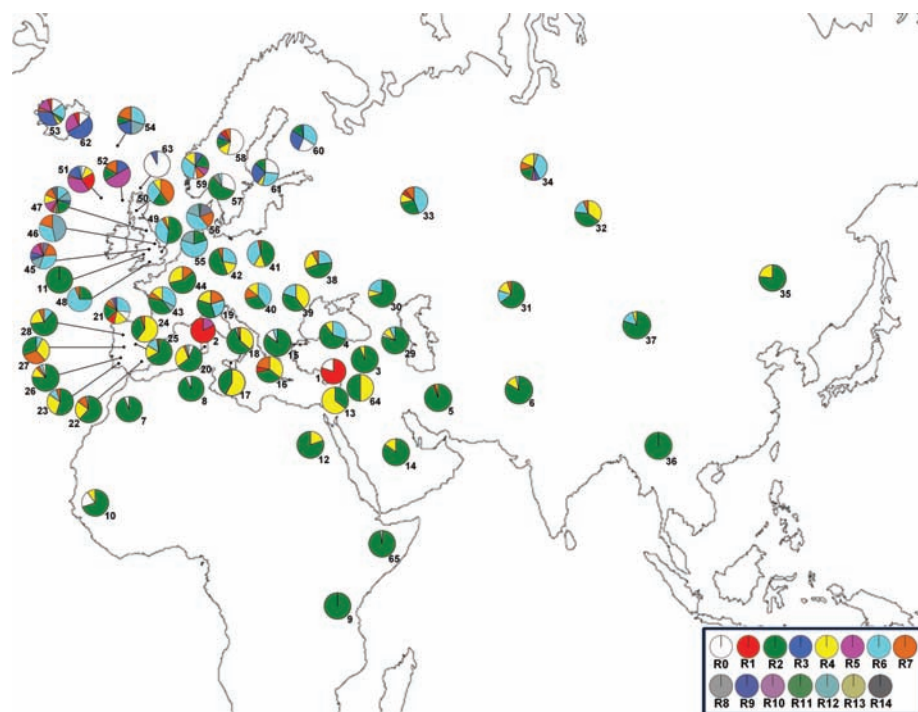
\*To whom correspondence should be addressed. E-mail: m.palmarini@vet.gla.ac.uk



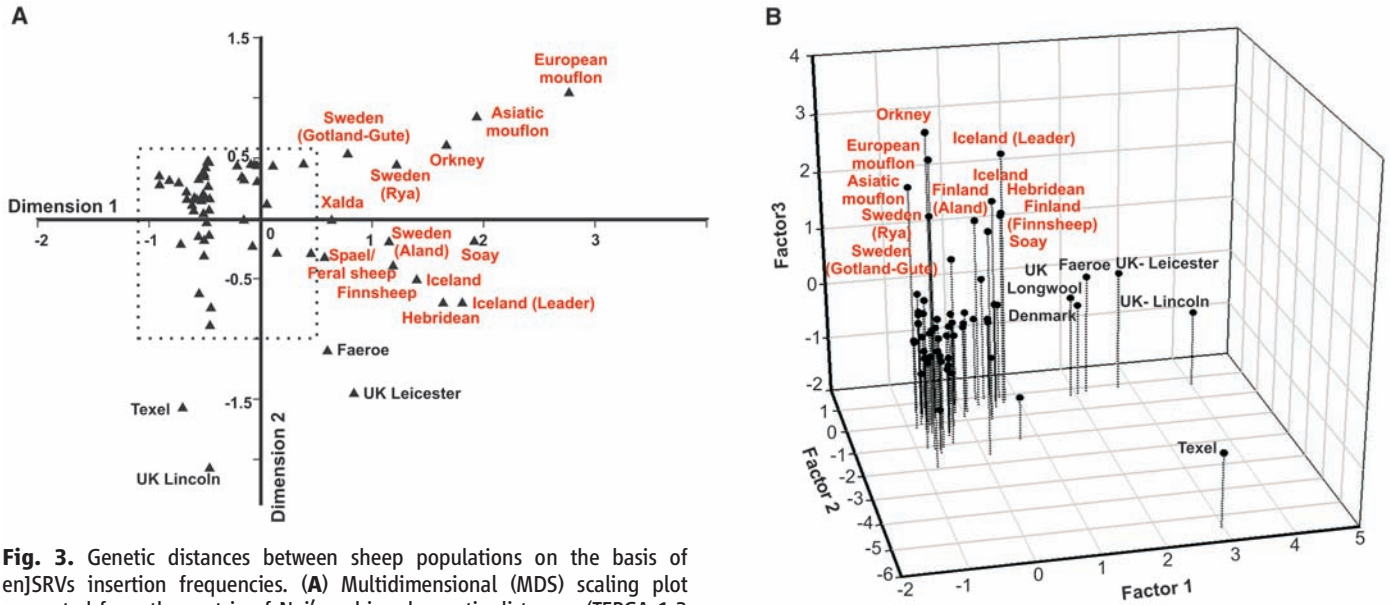
**Fig. 1.** Worldwide distribution of insertionally polymorphic enJSRVs. Distribution of the insertionally polymorphic enJSRV loci analyzed in this study in 65 sheep populations representing local breeds from the Old World. (A) Frequencies of each enJSRV locus in each population are represented by a vertical bar and arranged in descending order. Insertion frequencies were obtained with the software Arlequin 3.11 (27); the absence of a specific enJSRV provirus was treated as a recessive allele. (B) Locations of sheep populations sampled. (C to F) Interpolation maps

displaying the spatial distribution of estimated enJSRVs frequencies. The geographical variation was visualized with the “Spatial Analyst Extension” of ArcView GIS 3.2 software (ESRI, Redlands, California, USA). Interpolated map values were calculated by using the inverse distance weighted with 12 nearest neighbors and a power of 2, 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 R0 to R14 were defined 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 color (and pattern) as indicated in the figure. Numbers beside each pie chart indicate each of the 65 populations tested as indicated in table S1. Most of the populations in Southwest 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 STATISTICA '99 software (StatSoft, Tulsa, Oklahoma, USA). Each triangle represents one of the 65 populations tested. In the graph, only those populations outside the main cluster (enclosed within the square with the broken line and including most breeds from Africa, Asia, and Europe) have been named. **(B)** Tridimensional plot summarizing data obtained by PCA of the insertionally polymorphic enJSRV proviruses in the 65 sheep populations tested with the Proc Factor of the statistical package SAS/STAT (SAS Institute, Cary, North Carolina, USA) according to the recommendations by Cavalli-Sforza *et al.* (29). Four factors, accounting for 86.66% of variation,

with eigenvalue  $\geq 1$  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 United Kingdom and continental European (including Denmark and Texel) sheep populations and the others. Factor 2 (on the *y* axis) explained 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 the group of populations formed by the Mouflon and Scandinavian populations (including the Hebridean, Orkney, and Soay populations) from the rest. For clarity, 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 those of 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; and **(I)** Hebridean sheep.



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*). Most of the 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 with two sets of primers that amplify, respectively, the 5' and 3' long terminal repeats (LTRs) of each provirus (including the flanking genomic DNA sequences of the host) as described (14, 17). Provirus enJSRV-18 had by far the highest frequency in our data set (85%); enJSRV-7 and enJS5F16 were detected in 27% and 30% of the samples, respectively; and enJSRV-15, enJSRV-16, and enJSRV-8 were present in only 3 to 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 ~800,000 yr B.P. (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), whereas 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') LTRs of enJSRV-7 confirm its antecedence. The divergence between the 5' and 3' LTR 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 noncoding sequences ( $\sim 2.3 \times 10^{-9}$  to  $5 \times 10^{-9}$  substitutions per site per year). enJSRV-7 appears to be the oldest provirus in our samples because it displays five nucleotide (nt) substitutions between 5' and 3' LTRs (445 nt long), whereas all the other insertionally polymorphic proviruses (including enJSRV-18) have 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 enJSRV-7.

To visualize the geographical variation of all enJSRV loci, we constructed interpolation maps from their insertion frequency values (Fig. 1, B to 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 northwest 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 northwest Europe (Fig. 1D). Two enJSRV proviruses, enJS5F16 and enJSRV-8, showed a similar geographical pattern with a high frequency in the British Isles and Scandinavia (Fig. 1, E 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 analyzed (Fig. 2). The R2 retrotype (representing the presence of enJSRV-18 only) was the predominant retrotype in most of the populations tested. 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, as 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 homogeneous R2 retrotype pattern common to the populations in Southwest Asia, suggesting direct migratory links of domestic sheep between these areas. Most of the populations from Scandinavia displayed retrotypes similar to those of Icelandic and the Faeroe Island populations, supporting the historically registered movements of the Norse settlers during the later first millennium C.E. (26). To visualize the genetic relationship of the tested populations, we analyzed the data using two different ap-

proaches: 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 (Fig. 3A). Similar results were obtained by PCA (Fig. 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 homogeneous retrotypes (R2 only, or both R2 and 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 Southwest 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 Southwest Asia and then spread throughout Europe (11). The primitive breeds survived the second migrations of improved breeds from Southwest 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 analyzed from the British Isles, the Jacob sheep, displayed a homogeneous R2 retrotype very different from that of the other British populations and more similar to that of the southwestern 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 retrotype analysis supports a direct link between the Jacob



sheep and breeds in Southwest Asia or Africa rather than other British breeds. Our study also firmly links the Soay sheep with the Mediterranean and Asiatic Mouflon.

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

# References and Notes

1. M. A. Zeder, *Proc. Natl. Acad. Sci. U.S.A.* **105**, 11597 (2008).
2. S. Colledge, J. Conolly, S. Shennan, *Eur. J. Archaeol.* **8**, 137 (2005).
3. T. D. Price, *Europe's First farmers* (Cambridge Univ. Press, Cambridge, 2000).
4. G. Barker, in *Examining the Farming/Language Dispersal Hypothesis*, P. Bellwood, C. Renfrew, Eds. (McDonald Institute Monographs, Cambridge, 2002), pp. 151–162.
5. D. R. Harris, C. Gosden, in *The Origins and Spread in Agriculture and Pastoralism in Eurasia*, D. Harris, Ed. (UCL Press, London, 1996), pp. 370–389.
6. M. L. Ryder, *Sheep & Man* (Duckworth, London, 1983).
7. S. Pedrosa *et al.*, *Proc. Biol. Sci.* **272**, 2211 (2005).
8. M. Tapio *et al.*, *Mol. Biol. Evol.* **23**, 1776 (2006).
9. J. R. Meadows, I. Cemal, O. Karaca, E. Gootwine, J. W. Kijas, *Genetics* **175**, 1371 (2007).
10. S. Naderi *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **105**, 17659 (2008).
11. A. Sherratt, in *Patterns of the Past: Studies in Honour of David Clark*, I. Hodder, G. Isaacs, N. Hammond, Eds. (Cambridge Univ. Press, Cambridge, 1981), pp. 261–306.
12. D. Helmer, L. Gourichon, E. Vila, *Anthropozoologica* **42**, 41 (2007).
13. J. D. Boeke, J. P. Stoye, in *Retroviruses*, J. M. Coffin, S. H. Hughes, H. E. Varmus, Eds. (Cold Spring Harbor Laboratory Press, Plainview, NY, 1997), pp. 343–436.
14. F. Arnaud *et al.*, *PLoS Pathog.* **3**, e170 (2007).
15. M. Palmirini, J. M. Sharp, M. De las Heras, H. Fan, *J. Virol.* **73**, 6964 (1999).
16. F. Arnaud, M. Varela, T. E. Spencer, M. Palmirini, *Cell. Mol. Life. Sci.* **65**, 3422 (2008).
17. Materials and Methods are available as supporting material on Science Online.
18. M. Hernández Fernández, E. S. Vrba, *Biol. Rev. Camb. Philos. Soc.* **80**, 269 (2005).
19. S. Hiendleder, B. Kaupe, R. Wassmuth, A. Janke, *Proc Biol Sci* **269**, 893 (2002).
20. J. Clutton-Brock, *A Natural History of Domesticated Mammals* (Cambridge Univ. Press, Cambridge, ed. 2, 1999).
21. D. Zohary, E. Tchernov, L. Horwitz, *J. Zool.* **245**, 129 (1998).
22. F. Poplin, *Ann. Genet. Sel. Anim.* **11**, 133 (1979).
23. J.-D. Vigne, in *The Holocene History of the European Vertebrate Fauna. Modern Aspects of Research*, N. Benecke, Ed. (Verlag Marie Leidorf GmbH, Rahden/Westf., Germany, 1999), pp. 295–322.
24. S. Pedrosa *et al.*, *Genet. Sel. Evol.* **39**, 91 (2007).
25. F. Pereira *et al.*, *Mol. Biol. Evol.* **23**, 1420 (2006).
26. T. H. McGovern, *Annu. Rev. Anthropol.* **19**, 331 (1990).
27. L. Excoffier, G. Laval, S. Schneider, *Evol. Bioinform. Online* **1**, 47 (2005).
28. M. P. Miller, Department of Biological Sciences, Northern Arizona University, Flagstaff, AZ (1997).
29. L. L. Cavalli-Sforza, P. Menozzi, A. Piazza, *The History and Geography of Human Genes* (Princeton Univ. Press, Princeton, NJ, 1994).
30. We thank the ECONOGENE, North-SheD, and NordGen consortia, and C. Leroux, J. DeMartini, J. Trafford, S. Haywood, V. Andresdottir, S. Thorgeirsdottir, M. Ganter, M. Reichert, Y. Bayon Gonzalez, K. Voigt, V. Feinstein, F. San Primitivo, and P. Halstead for help in obtaining some of the samples used in this study and for useful comments. We thank P. Murcia for coining the term “retrotypes.” We are grateful to G. P. Di Meo and A. Perucatti for fluorescence in situ hybridization analysis. We also thank B. Huffman ([www.ultimateungulate.com](http://www.ultimateungulate.com)), NordGen, A. Ozgul, S. Jeppson, and K. Headspeath for images reproduced in Fig. 4. This study was supported by the Biotechnology and Biological Sciences Research Council, the Wellcome Trust and in part by NIH grant HD05274, the Scottish Funding Council through a Strategic Research Developmental Grant, Fundação para a Ciência e a Tecnologia, “Misura P5 Biodiversità animale,” from the Regione autonoma della Sardegna, the Chinese Academy of Sciences, the National Natural Science Foundation of China, and the European Regional Fund (Centre of Excellence FIBIR). The Soay sheep project in St. Kilda is supported by the Natural Environment Research Council. IPATIMUP is partially supported by “Programa Operacional Ciência e Inovação 2010” (POCI 2010), VI Programa Quadro (2002–2006). M.P. is a Wolfson–Royal Society Research Merit awardee. The GenBank accession numbers of the enJSRV loci described in this paper are as follows: EE680319 (enJSRV-6), EF680306 (enJSRV-8), EF680298 (enJSRV-7), EF680299 (enJSRV-15), EF680300 (enJSRV-16), EF680301 (enJSRV-18), and AF136224 (enJS5F16).

# Supporting Online Material

[www.sciencemag.org/cgi/content/full/324/5926/532/DC1](http://www.sciencemag.org/cgi/content/full/324/5926/532/DC1)

Materials and Methods

SOM Text

Figs. S1 to S3

Tables S1 and S2

References

7 January 2009; accepted 17 March 2009

10.1126/science.1170587