making the end lozenge-shaped.

Because feathers are the only integumental covering in vertebrates that have a tufted or branched structure[27] the occurrence of similar structures in NGMC 91, coupled with its phylogenetic position near the base of birds, is strong evidence that these structures are feather homologues. The myriad findings of flightless dinosaurs from Liaoning with similar integumentary structures that have been shown by independent phylogenetic studies[28,29] to be outside of the Avialae provide important evidence that the origin of feathers is unrelated to the origin of flight in Avialae. □

Received 29 November 2000; accepted 16 February 2001.

1. Ji, Q. & Ji, S. On discovery of the earliest bird fossil in China and the origin of birds. *Chin. Geol.* **17,** 30–33 (1996).
2. Ji, Q., Currie, P. J., Ji, S. & Norell, M. A. Two feathered dinosaurs from northeastern China. *Nature* **393,** 753–761 (1998).
3. Xu, X., Tang, Z. & Wang, X. A therizinosaurid dinosaur with integumentary structures from China. *Nature* **399,** 350–354 (1999).
4. Xu, X., Wang, X. & Wu, X. A dromaeosaurid dinosaur with a filamentous integument from the Yixian Formation of China. *Nature* **401,** 262–266 (1999).
5. Xu, X., Zhou, Z. & Wang, X. The smallest known non-avian theropod dinosaur. *Nature* **408,** 705–708 (2000).
6. Chen, P.-J., Dong, Z.-M. & Zhen, S.-N. An exceptionally well preserved theropod dinosaur from the Yixian Formation of China. *Nature* **391,** 147–152 (1998).
7. Zhou, Z.-H., Wang, X.-L., Zhang, F.-C. & Xu, X. Important features of *Caudipteryx*-evidence from two nearly complete new specimens. *Vert. PalAsiat.* **38,** 241–254 (2000).
8. Ji, Q., Luo, Z.-X. & Ji, S.-A. A Chinese triconodont mammal and mosaic evolution of the mammalian skeleton. *Nature* **398,** 326–330 (1999).
9. Hou, L. & Chen, P. *Liaoxiornis delicatus* gen. et sp. nov., the smallest Mesozoic bird. *Chin. Sci. Bull.* **44,** 311–314 (1999).
10. Hou, L. *Mesozoic Birds of China* (Phoenix Valley Provincial Aviary of Taiwan, 1997).
11. Gao, K., Tang, Z. & Wang, X. A long-necked diapsid reptile from the Upper Jurassic/Lower Cretaceous of Liaoning Province, northeastern China. *Vert. PalAsiat.* **37,** 1–8 (1999).
12. Gao, K., Evans, S., Ji, Q., Norell, M. & Ji, S. Exceptional fossil material of a semi-aquatic reptile from China: the resolution of an enigma. *J. Vert. Paleontol.* **20,** 417–421 (2000).
13. Simmons, L. M. Archaeoraptor fossil trail. *Nat. Geogr. Mag.* October 128–132 (2000).
14. Chiappe, L. M., Ji, S. A., Ji, Q. & Norell, M. A. Anatomy and systematics of the Confuciusornithidae (Theropoda: Aves) from the Late Mesozoic of northeastern China. *Bull. Am. Mus. Nat. Hist.* **242,** 1–89 (1999).
15. Ji, Q. *et al.* On the sequence and age of the protobird-bearing deposits in the Sihetun-Jianshangou area, Beipiao, western Liaoning. *Prof. Papers Stratigr. Paleontol.* **27,** 74–80 (1999).
16. Swisher, C. C., Wang, Y.-Q., Wang, X.-L., Xu, X. & Wang, Y. Cretaceous age for the feathered dinosaurs of Liaoning, China. *Nature* **400,** 58–61 (1999).
17. Lo, C.-H., Chen, P.-J., Tsou, T.-Y., Sun, S.-S. & Lee, C.-Y. in *Jehol Biota* (eds Chen, P.-J. & Jin, F.) 328–340 (Science Press, Beijing, 1999).
18. Varricchio, D. J. in *Encyclopedia of Dinosaurs* (eds Currie, P. J. & Padian, K.) 749–754 (Academic, San Diego, 1997).
19. Ostrom, J. H. in *The Dinosauria* (eds Weishampel, D. B., Dodson, P. & Osmólska, H.) 269–279 (Univ. California Press, Berkeley, 1990).
20. Houck, M. A. *et al.* Allometric scaling in the earliest fossil bird. *Archaeopteryx lithographica. Science* **247,** 195–198 (1990).
21. Currie, P. J. New information on the anatomy and relationships of *Dromaeosaurus albertaensis* (Dinosauria: Theropoda). *J. Vert. Paleontol.* **15,** 576–591 (1995).
22. Barsbold, R. & Osmólska, H. The skull of *Velociraptor* (Theropoda) from the Late Cretaceous of Mongolia. *Acta Palaeontol. Polonica* **44,**189–219 (1999).
23. Norell, M. A., Makovicky, P. J. & Clark, J. M. A new troodontid theropod from Ukhaa Tolgod, Mongolia. *J. Vert. Paleontol.* **20,** 7–11 (2000).
24. Norell, M. A. & Makovicky, P. J. Important features of the dromaeosaurid skeleton: information from a new specimen. *Am. Mus. Novit.* **3215,** 1–28 (1997).
25. Norell, M. A. & Makovicky, P. J. Important features of the dromaeosaurid skeleton II: information from newly collected specimens of *Velociraptor mongoliensis*. *Am. Mus. Novit.* **3282,** 1–45 (1999).
26. Gauthier, J. in *The Origin of Birds and the Evolution of Flight* (ed. Padian, K.) 1–55 (California Academy of Sciences, San Francisco, 1986).
27. Prum, R. O. The development and evolutionary origin of feathers. *J. Exp. Zool.* **285,** 291–306 (1999).
28. Sereno, P. C. The evolution of dinosaurs. *Science* **284,** 2137–2147 (1999).
29. Holtz, T. R., Jr A new phylogeny of the carnivorous dinosaurs. *Gaia* **15,** 5–61 (2000).

Supplementary information is available on *Nature*'s World-Wide Web site (http://www.nature.com) or as paper copy from the London editorial office of *Nature*.

Acknowledgements

Comments from J. Clark, P. Makovicky, K. Padian, R. Prum and M. Siddall improved the manuscript. M. Ellison provided the Figures. We thank X. Xing and Z. Zhonghe for access to *Sinornithosaurus*. Support was provided by the Division of Paleontology (AMNH); the Ministry of Land Resources of the People's Republic of China; the National Natural Science Foundation of China; the National Geographic Society; R. Byron and L. Jaffe; and V. Pan.

Correspondence and requests for materials should be addressed to M.A.N. (e-mail: norell@amnh.org).

Genetic evidence for Near-Eastern origins of European cattle

Christopher S. Troy*†, David E. MacHugh‡†, Jillian F. Bailey*, David A. Magee*, Ronan T. Loftus*, Patrick Cunningham*, Andrew T. Chamberlain§, Bryan C. Sykes‖ & Daniel G. Bradley*

** Department of Genetics, Smurfit Institute, Trinity College, Dublin 2, Ireland*
‡ Department of Animal Science and Production, Faculty of Agriculture, University College Dublin, Belfield, Dublin 4, Ireland
§ Department of Archaeology and Prehistory, University of Sheffield, Sheffield S1 4ET, UK
‖ Institute of Molecular Medicine, University of Oxford, John Radcliffe Hospital, Headington, Oxford OX3 9DS, UK
† These authors contributed equally to this work

The limited ranges of the wild progenitors of many of the primary European domestic species point to their origins further east in Anatolia or the fertile crescent[1,2]. The wild ox *(Bos primigenius)*, however, ranged widely[3] and it is unknown whether it was domesticated within Europe as one feature of a local contribution to the farming economy[1,2,4]. Here we examine mitochondrial DNA control-region sequence variation from 392 extant animals sampled from Europe, Africa and the Near East, and compare this with data from four extinct British wild oxen. The ancient sequences cluster tightly in a phylogenetic analysis and are clearly distinct from modern cattle. Network analysis of modern *Bos taurus* identifies four star-like clusters of haplotypes, with intra-cluster diversities that approximate to that expected from the time depth of domestic history. Notably, one of these clusters predominates in Europe and is one of three encountered at substantial frequency in the Near East. In contrast, African diversity is almost exclusively composed of a separate haplogroup, which is encountered only rarely elsewhere. These data provide strong support for a derived Near-Eastern origin for European cattle.

A deep bifurcation in bovine mitochondrial DNA (mtDNA) phylogeny has been described[5,6] and is indicative of a pre-domestic divergence well in excess of 100,000 years between the two cattle taxa, *Bos indicus* and *B. taurus*. Both clades were observable in these data; 383 samples were of *B. taurus* mtDNA type whereas 9 may be classed as variants of *B. indicus*. The latter were encountered as minority types in five morphologically taurine Near-Eastern populations, and fall within Indian cattle haplotypes in the left-hand cluster of the phylogeny (Fig. 1). The aurochs sequences are more closely related to modern *B. taurus*, but it is significant that they cluster tightly in isolation, well outside the range of sequence variation among extant taurine. This phylogenetic consistency is an indication of authenticity of these ancient data. This is shared with two sequences[7] (analysed in a separate laboratory), which also display the eight transitions that separate each wild ox sequence from the modern *B. taurus* root sequences. Additionally, each sequence has been confirmed in repeated extractions from several samples and was ascertained directly from polymerase chain reaction (PCR) products, without a cloning step.

Within the 383 *B. taurus* mtDNA sequences examined, 152 haplotypes were identified that were defined by 77 polymorphic sites. Among the 152 haplotypes detected, the most predominant (T3) occurs 99 times; T1 occurs 39 times; T2 occurs eight times; another two occur seven times; two occur six times; three (including haplotype T) occur five times; two occur four times; 14 occur in triplicate; 20 are duplicated; and 106 haplotypes are unique.

Genetic loci from a centre of origin are expected to retain more ancestral variation and show higher haplotypic and nucleotide diversity, with lineage pruning through successive colonization events leading to a reduction in derived populations. Accordingly,

diversity is visibly highest here within breeds from the Middle East and Anatolia (Fig. 2). Mean pairwise differences observed between 240 base-pair (bp) haplotypes from the Middle East are 3.97 (s.d. = 2.03) and 3.49 (s.d. = 1.81) from Anatolia. Values for the European mainland (1.92; s.d. = 1.10), Britain (2.68; s.d. = 1.45), western-fringe Europe (1.47; s.d. = 0.91) and Africa (2.09; s.d. = 1.18) are consistently lower.

On examination of reduced median networks of the 383 *B. taurus* mtDNA sequences (Fig. 2) we found that almost all sequences root back to the phylogeny through one of four main haplotypes that are also numerically prominent. Figure 2 includes a skeleton network that estimates their phylogenetic relationship. These main haplotypes are designated as T1, T2 and T3, which coalesce to the central sequence T—identified as the most probable *B. taurus* root sequence when the central *Bos primigenius* haplotype is added to the skeleton network (not shown).

The numerical and topological predominance of these four haplotypes suggests that they are ancestral; furthermore, the star-like pattern of derived variants surrounding each haplotype is consistent with a history of population expansion[8]. Two other treatments of the data support this demographic model. First, each geographical grouping yields smooth pairwise mismatch distributions that are suggestive of expansions[9]. Second, Fu's *F*s statistic, which is particularly sensitive to population growth[10,11], yields a highly significant departure from neutrality in each sample ($P < 0.0001$).

In Fig. 2 all sequences are colour coded into four haplogroups, indicating which root haplotype they coalesce to. Notably, these haplogroups are geographically distributed. Both the Anatolian and Middle-Eastern networks are primarily composed of three, star-like phylogenies that are centred on haplotypes T, T2 and T3. The expansion signatures of these haplotypes within the putative centre of domestication may indicate that multiple primary haplotypes were established during this phase of aurochs capture. Members of the fourth haplogroup, T1, occur at a relatively low frequency in Anatolian and Middle-Eastern regions (twice and three times, respectively) but predominate almost absolutely in Africa where 89 out of 95 sequences coalesce together in a markedly star-like phylogeny. This haplogroup is absent from the European regions presented here, although African sequences have been reported in cattle from southern Portugal[12] where they are probably a legacy of North African historical influence. African haplotypes have also been documented in populations from the Americas[13]—perhaps as a secondary consequence of slave trade routes.

In a marked display of symmetry with the African pattern, each regional European sample produces an essentially star-like phylogeny centred, in this case, on sequence T3. Such phylogenetic topologies are suggestive of past population expansions and estimates of time depths to such events may be made using ρ, the average mutational distance from the central sequence in a haplogroup, and an estimate of mutation rate (38% per Myr)[14]. Of note, the 95% CI range of estimates for expansions around the four haplogroup clusters are consistent with the time depth of cattle domestication (T, T1, T2 and T3 yield 5,600–21,100, 4,400–16,500, 5,500–20,700 and 4,300–16,100 yr BP, respectively). Under the assumption that the advent of agropastoralism would have resulted in dramatic and sustained population increases of both the early herders and their flocks, these patterns are probably genetic signatures of population expansion after domestication.

The relationship between African haplotypic variation and that of the Near East seems to be qualitatively different from that between the latter and Europe. Most of the African diversity is clustered around a haplotype that is absent from our European samples and encountered only at very low frequency in Anatolia and the Near East. This observation and the pattern and extent of this T1 cluster diversity suggest that it is the result of a domestication-induced expansion. The near absence of this haplogroup in the Middle East and Anatolia suggests that the expansion did not occur within the fertile crescent. This provides some support for archaeological assertions of a separate African domestication[15]. In contrast, reduced diversity and the predominance of one haplogroup T3, which is also encountered as a subset of the variation encountered in Anatolia and the Middle East, strongly suggests that European cattle are derivatives of the Near-Eastern Neolithic complex. This evidence, together with the marked phylogenetic distinction of British aurochs sampled over a wide time depth, does not support the hypothesis that local domestication made any significant contribution to the establishment of agropastoralism in Europe.

Notably, whereas African variants coalesce to a different root haplotype, the difference between that and the other *B. taurus* ancestral sequences is small. The overall coalescence of all sampled domestic *B. taurus* mtDNA sequences around the central sequence T is estimated as 10,100–37,600 yr BP. This shallow divergence may indicate an ultimate origin for all extant *B. taurus* in an ancestral

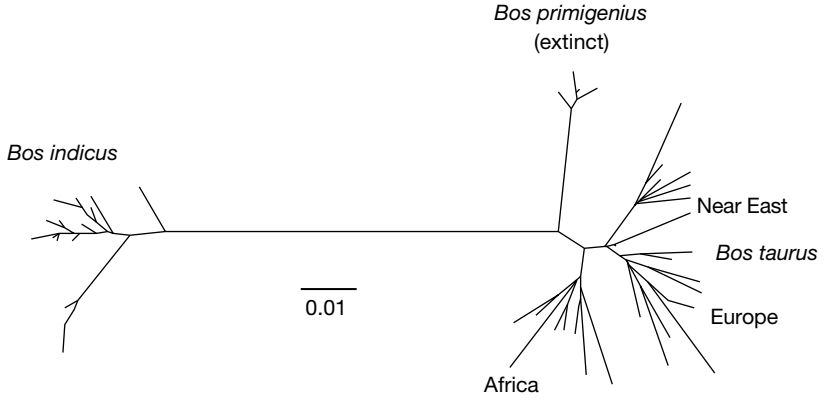

Figure 1 Neighbour-joining phylogeny of *B. indicus, B. taurus* and extinct British *B. primigenius* mtDNA control-region haplotypes. The sequences compared are 201 bp in length and include positions 16,042–16,158 and 16,179–16,262. Gaps were excluded and no distance correction was employed. The four ancient sequences cluster tightly (bootstrap value of 94%; 1,000 replications) and form an outgroup to the modern *B. taurus* haplotypes. This is despite their origin in animals dispersed geographically and chronologically. These ancient sequences all share eight transitions that separate them from the modern *B. taurus* root sequence. Notably, these substitutions are also shared with two previous British aurochs sequences[7]. The *B. indicus* cluster yielded a bootstrap value of 100% and the phylogeny root, estimated using bison as an outgroup (data not shown), occurs between this clade and all others. A scale bar (divergence of 0.01) is shown.

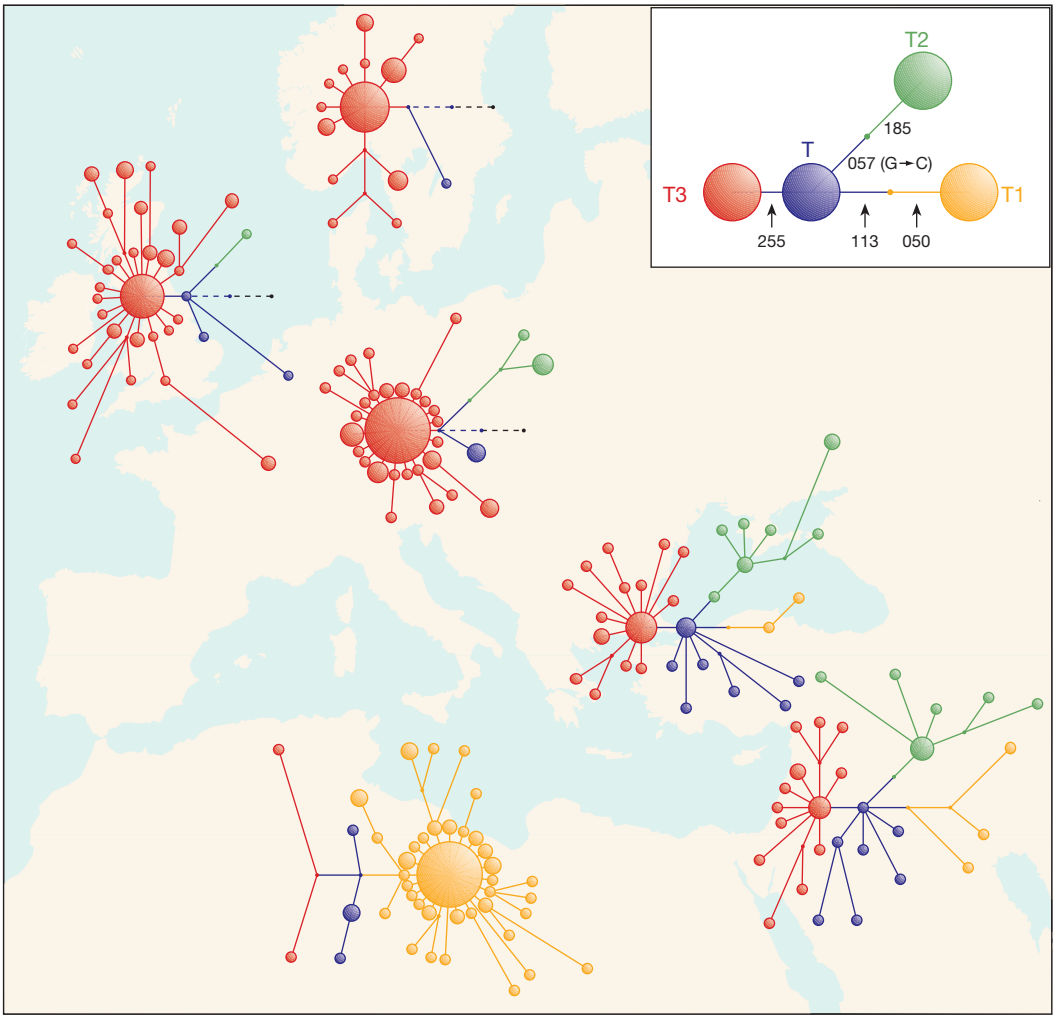

Figure 2 *B. taurus* mtDNA reduced median networks constructed from six regional haplotype groups. Inset, The relationships of the four central, primary *B. taurus* haplotypes, T, T1, T2 and T3. T is defined by a transition at position 16,255 from the reference sequence[18]; T1 by transitions at 16,050, 16,113 and 16,255; T2 by transitions at 16,185 and 16,255 plus a G to C transversion at 16,057; and T3 is identical to the reference sequence. The spatial arrangement of the skeleton network and the colour codes are preserved in the full data networks (placed in the region of origin on the background map). The samples are grouped as originating in the Middle East, Anatolia, mainland Europe, Britain, western-fringe Europe and Africa. Reduced median phylogenetic networks were constructed manually. Haplotypes encountered in each region (coloured circles) and unsampled intermediate nodes or unsampled primary haplotypes (small points) are shown. Circle areas are proportional to haplotype frequencies and colour coding indicates which of the four skeleton network haplotypes they root to. Most (71%) sites that were reduced correspond to the hypermutable sites identified above, and are underlined below. Networks were reduced at the following positions: Anatolian sequences, 16,050, 16,057, 16,074, 16,110, 16,113, 16,138, 16,247 and 16,248; Middle-Eastern sequences, 16,049, 16,050, 16,058, 16,074, 16,085, 16,109, 16,113, 16,121, 16,122, 16,200, 16,231, 16,247 and 16,260; Continental Europe sequences, 16,110, 16,164 and 16,260; British sequences, 16,049, 16,050, 16,057, 16,074, 16,109, 16,113, 16,122, 16,127 and 16,138. The Western Europe network did not require resolving. Geographical distribution of the four haplogroups is clear. T3 predominates in Europe and along with T and T2 comprises almost all Near-Eastern variation. Haplogroup T1 dominates African diversity but is scarcely represented elsewhere.

population of the wild ox, which was itself limited in diversity, perhaps as a result of climatically mediated range restriction. □

Methods

Modern samples

We examined 392 animals from 34 breeds: 208 from Europe, 89 from the Near East and 95 from Africa. For network analysis, populations were subdivided into six geographical groupings: western-fringe Europe (Icelandic, Kerry, Norwegian Red, Telemark, Westland Fjord); Britain (Galloway, Highland, Aberdeen Angus, Hereford, Jersey); European mainland (German Black, Friesian, Limousin, Charolais, Simmental, Romagnola, Barrenda, Sykia); Anatolia (Anatolian Black, East Anatolian Red, South Anatolian Red, Turkish Grey); Middle East (Damascus, Kurdish, North Iraq, South Iraq); and Africa (Egyptian, N'Dama, Somba, Kapsiki, Namchi, Kuri, Butana, Kenana, White Fulani).

Archaeological samples

The six aurochs mtDNA sequences examined as part of this survey consisted of two previously published[7] sequences, D740 and D812, and four new sequences from *B. primigenius* skeletal remains excavated from chronologically distinct sites widely dispersed across present-day England. The four new samples are (sample code and estimated age in brackets): Charterhouse Warren Farm Swallet (CHWF, 4,090–3,720 BP); Totty Pot (TP65, 7,570–7,320 BP); Carsington Pasture Cave (CPC98, 6,200–5,650 BP); and North Ferriby (NORF, 3,990–3,720 BP).

Extraction, DNA amplification and sequencing

DNA samples were extracted from blood[5], semen[13] and hair (see below). We used about 10–12 hair follicles per animal. Follicles were treated with 50 μl of 200 mM NaOH at 97 °C for 30 min. Then 50 μl of 200 mM HCl, 100 mM Tris-HCl, pH 8.5 was added. PCR and sequencing conditions were as described[12]. The region analysed was a defined, highly variable region of the mtDNA control region between bases 16,023 and 16,262 (ref. 5). We extracted and purified archaeological samples essentially as described[16]. General ancient DNA laboratory practices, PCR amplification procedures and DNA sequencing methods have been detailed before[17]. Two overlapping PCR products were amplified from each extract ($AN2_{FOR}$–$AN1_{REV}$ and $AN1_{FOR}$–$AN3_{REV}$). The PCR primers and their location and orientation in the reference sequence[18] are: $AN2_{FOR}$ (16,022–16,041); $AN1_{REV}$ (16,178–16,159); $AN1_{FOR}$ (16,159–16,178); and $AN3_{REV}$ (16,334–16,314). At least two independent samples were taken from each bone in the panel. DNA was extracted and purified from each independent sample in duplicate. A range of PCR amplifications and

subsequent direct sequence assays were performed from each purified DNA extract. The mtDNA sequences derived from each bone were therefore multiply verified through independent samples, extractions, amplifications and sequence determinations. In all cases for each bone in the panel, the replicated mtDNA sequences were consistent across all samples and extractions. The veracity and integrity of the aurochs sequence haplogroup is strongly supported by the congruity of the bones analysed in Dublin (CHWF, TP65, CPC98 and NORF) with those previously examined in Oxford (D740 and D812).

Analysis of mitochondrial sequence data

We aligned sequences by eye. Data from previous studies[5–7,17] were also included and novel sequences were submitted to GenBank (accession numbers AF336383–AF3366748). We constructed a neighbour-joining phylogeny using uncorrected distances with gapped positions excluded and the Neighbour program in the Phylip package[19]. Arlequin 2.0 (ref. 20) was used to estimate nucleotide diversity values for each population, to estimate parameters and goodness of fit from mismatch distributions and to compute Fu's *Fs* test of selective neutrality[10]. Phylogenetic analysis was also performed using reduced median networks, constructed manually according to ref. 21. Sites prone to hypermutablity were identified using a network of zebu and taurine samples examined in ref. 6. For the reduced median networks shown, most of the reductions (71%) were at these sites. The mutation rate for the 240-bp region of the cattle D-loop was calculated using the estimate of the transition/transversion ratio for the sequence data presented here (61/1). This agrees well with previous estimates on smaller data sets 57/1 (ref. 6) and 41/1 (ref. 7), and as the new rate incorporates data presented in those papers it is used in preference to either. The *Bison–Bos* divergence is considered to be around 1 Myr. The three transversions observed in the 240-bp region between the *Bison–Bos* groups constitute the equivalent of 183 transitions. The one-lineage rate was estimated as 38% per Myr or one substitution per 10,928 yr. The central 95% credible region for the expansion time was calculated for each of the four main cattle haplogroups using the program CRED[22].

Received 10 August 2000; accepted 19 January 2001.

1. Smith, B. D. *The Emergence of Agriculture* (Scientific American Library, New York, 1995).
2. Harris, D. R. in *The Origins and Spread of Agriculture and Pastoralism in Eurasia* (ed. Harris, D. R.) (UCL, London, 1996).
3. Epstein, H. *The Origin of the Domestic Animals of Africa* (Africana, New York, 1971).
4. Bökönyi, S. *The History of Domestic Animals in Central and Eastern Europe* (Akadémiai Kiadó, Budapest, 1974).
5. Loftus, R. T., MacHugh, D. E., Bradley, D. G., Sharp, P. M. & Cunningham, P. Evidence for two independent domestications of cattle. *Proc. Natl Acad. Sci. USA* **91,** 2757–2761 (1994).
6. Bradley, D. G., MacHugh, D. E., Cunningham, P. & Loftus, R. T. Mitochondrial diversity and the origins of African and European cattle. *Proc. Natl Acad. Sci. USA* **93,** 5131–5135 (1996).
7. Bailey, J. F. *et al.* Ancient DNA suggests a recent expansion of European cattle from a diverse wild progenitor species. *Proc. R. Soc. Lond. B* **263,** 1467–1473 (1996).
8. Richards, M. B., Macaulay, V. A., Bandelt, H.-J. & Sykes, B. C. Phylogeography of mitochondrial DNA in western Europe. *Ann. Hum. Genet.* **62,** 241–260 (1998).
9. Rogers, A. R. & Harpending, H. Population growth makes waves in the distribution of pairwise genetic differences. *Mol. Biol. Evol.* **9,** 552–569 (1992).
10. Fu, Y.-X. Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. *Genetics* **147,** 915–925 (1997).
11. Excoffier, L. & Schneider, S. Why hunter-gatherer populations do not show signs of Pleistocene demographic expansions. *Proc. Natl Acad. Sci. USA* **96,** 10597–10602 (1999).
12. Cymbron, T., Loftus, R. T., Malheiro, M. I. & Bradley, D. G. Mitochondrial sequence variation suggests an African influence in Portuguese cattle. *Proc. R. Soc. Lond. B* **266,** 597–603 (1999).
13. Troy, C. S. *Mitochondrial DNA Phylogeny and Biogeography of Domestic Cattle.* Ph.D Thesis, Univ.-Dublin (1998).
14. Macaulay, V. *et al.* The emerging tree of West Eurasian mtDNAs: a synthesis of control-region sequences and RFLPs. *Am. J. Hum. Genet.* **64,** 232–249 (1999).
15. Wendorf, F. & Schild, R. Nabta Playa and its role in northeastern African prehistory. *J. Anthropol. Archaeol.* **17,** 97–123 (1998).
16. Yang, D. Y., Eng, B., Waye, J. S., Dudar, J. C. & Saunders, S. R. Technical note: improved DNA extraction from ancient bones using silica-based spin columns. *Am. J. Phys. Anthropol.* **105,** 539–543 (1998).
17. MacHugh, D. E. *et al.* Early mediaeval cattle remains from a Scandinavian settlement in Dublin: genetic analysis and comparison with extant breeds. *Phil. Trans. R. Soc. Lond. B* **354,** 99–109 (1999).
18. Anderson, S. *et al.* Complete sequence of bovine mitochondrial DNA. *J. Mol. Biol.* **156,** 683–717 (1982).
19. Felsenstein, J. PHYLIP: Phylogeny Inference Package. (University of Washington, Seattle, 1993).
20. Schneider, S., Roessli, D. & Excoffier, L. *ARLEQUIN Ver. 2.0: a Software for Population Genetics Data Analysis* (Department of Anthropology, University of Geneva, Switzerland, 1999).
21. Bandelt, H. J., Forster, P., Sykes, B. C. & Richards, M. B. Mitochondrial portraits of human populations using median networks. *Genetics* **141,** 743–753 (1995).
22. Macaulay, V. A. CRED: Credible Regions for Coalescence Times. (Univ. Oxford, 1998).

Acknowledgements

We thank C. Hawkes, O. Ertugrul, A. H. Al Haboby, A. H. Harba, M. A. A. El-Barody, E. Thompson, T. Goodchild, H. Halila, A. Swaid, G. Guneren, B. Tekbas, M. Bruford, B. Sauveroche, G. Kana, D. Achu-Kwi, M. Diallo, L. Gnaho, K. Papadopolous, A. G. Georgoudis, C. Gaillard, O. Hanotte, E. Rege, the Nordic GeneBank and C. Hawes for assistance or provision of samples. We also thank the Highland Cattle Society and the Black and Belted Galloway Societies for sample information. This work was partly funded by a European Commission contract. D.E.M. was supported by a Wellcome Trust Fellowship in Bioarchaeology. J.F.B. is a European Commission Marie Curie Fellow. Radiocarbon dating at the Oxford Radiocarbon Accelerator Unit was funded by NERC.

Correspondence and requests for materials should be addressed to D.G.B. (e-mail: dbradley@tcd.ie).

The highly reduced genome of an enslaved algal nucleus

Susan Douglas*, Stefan Zauner†, Martin Fraunholz†‡, Margaret Beaton‡§, Susanne Penny*, Lang-Tuo Deng§, Xiaonan Wu§, Michael Reith*, Thomas Cavalier-Smith‡§ & Uwe-G Maier†

** National Research Council of Canada Institute for Marine Biosciences and Program in Evolutionary Biology, Canadian Institute of Advanced Research, 1411 Oxford Street, Halifax, Nova Scotia B3H 3ZI, Canada*
† Cell Biology and Applied Botany, Philipps-University Marburg, Karl-von-Frisch-Strasse, D-35032 Marburg, Germany
§ Program in Evolutionary Biology, Canadian Institute of Advanced Research, Department of Botany, University of British Columbia, Vancouver, British Columbia V6T 1Z4, Canada

Chromophyte algae differ fundamentally from plants in possessing chloroplasts that contain chlorophyll *c* and that have a more complex bounding-membrane topology[1]. Although chromophytes are known to be evolutionary chimaeras of a red alga and a non-photosynthetic host[1], which gave rise to their exceptional membrane complexity, their cell biology is poorly understood. Cryptomonads are the only chromophytes that still retain the enslaved red algal nucleus as a minute nucleomorph[2–4]. Here we report complete sequences for all three nucleomorph chromosomes from the cryptomonad *Guillardia theta*. This tiny 551-kilobase eukaryotic genome is the most gene-dense known, with only 17 diminutive spliceosomal introns and 44 overlapping genes. Marked evolutionary compaction hundreds of millions of years ago[1,4,5] eliminated nearly all the nucleomorph genes for metabolic functions, but left 30 for chloroplast-located proteins. To allow expression of these proteins, nucleomorphs retain hundreds of genetic-housekeeping genes[5]. Nucleomorph DNA replication and periplastid protein synthesis require the import of many nuclear gene products across endoplasmic reticulum and periplastid membranes. The chromosomes have centromeres, but possibly only one loop domain, offering a means for studying eukaryotic chromosome replication, segregation and evolution.

Soon after the symbiogenetic origin of chloroplasts from cyanobacteria[1] to form the common ancestor of green plants, red and glaucophyte algae (kingdom Plantae[6,7]), even more complex eukaryotic cells arose by secondary symbiogenesis[1,3,4] (Fig. 1). Such chimaeric integration of two evolutionarily distant eukaryotic cells occurred independently in the common ancestor of cryptomonads and other chromophytes, in which the endosymbiont was a red alga, and in chlorarachneans, which acquired a green alga[1,3,4]. In both cryptomonads and chlorarachneans, a flagellate host contributed the nucleus, endomembranes and mitochondria to the chimaera, whereas the photosynthetic endosymbiont provided its chloroplast, plasma membrane (the periplastid membrane[1,3,4]) and a second nucleus (the nucleomorph), which became miniaturized[3–5]. The nucleomorph of both groups kept its envelope, nuclear pores[8] and three minute chromosomes[9]. In the ancestor of cryptomonads and chromobiotes (treated as kingdom Chromista[6,8]) but not chlorarachneans, the former food vacuole membrane originally enclosing the enslaved endosymbiont apparently fused with the nuclear envelope[10], placing it in the rough endoplasmic reticulum[8,10] (RER; Fig. 1). Cryptomonad cells depend on four genomes, each encoding distinct protein synthesis machineries in discrete

‡ Present addresses: Goddard Laboratories, University of Philadelphia, 415 South University Avenue, Philadelphia, Pennsylvannia 19104-6018, USA (M.F.); Biology Department, 63B York Street, Mount Allison University, Sackville, NB, Canada E4L 1G7 (M.B.); Department of Zoology, University of Oxford, South Parks Road, Oxford OX1 3PS, UK (T.C.-S.).