Molecular Biology and Evolution

Phylogeography and Origin of Indian Domestic Goats

Manjunath B. Joshi, Pramod K. Rout,† Ajoy K. Mandal,† Chris Tyler-Smith,‡[1] Lalji Singh,* and Kumarasamy Thangaraj**

*Centre for Cellular and Molecular Biology, Hyderabad, India; †Central Institute of Research on Goat, Mathura, India; and ‡Department of Biochemistry, University of Oxford, Oxford, U.K.

The Indian subcontinent contains 20 well-characterized goat breeds, which vary in their genetic potential for the production of milk, meat, and fibre; disease resistance; heat tolerance; and fecundity. Indian goats make up 20% of the world's goat population, but there has been no extensive study of these economically important animals. Therefore, we have undertaken the present investigation of 363 goats belonging to 10 different breeds from different geographic regions of India using mtDNA sequence data from the HVRI region. We find evidence for population structure and novel lineages in Indian goats and cannot reconcile the genetic diversity found within the major lineage with domestication starting 10,000 years ago from a single mtDNA ancestor. Thus, we propose a more complex origin for domestic goats.

Introduction

The domestic goat *Capra hircus* is an important livestock species in India and other developing countries. Because it provides a good source of meat, milk, fiber, and skin, it is popularly known as the ''poor man's cow'' (MacHugh and Bradley 2001). Goats have fulfilled agricultural, economic, cultural, and even religious roles from very early times in human civilization. They are the most adaptable and geographically widespread livestock species, ranging from the high altitude of the Himalayas to the deserts of Rajasthan and humid coastal areas of India. Archaeological evidence indicates that the goat was one of the first animals to be domesticated by humans around 10,000 years ago at the dawn of the Neolithic period in the Fertile Crescent (Porter 1996; Pringle 1998). Goats played a central role in the Neolithic agricultural revolution and the spread of human civilization around the globe (Legge 1996; Porter 1996; Pringle 1998; Zeder and Hesse 2000). It has been suggested that at least two wild species of the genus *Capra* have contributed to the gene pool of domestic goats (Mannen, Nagata, and Tsuji 2001), whereas some studies have suggested that an independent domestication in Pakistan gave rise to the Cashmere breeds (Meadow 1996; Porter 1996). However, the origin of the domestic goat remains uncertain and controversial, despite the archaeological evidence.

Mitochondrial DNA (mtDNA) contains highly informative polymorphic sites and its simple maternal inheritance without recombination makes it useful for population studies in many organisms (Brown et al. 1986; Vigilant et al. 1991; Loftus et al. 1994; Bradley et al. 1996; Manceau et al. 1999; Luikart et al. 2001). Luikart et al. (2001) carried out a worldwide survey of domestic goat mtDNA diversity and identified three major mtDNA lineages. Lineage A was the most common in all continents. Lineage B was found in the Indian subcontinent, Mongolia, and Southeast Asia. Lineage C was observed in a few samples from Mongolia, Switzerland, and Slovenia. These three lineages were judged to have diverged over 200,000 years ago; this ancient divergence time and the different geographical localizations of the lineages suggested the likelihood of either multiple domestication events or introgression of additional lineages after the original domestication. The predominance of a single lineage in goats was in contradiction to the dual origin of livestock species reported in cattle, sheep, and pigs (MacHugh and Bradley 2001).

This initial survey provides a context for more detailed regional studies. India is a vast subcontinent with about 123 million goats comprising 20 recognized breeds and nondescript (Local) goats, which together make up approximately 20% of the world's goat population (http://fao.org). Indian goat breeds exhibit enormous variations in fecundity; production of meat, milk, and fibre; draughtability; disease resistance; and heat tolerance. However, a previous analysis of Indian goats was confined to a single study by Luikart et al. (2001) using a limited number of samples (14 individuals from five breeds). There is, therefore, a need for an extensive study of Indian goat breeds to understand their origin, divergence, and past migration patterns. Hence, we have undertaken the present investigation of 363 goats belonging to 10 different breeds from different geographic regions (fig. 1). We find evidence for population structure and additional lineages in Indian goats and cannot reconcile the genetic diversity found within the major lineage with domestication starting about 10,000 years ago from a single mtDNA ancestor. Thus, we propose a more complex origin for domestic goats.

Materials and Methods

Sample Collection and DNA Isolation

A total of 363 unrelated goat blood samples were collected from natural habitat belonging to four major geographical agroclimatic zones of India, including at least one breed from each major geographical region (fig. 1). Jamunapari goats are found in isolated pockets and sampling was carried out in almost all the villages where breeding takes place. Similarly, Jakhrana goats have a very narrow breeding tract and are confined to few villages in Rajasthan. The breeds Barbari, Black Bengal, Sirohi, Marwari, Osmanabadi, and Kutchi are distributed over

[1] Present address: The Wellcome Trust Sanger Institute, Hinxton, U.K.

Key words: goats, mtDNA, India, phylogeography, livestock.

E-mail: thangs@ccmb.res.in.

Sequence data from this article have been deposited with GenBank under accession numbers AY155674 to AY156039.

Mol. Biol. Evol. 21(3):454–462. 2004
DOI: 10.1093/molbev/msh038
Advance Access publication December 5, 2003
Molecular Biology and Evolution vol. 21 no. 3

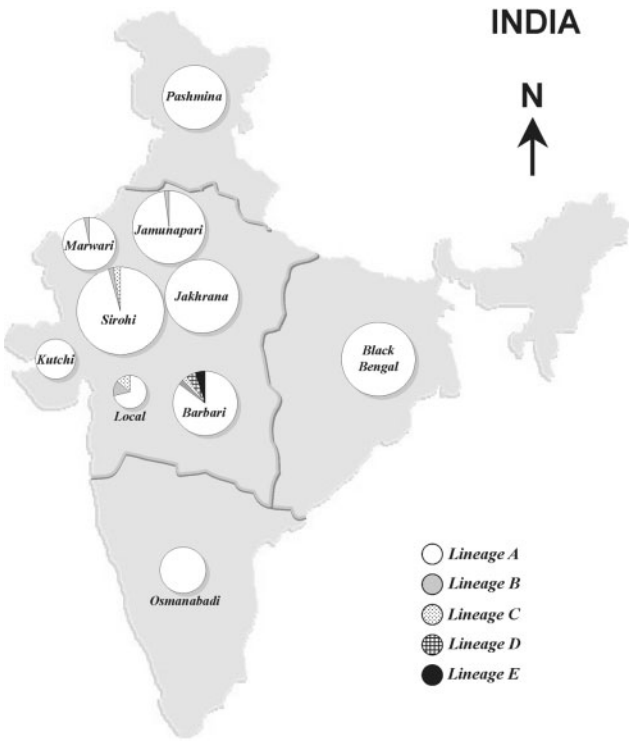

FIG. 1.—Geographical and lineage distributions of the Indian goat breeds sampled. Circle area is proportional to sample size.

a number of villages in their breeding area. Local goats are nondescript goats distributed over the Barbari home tract and are sometimes considered as the founder population for the breed. An effort was made to collect samples from unrelated individuals based on the information provided by farmers. Details of the breeds, regions, and sample sizes are given in table 1. DNA was isolated from the samples using the standard protocol published elsewhere (Thangaraj et al. 2002).

Analysis of Mitochondrial DNA

PCR Amplification

To amplify hypervariable region I (HVRI) of goat mtDNA, a pair of primers was designed using the known goat mtDNA sequence (AB044304, [Mannen, Nagata, and Tsuji 2001]). Primers were designed using the GENETOOL package and synthesized in an ABI392 oligosynthesizer (PerkinElmer, Foster City, Calif.). Primer sequences were as follows: forward 5′-GCCTTCATGTAGTTTACTGT-3′ and reverse 5′-GGGCCATCTCACCTAAAATC-3′. PCR amplification was carried out in a 10 μl reaction volume containing 5.0 ng of DNA, 10 pM of each primer, 200 μM of dNTPs, 1X PCR buffer containing 2.0 mM $MgCl_2$, and 2 U of AmpliTaq Gold (PerkinElmer). Amplification was carried out in a GeneAmp9600 thermal cycler (PerkinElmer) employing the following conditions: 94°C for 10 min; 30 cycles at 94°C for 1 min, 60°C for 1 min, and 72°C for 1 min; and final extension at 72°C for 5 min.

Sequencing of Amplicons

PCR products were sequenced directly on both strands by adopting the strategy described elsewhere (Thangaraj et al. 2003) using 50 ng (2.0 μl) of PCR product and 4 pM (1.0 μl) of primer, 4 μl of BigDye Terminator ready reaction kit (Perkin Elmer), and 3.0 μl of double distilled water to adjust the volume to 10.0 μl. Cycle sequencing was carried out in a GeneAmp9600 thermal cycler (Perkin Elmer) employing 30 cycles at 96°C for 10 s, 50°C for 5 s, and 60° for 4 min. Extended products were purified by alcohol precipitation followed by washing with 70% alcohol. Purified samples were dissolved in 10 μl of 50% Hi-Di formamide and analyzed in an ABI3700 automated DNA Analyzer (Perkin Elmer).

Data Analysis

Four hundred fifty-seven base pairs from the mtDNA HVRI region of 363 individuals of 10 Indian goat breeds and 14 sequences of six wild goat species (downloaded from NCBI; accession numbers AJ317864 to AJ31787), were aligned in the ClustalX package (Thompson, Higgins, and Gibson 1994). The GENEDOC package was used for formatting the sequences to make them compatible with the desired software. Reduced median networks (Bandelt, Forster, and Rohl 1999) were drawn using the program Network 3.1.1.1 (www.fluxus-engineering.com) with the parameters set to a weight of 2 and threshold 1; ρ estimates ± standard error (SE) were obtained using the same program. Haplotype diversity and its SE, Fu's Fs statistics (Fu 1997), mismatch analysis (Schneider et al. 1999), AMOVA (analysis of molecular variance), mean number pairwise differences, and population pairwise Φ_{ST} values were computed using ARLEQUIN version 2.001 (http://anthropologie.unige/arlequin) (Excoffier, Smouse, and Quattiro 1992). Φ_{ST} values were calculated using 100 bootstraps. These Φ_{ST} values were used to reconstruct an NJ/UPGMA tree in comparison with wild goat breeds using the PHYLIP package version 3.6 (Felsenstein 1993). A Maximum-likelihood tree was reconstructed with Kimura two- parameter model in PAUP* software version 4.0 (Swofford 1993). The substantial heterogeneity in substitution rates among nucleotide sites was estimated under Kimura two-parameter with gamma correction, which has been observed for other mammals (Vigilant et al. 1991; Bradley et al. 1996; Excoffier and Schnedier 1999; Luikart et al. 2001). The Tamura and Nei distances within goats and between goat and sheep were calculated in MEGA version 2.1 (Kumar et al. 2001) using a value for the alpha parameter of 0.29 (Luikart et al. 2001). Pairwise Φ_{ST} values between breeds were displayed by multidimensional scaling (MDS) using SPSS 11.0; the stress value was 0.12. A Mantel test was carried out to estimate the correlation between geographical distances and genetic distances between breeds (Adams 1999).

Results

mtDNA Variation in Indian Domestic Goats

We have examined 363 mtDNA HVRI sequences belonging to 10 different Indian goat breeds represent-

Table 1
Geographic Distribution of Indian Goat Breeds and Their Haplotype Diversity

Sample Number	Geographic Distribution	Breed	Sample Size	Number of Haplotypes	Diversity	SE of Diversity (±)
1	North Western temperate region	Pashmina	40	17	0.926	0.022
2	Eastern region	Black Bengal	50	28	0.960	0.014
3	Southern peninsular region	Osmanabadi	16	12	0.917	0.064
4	North Western arid and semi-arid region	Kutchi	10	5	0.844	0.080
5		Marwari	30	18	0.949	0.023
6		Sirohi	69	36	0.964	0.010
7		Barbari	41	28	0.964	0.018
8		Jakharana	50	28	0.940	0.023
9		Local	7	7	1.000	0.076
10		Jamunapari	50	34	0.984	0.007

ing different geographical regions of India, along with published wild goat sequences. Comparison of the 363 sequences revealed 200 different haplotypes. In addition to the base substitutions, deletion of two nucleotides was observed in two individuals each of Black Bengal and Barbari goats. The number of haplotypes found in each breed ranged from five to 36, and diversity values (for breeds with sample size ≥ 30) from 0.926 ± 0.022 in the Pashmina to 0.984 ± 0.007 in the Jamunapari (table 1). The overall ratio of transitions:transversions (16.7:1), calculated using a network, revealed a heavy transition bias in domestic goats similar to the 17:1 ratio seen before (Luikart et al. 2001), which has also been observed in bovine and human control regions.

The data were used to construct a reduced median network (fig. 2). The network shows considerable diversity, but one high-frequency haplotype, H1, (present in 27 individuals, 7.4% of the total) stands out at the center of a star-shaped phylogeny. It has 21 one-step neighboring haplotypes, of which four contained five or more individuals, so that the total number of individuals at this distance was 51 (14% of the total). Haplotypes two, three, four, and five steps away are represented by 58, 84, 73, and 12 individuals, respectively (16%, 23.1%, 20.1%, and 3.3%). We compared these haplotypes with the A, B, and C lineages described previously (Luikart et al. 2001). Luikart and colleagues reported the presence of lineage A and B in the Indian subcontinent but not lineage C. The H1 haplotype and its neighbors fell within the lineage A, as would be expected from their high frequency in both data sets. A few samples of Indian goats clustered with lineages B and C, demonstrating that C is present in India at low frequency; its previously reported absence was probably the result of the small sample size examined. We also observed additional lineages, D and E as distant from haplotype H1 as lineage B (16 steps, compared with 13 steps for lineage B [fig. 2]), showing that considerable additional diversity exists within Indian domestic goats. These two lineages are found only in the Barbari breed (fig. 1).

Population Structure

Sample sizes for individual breeds ranged from seven to 69 (table 1), with a mean of 36; seven of the 10 breeds consisted of 40 or more individuals, so it was possible to carry out comparisons of the breeds. We first constructed a neighbor-joining (NJ) tree of the Indian goat breeds and the wild species (fig. 3). All domestic goats clustered together, as expected, but within them, three groups could be distinguished. Local goats (nondescript) formed a separate group; this may reflect the small sample size, but they do contain a high proportion of divergent lineages illustrated by the long branches in figure 2. The other two clusters correspond to the geographical regions of origin: the four breeds from the northwest (NW) of the country are distinct from breeds from the central (C), eastern (E), and southern (S) regions. There thus appears to be geographical structure within India. We also constructed an NJ tree of individual Indian goat sequences incorporating the previously reported HVRI sequences belonging to different lineages (A, B, and C) and wild goats (fig. 4). As in the network, we observed a few haplotypes clustering with lineages B and C and some of the haplotypes falling into two separate lineages (lineages D and E). A similar clustering pattern was found when the NJ tree was constructed, excluding divergent haplotypes and haplotypes belonging to lineages B and C from Indian goats.

We investigated the geographical structure further by AMOVA analysis. Overall, 83% of the variation was within breeds and 17% among breeds (table 2); this was significantly different from zero ($P < 0.01$). When a hierarchical approach was taken and goat breeds were grouped according to their geographical location (NW versus C+S+E), 14% ($P < 0.01$) of the variation was among groups, and 8% ($P < 0.01$) was within groups but among breeds (table 2, two groups), demonstrating significant geographical structuring. If Local goats were considered as a distinct third group in this analysis, because of their separate position in the NJ tree, the results obtained were similar (table 2, three groups).

Genetic distances between breeds were calculated taking into account the molecular distances between haplotypes, and the resulting Φ_{ST} values were positively correlated with the geographical distances between the sampling sites ($r = 0.31$, $P < 0.05$ using a Mantel test). Φ_{ST} values are displayed as an MDS plot in figure 5. Here, the same geographical clustering was seen again: NW and C+S+E clusters, with Local goats as an outlier. When Φ_{ST} values were calculated using only the A lineage hap-

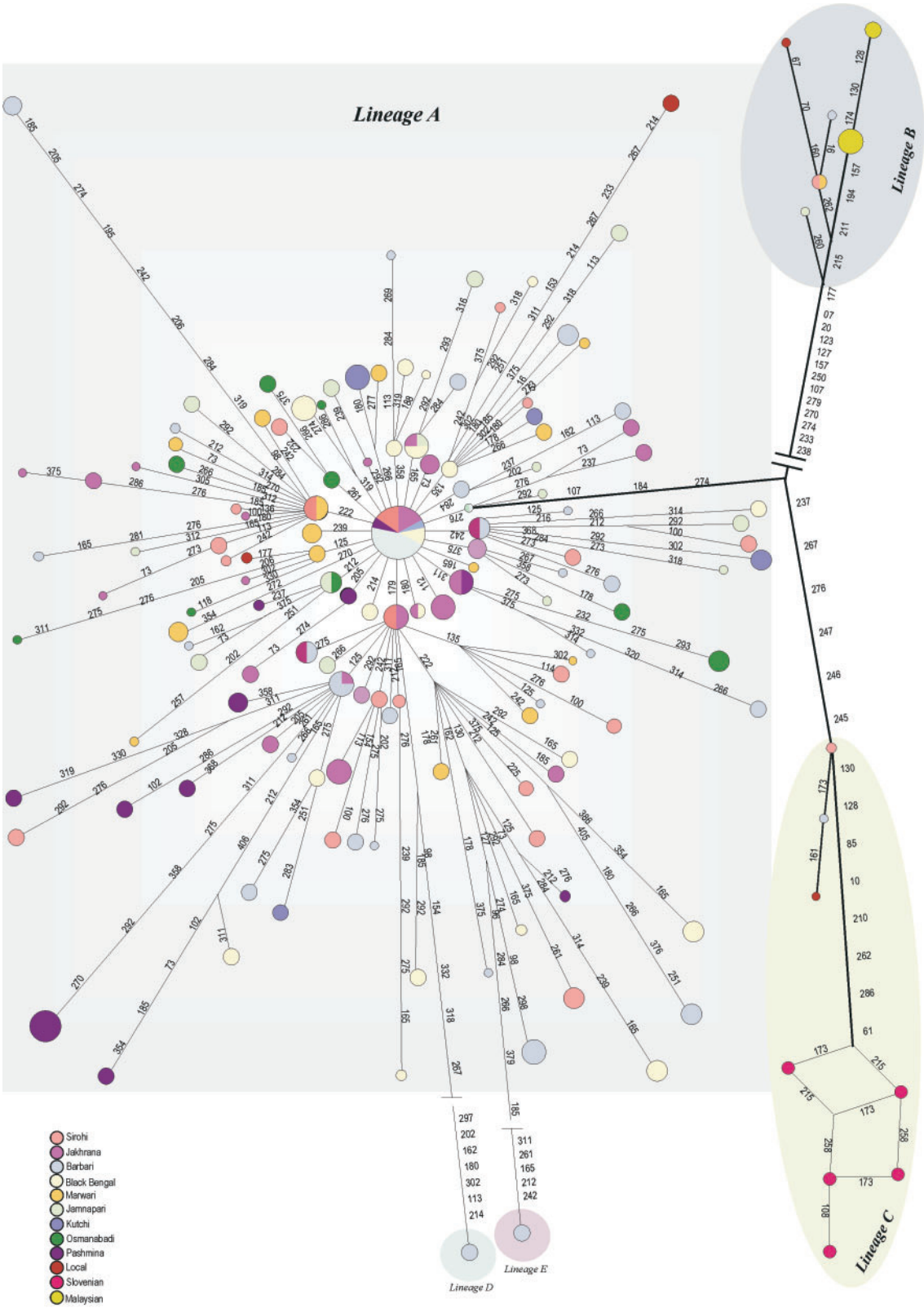

FIG. 2.—Reduced median network of Indian goat mtDNAs. Circles represent haplotypes and have a size proportional to frequency. Mutational differences are shown on lines. Shaded areas indicate the lineages A to E.

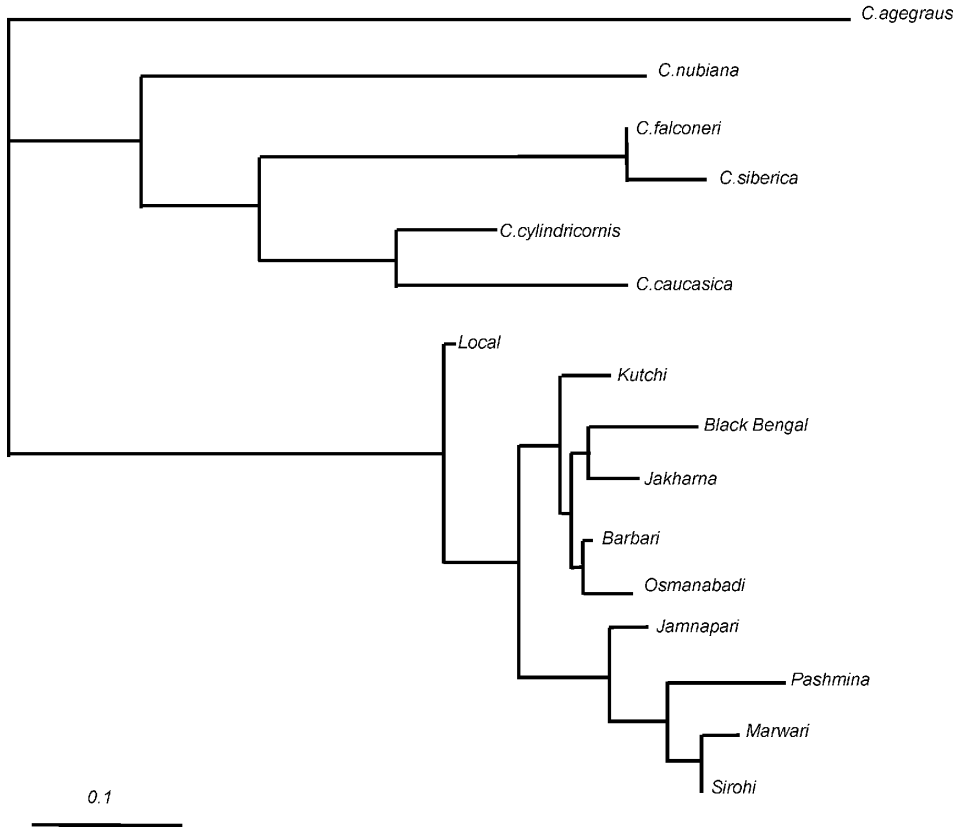

FIG. 3.—Neighbor-joining tree of Indian goat breeds and wild goat mtDNA sequences.

lotypes, a similar clustering pattern (NW versus C+S+E) was seen, except that Local goats fell into the C+S+E cluster (results not shown). The mean number of pairwise differences was calculated for the northwest breeds in comparison with the remaining breeds with and without local goats. For the northwest breeds, this number was 10.62 ± 4.2, whereas for the remaining breeds without local goats it was 8.87 ± 4.1 and with local goats it was 9.46 ± 4.6, still less than northwest breeds. Diversity within the NW group is thus greater than within the C+S+E group. We, therefore, conclude that there is substantial geographical structure to the goat mtDNA variation within India.

Population Expansion

Data on genetic variation contain information about demographic history, and mismatch distributions (pairwise comparisons) of mtDNA have been widely used to explore demography (Rogers and Harpending 1992). A constant-sized population is expected to show a ragged, multimodal distribution, while an expanding population shows a smooth, unimodal distribution. The match of a real data set to these models can be assessed by the ''raggedness,'' Fu's Fs statistics and the SSD values. Fu's Fs statistic provides a test for population growth. It is based on the probability of having a number of alleles greater or equal to the observed number in sample drawn from a stationary population. We carried out the analysis with 5,000 simulations. The overall validity of the estimated demographic model was tested by comparing the distribution of a test statistic, SSD (sum of squared differences), between the observed and the estimated mismatch distribution using a bootstrap approach. Significant SSD values can be taken as evidence for departure from the estimated demographic model, which can include either an expanding or a stationary population (Excoffier and Schneider 1999). Distributions derived from data may therefore be judged to match that expected from a constant-sized or expanding population; in the latter case, the position of the peak in the mismatch distribution curve provides information about the time when the expansion began. mtDNA mismatch distributions for the combined data set and the individual breeds were therefore determined and are shown in figure 6.

The combined data set shows a smooth curve (raggedness = 0.005) with a single major peak at around 10 differences and much smaller peaks at around 30 and 45 mismatches. These results suggested expansion of the goat population, starting about 10 mutational time units ago, with the presence of some divergent haplotypes that may have entered the population at a different time or not taken part in the expansion. The likely times, in years, are discussed in the next section. Results from the individual breeds (fig. 6) were, in general, similar to the complete data set, although stochastic variation was higher and the proportion of identical haplotypes within breeds (6% of mismatches) was also higher, probably reflecting recent drift within breeds and limited exchange between them. For the breeds with sample sizes greater than 20, smooth

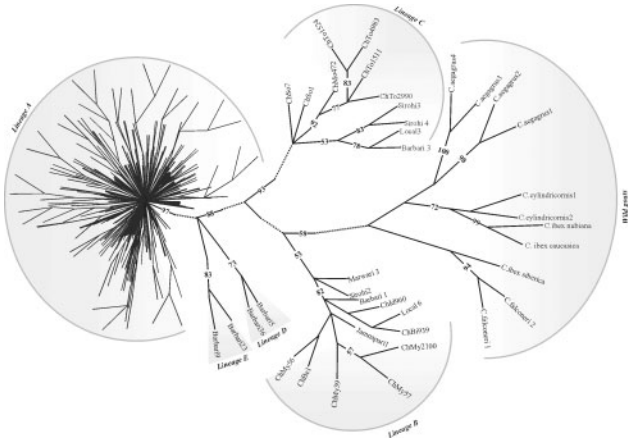

FIG. 4.—Neighbor-joining tree of individual Indian goat mtDNAs and wild goat sequences. Bootstrap values above 50 are shown in the figure. Branches containing lineages B, C, and D and wild goats are stretched for resolution and the expanded distances are represented by dotted lines.

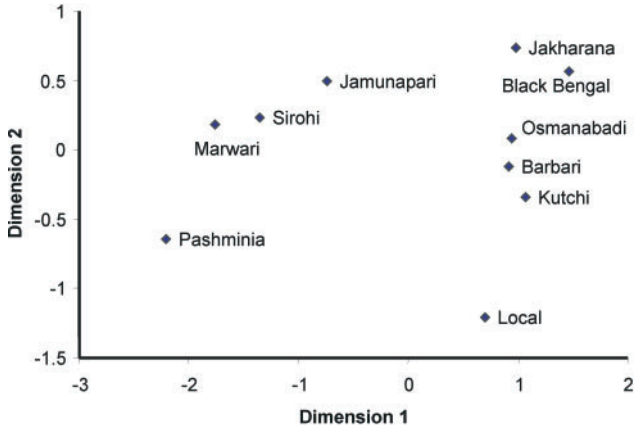

FIG. 5.—Multidimensional scaling plot of pairwise Φ_{ST} values between Indian goat breeds.

and predominantly unimodal curves with raggedness values less than 0.03 were seen, except for the Pashmina goats (raggedness = 0.04 [fig. 6 and table 3]). The Pashmina goat sample size was 40, and, thus, this value was unlikely to be the result of chance; they appear to have had a different demographic history from the other breeds. Similarly, Fu's Fs statistic showed a large negative value for all breeds with sample size greater than 16, except Pashmina (table 3), again indicating population expansion in most breeds and a different history for the Pashmina. The tau values of individual breeds differed somewhat, although there was considerable overlap of their 95% confidence intervals (table 3).

Dating

The time to the most recent common ancestor (TMRCA) of the goat mtDNA lineages can be estimated from a comparison of goat mtDNA diversity with sheep-goat mtDNA divergence, using a calibration of 5 to 7 Myr for the sheep-goat split derived from the fossil record (Savage and Russell 1983; Carroll 1988). Sheep and goat control regions differ at 94 positions within the 457-bp study region, corresponding to approximately 21% divergence. This, however, is a minimum estimate of the amount of change because some positions are likely to have mutated more than once. We have, therefore, adjusted this initial estimate in the ways used by Vigilant et al. (1991) or Tamura and Nei (1993) in their dating of the human mtDNA hypervariable region TMRCA by comparison with chimpanzee mtDNA. Using the method of Vigilant et al. (1991), the goat mtDNA TMRCA is dated to 103,000 to 143,000 years ago, whereas using the method of Tamura and Nei, we obtained a somewhat earlier figure of 201,000 to 280,000 years.

This estimate includes all of the diversity within Indian goat mtDNAs. Several distinct lineages are present (fig. 2), and the times associated with individual lineages, particularly the major one, are also of interest. We therefore estimated the TMRCA for the major lineage using ρ, the average number of mutational steps from the ancestral sequence, assuming that the ancestor was H1. Dates of 42,000 ± 7,000 to 59,000 ± 10,000 years were obtained.

Table 2
AMOVA Analysis of Indian Goat Breeds Based on mtDNA Variation

	Percentage of Variation*		
Source of Variation	No Grouping	Two Groups[a]	Three groups[b]
Among groups	16.6	13.7	13.8
Among breeds, within groups		7.7	7.4
Within breeds	83.4	78.6	78.8

NOTE.—*$P < 0.01$.
[a] Northwest versus remaining breeds.
[b] Northwest versus local versus remaining breeds.

Discussion

The present work presents the first substantial analysis of Indian goat mtDNA diversity and provides information about the genetic structure of goat breeds within this important region, and thus insights into their genetic history.

Indian goat mtDNA sequences exhibited considerable diversity, and the TMRCA calibrated against the fossil record was 103,000 to 143,000 or 201,000 to 280,000 years. This figure, as noted by Luikart et al. (2001) for their worldwide data set, is considerably older than the date of domestication of goats, approximately 10,000 years ago. Their figure of 201,380 to 281,932 years was derived from changes at the third position in codons of the mtDNA cytochrome b gene. They used different DNA samples and thus their estimate was partly independent, although the same fossil record calibration point was used. The two estimates are in good agreement.

The TMRCA for the most frequent lineage was estimated to lie between 35,000 and 69,000 years, using the same mutation rate and taking into account the uncertainty in the timing of the sheep-goat divergence and the standard (sampling) error of ρ. These periods, despite the uncertainty, are considerably older than the time

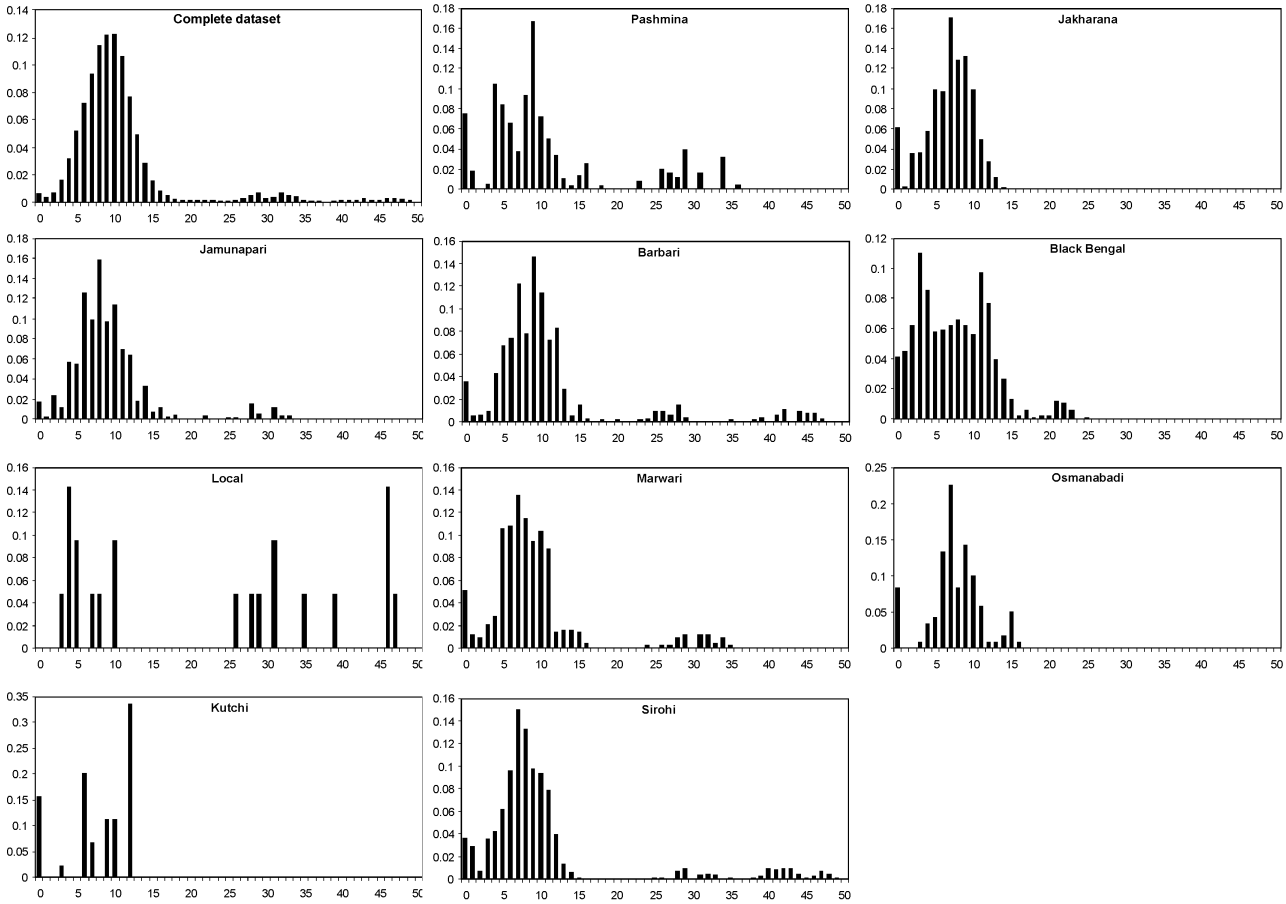

FIG. 6.—Mismatch distributions for mtDNA types of Indian goat breeds.

of domestication. Four explanations are possible: (1) the TMRCA estimate could be too old by a factor of 3.5 or more because of errors in the calculation, (2) domestication could have started earlier than revealed by the archaeological record or (3) involved multiple related lineages and thus incorporated preexisting variation that dates back more than 35,000 years, or, alternatively, (4) mtDNA could have been subject to diversifying selection, so calculations based on the assumption of neutrality are invalid. We note that our TMRCA does not conflict with the expansion time of lineage A of Luikart et al. (2001) because these authors simply assumed, and did not calculate, their expansion time of 10,000 years.

Do errors in the calculation provide a plausible explanation? Four sources of error are possible: diversity within goats, the number of mutational changes between sheep and goats, sheep-goat divergence time, and selection.

Our estimate of 10.0/457 bp mean pairwise differences is in good agreement with the 10.9/481 bp of Luikart et al. (for lineage A) and is thus unlikely to be greatly in error.

The adjustment for multiple hits used by Vigilant et al. (1991) led to a TMRCA of 166,000 to 249,000 years for human mtDNA and 80,000 to 480,000 years by Tamura and Nei (1993) based on a 4 to 6 Myr chimpanzee-human divergence time. A recent estimate for the human mtDNA TMRCA based on diversity within the entire sequence and a 5-Myr divergence time was 171,500 $\pm$ 50,000 years (Ingman et al. 2000), so these methods of adjustment for recurrent mutation in the control region lead to reliable time estimates.

Sheep-goat divergence time is established from the fossil record (Savage and Russell 1983, Carroll 1988) and would have to be approximately 1.4 Myr to lead to a TMRCA that overlaps with the 10,000-year domestication time. It seems unlikely that the fossil dating could be so much in error. Similarly, although domestication may have begun a few hundred or even a few thousand years before the first archaeological records, a time of 35,000 years ago in the Paleolithic period is not credible.

Human selection for increased mtDNA diversity could, in principle, have led to higher than expected diversity during the 10,000 years since domestication, but there is no plausible way in which the likely selection for phenotypic characteristics of interest to farmers could have led to such an increase in mtDNA variation. We, thus, conclude that domestication probably involved multiple maternally related goats carrying considerable preexisting mtDNA diversity.

Table 3
Summary Statistics of Indian Goat Breeds Showing Raggedness, Tau, SSD, and Fu's Fs Values

Breed	Raggedness	Tau Value (95% CI)	SSD	Fs Value	*P* (Fs)
Pashmina	0.040	8.96 (3.41–13.08)	0.0022	0.09	0.562
BlackBengal	0.008	10.75 (4.19–16.51)	0.0073	−8.37	0.010
Osmanabadi	0.057	8.05 (5.22–10.15)	0.0250	−2.42	0.122
Marwari	0.016	8.20 (5.10–10.49)	0.0061	−2.83	0.154
Sirohi	0.010	8.35 (5.02–9.99)	0.0040	−10.62	0.010
Barbari	0.018	9.07 (5.79–11.50)	0.0063	−8.54	0.009
Jakharna	0.018	7.90 (4.67–10.01)	0.0056	−10.68	0.002
Jamunapari	0.022	8.20 (5.22–10.68)	0.0048	−15.53	0.000

All the domestic goat lineages examined fall into a single monophyletic group that is distinct from all available wild goat sequences (fig. 4). The lineages contributing to domestic goats were therefore derived from an unknown population that may now be rare or extinct. Further investigations of wild goats and archaeological specimens are therefore needed to investigate these ancestors.

Conclusions

Significant genetic structure was found among Indian goat breeds. Whereas previous studies have emphasized the weak phylogeographic structure in goats compared with other domestic animals, it should be appreciated that there is still as much mtDNA variation among Indian goat breeds as among mtDNA sequences from human populations living on different continents (Seielstad, Minch, and Sforza 1998). Divisions did not fall between the different geographical regions, but instead revealed a NW versus C+S+E distinction. Goat genetic history is likely to be linked to human history. There is a major linguistic and geographic distinction between the Dravidian-speaking south of India and Indo-Aryan–speaking northern India. Although the first modern humans may have migrated to India about 50,000 years ago, Dravidian speakers probably entered about 10,000 years ago and Indo-Aryan speakers about 3,500 years ago (Cavalli-Sforza, Menozzi, and Piazza 1994; Misra 2001). Although the current Dravidian/Indo-Aryan boundary does not correspond to the goat mtDNA boundary, the two could have a common historical basis in the sequential human migrations into this region, with the goat mtDNA lineages accompanying the Indo-Aryan speakers penetrating less far than the humans.

The Pashmina goats appear to have had a different demographic history from the other breeds. These goats are adapted to living in a cold environment at high altitude in the Himalayas where the human population density is low. Perhaps early on their numbers reached a limit set by the environment, and cannot expand further. Whether the expansion seen in the other breeds has affected only mtDNA lineages or the entire genome could be assessed by analyzing autosomal and Y-chromosomal markers. If all loci show the same pattern, it is likely that the population has increased in size, but if other loci show different patterns, there could have been selection on the mtDNA.

We find that goats have had a more complex history of domestication than indicated by previous studies. We suggest that the diversity within the major A lineage dates back more than 35,000 years, indicating that domestication involved a considerable number of females 10,000 years ago and that diverse lineages in addition to B and C have been incorporated at low frequency.

Acknowledgments

This research work was supported by a grant from Department of Biotechnology, Government of India. We thank Peter Underhill for his valuable suggestions and Laurent Excoffier for providing HVRI sequence of goats belonging to lineages B and C.

Literature Cited

Adams, L. 1999. Mantel nonparametric test calculator. Version 2.00. http://www.sci.qut.edu.au/NRS/mantel.html

Bandelt, H., P. Forster, and A. Rohl. 1999. Median joining networks for inferring intraspecific phylogenies. Mol. Biol. Evol. **16**:37–48.

Bradley, D. G., D. E. MacHugh, P. Cunningham, and R. T. Loftus. 1996. Mitochondrial diversity and the origin of the African and European cattle. Proc. Natl. Acad. Sci. USA **93**: 5131–5135.

Brown, G. G., G. Gadaleta, G. Pepe, and C. Saccone. 1986. Structral conservation and variation in the D loop containing region of vertebrate mitochondrial DNA. J. Mol. Biol. **192**:503–511.

Carroll, R. L. 1988. Vertebrate palaentology and evolution. Freeman, New York.

Cavalli-Sforza, L. L., P. Menozzi, and A. Piazza. 1994. The history and geography of human genes. Princeton University Press, Princeton.

Excoffier, L., and S. Schneider. 1999. Why hunter-gatherer populations do not show signs of Pleistocene demographic expansions. Proc. Natl. Acad. Sci. USA **96**:10597–10602.

Excoffier, L., P. E. Smouse, and J. M. Quattiro. 1992. Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. Genetics **131**:479–491.

Felsenstein, J. 1993. PHYLIP (phylogeny inference package). Versions 3.5c. Distributed by the author, Department of genetics, University of Washington, Seattle.

Fu, X-Y. 1997. Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. Genetics **147**:915–925.

Ingman, M., H. Kaessmann, S. Pääbo, and U. Gyllensten. 2000. Mitochondrial genome variation and the origin of modern humans. Nature **408**:708–713.

Kumar, S., K. Tamura, I. B. Jakobsen, and M. Nei. 2001. MEGA2: molecular evolutionary genetics analysis. Version 2.1. Bioinformatics **17**:1244–1245.

Legge, T. 1996. The beginning of caprine domestication in Southwest Asia. Pp. 236–262 *in* D. R. Harris, ed. The Origins and Spread of Agriculture and Pastoralism in Eurasia. Smithsonian Institute, Washington DC.

Loftus, R. T., D. E. MacHugh, D. G. Bradley, P. M. Sharp, and P. Cunningham. 1994. Evidence for two independent domestications of cattle. Proc. Natl. Acad. Sci. USA **91**:2757–2761.

Luikart, G., L. Giellly, L. Excoffier, J. D. Vigne, J. Bouuvet, and P. Taberlet. 2001. Multiple maternal origins and weak phylogeographic structure in domestic goats. Proc. Natl. Acad. Sci. USA **98**:5927–5932.

MacHugh, D. E., and D. G. Bradley. 2001. Livestock genetic origins: goats buck the trend. Proc. Natl. Acad. Sci. USA **98**:5382–5384.

Manceau, V., P. Boursot, L. Despres, and P. Taberlrt. 1999. Systematics of genus *Capra* as inferred by mitochondrial DNA sequence data. Mol. Phylogenet. Evol. **13**:504–510.

Mannen, H., Y. Nagata, and S. Tsuji. 2001. Mitochondrial DNA reveal that domestic goat (*Capra hircus*) are genetically affected by two subspecies of bezoar (*Capra aegagurus*). J. Biochem. Genet. **39**:145–154.

Meadow, A. 1996. Animal domestication in the Middle East: a revised view from the eastern margin. Pp. 295–320 *in* G. Possehl, ed. Harappan civilisation. Oxford and IBH, New Delhi.

Misra, V. N. 2001. Prehistoric human colonization of India. J. Biosci. **26**:491–531.

Porter, V. 1996. Goats of the world. Farming Press, Ipswich, UK.

Pringle, H. 1998. Neolithic agriculture: reading the signs of ancient animal domestication. Science **282**:1448.

Rogers, A. R., and H. Harpending. 1992. Population growth curves in the distribution of pairwise genetic differences. Mol. Biol. Evol. **9**:552–569.

Savage, D. E., and D. E. Russell. 1983. Mammalian palaeofaunas of the world. Addison-Wesley, Reading, Mass.

Schneider, S., J. M. Kueffer, D. Roessli, and L. Excoffier. 1999. Arlequin 2.001: A software for population genetic data analysis. Genetics and Biometry Laboratory, University of Geneva, Switzerland.

Seielstad, M. T., E. Minch, and L. L. Cavalli-Sforza. 1998. Genetic evidence for a higher female migration rate in humans. Nat. Genet. **20**:278–280.

Swofford, D. L. 1993. PUAP*: phylogenetic analysis using parsimony (*and other methods). Version 4.0. Sinauer Associates, Sunderland, Mass.

Tamura, K., and M. Nei. 1993. Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. Mol. Biol. Evol. **10**: 512–526.

Thangaraj, K., M. B. Joshi, A. G. Reddy, N. J. Gupta, B. Chakraborthy, and L. Singh. 2002. CAG repeat expansion in androgen receptor gene is not associated with male infertility in Indian populations. J. Androl. **23**:815–818.

Thangaraj, K., L. Singh, A. G. Reddy, V. R. Rao, S. C. Sehgal, P. A. Underhill, M. Pierson, I. G. Frame, and E. Hagelberg. 2003. Genetic affinities of the Andaman Islanders, a vanishing human population. Curr. Biol. **13**:86–93.

Thompson, J. D., D. G. Higgins, and T. J. Gibson. 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalities and weight matrix choice. Nucleic Acids Res. **22**:4673–4680.

Vigilant, L. M., M. Stoneking, H. Harpending, K. Hawkes, and A. C. Wilson. 1991. African populations and the evolution of mitochondrial DNA. Science **253**:1503–1507.

Zeder, M. A., and B. Hesse. 2000. The initial domestication of goats (*Capra hircus*) in the Zagros Mountain 10,000 years ago. Science **287**:2254–2257.

David Goldstein, Associate Editor

Accepted October 1, 2003