The Pharma Innovation Journal 2022; SP-11(12): 887-895

www.ThePharmaJournal.com

The Pharma Innovation

ISSN (E): 2277-7695
ISSN (P): 2349-8242
NAAS Rating: 5.23
TPI 2022; SP-11(12): 887-895

www.thepharmajournal.com
Received: 28-09-2022
Accepted: 31-10-2022

Mayank Darji
Department of Animal Genetics & Breeding, SK Nagar, Kamdhenu University, Gujarat, India

Jay Prakash Gupta
Department of Animal Genetics & Breeding, SK Nagar, Kamdhenu University, Gujarat, India

Jagdish Chaudhari
Department of Animal Genetics & Breeding, SK Nagar, Kamdhenu University, Gujarat, India

Arth Chaudhari
Department of Animal Genetics & Breeding, SK Nagar, Kamdhenu University, Gujarat, India

Ashok Chaudhari
Department of Animal Genetics & Breeding, SK Nagar, Kamdhenu University, Gujarat, India

Dhara Gor
Department of Animal Genetics & Breeding, SK Nagar, Kamdhenu University, Gujarat, India

Fotariya Abhi
Department of Veterinary Gynecology & Obstetrics, SK Nagar, Kamdhenu University, Gujarat, India

Dalsaniya Urmil
Department of Veterinary Gynecology & Obstetrics, SK Nagar, Kamdhenu University, Gujarat, India

Corresponding Author:
Mayank Darji
Department of Animal Genetics & Breeding, SK Nagar, Kamdhenu University, Gujarat, India

Mitochondrial DNA: A molecular and evolutionary marker

Mayank Darji, Jay Prakash Gupta, Jagdish Chaudhari, Arth Chaudhari, Ashok Chaudhari, Dhara Gor, Fotariya Abhi and Dalsaniya Urmil

Abstract
Species evolution is a continuous process, influenced greatly by the environment and surroundings and also the various genetic interactions. Starting from the first animal domestication events taking place around Fertile Crescent, today we are in a position to harness the full potential of livestock, and poultry species. In order to get the better understanding of the evolutionary relationships, various markers have been utilized including autosomal microsatellites, Y-chromosome sequence, and mitochondrial DNA sequencing. The strict maternal transmission of mtDNA results in homoplasmic individuals, who typically have a single mtDNA haplotype, the maternal one. Heteroplasmy can primarily occur through somatic mutagenesis during an individual's lifetime and through leakage of paternal mtDNA in the zygote during fertilization. The mitochondrial DNA study revealed that the haplotype diversity was relatively high in the modern livestock population of different species which classified in cattle and buffalo in five major haplogroups, in sheep classified into two major and three minor haplogroups, in goat it classified in six major haplogroup and in horse it classified into total 18 haplogroup. Mitochondrial DNA, by virtue of the number of studies available and various other advantages, is far more practical than nDNA for phylogeographic surveys. Heteroplasmy is proved to be important for the evolution and the survival of mtDNA itself. There is need of further in-depth research in order to understand the uniparental transmission of mtDNA and the functional and evolutionary role of heteroplasmy.

Keywords: Maternal Inheritance, heteroplasmy, evolutionary growth, *mtDNA*

Introduction
Speciation is considered a slow process, as a result of complex forces acting altogether leading to slower, and stable changes in living beings which occurs when populations, usually evolving in geographic isolation for extended periods, accumulate genetic differences. Species evolution is considered a continuous process, influenced greatly by the environment, and surroundings, besides which genetic interactions, like horizontal gene transfer, might also affect evolutionary outcomes. Since evolution, animals have been an important part of human life for various utilities, apart from sources of food, and nutrition. Based on molecular studies, it is inferred that chicken evolved at almost 110 MYA, pig 28 MYA, bovines 24 MYA, whereas goat, and sheep 18-20 MYA. Domestication of various livestock species began with the settlement of human beings as communities after giving up hunter-gatherers' lifestyle. Progressing from the first animal domestication events taking place around Fertile Crescent, today we are in a position to harness the full potential of livestock and poultry species. Among large animal species, only 14 have been domesticated: sheep, goat, cow, pig, horse, Arabian camel, Bactrian camel, llama, and alpaca, donkey, reindeer, water buffalo, yak, Bali cattle, and mithun (gayal, domesticated gaur) (Kataria *et al.*,2021) [15]. Most of the livestock species during the process of domestication have undergone significant genetic, behavioral, and morphological changes from their wild ancestors. It is therefore important to understand the evolution, and phylogenetic relationships of domestic animal species, and their wild ancestors as well as closely associated species.
Understanding the evolution, and genetic diversity of Indian livestock, and classifying the populations by their evolutionary significance is essential for an appropriate conservation plan to be conceived and carried out. To comprehend evolutionary relationships, various markers have been utilized including autosomal microsatellites, Y-chromosome sequence, and mitochondrial DNA sequencing. In this review, we focused only on the mitochondrial DNA marker to know how and why it's important in evolutionary studies?

Mitochondria

Mitochondria are organelles, remnants of ancestral bacterial endosymbiosis, found in nearly all eukaryotic cells. It is membrane-bound cell organelles (mitochondrion, singular) that generate most of the chemical energy needed to power the cell's biochemical reactions. Chemical energy produced by the mitochondria is stored in a small molecule called adenosine triphosphate (ATP), because the phosphate is a high-energy bond and provides energy for other reactions within the cell, so the mitochondria's purpose is to produce that energy.

Structure of mitochondrial DNA

In 1963, DNA was first detected within mitochondria (Nass and Nass, 1963) [27]. In the next 30 years, the complete mitochondrial DNA (mtDNA) sequence [approximately 17,000 base pairs (bp)] was determined in more than a dozen species, including humans (Anderson *et al.,* 1981, Payne *et al*., 2013) [31]. Most vertebrate cells in culture appear to have approximately 1000-5000 molecules of the circular mitochondrial genome (Bogenhagen and Clayton, 1974; Shmookler and Goldstein, 1983) [5, 33]. The mtDNA localizes to the mitochondrial matrix and seems to be associated with proteins and lipids.

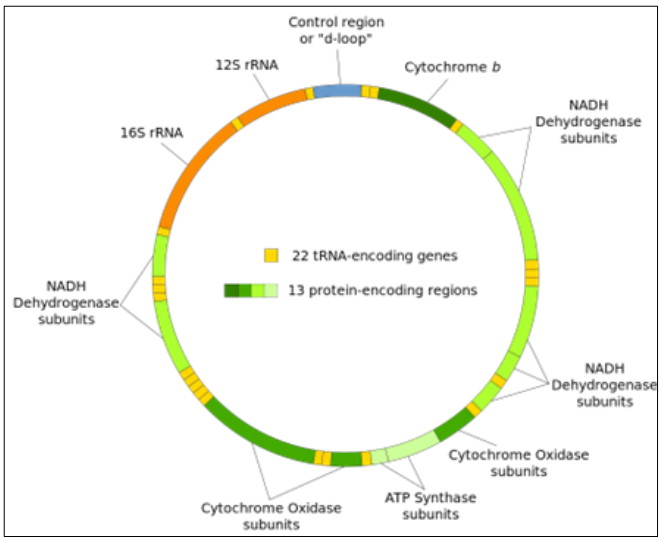

Fig 1: Structure of mtDNA

Electron microscopy analyses showed that mammalian mtDNA can be arranged as uncircular monomers, but also as uncircular dimers or catenated forms (Clayton, 1982) [9]. These early studies also showed what has been termed the displacement loop or "D-loop" as a separation of strands in a specific region of the mtDNA. It is now known that most sequences associated with initiation of mtDNA replication or transcription are in the proximity of the D-loop region (Clayton, 1982) [9]. Both transcription and replication of one strand and transcription of the complementary strand initiate in the proximity of the D loop. This l123-bp stretch of DNA is often in a single stranded configuration, and contains sites for DNA-binding proteins that control mtDNA replication and transcription (Moraes *et al.,* 2002) [24].

In most animals the mtDNA is a short, circular molecule that contains about 13 introns less, protein-coding genes, all of which are involved in oxidative phosphorylation (OXPHOS), a process also known as aerobic respiration. Animal mtDNA also contains two rRNA genes and 22 or 23 tRNA genes, which are part of the translational machinery of mitochondrion. With few exceptions, non-coding regions in the mtDNA molecule are few and short, apart from the region that contains the regulatory elements for replication and transcription. This region is known as a large non-coding region and its length and position within the molecule vary greatly among species. The prevailing view is that animal mtDNA is maternally transmitted, non-recombining and with elevated mutation rate compared to nuclear DNA. As noted, this general description of animal mtDNA cannot be extended to all eukaryotes. If we compare animals, fungi, protists and plants, we will find that their mtDNAs differ drastically in all major characteristics. MtDNA is an extremely variable genome, perhaps more variable than the nuclear genome. The variability is not surprising, given the 2 billion years of mtDNA evolution (Lang *et al.,* 1999; Lane *et al.,* 2010) [18, 17]. Even within animals, the variation is much more than the traditional view of animal mtDNA conservatism would imply.

Maternal transmission and heteroplasmy

The strict maternal transmission of mtDNA results in homoplastic individuals, who typically have a single mtDNA haplotype, the maternal one. However, heteroplasmy (the simultaneous presence of two or more types of mtDNA in the same individual) has often been reported in several animal species (Mastrantonio *et al.,* 2019) [22]. Given that the uniparental transmission of the mtDNA is one of the most general rules in biology and that mtDNA has been extensively used as a genetic marker for phylogenetic studies due to its maternal transmission, the scarce evidence for mtDNA heteroplasmy in the late 1980s and 1990s attracted attention from the scientific community. At that time, heteroplasmy was considered as an interesting exception of the strict maternal mtDNA inheritance. Recently, heteroplasmy has been extensively studied thanks to the modern sequencing techniques. These studies, most of which were conducted in model organisms, revealed that heteroplasmy was more widespread than it was previously believed, particularly as low frequency variants (Payne *et al*., 2013) [31], and that both drift and selection play a role in its dynamics within individuals and among generations.

Heteroplasmy and mtDNA as Genetic Marker

The asexual transmission of mtDNA is one of its fundamental assets as a molecular genetic marker (Avise, 1994) [4]. Heteroplasmy could potentially affect the validity of mtDNA as a marker with three ways. First, in case alternative haplotypes are sequenced from different individuals, then the comparison between individuals would erroneously increase their genetic difference. To make this clearer, let us imagine that two individuals are heteroplasmic for two mtDNA haplotypes A and B. In addition, during the experimental process, the A haplotype by chance is amplified and sequenced in the one individual and the B in the other. The comparison of the haplotypes would erroneously show that the two individuals are genetically different. Second, heteroplasmy can lead to interlineate recombination, which cannot exist under the strict asexual mtDNA transmission. Recombination is expected to mix the evolutionary histories of the different parts of mtDNA molecule and introduce noise in phylogenies which are based on mtDNA. Finally, false heteroplasmy due to nuclear mtDNA (NUMTs), and biased amplification and sequencing can lead to comparisons between nuclear (NUMT) and real mtDNA sequences or

between solely NUMT sequences, which obviously will lead to incorrect results.

The first two ways are not expected to significantly affect the use of mtDNA as a marker because heteroplasmy is rare both as percentage of individuals within a population and as levels of heteroplasmy within an individual. Furthermore, mtDNA recombination, albeit observable (Tsaousis *et al.*, 2005) [34], seems to happen at very low rates. The false heteroplasmy due to NUMTs on the other hand can significantly affect the information of mtDNA as a marker because NUMTs are common and easily amplified.

Evolutionary significance of heteroplasmy

The maternal transmission of mtDNA results in homoplastic individuals, which implies a lack of inter-lineage recombination. Non-recombining genomes are subject to deleterious mutation accumulation, a process known as Muller's ratchet (Muller, 1964) [25]. MtDNA seems to have escaped the ratchet because it has survived for more than two billion years and selection acts efficiently, as suggested by the very low dN/dS ratio, particularly in animals' mtDNA. The survival of the uniparentally transmitted mtDNA for such a long time has been called "the mitochondrial paradox" (Loewe *et al.*, 2006) [20]. Many mechanisms have been proposed to act in order for mtDNA to escape Muller's ratchet. One such mechanism, with experimental support, is the genetic bottleneck that was described above. The genetic bottleneck decreases the variation of the mtDNA within cells producing cells with marginal levels of heteroplasmy, but increases the variation among cells, increasing the efficacy of selection, removing cells which contain deleterious mtDNA variants. Given that the bottleneck has been observed in many organisms beyond mammals, it is expected to play a fundamental role in the mitochondrial paradox. The genetic bottleneck hypothesis though has the soft points that many organisms cannot have a bottleneck (Edwards *et al.*, 2021) [12], that it can be active only against relatively strongly deleterious mutations (slightly deleterious mutations can be fixed in each oocyte by genetic drift), and it can be applicable only to mutations that are expressed in the eggs or in the zygotes. An alternative, non-mutually exclusive hypothesis suggests that mtDNA overcomes Muller's ratchet by allowing a low rate of paternal leakage, which leads to recombination with the maternal lineage. The level of heteroplasmy needed for this process is low because the recombination rate sufficient for cancelling Muller's ratchet is extremely low (Gordo and Charlesworth,2000; Nunes *et al.*, 2013; DeLuca and O'Farrell,2012) [13, 28, 10].

Potential mechanisms behind the bottleneck

Although the simplest explanation is to imagine a physical restriction in the number of molecules sampled from during the maternal transmission of mtDNA, other mechanisms have been proposed. These include the focal replication of mtDNA, the focal destruction of mtDNA (e.g., through autophagy/mitophagy), and the packaging of mtDNA into larger segregating units. All these mechanisms could lead to rapid changes in heteroplasmy level and are not mutually exclusive (Zhang *et al.*, 2018) [36].

Several studies have provided evidence that there is a physical reduction in mtDNA during transmission. Initially studied in mice, similar findings in different fish species, sheep and recently in humans, point towards a dramatic reduction in the amount of mtDNA occurring at the earliest stage in female germ line development (primordial germ cells, PGCs). However, detailed studies of the variation in heteroplasmy level at different stages of germ cell development has shown that the physical reduction may not be sufficient to explain differences in heteroplasmy level between the offspring of a single mother. There may, therefore, be other mechanisms coming into play. This is an area of active research at present. Finally, although the same bottleneck process seems to occur in several different species, the precise mechanisms may differ (Payne, 2011) [30]. For example, in Drosophila, cytoplasmic bridges between developing germ cells could allow mitochondrial (and mtDNA) exchange, and there is evidence that this could also occur in mice. Finally, although mice, humans, fish and cows show similar 'bottlenecks', the precise size and the timing of each bottleneck appears to vary (Otten *et al.*,2016) [29]. All these features make mtDNA a useful, and one of the most frequently used markers in molecular systematics and has been widely employed to address questions of genetic diversity, and evolution (Ladoukakis and Zouros, 2017) [16].

Evolutionary growth of different livestock species by mtDNA

Evolutionary changes in riverine and swamp buffalo

Workers have utilized mitochondrial D-loop sequencing for understanding the evolution of riverine as well as swamp buffaloes among various livestock species. Based on studies outcome it is pinpointed that riverine buffalo after domestication in the western region of the Indian subcontinent almost 6300 years BP, spread west up to Egypt, the Balkans, and Italy. Whereas, after domestication in the China/Indochina border region almost 3000–7000 years BP, swamp buffaloes spread towards south-east Asia, and China as far as the Yangtze River valley. Based on karyotype, and mitochondrial D-loop sequencing studies, it has been concluded that the Assam region being the evolution point of both types of buffaloes, differing in their chromosome numbers. Both riverine (2n=50), and swamp (2n=48) type buffaloes along with their hybrids (2n=49) have been documented in the region. Existence of Asiatic wild buffalo (Bubalus arnee), progenitor of both riverine, and swamp buffaloes in the Assam region of India confirms the events. Recent studies on Chilika buffaloes of Odisha have shed a new light on the domestication of buffaloes, wherein hybrids of riverine, and swamp have been identified sharing haplotypes with both types.

Phylogenetic clustering of north-east Indian swamp buffaloes with both the lineages of Chinese swamp buffalo have been observed, revealing north-east region of India forming part of the wider hybrid zone of water buffalo that may probably extend from north-east India to Southeast Asia. Studies on mitochondrial D-loop sequence data of Odisha buffaloes, and other Indian riverine, swamp, and hybrid buffaloes have been carried out comparing with other reported Indian riverine, Chinese, and Bangladeshi swamp buffalo populations. Haplotype sharing between Chilika - an Odisha buffalo population with Indian swamp as well as Chinese swamp buffalo populations in the network analysis confirmed the presence of hybrids documented by cytogenetic analysis having 49 chromosomes in Chilika population. Phylogenetic analysis of Chilika, and swamp buffaloes of India, China, and Bangladesh has shown haplotypes with predomesticated haplogroup 'SA1', 'SA2', 'SA3', and 'SB1'. These findings thus indicate the migration of swamp buffaloes towards

Bangladesh, and adjoining lower parts of India, and north towards the Chinese domestication sites confirming the previous reports that the northeast region of India, close to the Indo-China border, is the point of evolution of swamp buffaloes with multiple sites of domestication (Kataria *et al.*, 2021) [15].

Evolutionary changes in taurine, zebu cattle and yak

All modern cattle derive from wild ancestral aurochs (Bos primigenius), which were distributed throughout large parts of Eurasia and Northern Africa during the Pleistocene and the early Holocene. Modern Bos taurus and Bos indicus are the result of two independent domestication events; the first occurred 10,000–11,000 years ago in the Upper Euphrates Valley, the second about 2000 years later in the Indus Valley. It is also believed that divergent African taurine cattle arose due to third independent domestication even later happening in Northeast African region. Based on mitochondrial DNA sequence, domestication, and dispersal of three major Bos species – taurine cattle, zebu cattle, and yak and their genetic interactions have been reviewed recently, indicating population bottlenecks leading to phylogeographic differentiation of three species. Breed formation of European taurine cattle is thought to be result of isolation by distance, selective breeding and expansion of limited popular populations. Wider distribution of zebu and taurine cattle and other bovine species led to hybridization with each other showing, introgression playing a major role in environmental adaptation of these species (Kataria *et al.*, 2021) [15].

Domestication of cattle

Surveys of mitochondrial DNA (mtDNA) variation in modern cattle revealed one well-diverged major haplogroup in B. taurus (haplogroup T) and one in B. indicus (haplogroup I) as well as the much rarer haplogroups P, Q and R (Bonfiglio *et al.*,2010) [6]. In contrast, haplogroup C and E mtDNAs have been detected only in ancient samples (Cai *et al.*, 2018) [7].

Mitochondrial DNA (mtDNA) sequence can be a viable tool to the establishment of the sites of domestication. A hotspot of zebu cattle mtDNA diversity (haplogroups I1 and I2) in the Indian peninsula with respect to the encompassing regions denotes domestication of Bos indicus happened in India, followed by migrations toward Southwest Asia, China and Southeast Asia (Chen *et al.*,2010) [8]. Similarly, Bos taurus has the highest diversity in Southwest Asia, with four different mtDNA haplogroups (T, T1, T2 and T3) (Zeder *et al.*, 2006) [35]. Migrations in other continents frequently brought about contact among taurine and zebu cattle, which helped to the appearance of hybrid breeds of southwest Africa, Asia, China and, much later, the Americas (Pramod *et al.*, 2019).

Recent mtDNA studies indicate that, presumably the Indus Valley was the centre of origin for the mtDNA I1 haplogroup and a primary centre of zebu domestication (Chen *et al.*,2010) [8]. Morphological contrasts between cattle illustrated in the rock art of South India and iconography of Indus Valley civilizations, (Allchin *et al.*, 1974) [37] presence of peculiar cattle-oriented Neolithic culture and zoo archaeological data propose that South India was a secondary center for zebu domestication. The frequency and dissemination of the I2 haplogroup within Uttar Pradesh and the Ganges region suggest a feasible secondary recruitment centre of local wild female aurochs into ancestors of domestic zebu cattle within Northern India.

Mitochondrial DNA, as well as nuclear polymorphisms, have revealed several other aspects of the early differentiation of taurine cattle. The predominance of one taurine mtDNA haplogroup (T1) in Africa and a new haplogroup in Eastern Asia (T4*:* suggested two other regions of domestication. However, complete mtDNA sequences showed that T1 and T4 are closely related to the major T3 haplogroup, so their predominance probably reflects founder effects in Africa and Eastern Asia respectively.

The T3 mtDNA haplogroup is predominant in most European breeds and Northern Asia and is one of the four major haplogroups (T, T_1, T_2 and T_3) in Southwestern Asia. By contrast, in the African taurine cattle haplogroup T_1 is dominant, which is rare in Southwestern Asia. These observations are in line with a Southwest-Asian origin of European cattle, confirming the paleontological evidence of a gradual introduction of domestic cattle in Europe from Southwestern Asia (Zeder *et al.*, 2006) [35]. There are two interesting exceptions to the T3 dominance in Europe. First, four ancient breeds from Tuscany have almost the same mtDNA diversity as found in Southwestern Asia, suggesting an ancient maternal origin and a direct link between Tuscan and Western Asian cattle. For the Chianina breed this was confirmed by microsatellite data (European Cattle Genetic Diversity Consortium, unpublished results).

Second, the T_1 haplogroup has appreciable frequencies in several Spanish and Portuguese breeds, indicating migration from Africa to the north. This may have occurred either during the Neolithic movement of cattle or later, for instance during the Islamic occupation. Importation of Iberian cattle into the newly discovered American continent explains the relatively high frequency of the T_1 haplogroup in Caribbean and South American cattle.

Evolutionary changes in sheep

India is the second country with a large population of approximately 74.26 million sheep belonging to 44 described breeds and several non-descript sheep populations. Mitochondrial DNA (mtDNA) studies of sheep breeds across the world have reported two major lineages of A and B, and three minor lineages of C, D and E. The presence of multiple lineages suggests the possibility of their multiple independent domestication events. Lineage A, B and C have been identified in Indian sheep with lineage A to be most abundant. A recent study on the mtDNA control region sequences of Indian sheep proposed that lineage A might have been domesticated in the Indian subcontinent while lineage B might have arrived in the Indian subcontinent through sea route. In contrast, based on the high levels of genetic diversity in sheep breeds of north China and Mongolian Plateau, LV, *et al.* (2015) [21] suggested that lineage A was brought into the Indian subcontinent from the Middle East via Arabia whereas the lineage B and C entered into the Indian subcontinent from Middle East via Mongolian Plateau. However, these studies have not included south Indian breeds.

Phylogenetic structure of south Indian sheep breeds

To understand the phylogenetic relationship of south Indian sheep breeds, we constructed NJ trees using D-loop and CYTB gene sequences along with representative lineage sequences. The NJ trees showed two distinct major clades for south Indian sheep which corresponded to the lineages of A and B. The lineage A was found to be predominant, which consisted of 92.2% of south Indian sheep while the lineage B

consisted of only 7.8% of south Indian sheep.

Median-joining networks were constructed using D-loop and CYTB gene sequences to understand the distribution pattern of south Indian sheep haplotypes. The MJ network of D-loop was found to be complex but clearly distinguished the lineages of A and B. The lineage A was predominant in all the breeds except Mandya. Notably, seven breeds of Bellary, Coimbatore, Hassan, Katchaikatty Black, Nilgiri, Ramnad White and Vembur were fully encompassed with lineage A while Kenguri, Kilakarsal, Madras Red, Mecheri and Tiruchy Black breeds all had very low occurrences (0.28–0.84%) of lineage B. In contrast, majority of the Mandya sheep carried lineage B (79%). Lineage A displayed five majors expanding haplogroups (A1, A2, A3, A4 and A5), indicating extensive expansion of lineage A among south Indian sheep. Also, there was a large number of singletons in the lineage A. On the other hand, lineage B showed no star like expansion among south Indian sheep breeds, indicating the lack of expansion of lineage B.

Origin and migration of lineage A and B

The ABC analysis was capable to reconstruct the most probable colonization scenarios for lineages of A and B of the domestic sheep using D-loop sequences. For lineage A, the scenario with the highest posterior probability (0.9500; 95% CI 0.9375–0.9625) identified the Indian subcontinent as the source population, which first spread to the Middle East region and then expand to the Mongolian Plateau. It indicated that the domestication of sheep occurred in Indian subcontinent much earlier than previously reported. The posterior predictive error was considerably low (0.285). The posterior distributions of demographic parameters were inferred under the best colonization scenario. Most parameter estimates showed high RMAE values (> 0.2) and cannot be considered fully reliable. However, because our main goal was to understand the origin of domestic sheep mtDNA lineages and colonization routes, the high RMAE values did not affect our evaluation of the drawn and tested scenarios. For lineage B, ABC analysis identified the Arabian Sea route as the most probable colonization route to the Indian subcontinent (posterior probability: 0.6801; 95% CI: 0.6442–0.7159). This scenario also showed low posterior predictive error (0.254), which represented high confidence in the chosen scenario.

The NJ tree of D-loop showed that Indian domestic sheep lineage A shared ancestry with wild species O. orientalism Anatolia whereas lineage B shared ancestry with O. Musimon and O. Orientalism which is in accordance with previous studies. In case of CYTB gene, the O. Orientalis subspecies (O. Orientalis gmelini, O. orientalism Anatolia and O. orientalis isphahanica) clustered with lineage A while O. Aries musimon and O. Orientalism gmelini clustered with lineage B as has been reported previously.

Further, the ABC analysis also provides strength to the argument of Indian subcontinent as one of the domestication centers for lineage A by predicting independent origin of Indian sheep from the Middle East sheep. Therefore, based on Kamalakkannan *et al.*, (2021) [14] findings, previous molecular studies and archaeological evidence, lineage A might have been domesticated in the Indian subcontinent. In conclusion, Kamalakkannan *et al.*, (2021) [14] study suggests Indian subcontinent as one of the domestication centres for the lineage A sheep. The study also supports the earlier conclusion that the lineage B might have entered into India through sea route.

The breed effect on mitochondrial genetic variation was analyzed by AMOVA. The breed component contributed 13.41% to the total mitochondrial genetic variation ($p < 0.01$) indicating significant breed differentiation with respect to maternal lineages of Indian sheep. To further understand the genetic differentiation among breeds, F_{ST} values were calculated between all breed pairs based on the nucleotide differences. Maximum differentiation was observed between Deccani and Patanwadi (44.90%), while no significant differentiation was observed among a group of breeds such as Marwari, Garole, Chokla, Nali, Patanwadi, Muzaffarnagar and Nellore. Calculated pairwise F_{ST} values were also analyzed by multidimensional scaling (MDS) plot with stress value of 0.12. MDS plot revealed three major groups: 1) Deccani and Ganjam; 2) Chokla, Marwari, Muzzafarnagri, Nali, Nellore, Patanwadi, Jaisalmeri, Garole and Kheri; 3) Sonadi. Consistent with these results AMOVA analysis after classification of Deccani and Ganjam breeds as one group versus the remaining breeds as a single group explained 21.15% ($p < 0.01$) of total genetic variation. Similarly, when Sonadi breed was treated as an additional group 23.33 % ($p < 0.01$) of the total genetic variation was accounted by this grouping. While majority of the breeds had predominantly lineage A (66 to 93%), Deccani and Ganjam breeds had only lineage A while in Sonadi breed the two lineages were almost equally represented.

Based upon control region mtDNA sequences Hindered and colleagues thought that the lineage B type sheep might have originated from European mouflon (Ovis musimon). On the other hand, Groeneveld and coworkers argued that the latter might represent the feral form of lineage B type sheep rather than being the wild originator of domestic sheep in Europe. The wild ancestors of domestic sheep remain to be identified. Although mitochondrial and other DNA markers studies on sheep breeds from different parts of the world have revealed very poor geographical structuring mitochondrial diversity studies have shown that lineage B is predominantly found in European regions and lineage A is prevalent in Asia. In a limited study, only lineage A has been reported in Indian sheep.

Evolutionary changes in goat

First attempt to delineate mitochondrial DNA diversity in goats was made there were observed three mitochondrial haplogroup (A, B and C) suggesting either multiple maternal origins of domestic goats or introgression of haplotypes after initial domestication. This pioneering study subsequently witnessed a spurt of similar studies exploring genetic diversity of goats in different countries such as Pakistan, Spain, India, China, Portugal, Sicily and South Korea resulting in delineation of three additional haplogroups (D, E and F).

Naderi *et al.*, 2007 [26] presented a wholesome perspective of caprine genetic diversity across the globe and identified 6 mitochondrial haplogroups (A, B, C, D, F and G) in domestic goats. They identified 22 haplotypes that are representative of these 6 haplogroups and can be used by researchers to identify new haplogroups or to classify new goat haplotypes to existing haplogroups. Assessment of genetic relationship between Indian goats and the reference haplogroups revealed that caprine genetic resources of India belong to 4 haplogroups (A, B, C and D). Haplogroup A was the predominant and exhaustively represented haplogroup, which included 319 haplotypes out of 341 haplotypes observed in

this study. Haplogroups B, C and D consisted of 14, 4 and 4 haplotypes, respectively. Overall, the two most frequent haplogroups were A and B, constituting 93.5% and 4.1% haplotypes, respectively.

To get insights into genetic relationships of Indian goats in the context of global caprine genetic diversity, NJ tree and MJ network were constructed by including reference haplogroups that are recognized worldwide. Four haplogroups that represent the goat genetic resources of India were observed to be A, B, C and D. Consistent with reports from different continents and countries that haplogroup A has the widest distribution and is present in > 90% of domesticated goats, Diwedi *et al.*, (2020) [11] also observed it be the predominant haplogroup across the length and breadth of India. Haplogroup B was the second most represented haplogroup in Indian goats. Its presence in many Asian countries has been substantiated by researchers in the past which include India, Pakistan and Iran (Southern Asia), Mongolia and China (Eastern Asia), Cambodia and Myanmar (Southeastern Asia and Oman (Western Asia). In addition, haplogroup B is also prevalent in goats from Sub-Saharan Africa (South Africa and Namibia) and Europe. Diwedi *et al.*, (2020) [11] showed that haplogroup C was observed to be present at low frequency in their study. This observation was also in concert with previous findings in many Asian and European countries including India, Mongolia, Pakistan, Spain, Slovenia and Switzerland, wherein the proportion of animals belonging to haplogroup C was low. Haplogroup D is considered rare since it is exclusively found in Asian countries such as Pakistan, China and India. Although the number of animals belonging to haplogroup D was low, but conspicuous clustering of these samples with haplogroup D was observed in both NJ tree and MJ network. Haplogroup F, which has so far been reported only from Sicily and haplogroup G that is present in goats from Egypt, Ethiopia, Iran, Iraq, Kenya, Oman, Saudi Arabia and Turkey were not detected in Indian goats. Presence of 4 haplogroups, out of a total of 6 haplogroups recognized at the global level supports complex maternal genetic history of Indian goats.

The reported weak geographical structuring of goat mitochondrial variability was often interpreted as a consequence of the frequent transportation of goats along terrestrial and maritime routes of migration and commerce, probably during the early domestication phases. A subsequent comparison with wild stocks confirmed the presence of all "domestic" haplogroups in the current *C. aegagrus* populations, which could result from early translocations of animals and/or feralization before the worldwide spread of goats. The haplogroup A is largely predominant (> 90%) among domestic goats, but rare (6%) in the bezoar and never observed in the Iranian Zagros Mountains. The probable origin of haplogroup A occurred in Eastern Anatolia, where it is still present among wild populations, and its presence in Eastern Iran probably is the result of a subsequent finalization of domestic goats. The most frequent haplogroup in wild populations is C (39%) detected in most of the bezoar distribution area and more common in Southern Zagros/Central Iranian Plateau. The evidence that C control-region haplotypes from Pakistan are the farthest from the domestic-related ones disproved the Indus Valley domestication hypothesis suggested by archaeological remains from Mehrgarh (Baluchistan, Pakistan). Haplogroup F is still found in wild populations (from Northern Caucasus to lower Indus Valley), but it is very rare in domestic goats (< 0.2%) as it was identified only in three Sicilian samples. The other haplogroups were found only in Iranian (D and G) or in both Northern Iranian and Eastern Anatolian bezoars (B). It has been proposed that these haplotypes might have entered the domestic goat gene pool either during the early spread of domestic goats, or due to small-scale domestication events. These findings indicate that the process of goat domestication occurred not only in Eastern Anatolia, as marked by haplogroup A and supported by zoo archaeological data, but possibly also in Central Iran (Zagros Mountains and Iranian Plateau). This additional easternmost domestication event has been marked by haplogroup C, although it led only to a small contribution detectable in the mitochondrial gene pool of current domestic goats (1.5 %) and no archaeological substantiations.

MtDNA haplotypes belonging to haplogroups A and C have been found in ancient goat samples retrieved from an early Neolithic site in Southern France. The two haplogroups occurred with almost the same frequency among the analyzed bones (i.e., 8 samples carrying A haplotypes and 11 samples carrying C haplotypes), suggesting that domestic goat populations were already characterized by the mtDNA variants A and C during the first colonization waves that brought Neolithic farmers into the Mediterranean area about 7.5ka ago.

Evolutionary changes in Horse

DNA analyses for other domesticated species (e.g., cattle, sheep, and goat) revealed that modern livestock has derived from a limited number of animals that were domesticated in just a few places 8–10kya. This finding is also substantiated by today's reduced genetic variation in these animals compared with their ancient forebears. For instance, mtDNA of modern taurine cattle falls into a few distinct haplogroups, suggesting that (almost) only the offspring of original livestock were used to establish herds elsewhere. However, horse mtDNA tells a different story. Modern horse mitochondrial genomes, when analyzed at the highest level of molecular resolution, show a high diversity in terms of haplogroups.

According to the most widely accepted hypothesis, the earliest domestication of the horse happened in the western parts of the Eurasian steppes, between the Northern Black Sea region and present-day Kazakhstan and Turkmenistan. The Bering Land Bridge (BLB) during the Late Pleistocene played a crucial role for the genomic exchange between the New World and the Old World caballine horse populations. The paleontological data showed that caballine horses (*Equus spp.*) living in the Pleistocene were the most closely related to the extant domestic horse, Equus caballus. Sequence analysis of the mitochondrial DNA (mtDNA) control region revealed that caballine horses showed two major clades. One of them (clade A) was distributed simultaneously in Eurasia and North America, while the second clade B has been restricted to North America. Additionally, the clade A in North American horses has been split into two subclades – A1 and A2. Another clade in ancient Eurasia that has been detected is clade C. The latter is specific for ancient Siberian horse populations, but its phylogenetic position of this clade is still under debate (Atsenova *et al.*, 2022) [3].

It has been proposed that at least two intercontinental dispersal events between Eurasian and American horse populations across the Bering Land Bridge (BLB) occurred during the Late Pleistocene. The first of them took place

predominantly from North America to ancient Central and European Eurasia and Asian Far East between ~0.95 and ~0.45 Ma and played a crucial role in horse domestication, since it presented the earliest cabal line horses in Eurasia. The second intercontinental event occurred between ~0.2 and ~0.05 Ma in a bi-directional manner, but predominantly from Eurasia to North America. This allowed genetic admixture between the Eurasian Clade A horses and the members of the North American Clade B.

As for the horse, mitogenomic analysis shows the presence of 18 haplogroups (A–R) in the modern population, presented with different frequency in different breeds (Achilli *et al*., 2012) [1]. The only exception is haplogroup F, which was found only in Equus przewalskii (an endangered wild horse subspecies) and is not observed in modern horses. The large number of haplogroups in breeds of different geographical origin indicates the presence of several centers of domestication in ancient Eurasia. In addition, the great variety of mitochondrial maternal lineages in modern domesticated horse populations may be associated with the presence of additional processes of domestication after the first domestication events. This diverse mitochondrial profile in modern horses is clearly indicative of multiple origins. Otherwise, if there was one center of origin, mitochondrial diversity would be much more limited and uniform.

The shotgun sequencing of 264 ancient horse genomes threw a new light on horse domestication. Yamnaya pastoralists employs early horse management and herding practices, but it is interesting to note that they did not spread horses far outside their native range, similar to the Botai horse domestication, due to their sedentary lifestyle. The development of technological innovation made it possible to create spoke-wheeled chariots around 2000–1800 BC in the Trans-Ural Sintashta culture, which in turn significantly increased mobility and created preconditions for the spread of DOM2 outside the area of domestication.

The results of a sequence analysis of mitochondrial genomes for the 264 ancient horse remains have identified three horses from the western lower Volga-Don region as genetically closest to DOM2. One of them, which is the oldest, dating back to about 3400 BC, comes from a horse's leg bone, buried with a small child in a kurgan built by the Maykop culture (a major Bronze Age archaeological culture in the western Caucasus region, ca 3700 BC–3000 BC) at Aygurskiy, in southern Russia. Another early specimen closely related to the DOM2 horse was found at Repin Khutor, a site on the Don River (ca 3300–2600 BC) a typical representative of the Pit-Grave (Yamnaya) culture. The last horse remains were found in a settlement situated near the village Sosnovka, a part of Poltavka culture's complexes. The Poltavka culture was an early to middle Bronze Age archaeological culture, which flourished on the Volga-Ural steppe and the forest steppe in 2700–2100 BC (Librado *et al*., 2021) [19]. Moreover, genetic similarity has been found between DOM2 and two late Yamnaya specimens from approximately 2900 to 2600 BC (settlement of Turganik), located further east than the western lower Volga-Don region (Southern Cis-Urals steppe)

These results clearly demonstrate that DOM2 ancestors lived in the Western Eurasia steppes, especially the lower Volga-Don during the late fourth and early 3rd millennia BC and that this geographical region is the most probable center of horse domestication.

Indian horse breeds were grouped into seven (A–G) haplogroups. Thirteen sequences from Gen Bank representing the seven haplogroups described for horse mtDNA D-loop were included as references and a neighbor-joining tree was constructed with the chosen reference sequences so as to make sure that the original clade structure was maintained (Fig. 1). Median joining network revealed that the haplogroups A, D and F were predominant (Fig. 1), comprising 17.6, 33.3 and 23.5% of the 35 haplotypes respectively while haplogroups B, C, E and G comprised of only 9.8, 5.9, 3.9 and 5.9% respectively. The results specified that the majority of the sequences in Indian horse breeds belong to haplogroup D. Manipuri breed shared only three haplogroups viz. A, D and G. Majority of the sequences of this breed was in haplogroup D that also comprised of other breeds except Zanskari. Haplogroup E comprised of only the Kathiawari breed. Seven haplotypes namely Hap1, Hap2, Hap12, Hap16, Hap23, Hap28 and Hap33 were independent of the clustering of the major clades.

Conclusion

In the early 80s of past century animal mtDNA was a very accommodating molecule, God's gift to those interested in phylogeny, taxonomy and population genetics: small, conservative in gene arrangement, with slow and fast diverging segments, uniparentally inherited, homoplasmic, non-recombining, with one function (the OXPHOS), easy to study. Heteroplasmy might be important not only for the mitochondrial function but also for the evolution and the survival of mtDNA itself, as a first step for interlineage recombination, and the escape of mtDNA from Muller's ratchet. For the near future, we need research in-breadth and in-depth in order to understand the uniparental transmission of mtDNA and the functional and evolutionary role of heteroplasmy. The genetic distance, phylogenetic tree and divergence time analyses definitely differentiated cattle populations as per their historical origins and represented the genetic distinctiveness of Indian breeds. Indigenous cattle breeds picked up the genetic diversity through selection over centuries and depict the most efficient systems in their respective breed tracts. However, in the recent decades, the population size of some of the indigenous breeds has diminished due to several reasons including disregard of their genetic strengths and their genetic dilution through uncontrolled crossbreeding and interbreeding. This investigation gives an understanding of the history and genetic structure and, in future which may help in prioritization and designing of the conservation plans.

References

1. Achilli A, Olivieri A, Soares P, Lancioni H, Kashani BH, Perego UA, *et al*. Mitochondrial genomes from modern horses reveal the major haplogroups that underwent domestication. Proceedings of the National Academy of Sciences. 2012 Feb 14;109(7):2449-2454.
2. Anderson S, Bankier AT, Barrell BG, De Bruijn MH, Coulson AR, Drouin J, *et al*. Sequence and organization of the human mitochondrial genome. Nature. 1981 Apr;290(5806):457-65.
3. Atsenova N, Palova N, Mehandjyiski I, Neov B, Radoslavov G, Hristov P. The Sequence Analysis of Mitochondrial DNA Revealed Some Major Centers of Horse Domestications: The Archaeologist's Cut. Journal of Equine Veterinary Science. 2022 Feb 1;109:103830.

4. Avise JC. Conservation genetics. In Molecular Markers, Natural History and Evolution. Springer, Boston, MA; c1994; p. 361-398.
5. Bogenhagen D, Clayton DA. The number of mitochondrial deoxyribonucleic acid genomes in mouse L and human HeLa cells: Quantitative isolation of mitochondrial deoxyribonucleic acid. Journal of Biological Chemistry. 1974;249(24):7991-5.
6. Bonfiglio S, Achilli A, Olivieri A, Negrini R, Colli L, Liotta L *et al*. The enigmatic origin of bovine mtDNA haplogroup R: sporadic interbreeding or an independent event of Bos prim genius domestication in Italy? PLoS One. 2010 Dec 28;5(12):e15760.
7. Cai D, Zhang N, Zhu S, Chen Q, Wang L, Zhao X *et al*. Ancient DNA reveals evidence of abundant aurochs (Bos primigenius) in Neolithic Northeast China. Journal of Archaeological Science. 2018 Oct 1;98:72-80.
8. Chen S, Lin BZ, Baig M, Mitra B, Lopes RJ, Santos AM *et al*. Zebu cattle are an exclusive legacy of the South Asia Neolithic. Molecular biology and evolution. 2010;27(1):1-6.
9. Clayton DA. Replication of animal mitochondrial DNA. Cell. 1982;28(4):693-705.
10. DeLuca SZ, O'farrell PH. Barriers to male transmission of mitochondrial DNA in sperm development. Developmental cell. 2012 Mar 13;22(3):660-8.
11. Diwedi J, Singh AW, Ahlawat S, Sharma R, Arora R, Sharma H *et al*. Comprehensive analysis of mitochondrial DNA based genetic diversity in Indian goats. Gene. 2020 Sep 25;756:144910.
12. Edwards DM, Røyrvik EC, Chustecki JM, Giannakis K, Glastad RC, Radzvilavicius AL *et al*. Avoiding organelle mutational meltdown across eukaryotes with or without a germ line bottleneck. PLoS biology. 2021 Apr 23;19(4):e3001153.
13. Gordo I, Charlesworth B. The degeneration of asexual haploid populations and the speed of Muller's ratchet. Genetics. 2000 Mar 1;154(3):1379-87.
14. Kamalakkannan R, Kumar S, Bhavana K, Prabhu VR, Machado CB, Singha HS *et al*. Evidence for independent domestication of sheep mtDNA lineage A in India and introduction of lineage B through Arabian sea route. Scientific reports. 2020;11(1):1-6.
15. Kataria RS, Lavakumar S, Gurao Ankita, Sodhi Monika. Mitochondrial DNA sequencing as a tool for exploring evolution, and domestication of livestock species. DBT Sponsored e-Training at GADVASU; *c*2021; p. 151-158.
16. Ladoukakis ED, Zouros E. Evolution and inheritance of animal mitochondrial DNA: rules and exceptions. Journal of Biological Research-Thessaloniki. 2017 Dec;24(1):1-7.
17. Lane N, Martin W. The energetics of genome complexity. Nature. 2010;467(7318):929-34.
18. Lang BF, Gray MW, Burger G. Mitochondrial genome evolution and the origin of eukaryotes. Annual review of genetics. 1999 Dec;33(1):351-97.
19. Librado P, Khan N, Fages A, Kusliy MA, Suchan T, Tonasso-Calvière L, *et al*. The origins and spread of domestic horses from the Western Eurasian steppes. Nature. 2021 Oct;598(7882):634-40.
20. Loewe L. Quantifying the genomic decay paradox due to Muller's ratchet in human mitochondrial DNA. Genetics Research. 2006 Apr;87(2):133-159.
21. Lv FH, Peng WF, Yang J, Zhao YX, Li WR, Liu MJ, *et al*. Mitogenomic meta-analysis identifies two phases of migration in the history of eastern Eurasian sheep. Molecular biology and evolution. 2015;32(10):2515-2533.
22. Mastrantonio V, Latrofa MS, Porretta D, Lia RP, Parisi A, Iatta R, *et al*. Paternal leakage and mtDNA heteroplasmy in Rhipicephalus spp. ticks. Scientific reports. 2019;9(1):1-8.
23. Millar M, Rangayya V, Jafar BB, Chambers D, Vitzu M, Wyborny LE. Membrane-bound mitochondrial DNA: isolation, transcription and protein composition. Archives Internationales de Physiologie ET DE Biochimie. 1979 Jan 1;87(1):29-49.
24. Moraes CT, Srivastava S, Kirkinezos I, Oca-Cossio J, Woischnick M, Diaz F, *et al*. Mitochondrial DNA structure and function. International review of neurobiology. 2002;53:3-23.
25. Muller HJ. The relation of recombination to mutational advance. Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis. 1964 May 1;1(1):2-9.
26. Naderi S, Rezaei HR, Taberlet P, Zundel S, Rafat SA, Naghash HR *et al*. Large-scale mitochondrial DNA analysis of the domestic goat reveals six haplogroups with high diversity. Plos one. 2007;2(10):e1012.
27. Nass MM, Nass S. Intramitochondrial fibers with DNA characteristics: I. Fixation and electron staining reactions. The Journal of cell biology. 1963 Dec 1;19(3):593-611.
28. Nunes MD, Dolezal M, Schlötterer C. Extensive paternal mt DNA leakage in natural populations of Drosophila melanogaster. Molecular ecology. 2013;22(8):2106-2117.
29. Otten AB, Theunissen TE, Derhaag JG, Lambrichs EH, Boesten IB, Winandy M *et al*. Differences in strength and timing of the mtDNA bottleneck between zebrafish germline and non-germline cells. Cell reports. 2016 Jul 19;16(3):622-630.
30. Payne BA, Wilson IJ, Hateley CA, Horvath R, Santibanez-Koref M, Samuels DC, *et al*. Mitochondrial aging is accelerated by anti-retroviral therapy through the clonal expansion of mtDNA mutations. Nature genetics. 2011;43(8):806-810.
31. Payne BA, Wilson IJ, Yu-Wai-Man P, Coxhead J, Deehan D, Horvath R, *et al*. Universal heteroplasmy of human mitochondrial DNA. Human molecular genetics. 2013 Jan 15;22(2):384-390.
32. Pramod RK, Velayutham D, PK S, PS B, Zachariah A, Zachariah A, *et al*. Complete mitogenome reveals genetic divergence and phylogenetic relationships among Indian cattle (Bos indicus) breeds. Animal biotechnology. 2019;30(3):219-232.
33. Reis RS, Goldstein S. Mitochondrial DNA in mortal and immortal human cells. Genome number, integrity, and methylation. Journal of Biological Chemistry. 1983;258(15):9078-9085.
34. Tsaousis AD, Martin DP, Ladoukakis ED, Posada D, Zouros E. Widespread recombination in published animal mtDNA sequences. Molecular Biology and Evolution. 2005; 22(4): 925-933.
35. Zeder MA, Emshwiller E, Smith BD, Bradley DG. Documenting domestication: The intersection of genetics and archaeology. TRENDS in Genetics. 2006 Mar 1;22(3):139-155.
36. Zhang H, Burr SP, Chinnery PF. The mitochondrial DNA genetic bottleneck: inheritance and beyond. Essays in

biochemistry. 2018 Jun 7;62(3):225-234.
37. Allchin B, Goudie A. Pushkar: Prehistory and climatic change in Western India. World Archaeology. 1974 Feb 1;5(3):358-68.