Scientific Research Publishing

Open Journal of Animal Sciences, 2023, 13, 179-198
https://www.scirp.org/journal/ojas
ISSN Online: 2161-7627
ISSN Print: 2161-7597

Phylogenetic Relationship and Molecular Divergence Dating Using SRY Gene Polymorphism about Four Ladoum Sheep Lineages in Senegal

Pape Madiop Sembène*, Fatimata Mbaye, Mama Racky Ndiaye, Mbacké Sembène

Genomic Laboratory, Department of Animal Biology, Cheikh Anta Diop University, Dakar, Senegal
Email: *madiop.sembene@gmail.com

How to cite this paper: Sembène, P.M., Mbaye, F., Ndiaye, M.R. and Sembène, M. (2023) Phylogenetic Relationship and Molecular Divergence Dating Using SRY Gene Polymorphism about Four Ladoum Sheep Lineages in Senegal. *Open Journal of Animal Sciences*, **13**, 179-198.
https://doi.org/10.4236/ojas.2023.132013

Received: February 4, 2023
Accepted: April 3, 2023
Published: April 6, 2023

Abstract

Animal genetic resources are playing a vital role in livestock production and are essential to food security. The present study aims to contribute to a better understanding genetic local sheep breeds and to elucidate the phylogenetic relationships through the evolution of the *SRY* gene in four different lineages of Ladoum sheep raised in Senegal. After a brief analysis of genetic diversity, the phylogenetic relationships and molecular dating were inferred through haplotype networks and four phylogenetic reconstruction methods. The different haplotype networks are constructed with NETWORK ver. 5.0.0.0 using the Median-Joining method. Phylogenetic trees were reconstructed using neighbor-joining, maximum parsimony, maximum likelihood and Bayesian inference. The robustness of the nodes in phylogenetic trees of the three first methods was assessed by 1000 bootstraps. For Bayesian inference, the posterior probability distribution of the trees was estimated by 4 MCMC chains. 5,000,000 generations were performed for each of the chains by sampling the different parameters every 1000 generations. Results show a low polymorphism. Haplotypic diversity is much higher than the average nucleotide divergence between all pairs of haplotypes. The majority and central haplotype indicates a close relationship between "Batling" and "Tyson" individuals. "Birahim" lineage is very distinct from the rest. Phylogenetic trees confirm two genetically separate clades between "Birahim" and the other lineages. The period of divergence between "Birahim" lineage *versus* the common ancestor of the other three lineages was 2504 years ago. The polyphyly revealed in "Birahim" lindicates that this lineage does not contain the common ancestor of all individuals who compose it. It could therefore be derived from two or more sheep breeds with a common ancestor, *Ovis aries*. The monophyletic

clade appears to be a group including a common ancestor and all of its genetic descendants. This group, bringing together the other three lineages, is in the process of being structured into sub-lineages. This study is the first to show that there are only two genetic lines within ladoum sheep in Senegal.

Keywords

Genetic Diversity, Phylogenetic Relationship, Molecular Dating, *SRY* Gene, Sheeps, Ladoum, Senegal

1. Introduction

The efforts of conservation of animal genetic resources in developing countries are minimal. The conservation programs that had been launched were running without a full spectrum of conservation strategies. With the advent and development of artificial insemination during the last 50 years, the replacement of locally adapted breeds by a narrow range of high-yielding international transboundary breeds, is becoming an area of concern affecting animal genetic diversity in developing countries [1]. This development has led to mounting concerns about the genetic erosion. Consequently, livestock breeds are threatened by the devastating effects of genetic erosion, the most significant of which are the dilution of genetic pool and the disappearance of important adaptive traits [1]. Local breeds represent an original and unique heritage due to the processes that have shaped them in relation to the history of the region. They are also deeply embedded in the local production systems, through their breeding methods and uses. They have thus developed particularly useful zootechnical aptitudes in terms of production performance and quality of adaptation [1]. However, zoogenetic resources value goes far beyond their current use, as they provide options for the future. In 2010, the Senegalese herd included 10 million small ruminants [2], and a wide variety of sheep breeds reared in different agro-ecological zones according to different management methods. With the advent of urban breeding and the craze for the Ladoum sheep, several associations have been created to promote and develop this breed. The process of creating the Ladoum breed, which in its early days was driven by passion and prestige, is now of great importance on the national economic level, but also on the scientific level [3]. In the breeders' jargon, there are several lineages within this breed, each headed by a male whose reproductive performance has been confirmed by his phenotypic criteria, which are appreciated and measured during the annual competitions organised by the breeders' associations, but also by the performance of his offspring. This work, which today has resulted in this beautiful and great breed of sheep that is the Ladoum, was done empirically and did not follow the conventional process of creating a breed [4]. There remains today a scientific void as to its origin and the different breeds involved in this process. Indeed, some authors [3] [4] [5] describe it as a subpopulation of Touabire obtained after a long

selection on the latter. However, the results obtained on live weight and morphobiometric parameters of the Ladoum breeds described in Senegal by Sall [6], seem to indicate a racial difference with the Touabire described by Kane [7] in Mauritania. The Ladoum is thought to belong to the group of short-haired Moorish sheep whose origin is much debated [8] and is thought to have been introduced to Senegal from Kayes (Mali) from the neighboring Mauritanian region of Hodh el Gharbi, inhabited by the Ladem tribe [6] [9]. The use of this breed, which is unstable and recent from a geological point of view, in the improvement of the local herd, requires a better knowledge of its genetic structure and evolution. However, to understand the structure of biological communities observed today, it is necessary to trace their history in order to separate the historical components (in the sense of an evolutionary time scale) from the recent ecological components. Indeed, because of their similarity, individuals of the same clade do not constitute independent samples in the statistical sense of the term. The solution is to reconstruct the phylogeny of the studied population to appreciate the relationships of the studied species [10]. Molecular analysis is based on the neutralist theory of evolution, and postulates that the evolutionary time separating different taxa is reflected by the divergence of DNA sequences. The improvement of molecular techniques has greatly contributed to the multiplication of molecular phylogenies, thus simplifying the access to information contained in the genome. Whatever method of phylogenetic reconstruction used, its principle is based on the assumption that the autonomous functioning of a population, *i.e.* its reproductive isolation, may result in genetic differentiation, thus providing information on the structure of this population as well as the factors that determine it. Although the importance of population structure in genetics has been recognised since the beginning of the century, the theoretical definition of the population and its subdivisions is not obvious. The subgroups of the species within which individuals have the opportunity to exchange genes are defined as populations. In each of these populations, individuals share a common gene pool that is perfectly characterized by the gene frequencies on which the mutations that underlie all genetic differentiation act. It should be noted, however, that in many developing countries, phenotypic characteristics are emphasized in the characterization of breeds. In Senegal, for example, the Ladoum, Peul-peul, and Djallonké races, among others, are only recognized on the basis of phenotypic characteristics [6]. This is why molecular analysis, based on the neutralist theory of evolution and which postulates that the evolutionary time separating the different taxa is reflected by the divergence of DNA sequences, is an essential tool for elucidating the evolution of intermediate structures. The present study aims to contribute to a better understanding of local sheep breeds at the genetic level and to elucidate the phylogenetic relationships through the evolution of the *SRY* gene in the different lineages of Ladoum sheep raised in Senegal. The *SRY* gene is located on the Y chromosome. Its mutation rate on the *SRY* gene appears to be higher than that of other human genes [11]. This gene is a valuable tool in the search for paternal relationships, but also for

evolutionary divergences between species and within the same species morphologically or ecologically structured in populations. Is the gene under purifying selection to maintain its function in a pattern correlated with speciation and divergence events in sheep? What are the phylogenetic relationships and their evolution between the different lineages? This study, beyond characterising the genetic diversity and evolution of the four predefined lineages [12], will attempt to answer these questions more specifically.

2. Material and Methods

2.1. Study Site and Choice of Animals

The study was conducted in the Dakar region, where all the Ladoum "lineages" converge during the fairs. Located on the Cape Verde peninsula, Dakar region covers an area of 550 km^2, or 0.28% of the national territory. It lies between 17˚10 and 17˚32 West longitude and 14˚53 and 14˚35 North latitude. Sheep farming in this area is intensive and mainly involves Ladoum sheep [2] [6] [8]. Animals are raised in enclosures in front of houses or indoors on stall terraces. The choice of flocks was guided by the collaboration of the owners. The males sampled are between 6 months and a year old and are in good health. A total of 47 Ladoum sheep were selected, of which 22 were from the "Batling" lineage, 12 from the "Birahim" lineage, 5 from the "Gorgui" lineage and 8 from the "Tyson" lineage.

2.2. Collection and Storage

Samples were collected from January to April 2022. The biological material for this study consisted of whole blood obtained from the animal's jugular vein and collected in EDTA (ethylene diamine tetra acetic acid) tubes of 4 ml capacity, which allow a better long-term conservation of the nucleic acids in the blood. After collection, the blood samples were directly sent to the genomics laboratory of the Department of Animal Biology of the Faculty of Science and Technology of Cheikh Anta Diop University in Dakar. The blood samples obtained were then stored at −18˚C until DNA extraction. See the detail of the sampling *in* Sembène *et al.* [12].

2.3. DNA Extraction, Amplification and Sequencing *SRY* Gene

Total DNA from the blood was extracted using the standard protocol of the Zymo research kit. 200 µl of blood from each sample were placed in 1.5 ml eppendorf tubes, mixed with 200 µl BioFluid Cell Buffer (digestion buffer) and 30 µl of proteinase K. The whole was incubated in a Bain Marie at 70˚C for one hour. After incubation, the mixture was transferred to a Zymo-Spin column, previously placed in a collection tube. Indeed, the negatively charged DNA is fixed, by ionic interactions, on the positively charged silica membrane. On the other hand, proteins, lipids and polysaccharides are eliminated. The DNA fixed on the Zymo-Spin column is then purified to eliminate all traces of contaminants. This

washing was carried out by successive additions of 400 µl of DNA Pre-Wash Buffer, 700 µl and 200 µl of g-DNA Wash Buffer. The debris passes through the membrane by centrifugation at 12,000 rpm for one minute. After washing, the column is then placed on a new 1.5 ml eppendorf tube and 50 µl of DNA Elution Buffer previously preheated to 70˚C are placed on the silica membrane to collect the DNA. The DNA extracts obtained were stored at −20˚C. After extraction, Fragments of *SRY* gene was amplified in a reaction volume of 25 µl with two primers: F 5'CAACTTTCAAGTTTGCCTTATGG-3' and R 5'ACAGCCCAATCCTG TTATATA-3'. The PCR was performed in an Eppendorf thermal cycler with the initial denaturation performed at 94˚C for five minutes, followed by a repeat of 35 cycles with denaturation at 94˚C for 30 seconds, hybridization of the primers at 48˚C for 40 seconds, elongation of the complementary DNA strands at 72˚C for 40 seconds, and terminated by a final elongation at 72˚C for 10 minutes. Each target gene was amplified by performing an electrophoretic migration on a 2% agarose gel. Sequencing has been outsourced for optimum results. The Sanger method was used to identify the nucleotide sequence of *SRY* gene.

2.4. Sequence Cleaning and Alignment

The resulting sequences were corrected, manually cleaned and aligned with BioEdit 7.0.8.0 software [13] using the Clustal-W algorithm [14]. The nucleotide sequences were checked and corrected thoroughly with reference to the electrophoregram. They were subjected to a BLAST (Basic Local Alignment Search Tool) to check the similarity with the reference sequence in GenBank (https://blast.ncbi.nlm.nih.gov).

2.5. Analysis of Genetic Variability and Diversity

The nucleotide composition of the sequences was calculated with the editor BioEdit. The standard indices of genetic variations (number of polymorphic sites, number of informative sites, number of total haplotypes and by lines) are explained with the MEGA 7 software [15]. The ratio between transitions and transversions and the frequency of nucleotides were also calculated using this same software by the Pattern substitution test. For sequence polymorphism analysis, haplotypic diversity (Hd) and nucleotide diversity (π) [16] [17] were calculated using DNAsp version 5.10.01 [18]. All these analyses were performed in the study population and for each lineage.

2.6. Phylogenetic and Statistical Analyses

2.6.1. Haplotype Networks

The haplotype network was constructed. Haplotype networks represent an application of the median linkage method to show phylogenetic relationships between different haplotypes. A minimum haplotype network is characterised by nodes (circles) and branches (links) that connect the nodes. Each node corresponds to a haplotype whose size is proportional to the frequency of the haplotype in the dataset. They are either sequences of the dataset (haplotypes), or me-

dian vectors. The links are the differences in characters. A median vector is a hypothetical (often hereditary) sequence that is needed to connect existing sequences in the network with maximum parsimony. Links and median vectors are respectively the number of mutational steps and the gene flow barrier that separate the different nodes. The different haplotype networks are constructed with NETWORK ver. 5.0.0.0 using the Median-Joining method [19] to show the phylogenetic relationships between the different haplotypes. The construction of these networks is based on coalescent theory [20] which is a retrospective approach that mathematically describes the process of binary merging of all genealogical lineages of a sample of genes back to their closest common ancestor.

2.6.2. Phylogenetic Reconstructions

Phylogenetic trees were reconstructed using four methods: neighbor-joining, maximum parsimony, maximum likelihood and Bayesian inference. Neighbor-joining [21] is based on a genetic distance matrix calculation that aims to minimise the total tree length and uses the Kimura-2-P model (=2 parameters) which postulates that transitions and transversions are not equiprobable during the evolution of sequences. The maximum parsimony method [22], which consists of searching among all possible trees and all possible sequences of ancestral nodes for the combination that requires the least number of evolutionary changes in the phylogenetic tree. The maximum likelihood method [23] [24], tests all histories that could have generated the current dataset being analysed. It is based on individual traits and uses an explicit evolutionary model. The likelihood of each tree is thus calculated following the chosen model, and the phylogenetic reconstruction chosen is the one for which the likelihood is maximum. The trees developed by the above methods were obtained by MEGA 7 software [15]. The robustness of the nodes was assessed by bootstrapping which is a statistical technique used to assess the robustness of phylogenetic reconstruction [25]. A node with a support of 70% or more is considered significant [26]. Mahé [27] considers that a threshold of 1000 replicates is a good compromise between accuracy and computation time. Bayesian inference is estimated using the MrBAYES ver. 3.2.5 program [28]. The latter method and maximum likelihood are described as probabilistic and achieve the best results on test sets [27]. For Bayesian inference, the posterior probability distribution of the trees was estimated by 4 MCMC chains (three of which were gradually "warmed up" and one cold). 5,000,000 generations were performed for each of the chains by sampling the different parameters every 1000 generations. The degree of convergence of the chains can be checked by examining the evolution of the likelihood function during the run of the "cold" chain to determine the ignition period. Generations that occur during the ignition period are eliminated from the analyses. Thus, conservatively, the first 1, 250,000 generations have been eliminated (25%) and inferences are then made on the following 3,750,000 generations. The reliability of a tree is thus reduced to the reliability of its internal branches (nodes). For the phylogenetic

reconstructions, each of the trees was rooted by a referenced gene sequence (AF026566.1) obtained from the GENBANK database.

2.7. Degree of Similarity with Referenced Species (BLAST)

In order to confirm that the identified groups belong to *Ovis aries* (Linne, 1758), a search of the degree of similarity of the sequences of the studied individuals with referenced species was made. The search strategy consists in finding all the similar segments (HSPs: High-Scoring Segment Pair) between the searched sequence and the sequences of the database. To determine an HSP, words of fixed length are identified in a first step between the searched sequence and the sequence in the data bank. In the case of nucleic acids, this amounts to identity searches between the two sequences on fixed-length segments (generally 11). In a second step, the similarity is extended in both directions along each sequence, starting from the common word, so that the cumulative score can be improved. In a third step, the significance of the similar segments obtained is statistically evaluated. The similarity score is normalised and evaluated in standard units of information (hit). Then, the probability (E-value) of having such a score at random is calculated for this segment length (m) in a data bank containing a total of (n) nucleotides or amino acids. Only significant HSPs, *i.e.* those with the lowest probability, are retained and ranked. Following this procedure, the sequences obtained from the samples of the four lineages were compared by alignment with other sequences in the nucleotide database of NCBI (Nucleotide Center for Biotechnology Informatics) using the BLASTn tool.

2.8. Molecular Dating

Good molecular dating depends on the trilogy: careful choice of nucleotide substitution model, suitable molecular clock model and good calibration points. In order to date the emergence and diversification of lineages through phylogenetic analysis, molecular dating analyses were performed using a Bayesian approach with the BEAST ver. 1.8.2 package [29]. These analyses were performed under the GTR evolution model [30] with a gamma distribution (Γ) and a proportion of invariant sites (+I). The MCMC analysis was run with 10,000,000 generations sampling every 1000 generations; then the first 1,000,000 generations were eliminated (10%) and inferences are then made on the following generations. The molecular clock hypothesis assumes that substitutions in a sequence occur at a constant rate. The number of observed substitutions would therefore allow us to know the date of divergence between two species, but molecular clocks are more or less rapid between organisms, or even within an organism, under the influence of generation time, repair systems and selection pressure. The molecular clock hypothesis with MEGA using the Maximum Likelihood method is then tested under 2 types of estimates: 1) the constant molecular clock method which postulates that the rate of evolution of lineages is constant. Therefore, we dated the lineages by considering an evolutionary rate equivalent to $2.09 \times 10^{-3} \pm 1.08$

$\times 10^{-5}$ substitution/site/million years in the genus *Rupicapra* for the *SRY* gene [31] [32]. 2) the relaxed molecular clock method of a model of variation within a dataset that assumes that the substitution rates associated with each branch are drawn independently from a single uncorrelated lognormal distribution [33] to estimate the rate of molecular evolution and divergence times between different detected clades. The relaxed molecular clock method has three advantages: 1) No need to test the hypothesis of a constant evolutionary rate throughout eukaryotic history; 2) The incorporation of a priori constraints on divergence time is preferred over the use of calibration points in order to handle uncertainties inherent in the palaeontological data (fossils never match exactly at the nodes of the phylogenetic trees); 3) several independent calibration points can be used simultaneously [34].

3. Results

3.1. Genetic Diversity

The data set obtained after alignment and cleaning of the sequences revealed 558 sites which 96.77% (540/558) of them are monomorphic and 03.22% (18/558) are variable, of which 01.61% (09/558) of singleton sites and 01.61% (09/558) of informative sites in parsimony (**Table 1**) are obtained after alignment and cleaning of the sequences. Transversions (51.18%) appear to be more numerous than transitions (48.83%). The 42 sequences are divided into 13 haplotypes. The mutation rate is 0.83; the average number of nucleotide differences is 2.907 (**Table 1**).

The *SRY* gene shows a high haplotypic diversity and a low nucleotide diversity except for the "Tyson" population where haplotypic and nucleotide diversities are zero. In the "Gorgui" lineage, haplotypic diversity is maximal and nucleotide diversity is very low (**Table 1**).

3.2. Phylogenetic Relationships

3.2.1. Haplotype Network

The haplotype network (**Figure 1**) showed a differentiated relationship between the different lineages. The results reveal a haplogroup divided into three major haplotypes represented by the "Batling", "Birahim" and "Tyson" lineages. The haplotypes of the "Gorgui" population appear as transitional (secondary) haplotypes of the "Batling" lineages. The "Birahim" lineage which is distinguished by its haplotypic homogeneity derives almost all from the Amo haplotype by a single mutational step. The "Tyson" lineage, although belonging to the original haplogroup, is highly conserved and does not evolve (**Figure 1**).

3.2.2. Phylogenetic Trees

Figures 2(a)-(d) represent the topologies obtained with the Neighbor-Joining, Maximum Likelihood, Parsimony Maximun and MrBayes approaches respectively.

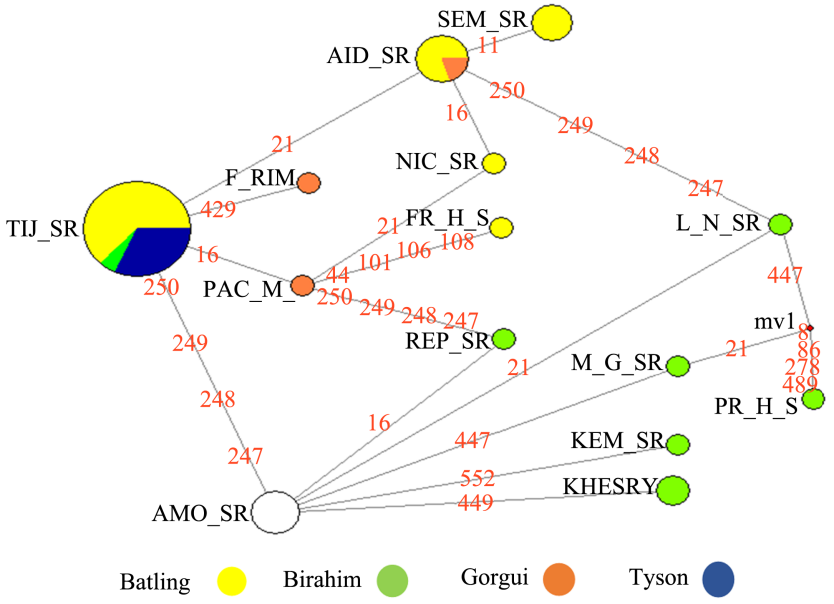
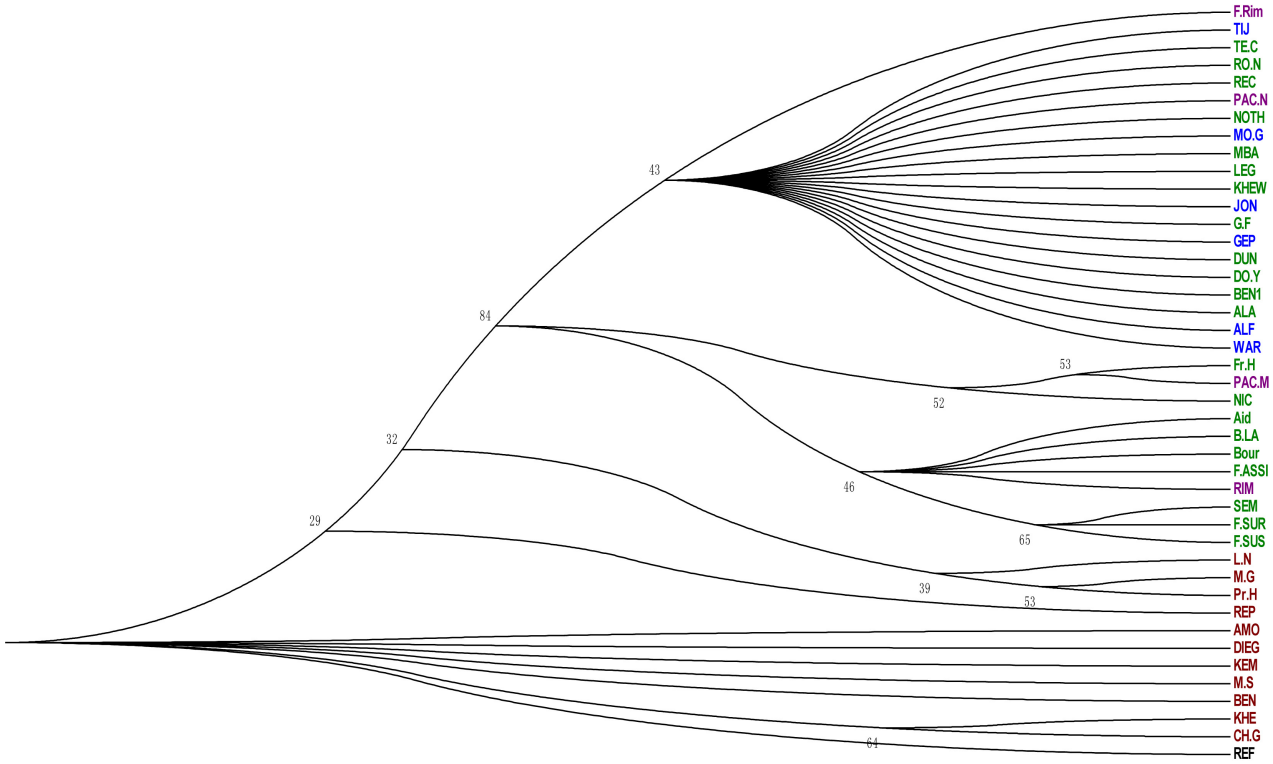


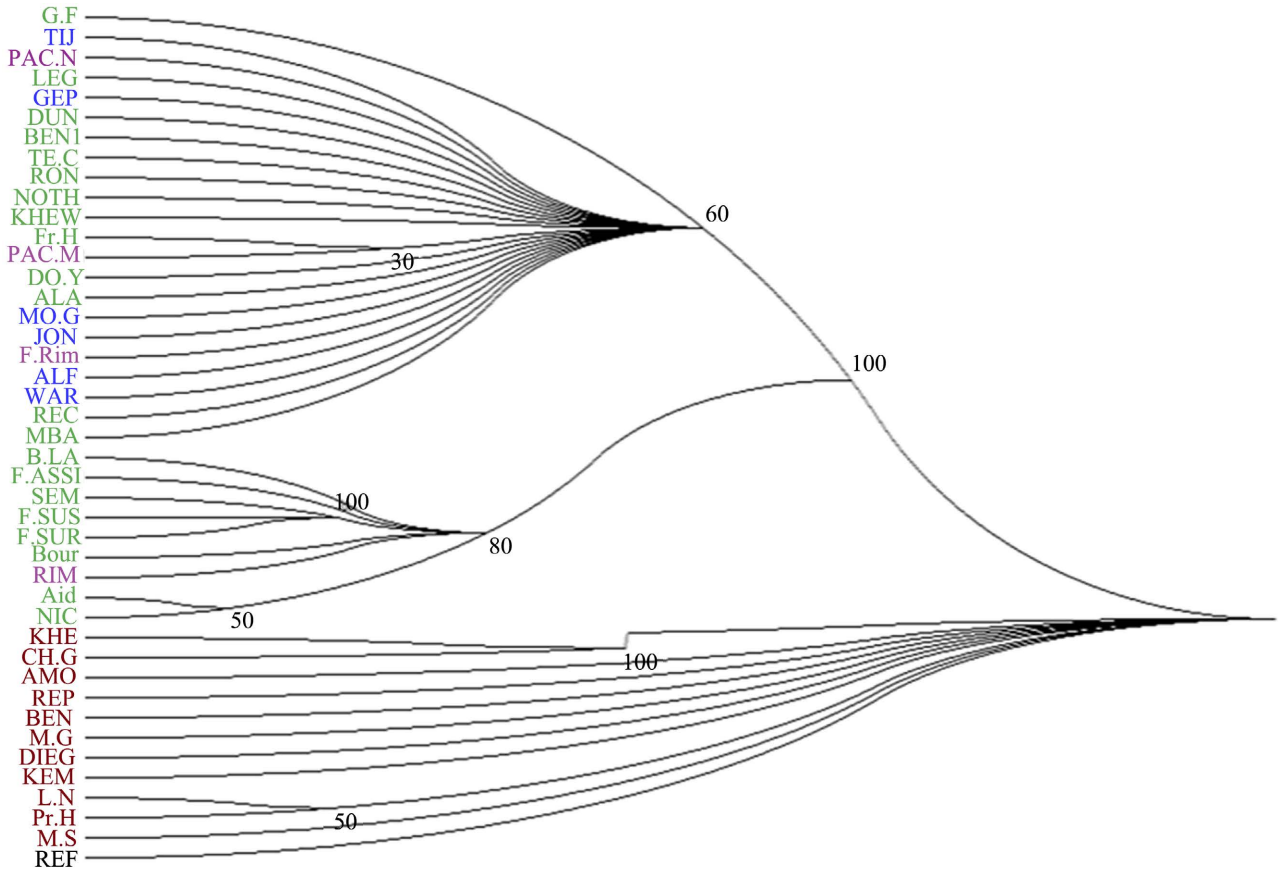
Figure 1. *SRY* gene haplotype network.

Table 1. Genetic variability parameters of *SRY* gene sequences.

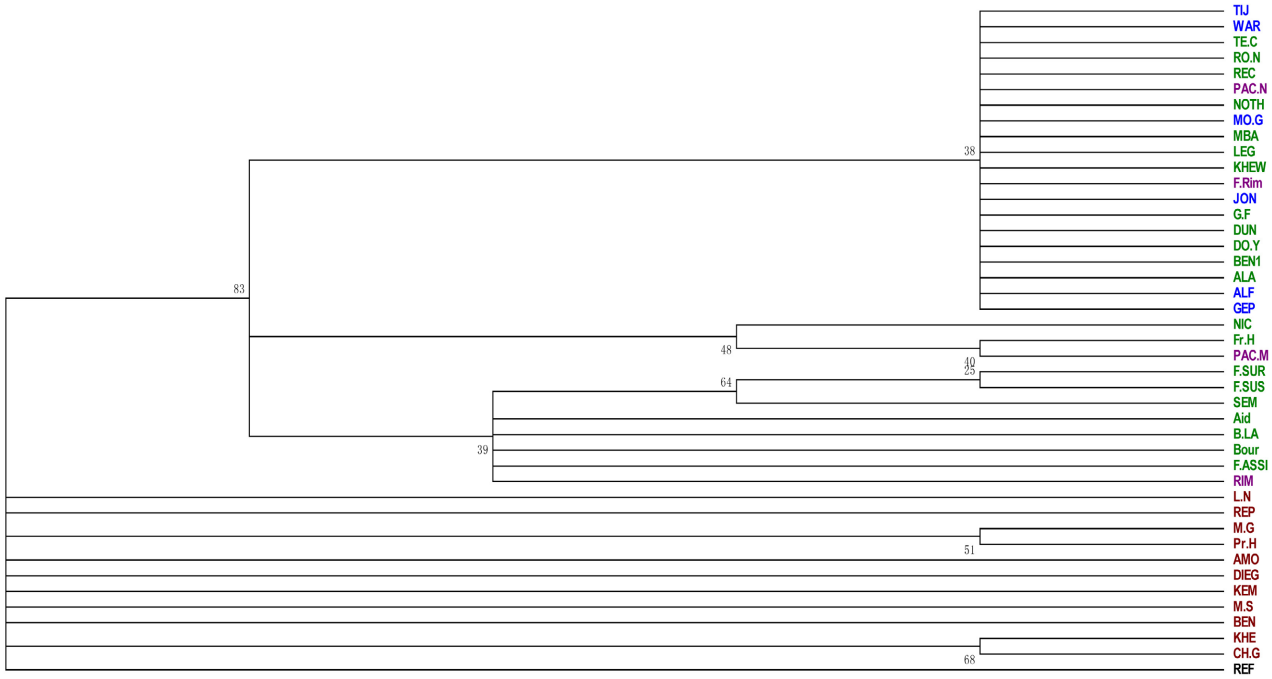
Genetics variability			
Parameters			Variabilities
Number of sites 558			558
Monomorphic sites			540
Variable sites	Polymorphic sites		18
	Non-informative (singleton)		09
	Informative (parsimony)		09
Total number of Eta mutations			18
Number of haplotypes (h)			13
Average number of nucleotide differences K			2.907
Transition rate (%)			48.83
Transversion rate (%)			51.18
Mutation rate R			0.83
Genetic Diversity indices			
		Haplotypic diversity (Hd)	nucleotide diversity (π)
By population	"Batling"	0.643	0.00236
	"Birahim"	0.800	0.00339
	"Gorgui"	1.000	0.00269
	"Tyson"	0.000	0.00000
Total population		0.774	0.00522


F.Rim
TIJ
TE.C
RO.N
REC
PAC.N
NOTH
MO.G
MBA
LEG
KHEW
JON
G.F
GEP
DUN
DO.Y
BEN1
ALA
ALF
WAR
Fr.H
PAC.M
NIC
Aid
B.LA
Bour
F.ASSI
RIM
SEM
F.SUR
F.SUS
L.N
M.G
Pr.H
REP
AMO
DIEG
KEM
M.S
BEN
KHE
CH.G
REF
43
84
32
29
53
52
46
65
39
53
64

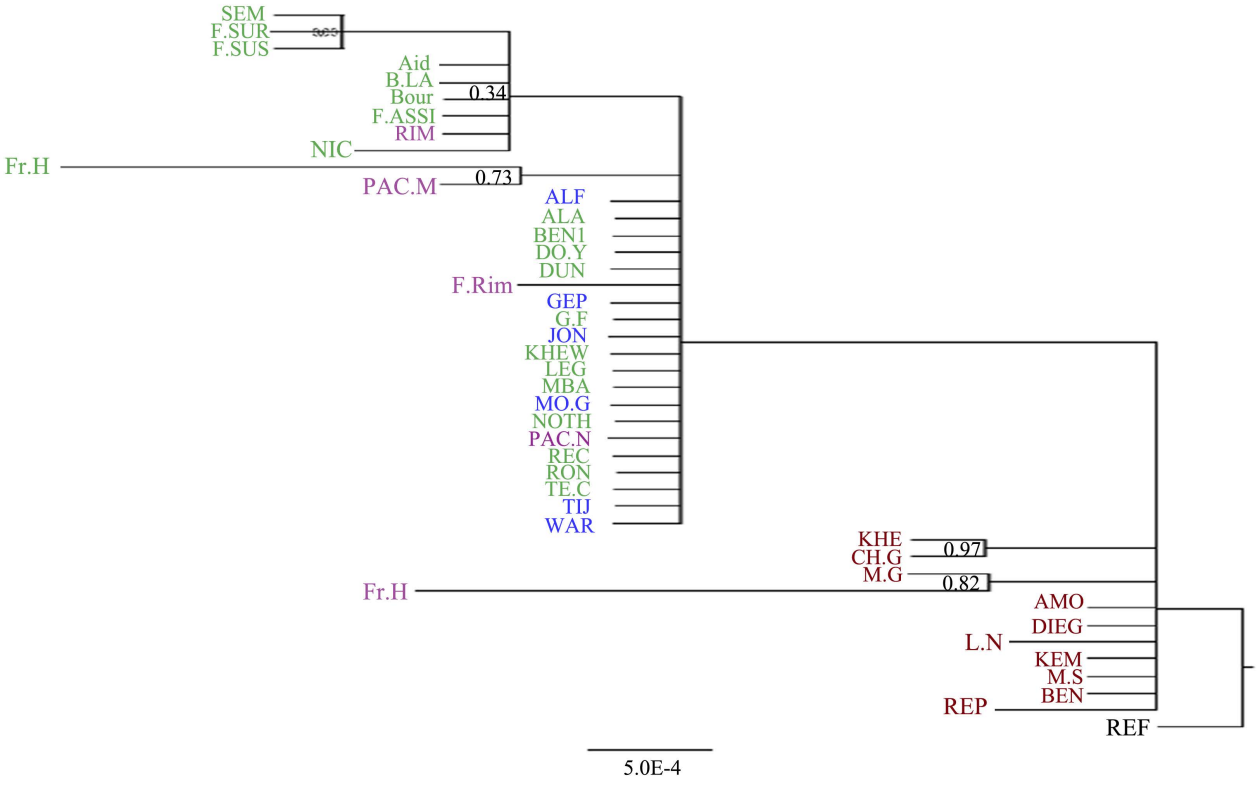
(a)


G.F
TIJ
PAC.N
LEG
GEP
DUN
BEN1
TE.C
RON
NOTH
KHEW
Fr.H
PAC.M
DO.Y
ALA
MO.G
JON
F.Rim
ALF
WAR
REC
MBA
B.LA
F.ASSI
SEM
F.SUS
F.SUR
Bour
RIM
Aid
NIC
KHE
CH.G
AMO
REP
BEN
M.G
DIEG
KEM
L.N
Pr.H
M.S
REF
60
30
100
100
80
50
100
50

(b)



(c)



(d)

Figure 2. Phylogenetic relationships inferred from *SRY* gene sequences of Ladoum individuals belonging to four lineages: "Batling", "Birahim", "Gorgui" and "Tyson". (a) Phylogenetic tree based on *SRY* gene sequences sequences by the Neighbour-Joining method (NJ, K2P, 1000 bootstrap); (b) Phylogenetic tree based on *SRY* gene by the Maximum Likelihood method (ML, 1000 bootstrap); (c) Phylogenetic tree from *SRY* gene sequences using Maximum Parsimony (MP, 1000 bootstrap); (d) Phylogenetic tree from *SRY* gene sequences using Bayesian inference.

Globally, the four approaches give the same topologies, all resolving, with slight differences concerning the internal nodes. Two clades are identified: one that appears monophyletic with the maximum likelihood method and polyphyletic with the other methods grouping individuals from the "Birahim" lineage and the monophyletic one grouping individuals from the other lineages. The monophyletic clade is composed of four subclades, each of which is made up of individuals from each of the three lineages. The individuals of the "Tyson" lineage are all found in the same clade which they share in majority with those of the "Batling" lineage. The bootstrap values are globally low.

3.3. Degree of Similarity with Referenced Species

Table 2 shows the individuals and their percentage identity with the homologous sequences.

3.4. Molecular Dating Results

The most recent common ancestor, in other words, the period of divergence between the two clades (the "Birahim" lineage *versus* the common ancestor of the other three lineages) was 2504 years ago (Figure 3). The results of the Bayesian dating of the molecular clock show a divergence time of between 2300 and 2600 years. Furthermore, the null hypothesis, that the rate of evolution is homogeneous between branches, is rejected. Substitution rates vary significantly between branches.

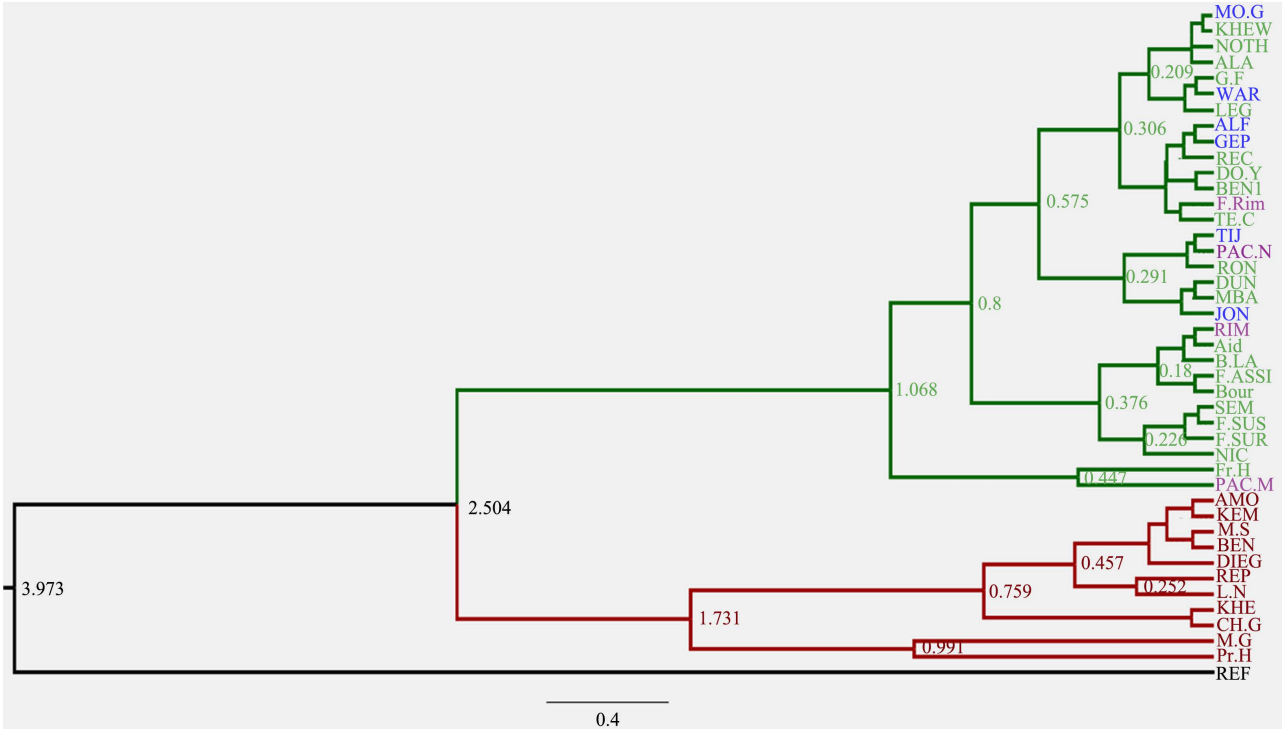


Figure 3. Phylogenetic tree from *SRY* gene sequences using Bayesian Inference combined with molecular dating (evolutionary rate $2.09 \times 10^{-3} \pm 1.08 \times 10^{-5}$ substitution/site/million years).

Table 2. Percentage of identity with the homologous sequences.

Samples	Homologous species	Lineage	Identity	Accession NCBI
REC	*Ovis aries*	"Batling"	98.92	AF026566.1
ALA	*Ovis aries*	"Batling"	98.92	AF026566.1
SEM	*Ovis aries*	"Batling"	98.92	AF026566.1
DUN	*Ovis aries*	"Batling"	98.92	AF026566.1
BEN1	*Ovis aries*	"Batling"	98.92	AF026566.1
DO.Y	*Ovis aries*	"Batling"	98.92	AF026566.1
Aid	*Ovis aries*	"Batling"	98.92	AF026566.1
NIC	*Ovis aries*	"Batling"	98.92	AF026566.1
NOTH	*Ovis aries*	"Batling"	98.92	AF026566.1
MBA	*Ovis aries*	"Batling"	98.92	AF026566.1
B.LA	*Ovis aries*	"Batling"	98.92	AF026566.1
Bour	*Ovis aries*	"Batling"	98.92	AF026566.1
LEG	*Ovis aries*	"Batling"	98.92	AF026566.1
Fr.H	*Ovis aries*	"Batling"	98.92	AF026566.1
TE.C	*Ovis aries*	"Batling"	98.92	AF026566.1
RO.N	*Ovis aries*	"Batling"	98.92	AF026566.1
F.ASSI	*Ovis aries*	"Batling"	98.92	AF026566.1
F.SURP	*Ovis aries*	"Batling"	98.92	AF026566.1
F.SUSP	*Ovis aries*	"Batling"	98.92	AF026566.1
GF	*Ovis aries*	"Batling"	98.92	AF026566.1
AMO	*Ovis aries*	"Birahim"	99.64	AF026566.1
Prof. H	*Ovis aries*	"Birahim"	99.27	AF026566.1
MG	*Ovis aries*	"Birahim"	99.27	AF026566.1
CH.G	*Ovis aries*	"Birahim"	99.64	AF026566.1
KEM	*Ovis aries*	"Birahim"	99.64	AF026566.1
KHE	*Ovis aries*	"Birahim"	99.64	AF026566.1
L.N	*Ovis aries*	"Birahim"	99.65	AF026566.1
M.S	*Ovis aries*	"Birahim"	99.64	AF026566.1
REP	*Ovis aries*	"Birahim"	99.60	AF026566.1
BEN	*Ovis aries*	"Birahim"	99.64	AF026566.1
DIEG	*Ovis aries*	"Birahim"	99.64	AF026566.1
RIM	*Ovis aries*	"Gorgui"	98.92	AF026566.1
PAC.M	*Ovis aries*	"Gorgui"	98.92	AF026566.1

Continued

PAC.N	*Ovis aries*	"Gorgui"	98.92	AF026566.1
F.RIM	*Ovis aries*	"Gorgui"	98.92	AF026566.1
Ado1	*Ovis aries*	"Gorgui"	98.92	AF026566.1
MO.G	*Ovis aries*	"Tyson"	98.92	AF026566.1
GEP	*Ovis aries*	"Tyson"	98.92	AF026566.1
JON	*Ovis aries*	"Tyson"	98.92	AF026566.1
ALF	*Ovis aries*	"Tyson"	98.92	AF026566.1
TIJ	*Ovis aries*	"Tyson"	98.92	AF026566.1
WAR	*Ovis aries*	"Tyson"	98.92	AF026566.1

4. Discussion

The *SRY* (*Sex-determining Region of Y chromosome*) gene is a portion of this DNA molecule, located on the Y chromosome responsible for male gonadal differentiation and conserved throughout evolution [35]. According to Meadows *et al.* [36], autosomal microsatellites and genetic variation especially in the 5' promoter region of the *SRY* gene are studied in discrimination research and origins of domestication in sheep [37]. In this study, the promoter region and a coding portion of the *SRY* gene were targeted to highlight genetic diversity, evolutionary mode as well as phylogenetic relationships between four lineages of Ladoum, a local sheep breed from Senegal. Any interpretation of the results will have to take into account the disproportion between the samples, explained by the fact that some breeders are reluctant to let their animals be sampled because of a strong suspicion of mysticism that reigns in this environment.

A low level of polymorphism is obtained. This low genetic variability of the *SRY* gene may be due to its specific function as a male determinant, which is transmissible from generation to generation [38] [39]. Fu *et al.* [40] on marsh buffalo was described that in the 5' region, the promoter sequences of transcriptional regulation were highly conserved confirming the observations made in our study population. Furthermore, it has been shown that mutations in the Y chromosome are very rare and that the *SRY* gene sequence is very stable within the same species, a group of potentially or actually interbreeding natural populations that are productively isolated from other similar groups [39]. Haplotypic diversity is much higher than the average nucleotide divergence between all pairs of haplotypes. Although Ndiaye [41] had shown that the Ladoum breed is the most stable compared to other sheep breeds in Senegal, the comparative trends in genetic diversity indices show a signal of rapid population growth from a small ancestral population that has been effective for a sufficient time for recovery of haplotypic diversity through mutation, but too short for the accumulation of strong sequence differences.

The patterns in structuring and genetic evolution highlighted by Sembène *et al.* [12] are confirmed by visual supports obtained. Networks and trees are two

categories of graphs used by biologists working in evolutionary science. Trees and networks differ fundamentally in their forms and meanings. From a mathematical point of view, networks can be defined as "cyclic related graphs". The adjective "related" indicates that all vertices are connected by links, and "cyclic" indicates that some nodes can be connected by different combinations of links [11]. In this study, the majority and central haplotype that groups in higher frequency the "Batling" population followed by the "Tyson" population indicates a close relationship between these two lineages and also suggests their belonging to a common ancestor. The small genetic distance [12] between "Batling" and "Tyson" and between "Balting" and "Gorgui" is supported by a small number of mutational steps. The "Birahim" lineage is very distinct from the rest of the population, making it a subpopulation in its own right. In any case, the analysis of the phylogenetic trees confirms two genetically separate clades. This very marked separation between "Birahim" and the other lineages could be explained by the evolutionary history in the selection of these local breeds which is therefore relative to the divergence between *Ovis aries* and *Ovis longipes*, about 200,000 years ago [42]. It is clear that there are currently two genetically identified lineages in Senegal: the "Birahim" lineage and another "lineage" that combines the other three defined earlier in this study. The specificity of the highly conserved "Tyson" lineage should also be noted, even though the single haplotype that constitutes it is found in individuals from "Batling". The polyphyly revealed by three of the four phylogenetic reconstruction methods in the "Birahim" lineage indicates that this group contains a number of individuals, but does not contain the common ancestor of all. The "Birahim" lineage could therefore be derived from two or more sheep breeds with a common ancestor, as shown by the similarity index of 99.64% to the ancestral reference *Ovis aries*. This hypothesis is further strengthened by its basal position in the phylogenetic trees. The monophyletic clade, composed of four subclades, appears to be a group that includes an ancestral species and all its descendants, but is rapidly structuring itself, certainly as a result of directed selection; this is evidenced by the homogeneity of the individuals of the "Tyson" lineage, even though they are phylogenetically related to those of "Batling". Although distinct from the "Birahim" lineage, this monophyletic group is thought to share a common ancestor with *Ovis aries*, but with a homology rate of 98.98%. As the reference sequence of *Ovis longipes does* not exist in the genetic databases, the group can only be related to the genetically closer reference sequence. In any case, in the rock engravings discovered in the Sahara, *Ovis longipes*, a sheep with a busted head, long and strong limbs, a narrow chest, and a short and oblique rump, is reminiscent of the large sheep that is the Ladoum [6] [8]. Inference of the demographic history of the clades identified as a result of phylogenetic analyses, which is essential to determine the factors that have shaped the evolution of populations, establishes that the divergence between two populations based on DNA sequences increases with time and the rate of evolution or mutation according to the molecular clock principle. This very recent moment of divergence on a geo-

logical step between the two clades suggests a rapid reorganisation into a highly mutational microsatellite zone characterised by cytosine-punctuated polyadenylation. Although the origin of the microsatellites remains unknown, a number of arguments suggest that they may result from the random or transpositional appearance of a sufficient number of repeated motifs to initiate the variability-generating process [43]. The origin of the Ladoum breed from selection between local and exotic breeds can therefore only be confirmed when we have a robust phylogeny involving the various current sheep breeds. The sheep was, with the goat, one of the first domesticated ungulates. It was then transferred by man to the whole world. The first traces of domestication date back 10,300 years. In general, the history of sheep domestication is not well known. The existence of several independent domestication events is suggested by the presence of multiple, highly divergent mitochondrial haplogroups in the domestic species, but the origin of domestic sheep is controversial [44]. On the basis of archaeological and genetic data, three taxa have been proposed as the origin of the domesticated wild species. These are the Argali (*Ovis ammon*), the Asiatic Mouflon (*O. orientalis*), and the Urial (*O. vignei*), according to the classification of Nadler *et al. in* Issam [45]. Archaeological data indicate that several regions would have been involved in locating domestication center(s), most notably eastern Anatolia, the Zagros, and the Indus Valley. Studies based on mitochondrial DNA (*MT-CYB*) and nuclear DNA (fragments distributed in 12 genes for a total of more than 4000 base pairs) have unambiguously shown that sheep were domesticated from *Ovis orientalis* once considered as a distinct species, but later recognized as a wild subspecies of the domestic sheep *Ovis aries* [44]. The location of the wild individuals that are genetically closest to the domesticated ones indicates that the closest taxon to the domesticated sheep is *O. orientalis gmelini*, which is located in western Anatolia and northern Zagros (mountains of Iran). Domestication would therefore have occurred in the Zagros [46]. In this study, the notion of population or race is still relevant, based on the existence of intermediate structures between the individual and the species. It also implies that the population is the place where new properties emerge that cannot be deduced from knowledge of the individuals that compose it.

5. Conclusion

The existence of populations implies that all individuals of a species are not randomly distributed throughout the species' range. On the contrary, they are distributed according to particular patterns, the spatial structures, linked in particular to habitat heterogeneity. For other species, the structuring is more of a social origin insofar as certain individuals maintain privileged relationships with each other, particularly for reproduction. It is clear that all organisms live in "populations" but the constitution of these populations, their origin, their development, their size and their functioning pose numerous and difficult problems, and lead researchers to lack a unified definition of this fundamental structure. This study, taking into account the disproportionate number of samples

per lineage, has nevertheless shown that: 1) there is a genetic structuring within the Ladoum breed in Senegal; 2) there are only two genetically identified groups within the Ladoum breed in Senegal: 3) the "Birahim" lineage is directly descended from *Ovis aries*; 4) the "Batling", "Gorgui" and "Tyson" lineages could be descended from a congener of *Ovis aries.*

Acknowledgements

The authors would like to thank **Ibrahima Sembène** for the help given to draw the phylogenetic trees. The authors would also like to thank **Anna Ndong** of Genomics laboratory for her help. They also thank the **Mbao breeders** who have accepted the blood sampling of their sheep.

Conflicts of Interest

The authors declare that the research was conducted in the absence of any commercial relationships that could be construed as a potential conflict of interest.

References

[1] Navès, M.N., Alexandre, G.G., Mahieu, M., Gourdine, J.L. and Mandonnet, N. (2011) Les races animales locales: Bases du développement innovant et durable de l'élevage aux Antilles. *Carrefours de l'innovation agronomique*, **16**, 220-221.

[2] Gueye, C. (2020) Caractérisation génétique des moutons de race Ladoum en fonction des locations au Sénégal, Mémoire de Master en Biologie animale, Spécialité Génétique des populations, UCAD, Dakar, 30 p.

[3] Paul, M.C.B. (2005) Sénégal: Un système de santé animale en voie de privatisation. Thèse de Médecine vétérinaire: Ecole Nationale Vétérinaire d'Alfort, à la Faculté de Médecine de Créteil, 145 p.

[4] Ousseini, H. (2011) Analyse socio-économique des élevages du mouton Ladoum dans la commune de Thiès/Sénégal, Mémoire de Master en Productions Animales et Développement Durable, EISMV, Dakar, 35 p.

[5] Diack, F., Traore, E.H., Thior, E.H.Y., Gueye, C., Sembène, M. and Seck, M. (2021) Etude de quelques caractéristiques morphobiométriques de la race Ladoum du Sénégal. *International Journal of Advanced Research*, **9**, 315-323.

[6] Sall Sada, M. (2007) Caractérisation morpho-biométrique et système d'élevage du Mouton Ladoum. Mémoire d'Ingénieur des travaux d'élevage (ENCR), Université de Thiès, Thiès, 37 p.

[7] Kane, M. (1995) Les races d'animaux élevés en Mauritanie. Bulletin d'information sur les ressources génétiques animales (FAO), 15 p.

[8] Thior, Y.E.H. (2013) Analyse des stratégies endogènes d'alimentation en élevage ovin Ladoum dans la région de Thiès et propositions d'amélioration. Thèse de doctorat en médecine vétérinaire; UCAD (EISMV), Dakar, 98 p.

[9] Sadio, M.C. (2010) Caractérisation génétique des races ovines sahéliennes: Etude du Ladoum et du Touabire. Mémoire de Master en Biologie animale, Spécialité Génétique des populations, UCAD, Dakar, 30 p.

[10] Hillis, D.M. (1997) Biology Recapitulates Phylogeny. *Science*, **276**, 218-219. https://doi.org/10.1126/science.276.5310.218

[11] Katsura, Y., Kondo, H.X., Ryan, J., Harley, V. and Satta, Y. (2018) The Evolutionary Process of Mammalian Sex Determination Genes Focusing on Marsupial *SRYs*. *BMC Evolutionary Biology*, **18**, Article No. 3. https://doi.org/10.1186/s12862-018-1119-z

[12] Sembène, P.M., Mbaye, F. and Sembène, P.M. (2023) Genetic Diversity and Demographic Evolution in Four Lineages of Ladoum Sheep in Senegal Using Polymorphisme of the *SRY* Gene. *International Journal of Genetics and Genomics*, **11**, 6-17.

[13] Hall, T.A. (1999) BioEdit: A User-Friendly Biological Sequence Alignment Editor & Analysis Program for Windows 95/98/NT. *Nucleic Acids Symposium Series*, **41**, 95-98.

[14] Thompson, J., Higgins, D. and Gibson, T. (1997) CLUSTAL W: Improving the Sensitivity of Progressive Multiple Sequence Alignment through Sequence Weighting, Position-Specific Gap Penalties & Weight Matrix Choice. *Nucleic Acids Research*, **22**, 4673-4690. https://doi.org/10.1093/nar/22.22.4673

[15] Kumar, S., Stecher, G. and Tamura, K. (2016) MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. *Molecular Biology & Evolution*, **33**, 1870-1874. https://doi.org/10.1093/molbev/msw054

[16] Nei, M. and Tajima, F. (1981) Polymorphisme de l'ADN détectable par les endonucléases de restriction. *Génétique*, **97**, 145-163. https://doi.org/10.1093/genetics/97.1.145

[17] Nei, M. (1987) Molecular Evolutionary Genetics. Columbia University Press, New York, 34 p. https://doi.org/10.7312/nei-92038

[18] Librado, P. and Rozas, J. (2009) DNASP V5: A Software for Comprehensive Analysis of DNA Polymorphism Data. *Bioinfromatics*, **25**, 1451-1452. https://doi.org/10.1093/bioinformatics/btp187

[19] Bandelt, H.J., Forster, P. and Röhl, A. (1999) Median-Joining Networks for Inferring Intraspecific Phylogenies. *Molecular Biology and Evolution*, **16**, 37-48. https://doi.org/10.1093/oxfordjournals.molbev.a026036

[20] Kingman J.F.C. (2000) Origins of the Coalescent: 1974-1982. *Genetics*, **156**, 1461-1463. https://doi.org/10.1093/genetics/156.4.1461

[21] Saitou, N. and Nei, M. (1987) The Neighbor-Joining Method: A New Method for Reconstructing Phylogenetic Trees. *Molecular Biology Evolution*, **4**, 406-425.

[22] Felsenstein, J. (1978) Cas dans lesquels la parcimonie ou les méthodes de compatibilité seront positivement trompeuses. *Zoologie Systématique*, **27**, 401-410.

[23] Cavalli-Sforza, L.L. and Edwards, A.W.F. (1967) Phylogenetic Analysis: Models and Estimation Procedures. *Evolution*, **32**, 550-570. https://doi.org/10.2307/2406616

[24] Felsenstein, J. (1981) Evolutionary Trees from DNA Sequences: A Maximum Likelihood Approach. *Journal of Molecular Evolution*, **17**, 368-376. https://doi.org/10.1007/BF01734359

[25] Felsenstein, J. (1985) Confidence of Limits on Phylogenies: An Approach Using Bootstrap. *Evolution*, **39**, 783-791. https://doi.org/10.2307/2408678

[26] Soltis, P.S. and Soltis, D.E. (2003) Applying the Bootstrap in Phylogeny Reconstruction. *Statistical Science*, **18**, 256-267. https://doi.org/10.1214/ss/1063994980

[27] Mahé, F. (2009) Phylogénie, éléments transposables et evolution de la taille des génomes chez les lupins. Université Rennes, Rennes, 253 p.

[28] Huelsenbeck, J.P. and Ronquist, F. (2001) Mrbayes: Bayesian Inference of Phyloge-

netic Trees. *Bioinformatics*, **17**, 754-755.
https://doi.org/10.1093/bioinformatics/17.8.754

[29] Drummond, A.J., Rambaut, A. and Suchard, M.A. (2015) BEAST: Bayesian Evolutionary Analysis by Sampling Trees. V. 1.8.2. http://beast.bio.ed.ac.uk

[30] Rodriguez, F., Oliver, J.L., Marin, A. and Medina, J.R. (1990) The General Stochastic Model of Nucleotide Substitution. *Journal of Theoretical Biology*, **142**, 485-501. https://doi.org/10.1016/S0022-5193(05)80104-3

[31] Pérez, T., Hammer, S., Albornoz, J. and Domínguez, A. (2011) Y-Chromosome Phylogeny in the Evolutionary Net of Chamois (Genus *Rupicapra*). *BMC Evolutionary Biology*, **11**, 272. https://doi.org/10.1186/1471-2148-11-272

[32] Brunel, S. (2019) Le chromosome Y et le gène *SRY*. Epidémiologie et Santé, 3 p.

[33] Drummond, A.J., Ho, S.Y.W., Phillips, M.J. and Rambaut, A. (2006) Relaxed Phylogenetics and Dating with Confidence. *PLOS Biology*, **4**, 88-91. https://doi.org/10.1371/journal.pbio.0040088

[34] Douzery, E.J.P., Snell, E.A., Bapteste, E., Delsuc, F. and Philippe, H. (2004) The Timing of Eukaryotic Evolution: Does a Relaxed Molecular Clock Reconcile Proteins and Fossils? *Proceedings of the National Academy of Sciences of the United States of America*, **101**, 15386-15391. https://doi.org/10.1073/pnas.0403984101

[35] Sinclair, A.H., Berta, P., Palmer, M.S., *et al.* (1990) Un gène de la région humaine déterminant le sexe code pour une protéine présentant une homologie avec un motif de liaison à l'ADN conservé. *Nature*, **346**, 240-244. https://doi.org/10.1038/346240a0

[36] Meadows, J.R.S., Hanotte, O., Drögemüller, C., *et al.* (2006) Globally Dispersed Y Chromosomal Haplotypes in Wild and Domestic Sheep. *Animal Genetics*, **37**, 444-453. https://doi.org/10.1111/j.1365-2052.2006.01496.x

[37] Niemi, M., Bläuer, A., Iso-Touru, T., Nyström, V., Harjula, J., Taavitsainen, J.P. and Kantanen, J. (2013) ADN mitochondrial et diversité chromosomique Y dans les anciennes populations de moutons domestiques (*Ovis aries*) en Finlande: Comparaison avec les races ovines contemporaines. *Évolution de la Sélection Génétique*, **45**, 1-14. https://doi.org/10.1186/1297-9686-45-2

[38] Meadows, J.R.S., Hawken, R.J. and Kijas, J.W. (2004) Nucleotide Diversity on the Ovine Y Chromosome. *Animal Genetics*, **35**, 379-385. https://doi.org/10.1111/j.1365-2052.2004.01180.x

[39] Hasanain, M.H., Mahmoud, K.G.M., Ahmed, Y.F., *et al.* (2022) Enquête sur le polymorphisme du gène du facteur déterminant le sexe (*SRY*) et l'association des critères de sperme avec la fertilité du champ chez les taureaux Buffalo égyptiens. *Journal égyptien de chimie*, **65**, 1-2.

[40] Fu, Q., Zhang, M., Qin, W.S., *et al.* (2007) Cloning the Swamp Buffalo *SRY* Gene for Embryo Sexing with Multiplex-Nested PCR. *Theriogenology*, **68**, 1211-1218. https://doi.org/10.1016/j.theriogenology.2007.07.007

[41] Ndiaye, B. (2019) L'élevage des petits ruminants dans le Ferlo: Pratiques d'élevage, dynamique des troupeaux et caractérisation génétique du mouton Peul-peul du Sénégal. Thèse de doctorat en Biologie animale, Spécialité Génétique des populations, UCAD, Dakar, 123 p.

[42] MacHugh, D.E., Shriver, M.D., Loftus, R.T., Cunningham, P. and Bradley, D.G. (1997) Microsatellite DNA Variation and the Evolution, Domestication and Phylogeography of Taurine and Zebu Cattle (*Bos taurus* and *Bos indicus*). *Genetics*, **146**, 1071-1086. https://doi.org/10.1093/genetics/146.3.1071

[43] Jeffrey, S.A., Webb, A. and Schulz, A. (2010) Factors Affecting Goal Difficulty and Performance When Employees Select Their Own Performance Goals: Evidence from the Field. *Journal of Management Accounting Research*, **22**, 209-232. https://doi.org/10.2308/jmar.2010.22.1.209

[44] Hamidreza, R. (2011) Phylogénie moléculaire du Genre Ovis (Mouton et Mouflons), Implications pour la Conservation du Genre et pour l'Origine de l'Espèce Domestique. 134 p. https://theses.hal.science/tel-00625145

[45] Issam Belaid (2012) Caractérisation morphologique des troupeaux ovins dans la région de Sétif. Mémoire: Magister Université Farhat Abbas-Setif (Algérie); Faculté des Sciences de la Nature et de la Vie, département des Sciences Agronomiques, 199 p.

[46] Cheng, Z.Y., Duan, T., Sha, J., *et al.* (2006) Origin, Genetic Diversity, and Population Structure of Chinese Domestic Sheep, *Gene*, **376**, 220-223. https://doi.org/10.1016/j.gene.2006.03.009