Full Length Article

Molecular Phylogeography of China Roe Deer Population Based on MtDNA D-Loop Gene

Yanhua Liu*, **Xue Liang and Jinyue Liu**
College of wildlife and Protected Area, Northeast Forestry University, Harbin, Heilongjiang Province, China
*For correspondence: wildlife_deer@126.com

Abstract

There are different opinions on the phylogeographical pattern of roe deer in the world. This study studied some sequence of the mitochondrial DNA (mtDNA) control region (420 bp) of 12 wild roe deer from the Greater Khingan Mountains, China, compared them with 39 Genebank downloaded mtDNA control region sequences around the world, and analyzed the phylogeographic status of roe deer. The phylogenetic tree and haplotype network analysis results showed that the roe deer were divided into two subspecies, European roe deer and Siberian roe deer. The geographical distribution of the European roe deer included Austria, France, Crimea, Hungary, Lithuania, Romania, Italy and Spain. Among them, the clade A and B were sister branches of clade C in parallel. The Siberian roe deer were geographical distributed in Russia, South Korea, China, Mongolia, Poland. The clade D and E were parallel sister branches. The results of this study indicated that the Spanish and Italian roe deer were different from the haplotypes of the other European roe deer because of the geographic isolation. While the roe deer from Poland and those from western Russia and Mongolia were clustered into one branch. It is speculated that this might due to introgression of an ancient gene by highly differentiated lineages during species expansion, and population spread in the ice age may be one of the important reasons for the existing distribution pattern. © 2020 Friends Science Publishers

Keywords: Roe deer; Phylogeography; Mitochondrial DNA D-loop

Introduction

Phylogeography is a discipline that studies the historical reasons and evolutionary processes of the contemporary geographical distributions of closely related species or populations of the same species (Avise 1989). In recent years, with the rapid development of molecular biotechnology, an interdisciplinary discipline, molecular phylogeography has emerged. Molecular phylogeography primarily uses molecular biology techniques to explore the formation mechanism of phylogeographical structures within and among species at the DNA level, successfully integrating intraspecific microevolution and interspecific macroevolution (Avise 1998; Arbogast 2001). Currently, the most widely applied fields include inference of the historical evolution of a population (Arbogast and Kenagy 2001; Päckert *et al.* 2010; Recuero and García-París 2011), determination of the glacier refuges (You *et al.* 2010; Beatty and Provan 2011; Recuero and García-París 2011), inference of the phylogeographical pattern of a population and its causes (Lin *et al.* 2014), studies on species differentiation and biodiversity protection (Buckley 2009; Xu *et al.* 2010). These studies can help understand the formation, distribution, dispersal pathways and even

extinction of populations, as well as the impacts of historical geological events on them.

Phylogeography studies are usually carried out based on phylogenetic and geographical distribution studies, and the mitochondrial genome-based phylogenetic methods are also adopted in various biological population studies. Mitochondrial DNA (mtDNA) is a covalently closed-circular double-stranded DNA molecule and has the characteristics of simple molecular structure, maternal inheritance, high nucleotide divergence and high evolution rate. As a well-behaved molecular marker, it has been widely used in the population genetic and phylogenetic relationship studies (Zhang and Shi 1992). Compared with the other mtDNA genes, the mitochondrial displacement loop (D-loop) gene has the advantages of rich in A/T bases, being a hypervariable region genetically that its nucleotide replacement rate is 5–10 times higher than the other mtDNA regions. And it has the fastest evolution rate and the highest diversity that the polymorphism is higher than mitochondrial fragments in the other regions. Therefore, it is often used as the primary genetic marker in phylogenetic research.

Currently, there are a lot of debates about the classification of species and subspecies of roe deer

---

worldwide. Some scholars concluded that the Siberian roe deer (*Capreolus pygargus*, *C. pygargus*) could be divided into two major subspecies, the eastern subspecies in the Russian Far East and the western subspecies in the Western Siberian region (Randi *et al.* 1997). Some scholars believe that there were genetic differences between the roe deer of Jeju Island and other populations, though with uncertain taxonomy (Lee *et al.* 2016). While Chinese scholars divided Chinese roe deer into four subspecies: Central Asia subspecies, North China subspecies, Northeast subspecies, and Northwest subspecies (Wang 2003). Another scholar temporarily classified the roe deer from Heilongjiang Province as the Northeast subspecies (Ma 1986). Whereas there are scholars believed that the Siberian roe deer found in northeastern China belonged to the *C. p. manchuricus* subspecies according to their morphological differences from other subspecies (Xiao *et al.* 2007). Therefore, the present study used feces of roe deer as experimental materials to conduct an exploratory study on the phylogeographical evolution of roe deer, expecting to provide some theoretical basis for the phylogeographical evolution study of the other wild animals.

Materials and Methods

Samples

These research materials were 12 fecal samples of roe deer collected by non-invasive sampling method from the Greater Khingan Mountains in Heilongjiang Province. The detailed sampling method was as follows: first, the samples were collected in winter to reduce the activity of bacteria in the feces, and the areas where roe deer often appear were searched along their footprints for feces after snowfall. Second, try to avoid sample contamination. The sample collector wears disposable gloves and sealed each fecal sample in a separate bag to ensure the quality of target DNA. Third, the samples were stored at low temperature. If a fresh fecal sample was found, it could be naturally frozen outdoors, transported to the laboratory in an insulated chest filled with crushed ice, and stored in a freezer below -20°C once arrived at the lab.

DNA extraction and amplification

The fecal DNA was extracted using the QIAamp DNA Stool Mini Kit with primers L-Pro: 5'-CGTC AGTC TCAC CATC AACC CCCA AAGC, and 3'H-Phe: 5'-GGGA GACT CATC TAGG CATT TTCA GTG (Randi *et al.* 1997), amplifying the D-loop sequence from the 3'-end. The PCR reaction system was: 15.3 μL DNA, 4 μL dNTPs, 1.6 μL F-primer, 1.6 μL R-primer, 0.5 μL Taq™ DNA polymerase, 4 μL 1 × PCR buffer, 15.3 μL H$_2$O, and 3 μL BSA. The PCR reaction was set as pre-denaturation at 95°C for 10 min, followed by 40 cycles of (denaturation at 95°C for 30 s, annealing at 50°C for 40 s, and extension at 72°C

for 60 s), a final extension at 72°C for 10 min, and then stored at 4°C. The PCR products were detected by 1.0% conventional agarose gel electrophoresis, with a sample load of 5 μL PCR product mixed with 1 μL 6 × Loading Buffer, subjected to 100 V electrophoresis for 30 min, and stained with Good Wall staining. The electrophoresis bands were examined using a gel imaging system and the bright and narrow bands were sent to Shanghai Sangon Biotech Co., for purification and sequencing.

Statistical analysis

The obtained sequences were subjected to alignment using Clustal X version 2.0 with manual correction (Larkin *et al.* 2007). By BLAST search (Alschul 1997) the sequences were aligned with the known roe deer mitochondrial homologous sequences to determine whether they were the target sequences. DnaSP 4.0 (Rozas *et al.* 2003) was used to determine haplotypes. Based on Kimura's two parameter model, MEGA7 (Kumar *et al.* 2016) was used to construct phylogenetic trees using the maximum likelihood, neighbor-joining, and maximum parsimony methods, and the branch confidence level was obtained by bootstrapping (1000 times). PopART 1.7 (Leigh and Bryant 2015) was used to estimate the phylogenetic relationship between geographical samples based on the TCS Network.

Results

A total of 12 sequences were successfully amplified from the 12 collected fecal samples, resulting in nine haplotypes. In order to allow all the obtained sequences to tell more genetic information, three haplotypes (Chi1:AY854040, Chi2:AY854041, and Chi3:AY854042) were selected for phylogenetic analysis. A number of 39 sequences were downloaded from GeneBank (Table 1) and elk (*Cervus elaphus*) was used as an outgroup to construct the phylogenetic trees.

The three different molecular phylogenetic trees basically shared the same topology, and the most clear and comprehensive topology was generated by neighbor-joining (Fig. 1). The topology clearly demonstrated two haplotype clades, *C. capreolus* and *C. pygargus*, both with a high confidence level.

The *C. capreolus* clade is composed of three haplogroups, named A, B, and C; where A is composed of haplotypes from Austria, France, Crimea, Hungary, and Lithuania; B is composed of haplotypes from Romania, eastern Italy and central-southern Italy, and the central-southern Italy itself has a single branch; and C is composed of haplotypes from northern Spain and central-southern Spain, showing distinct geographical features. A, B and C are sister branches in parallel. The *C. pygargus* clade consists of two haplogroups named D and E, where D consists of haplotypes from western Russia, Mongolia, and Poland, and E consists of haplotypes from eastern Russia,

Table 1: GeneBank download sequence information

Code	Location	GenBank accession nos
Aus1-2	Austria	KF724415.1 KF724419.1
Poland1-3	Poland	KJ558283 KJ558285 KJ558286
Fra1-2	France	KF700106 KF700111
Cri	Crimean	KF724416
Nsp1-2	Northern Spain	KF700100 KF700103
Csp1-2	Central South Spain	KF700102 KF700104
Eit1-2	Eastern Italy	KF700108 KF724418
Cit1-3	Central South Italy	KF724429 KF724430、KF724431
Hun1-2	Hungary	KP659209 KP659211
Lit	Lithuania	KM215767
Rom1-5	Romania	KF724427 KF724432 KF724436 KF724437 KF724438
Eru1-5	Eastern Russia	KF724444 KF724445 KF724446 KF724447 KF724448
Wru1-2	Western Russia	KF724442 KF724443
Kor1-4	South Korea	JX428900 JX428902 JX428903 JX428905
Mon1-3	Mongolia	JQ958973 JQ958976 JQ958975
Cervus elaphus	——	GU457434

South Korea, and northeastern China. D and E are sister branches in parallel too. It can be seen from the network analysis diagram of roe deer (Fig. 2) that roe deer was clearly composed of two clusters of haplotypes, *C. capreolus* and *C. pygargus*, consistent with the phylogenetic tree results.

Discussion

Taxonomically, the roe deer belongs to the class Mammalia, the clade Eutheria, the order Artiodactyla, the suborder Ruminantia, the family Cervidae, and the genus *Capreolus*, however, with debates on species and subspecies. At present, it is generally believed that *Capreolus* has two extant species: the smaller European roe deer (*C. capreolus*) and the bigger Siberian roe deer or eastern roe deer (*C. pygargus*), who has a shoulder height of 60 to 75 cm, a weight of 25 to 45 kg, and a body length of 110–120 cm, about twice the size of European roe deer. The former is widespread in central-western Europe, while the latter is found across Asia and Eastern Europe. The distribution of these two species is staggered in the Caucasus. *C. capreolus* is found on the southern slope of the Caucasus Mountains, while the territory of *C. pygargus* stretches to the northern Caucasus. Nevertheless, there were no hybrids found between these two species under natural conditions. At the subspecies level, some scholars believe that the Siberian roe deer was composed of three subspecies, (1) *C. pyargus pygargus*, distributed in some regions of western and eastern Siberia; (2) *C. pygargus tianschanicus*, distributed in the Tianshan Mountains; and (3) *C. pygargus manchuricus*, distributed in the Far East. Whereas other scholars divided *C. pygargus* into two subspecies: *C. pygargus pygargus* from some regions of western and eastern Siberia, and *C. pygargus tianschanicus* from Tianshan and East Asia. There are also scholars believe that Manchurian roe deer (*C. capreolus bedfordi*) included Northeast China subspecies *manchuricus* and South Korea subspecies *ochracea*, the subspecies *mantschuricus*, *melanotis*, and *ochracea* were

synonymous of *C. pygargus bedfordi*, which included roe deer populations from China, Korean Peninsula, Mongolia and southeastern Siberia. Some researchers have pointed out that the roe deer from Korean Peninsula, Northeast China, and regions near Russia belonged to *C. pygargus bedfordi*, and *mantschuricus* and *ochracea* were synonyms of *C. pygargus bedfordi*.

This study analyzed the genetic relationship of the roe deer population at the molecular level. The results showed that *C. capreolus* was composed of three haplogroups, Including, the haplogroup from southern Iberia (central-southern Spain and northern Spain), the haplogroup from Eastern Europe (Italian Alps, Romania, Greece, and Lithuania), and the Central European haplogroup (not geographically restricted, including haplotypes widely distributed throughout Europe except Lithuania and Crimea) (Lorenzini *et al.* 2014). It was only due to the last glaciation period that the southern Iberian Peninsula became a refuge for most mammals (Arribas 2004). The interglacial population expansion was blocked by the Pyrenees Mountains and could not reach the other regions of Europe; therefore, the haplotype there was different from those from the other parts of Europe (Randi *et al.* 2004; Royo *et al.* 2007). In the phylogenetic and haplotype network diagrams, the haplotypes in central-southern Italy are clustered into a single branch because of the barrier of Alps. The results of this study are consistent with the findings of other scholars that the Italian roe deer had two lineages, one was the Alps population, and the other lineage was the central-south Italy population (named *C. C. italicus*) (Lorenzini *et al.* 2002). Regarding *C. pygargus*, this study found that the mtDNA haplotype of the Polish roe deer favored *C. pygargus* with a higher support, while the Don and Volga rivers somehow farther from Poland were the hybrid zone of *C. capreolus* and *C. pygargus*. However, currently scholars generally believe that the roe deer found in Poland belonged to *C. pygargus* (Vorobieva *et al.* 2011). In this study, only mtDNA was used as the genetic marker, and there was not enough evidence to support that Polish roe deer belonged to
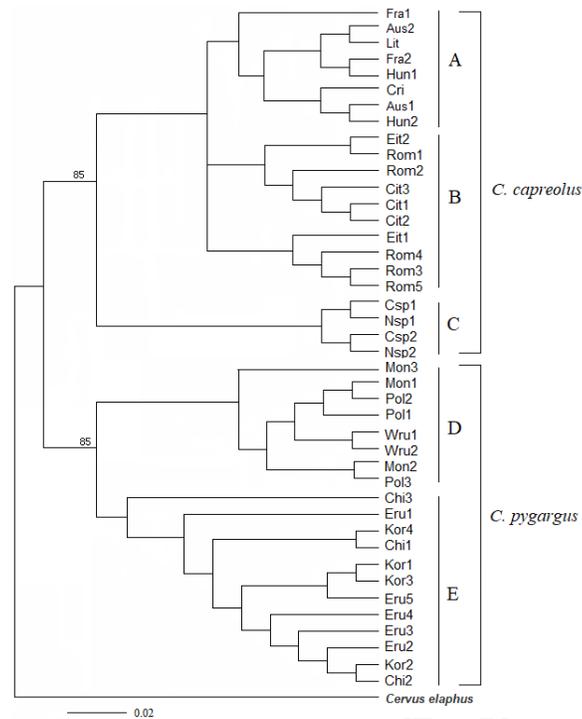
Fig. 1: Phylogenetic tree of mtDNA D-loop gene of *C. pygargus* constructed based on NJ method. Red deer (*Cervus elaphus*) was the outgroup of homologous sequence. The number at the node is the percentage of Bootstrap repeats with 1000 repetitions
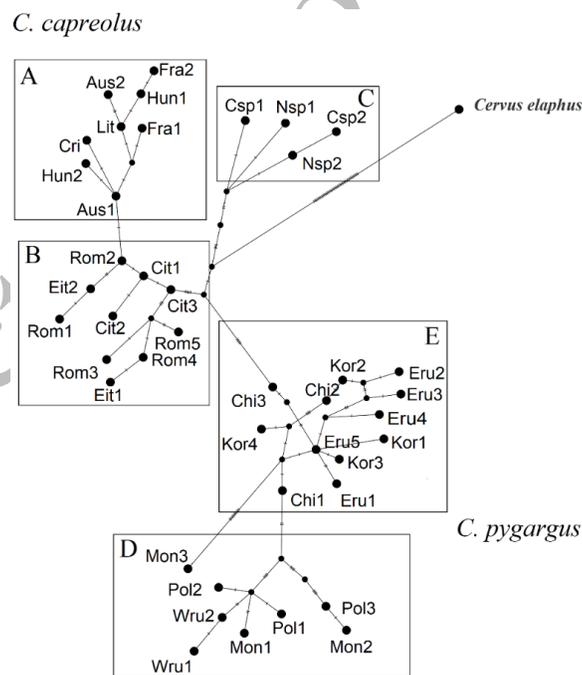

Fig. 2: Haplotype Network analysis of mtDNA D-loop gene of *C. pygargus* constructed based on TCS Network

C. pygargus due to the limitation of maternal inheritance. Then, the *C. pygargus* haplotype found in Poland might be because of introgression of an ancient gene by highly differentiated lineages during species expansion in its continuous distribution range (Lorenzini *et al.* 2014). From the phylogenetic and network analysis, the haplotypes of the Korean roe deer shared the same taxa as that of the Northeast China roe deer (Xiao *et al.* 2007). While the roe

deer from Yakutia of Russia, Primorsky Krai, northern Mongolia, and South Korea converged to a same branch (Lee *et al.* 2016), indicating their close relationship in the history of molecular evolution in addition to being geographically adjacent.

Population dispersal is one of the important reasons for the current distribution pattern of roe deer. Based on morphological characteristics, *Capreolus* may originate from *Procapreolus* and may occur in the late Pliocene Epoch (Groves 2007; Valli 2010). *C. capreolus* and C. pygargus may occur about 10 kya ago. These two species seemed to occupy their modern distribution, that is, *C. capreolus* was distributed in Europe and *C. pygargus* was distributed in Central Asia. During the cold period of the Pleistocene Epoch, *C. capreolus* was excluded from northern Europe. During the Ice Age, *C. capreolus* survived in the refuges in the Iberian and Apennine peninsulas and in the refugees in Eastern Europe (Sommer *et al.* 2009). During the interglacial period, *C. capreolus* returned to northern Europe and thrived there. Whereas the ancestors of *C. pygargus* may have survived in the refuge in Central Asia, and it was not until the interglacial period that *C. pygargus* resettled in Central Asia (Lee *et al.* 2016).

Conclusion

This study analyzed the genetic relationship of the roe deer population at the molecular level. The analysis results showed that the roe deer were divided into two subspecies, European roe deer and Siberian roe deer. The geographical distribution of the European roe deer included Austria, France, Crimea, Hungary, Lithuania, Romania, Italy, Spain. The Siberian roe deer were geographical distributed in Russia, South Korea, China, Mongolia, Poland.

References

Alschul SF (1997). Gapped BLAST and PSI-BLAST: A new generation of protein detabase search programs. *Nucl Acids Res* 25:389–402

Arbogast BS (2001). Phylogeography: The History and Formation of Species. *Integr Comp Biol* 41:134–135

Arbogast B, G Kenagy (2001). Comparative phylogeography as an integrative approach to historical biogeography. *J Biogeogr* 28:819–825

Arribas O (2004). *Fauna y paisaje de los Pirineos en la Era Glaciar*. Lynx Edicions

Avise JC (1998). The history and purview of phylogeography: A personal reflection. *Mol Ecol* 7:371–379

Avise JC (1989). Gene Trees and Organismal Histories: A Phylogenetic Approach to Population Biology. *Evolution* 43:1192–1208

Beatty GE, J Provan (2011). Phylogeographic analysis of North American populations of the parasitic herbaceous plant reveals a complex history of range expansion from multiple late glacial refugia. *J Biogeogr* 38:1585–1599

Buckley D (2009). Toward an organismal, integrative, and iterative phylogeography. *Bioessays* 31:784–793

Groves (2007). *The Evolution of Artiodactyls*. The John Hopkins University Press, Baltimore, Maryland, USA

Kumar S, G Stecher, K Tamura (2016). MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. *Mol Biol Evol* 7:1870–1874

Larkin MABG, NP Brown, R Chenna, PA McGettigan, H McWilliam, F Valentin, IM Wallace, A Wilm, RLopez, JD Thompson, TJ Gibson, DG Higgins (2007). Clustal W and Clustal X version 2.0. *Bioinformatics* 23:2947–2948

Lee YS, N Markov, A Argunov, I Voloshina, D Bayarlkhagva, BJ Ki, KS Kim (2016). Genetic diversity and phylogeography of Siberian roe deer, *Caproulus pygargus*, in central and peripheral populations. *Ecol Evol* 6:7286–7297

Leigh J, D Bryant (2015). PopART: Full-feature software for haplotype network construction. *Meth Ecol Evol* 6:2496–2497

Lin A, G Csorba, LF Lin, T Jiang (2014). Phylogeography of Hipposideros armiger (Chiroptera: Hipposideridae) in the Oriental Region: The contribution of multiple Pleistocene glacial refugia and intrinsic factors to contemporary population genetic structure. *J Biogeogr* 41:317–327

Lorenzini R, L Garofalo, X Qin, I Voloshina, S Lovari (2014). Global phylogeography of the genus *Capreolus* (Artiodactyla: Cervidae), a Palaearctic meso-mammal. *Zool J Linn Soc* 170:209–221

Lorenzini R, S Lovari, M Masseti (2002). The rediscovery of the Italian roe deer: Genetic differentiation and management implications. *Ital J Zool* 69:367–379

Ma YQ (1986). *Mammal List in Heilongjiang Province*, pp: 396–400. Heilongjiang Sci Technol Press, Harbin, China

Päckert M, J Martens, YH Sun (2010). Phylogeny of long-tailed tits and allies inferred from mitochondrial and nuclear markers (Aves: Passeriformes, Aegithalidae). *Mol Phylogenet Evol* 55:952–967

Randi E, PC Alves, J Carranza, S Milosevic-Zlatanovic, A Sfougaris, N Mucci (2004). Phylogeography of roe deer (*Capreolus capreolus*) populations: The effects of historical genetic subdivisions and recent nonequilibrium dynamics. *Mol Ecol* 13:3071–3083

Randi EM, A Pierpaoliand, Danikin (1997). Mitochondrial DNA polymorphism in populations of Siberian and European roe deer (*Capreolus pygargus* and *C. capreolus*). *Heredity* 80:429–441

Recuero E, M García-París (2011). Evolutionary history of *Lissotriton helveticus*: Multilocus assessment of ancestral *vs*. recent colonization of the Iberian Peninsula. *Mol Phylogenet Evol* 60:170–182

Royo LJ, G Pajares, I Alvarez, I Fernandez, F Goyache (2007). Genetic variability and differentiation in Spanish roe deer (*Capreolus capreolus*): A phylogeographic reassessment within the European framework. *Mol Phylogenet Evol* 42:47–61

Rozas J, JC Sánchezdelbarrio, X Messeguer, R Rozas (2003). DnaSP, DNA polymorphism analyses by the coalescent and other methods. *Bioinformatics* 19:2496–2497

Sommer RS, JM Fahlke, U SchmÖLcke, N Benecke, FE Zachos (2009). Quaternary history of the European roe deer *Capreolus capreolus*. *Mammal Rev* 39:1–16

Valli AMF (2010). Dispersion of the genus *Procapreolus* and the relationships between *Procapreolus cusanus* and the roe deer (Capreolus). *Quat Intl* 212:80–85

Vorobieva NV, DY Sherbakov, AS Druzhkova, R Stanyon, AA Tsybankov, SK Vasil'ev, AS Graphodatsky (2011). Genotyping of Capreolus pygargus fossil DNA from Denisova cave reveals phylogenetic relationships between ancient and modern populations. *PLoS One* 6:1–13

Wang YX (2003). *A Complete Check list of Mammal Species and Subspecies in China: A taxonomic and Geographic Reference*, p:127. China Forestry Publishing House, Beijing, China

Xiao CT, MH Zhang, Y Fu, HS Koh (2007). Mitochondrial DNA distinction of northeastern China roe deer, Siberian roe deer, and European roe deer, to clarify the taxonomic status of northeastern China roe deer. *Biochem Genet* 45:93–102

Xu L, C He, C Shen, T Jiang, L Shi, K Sun, J Feng (2010). Phylogeography and Population Genetic Structure of the Great Leaf-Nosed Bat (*Hipposideros armiger*) in China. *J Hered* 101:562–572

You Y, K Sun, L Xu, L Wang, T Jiang, S Liu, J Feng (2010). Pleistocene glacial cycle effects on the phylogeography of the Chinese endemic bat species, *Myotis davidii*. *BMC Evol Biol* 10; Article 208

Zhang YP, LM Shi (1992). Mitochondrial DNA polymorphisms in animals: A review. *Cell Res* 13:289–298