Population Genetic Analysis of Red Foxes (*Vulpes vulpes*) in Hedmark Country, Norway

- A Pilot Study

Aarathi Manivannan

Master Degree in Applied & Commercial Biotechnology

HEDMARK UNIVERSITY COLLEGE

Department of Natural Sciences and Technology

2013

Acknowledgement

Coming to the end of my journey through the master program in Applied and Commercial Biotechnology, I would like to express my gratitude to all those who have encouraged me through it. A special thanks to my parents, S. Manivannan and M. Revathi who trusted me and gave me this opportunity to study in a privileged university, Hedmark University College. Thank you Appa for giving me a priceless gift “my studies” no matter whatever obstacles I came across. Thank you Amma for taking care of me no matter the distance between us.

This achievement could not be possible without the guidance of my supervisor Robert C. Wilson. I am grateful for your never-ending support, excellent and enthusiastic supervision during this project. I would also like to thank my Associate Professor, Arne Linløkken for guiding me through the statistical part of my project.

I dedicate this thesis to my fiancé, Aravindh Suryamoorthy. Thank you for your support, assistance and motivation in completing this thesis. Love you!

Last but not the least, special thanks to all my friends in Norway and India who proved that “wherever you are, friends will never let you down”. Thank you for being with me at every possible situation!

Hamar, May 31st 2013

Aarathi Manivannan

Table of Contents

Abstract

Ecological and environmental studies commonly depended on the knowledge of genetic variation among individuals within a population and most of the population studies were carried out using genetic markers. In this thesis, individuals of *Vulpes vulpes* species were genotyped using microsatellite markers to study their population diversity. DNAs were isolated successfully from the muscle tissues of 33 individuals and in some cases, from the hair and scat of the foxes. Twenty microsatellite primers with 1 sex chromosome marker were successfully designed using MP primer downloaded program out of which 16 primers amplified the DNA template samples successfully. The Amplicons were successfully genotyped and checked for genotyping errors. The LOSITAN program identified two of the markers to be under selection and the Micro-Checker program adjusted the alleles, which were identified as 'null' alleles. The analysis was performed on three different types of data: original genotype data, adjusted genotype data with the null allele markers and genotype data without the null allele markers. The marker under balanced selection was also added to verify its effect on the study. The STRUCTURE program assembled the individuals into different clusters. The Arlequin program estimated that the individuals grouped using original genotype data proved genetically variable, whereas individuals grouped using adjusted null alleles did not show any significant variance. Original genotype data looks promising as the loci were not significantly in Linkage Disequilibrium but this can be justified by increasing the number of individuals for the study. Analysis using genotypic data without null allele markers did not vary from the analysis using original genotypic data. Addition of marker under balanced selection showed significant Linkage disequilibrium in the study of all the three types of data. Relationship between the individuals showed that many individuals were half- siblings and 4 of them were full siblings.

1. Introduction

1.1. Fox

Fox is a common name for many species of omnivorous mammals belonging to the Canidae family along with 37 other species. Among these only 12 species under the genus *Vulpes* are considered as "true foxes". Table 1 illustrates the genera of different types of foxes.

Table 1 Genera of the foxes

Fox members	**Genus**	**Population**
Artic fox	*Alopex*	Low
Semien fox	*Canis*	Low
Crab-eating fox	*Cerdocyon*	Moderate
Falkland Islands fox	*Dusicyon*	Rare (almost extinct)
Bat-eared fox	*Otocyon*	Fairly stable
Grey fox, Island fox and Cozumel fox	*Urocyon*	Low
Six South American species	*Lycalopex*	High
True foxes including red fox, silver fox and Tibetan fox	*Vulpes*	High

1.1.1. Physical characteristics of Foxes

Foxes are generally smaller than other members of the Canidae family like wolves, jackals, and domestic dogs. Foxes are characterized by their long narrow snout and bushy tail. But other physical characteristics they possess vary according to their habitat. For example the kit fox has large ears and tiny fur coat, whereas the Arctic fox has tiny ears and thick, insulating fur. Litter sizes depends on the species and environment – the Arctic fox, for example, has an average litter size ranging from four to five, with eleven as maximum (Hildebrand, 1952).

1.1.2. Distribution of Foxes

Foxes are usually cautious and thereby tend to hide at the sight of humans. Their dwelling place is usually a burrow underground. They can also live above ground in a cosy hollows. Sometimes, they are also sighted near an area where they feel secure or in areas close to cover. Wild foxes can live for up to 10 years, but most of the foxes die in 2 to 3 years due to road accidents, diseases or hunting, as they are commonly pursued for fur. Although the foxes are native to North America, Europe, Asia and North Africa, they were originally sighted in Australia. Foxes are primarily nocturnal but they are also often seen in urban areas during the day. Their eyes are highly adapted to night vision. The light sensitive cell that lies behind the tapetum lucidum reflects the light back through the eye which doubles the intensity of images received by the fox (http://www.onekind.org).

1.1.3. Diet of the Foxes

Foxes normally gather a wide variety of foods like snakes, scorpions, berries, fruit, fish, insects and all other kinds of small animals. Many species are generally predators, but some such as the crab-eating fox are more specialist predators. They are mostly opportunistic feeders that hunt live prey, especially rodents. Foxes are omnivores and mostly kill their prey using pouncing techniques. Normally the foxes do not chew their food; instead they use their carnassials to cut the meat into manageable chunks. Foxes consume about 1 kg of food per day. Foxes usually reserve their food by hiding the excess under leaves, snow or soil (Fedriani, et al., 2000).

1.2. *Vulpes* Species

Vulpes is a genus of the Canidae family whose members are usually referred to as 'true foxes'. True foxes are distinguished from other members of the genus *Canis* by their smaller body size and flatter skulls. They can be distinguished by their black, triangular markings between the eyes and nose, and the tips of their tails (Macdonald, 1984). The different species belonging to *Vulpes* are illustrated in table 2.

Table 2 Various Vulpes species and their distribution modified from http://en.wikipedia.org/wiki/Vulpes.

Species	Scientific Name	Distribution
Bengal fox	*Vulpes bengalensis*	Indian subcontinent,Nepal and eastern Pakistan
Blanford's fox	*Vulpes cana*	West Asia, central Iran
Cape fox	*Vulpes chama*	Kwazulu-natal (South Africa)
Corsa fox	*Vulpes corsac*	Central and northeast Asia, Mongolia, China, Kazakhstan, Uzbekistan, and Turkmenistan
Fennec fox	*Vulpes zerda*	North Africa
Kit fox	*Vulpes macrotis*	Mexican, southern California to western Colorado and western Texas, north into southern Oregon and Idaho.
Pale fox	*Vulpes pallida*	Semi-arid Sahelian region of Africa
Red fox (includes silver fox)	*Vulpes vulpes*	North America, Asia and Europe
Swift fox	*Vulpes xvelox*	Canada
Tibetan sand fox	*Vulpes ferrilata*	China, Pakistan,Tibet and border of Nepal and India
Artic fox	*Vulpes lagopus*	Artic Tundra region

1.2.1. Red Fox and its Distribution

The **red fox** (*Vulpes vulpes*) is the largest of the true foxes and has successfully adapted to a wide range of ecosystems (Barton and Zalewski, 2007). Among the true foxes, the red fox represents a more progressive form in the direction of carnivory. It is considered to be the most geographically distributed member of the Carnivora, spread across the entire Northern Hemisphere from the Arctic Circle to North Africa, Central America and Asia (Heptner and Naumov, 1998). Red foxes are assumed to have originated from Eurasia during the Middle

Villafranchian age and later colonised North America shortly after the Wisconsin glaciation. These are dominant over other fox species.

Normally, it was observed that when red foxes occupy an area other foxes like artic foxes tend migrate far away from them. Starting with the beginning of the twentieth century, they began to spread across Australian, European, Japanese and North American cities showing a holocratic distribution. The species were first sighted in British cities during the 1930s, and began to spread to Bristol and London during the 1940s and later establishing themselves in Cambridge and Norwich. Similarly in Australia, red foxes were found in Melbourne as early as the 1930s, while in other places like Zurich, Switzerland, they started to appear in the 1980s. These species are most commonly sighted in residential areas, but are rare to see in other places. It was reported that there are around 10,000 red foxes in London (Harris and Yalden, 2008). These species of foxes utilize a wide range of habitats, which ranges from temperate to terrestrial including woodland, desert, mountains, woodlands and city areas. They prefer mixed vegetation areas such as edge territories and mixed scrubland and forest. Typically they can be found anywhere from sea level to 4500 meters elevation across wide terrain (MacDonald and Reynolds, 2005).

1.2.2. Physical Characteristics of Red Foxes

Like other foxes, red foxes are endothermic, homoeothermic and possess bilateral symmetry. They are considered to be the largest of the *vulpes* species. The coat color of the red foxes vary from pale yellowish red to deep reddish brown on the dorsal side and are whitish on the ventral side of the body. Their body length can extend to 900 mm, head 455 mm and tail length ranges from 300 to 555 mm with the weight varying between 3 to 14 kg. They possess tail glands similar to other candid species but red foxes are located above the root of the tail on the dorsal side. The eye colour is normally yellow. The manus has 5 claws and the pes 4 claws where the first digit, or dewclaw, is rudimentary but clawed and does not contact the ground. The nose is normally dark brown or black and their teeth row is more than half the length of the skull, which emphasizes their feeding habit of crushing the meat (MacDonald and Reynolds, 2005). However, their body mass and length was observed to vary across populations with latitude and also between the male and female breeds. Normally the male foxes called reynards weigh on average 5.9 kilograms while female foxes called the vixens weigh less than the male, at around 5.2 kilograms (National Geographic; Animals; Mammals: Red Fox).

1.2.3. Diet of Red foxes

As other foxes, red foxes are also nocturnal and omnivores. Their diet varies from small, mouse-like rodents like mice, ground squirrels, hamsters and deer mice (Thompson and Chapman, 2003). They usually target animals up to 3 kg in weight, and require about half kg of food daily and also on very rare occasions; they may attack young or small mammals (Sillero-Zubiri et al., 2004). Red foxes can also readily feed on plant material and in some areas; they can live only on fruit especially in the autumn.

1.2.4. Reproduction

Unlike many canids, red foxes are solitary animals and do not form packs like other members of candis family do, like the wolves. A typical red fox family consists of an adult male and one or two adult females with their cubs. Regarding their mating behaviour, the males and females are often monogamous but they also follow polygamous and cooperative breeding system. Usually the male pairs up with non-breeding female helpers in raising their young. They mainly live in the dens of rodents or rabbits and dig big dens during winter, which is occupied by several generations. Red fox groups always have only one breeding male. However, the male does not necessarily breed within the pack (MacDonald and Reynolds, 2005). Once the cubs have grown and are able to find their own food, mother vixen chases them away to find their own territories thereby ensuring that any sort of local disaster cannot destroy their generation completely.

Figure 1: Skulk of foxes. Figure taken from by Matt Walker and Ella Davis (BBC Nature)

It has also been investigated and reported that there has been substantial gene pool mixing between different subspecies. For example, the British red foxes have crossbred extensively

with foxes imported from German, France, Belgium, Sardinia, and possibly Siberia and Scandinavia (Dale, 1906). European foxes introduced to some parts of the USA in the 18th century were crossbred with local North American populations. The eastern red foxes found in California may also be due interbreeding with local *Vulpes vulpes* necator populations for the development of fur-trade industries (Bryan, n.d., p. 514) .

1.2.5. Predation

The red fox is an efficient predator and has a long history of association with humans through hunting in large parts of its distribution (Heydon and Reynolds, 2000). The interest for fox hunting in Sweden has been low in the last 30 years, which probably has resulted in a historically high abundance of red foxes. Because of its widespread and large population, the red fox has become a key among the furbearing animals that were harvested for fur trade by young hunters in Norway, which led the foxes to extinction (Bachrach, 1953). The predation on the red foxes has been increasing over the recent years. Recent reports of the parasite, *Echinococcus multicolocularis* in the Scandinavian red foxes have put further focus on the possibilities to control, track and regulate the red fox population in order to maintain ecological diversity.

Echinococcus multicularis is a zoonotic parasite, involving primary and secondary hosts to complete their life cycle. Canids act as the primary host and wild rodents as secondary hosts for the parasite. Normally, the parasite attaches to the intestinal mucosa of the canids and produce hundreds of microscopic eggs that are later dispersed through the faeces. Wild rodents like mice consume the eggs. The eggs develop multilocular cysts in the liver, lungs and other organs of the rodent. When the canids like foxes or wolves consume the infected rodents, the larvae from the cysts develop into adult in the intestinal tract. The parasites from the faeces or from rodents spread to humans and causes severe trauma and liver cancer in them (Hegglin, et.al., 2007). Therefore, the foxes were tested for the parasite and the parasite bearing foxes were shot to prevent the spread of the parasite. The first parasite-infected fox was found in Denmark and then later it prevailed in Sweden. The annual report submitted by Davidson, Øine and Norström (2009) on E*chinococcus multilocularis* in red foxes (*Vulpes vulpes*) in Norway estimated low prevalence of the parasite among the red foxes.

1.3. DNA markers and Genotyping

Deoxyribonucleic acid known as DNA is the genetic material in every living organism. The genetic information in the DNA molecule is encoded of the nucleotides guanine, adenine, thymine and cytosine. These make sequences that encode for amino acids. The DNA has a double stranded helix structure. The backbone of the helix is made of deoxyribose sugar and phosphate groups. Nucleotides are attached to the backbone with hydrogen bindings. DNA is tightly packed into chromosomes in the nucleus of the cell. A functional segment of DNA is called a gene. In a gene the sequence of nucleotides vary in repeats, insertions/deletions and transitions/transversions from an individual to another. This leads to genetic variation in the same species (Eenennaam,2009).

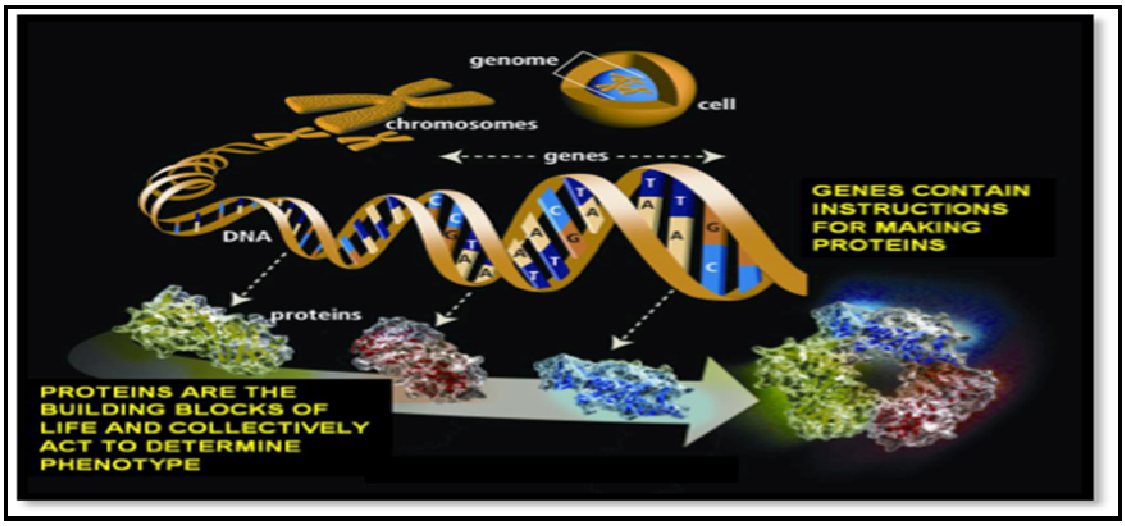


Figure 2: from genomic DNA to protein cluster. Figure from Eenennaam, 2009

The genetic variation has led to the study of effective population size, population history that includes migration and recent expansion, population structure and various genetic diseases (Mburu and Hannote, 2005).

Genotyping is a process of analysing the variation of an individual's genotype by examining the DNA sequence from its biological assay and also by comparing with it with other individual's sequences. This term is generally used to describe a method that determines the DNA-marker alleles that an individual carries at a particular genomic locus. This makes it easy to identify the inherited properties of the individual within its family and also how they differ between other species. Many alleles come in dominant and recessive forms, but there may also be many more ways that these alleles express a specific trait which lead to

hereditary variations. In other words, genotype is the molecular code of the DNA that can also lead to the study of phenotype of an organism (Eenennaam, 2009).

Several studies have been successful in identifying the DNA regions that enhance the production traits. Tests have been successfully developed to identify whether an animal carries a particular trait of interest with the help of several type of genetic markers.

Mitochondrial DNA (mtDNA) is a small circular molecule that comprises of approximately 37 genes coding for 22 tRNAs, 2 rRNAs and 13 mRNAs within the cytochrome b coding region that can be used in phylogenetic work. Phylogenetic studies can also be conducted in the non-coding region, where the displacement loop (D-loop) controls the mtDNA expression. mtDNA polymorphisms are widely used to determine the structure of population, variability among species, and evolutionary relationships (Mburu and Hannote, 2005).

RFLPs (Restriction Fragment Length Polymorphism) are based on the patterns derived from a DNA sequence digested by a known restriction enzyme. The length of the digested fragments differs as a result of point mutation, which can be created/destroyed at the restriction site or by insertion/deletion that can alter the length of restriction fragment. RFLPs are widely used as co-dominant markers. Only high quality of DNA should be used. Although RFLPs are suited for phylogenetic studies, due to the tedious methodology the use of RFLPs as markers become less reliable (Nagaoka and Ogihara, 1997).

AFLPs (Amplification Fragment Length Polymorphism) differed from RFLPs with detection of presence or absence of restriction fragments and not by the size of fragments. Normally AFLP primers are highly specific to the targeted restriction site of whole digested genome. As these markers are dominant markers they could only be used to estimate genetic variation for example in DNA finger printing and cannot be used in population genetic studies (Mueller and Wolfenbarger, 1999).

RAPD (Random Amplified Polymorphic DNA) are short arbitrary primers that can bind to any place in a sequence without any prior knowledge of the DNA. The target product is amplified and studied. In case of mutations, mismatches between the target and primer may not result in PCR product. Moreover these are dominant markers and detection of polymorphisms is limited. In general, the use of dominant markers like RAPD for

identifying heterozygotes will require twice the amount of co-dominant markers (Williams et.al, 1990).

Allozymes and Isozymes are variants of the same enzyme. The difference between them is that; allozymes are encoded by different alleles of a same gene whereas isozymes are encoded by different alleles from different genes. As they are product of gene duplication, if one of the variant passed through the generation other can be lost due to mutation; thereby leading to genetic variation. They are commonly used as molecular markers. These molecular markers are used for population genetic analysis and studied with the help of gel electrophoresis based on enzymatic electric charge. As most of the enzymes are invariant in a population, the use of allozymes or isozymes as molecular markers for genetic studies have reduced (Ulrich and LaReesa, 1999).

SNPs are single nucleotide polymorphisms, which include a single base change in the genomic sequence. SNPs are normally used as bi-allelic co-dominant genetic markers, as they reveal polymorphisms at the DNA level. Genotyping with SNPs depends on the comparison of locus-specific sequences arising from different chromosomes. SNP markers need locus-specific primers and are also limited in scope for heterozygotes when it comes to differentiating sequencing artefacts in double peaks making SNPs expensive to use in research (Alain et al., 2002).

Microsatellites are stretches of DNA consisting of short tandem repeats of nucleotides (usually 1 to 5bp). The tandem repeats varies among every individual such that no individual can have the same number of repeats. Microsatellites have been used commonly in forensics, disease diagnosis, population genetic studies, in conservation biology and linkage analysis due to their length polymorphism and abundance in all genomes. Rodrigues and Kumar (2006) used the major criteria of the microsatellites i.e. genetic variability to assess the significant genetic factors of *Paphiopedilum rothschildianum* orchid from the Sabah region that keeps it from becoming endangered. They illustrated 43 microsatellite loci from the orchid to determine the intra and interspecific genetic variations.

>gi|133872848|gb|EF462877.1|*Paphiopedilum rothschildianum* clone PRA12 microsatellite sequence

CACACACACACACACACAGCCAAATAGTCGCCCAGAGGGACACGGCGTTGG
GGTGCTTAACACGCACCCCCTCGCACACCCGCGTGCACACACGGGTCTCCGT
GGGCACTGCGTACACACGAACCACAGACACTAGCACCTGTGTGCCTGCATTC
ACCCACACAAACCGCCCTGCTTTCGGCCCTTGCCTCTCTCCGACACCTGCAG
CCTGGCAGACCCCGTGCCGTAGAGGTGACACAGAAACACACACTGGGTGAG
TTAACAGTCGGGGCTGGCATTTGGCACCACACTCTGGATAGACGGGAACAG
CCTGCACGCATGCACCTCCGTGTCTCTGCATCGTGTGTGTGTGTGTGTGTGT
GTTCGCTCA

Figure 3: Microsatellite sequence showing where locus-specific primers may be designed (Rodrigues and Kumar, 2006).

1.3.1. Application of Microsatellite markers

Genetic studies using microsatellites have been confirmed by the parentage assignment in aquaculture. In this study, individual *Haliotis asinine* was genotyped by analysing polymorphic microsatellites. 5 polymorphic loci were assayed to identify the parents produced in 3 different crosses, which was achieved by matching the alleles of a single locus. 96% of the parentage assignment was successful. In half-sib family crosses, only one locus was confirmed to determine the parentage. This study also concluded that microsatellites could be successfully used as genetic tags in breeding and enhancement programs to ensure the successful maintenance of genetic diversity (Maria et al., 2001).

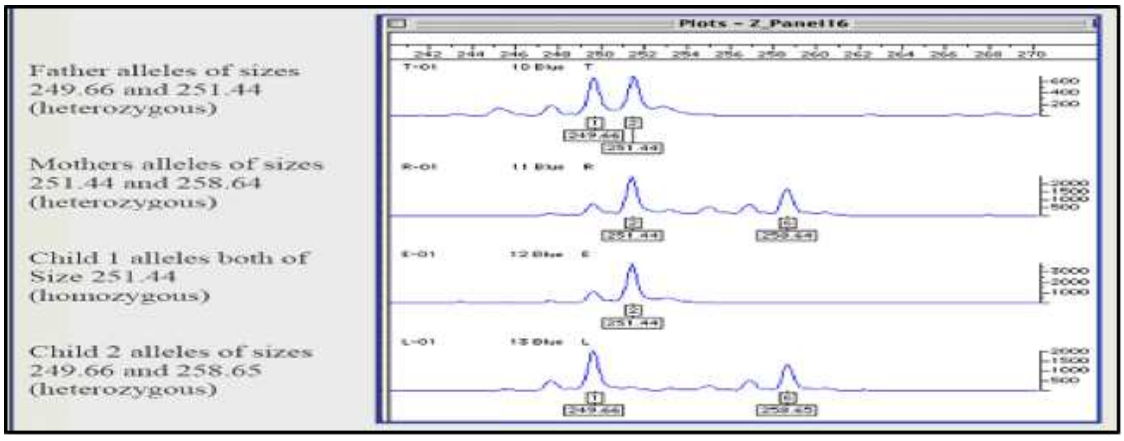


Figure 4: Raw genotyping data (Rosalind)

The genetic structure of domestic species, which aimed at the process of domestication through their genetic study analysis, was performed in Tunisia on rabbits with the help of microsatellite markers. Ben (2012) studied the first detailed analysis of genetic diversity of the Tunisian rabbit populations. In this study fifteen rabbit populations from villages of

Tozur and Gafsa were analysed using 36 microsatellite markers, out of which 294 were genotyped. The genetic differentiation among the population implies that 98.9 % of the total genetic variation was explained by individual variability with heterozygosity ranging from 0.3 to 0.53. This analysis not only helped in conservation of the population but also enabled identification of the loci involved in the economically important traits. The re-colonization process was studied by using the population genetics approach in wolves. The wolves from Alps and Apennines were genotyped using 12 microsatellite loci and the study concluded that the wolves from Alps showed lower genetic diversity compared to those from the Apennines (Fabbiri, 2007).

Another study was conducted on dogs to validate the efficiency of the microsatellite markers. Dogs are normally differentiated by their phenotypic traits such as size, shape, and coat colour, behaviour. Apart from these characteristics, 28 breeds of dogs were analysed to assess the genetic variation using 100 microsatellite markers. The resulting breed-specific allele frequencies were then used to interpret the genetic distances between the breeds. These results also concluded that the heterozygosis tended to decrease as the population sizes decreased (Irion, 2002).

As microsatellites markers are present on the sex chromosomes, they are also utilised for sex determination. The consistency of the markers to differentiate X and Y-chromosomes even in the immature individuals lead to the reliability of microsatellite markers in sex differentiation (Takahito et al., 2011).

1.3.2. Sex Determination

A fundamental process in most species concerns the sexual phenotype, which determines the future of an individual at its embryonic period. Sex determination normally extends the knowledge of genetics and its mechanism involved in two alternate embryonic routes: male or female phenotype. Sex identification is a very important criterion in evolution, environmentalism, prenatal diagnosis, forensic identification, and population genetics and also in the studies concerning endangered species. Sex determination is divided into two broad categories: Environmental-dependent Sex Determination (ESD) and Genetic Sex Determination (GSD). In ESD, sex determination occurs in response to external signals after fertilization, which means that the sex into which the zygote differentiates is independent of their genetic chromosomal composition (Garzon, Camacho, and Sanchez, 2012).

Sex determination in mammals mainly depends on chromosomal constitution of gametes, ranging from homomorphism chromosomes to heteromorphy chromosomes. Sex determination is made by the sex chromosome system XX for female and XY for male, where the Y chromosome is the dominant factor inducing male phenotype development. In mammalian genome, the Y chromosome contains highly preserved set of genes that defines the patrilineages. The presence of *SRY* (sex determination gene) on the short arm of the Y-chromosomes enables sex determination. As Y specific SNPs have proved to have slow mutation rate, use of Y-specific microsatellites have been considered. *Sry* gene has also been used as a phylogenetic marker directly (Mburu and Hannote, 2005).

The *Sry* gene belongs to the *Sox* gene family consisting of a protein that has greater than 62% identity in the high mobility group as *Sry*. There are more than 20 known *Sox* genes in mammals, although most of them are not involved in sex determination. The *Sox*9 gene is observed to be involved in sex determination as its expression was analysed to be specific to the male gonads of birds and mammals, with no expression in females. *Sox* gene is usually considered to be regulated by the *Sry* gene since it allows the development of the masculinity in the absence of *Sry* gene in addition to imitating the latter's function (Canning and Lovell Badge, 2002).

The functional analysis of *Sox*8 and *Sox*9 during sex determination was determined using specific knockouts in mice. Sex-determination in mammals is usually centred on the differentiation of the bi-potential gonads into testis and ovaries. Triggering of the sex-determination gene *Sry* lead to the activation of Sox9. In this experiment, Sex determination was analysed using conditional gene targeting followed by heterozygous deletion of Sox9 in XY gonads lead to the higher suppression of Mis (Mullerian-inhibiting substance, which has been suggested as direct target of *Sox*9) and Sox 8 than in XX tissues. Homozygous deletion of Sox9 in XY gonads exhibited interference with the sex cord development, with no signs of sex cords and inactivation of *Sry*. Moreover double knockout suggested that the Sox8 emphasizes Sox9 function in testis differentiation of mice (Marrie, 2004). *Sry* expression was also proved to direct the undifferentiated gonads to differentiate as Sertolli cells, which in turn proved that initiating the testis differentiation pathway in its absence or presence in decreased levels lead to ovarian development pathway (Sergei, 2002). As *Sry* gene test is observed to be problematic in distinguishing the failure of PCR amplification from the female samples, in some mammals, an alternative and dependable method appears

to be the zinc finger proteins/ Amylogenin genes method of sex determination (Takabayashi and Katoh, 2011)

Zinc finger domains are predicted to be the amino acid sequence of an open reading frame that can be encoded by nucleic acid sequence of a Y chromosome genomic fragment. ZFY is the zinc –finger protein encoded by the Y chromosome and ZFX by X chromosome (Mark, 1989). In Japan, sex identification was experimented using ZFX and ZFY genes on marmosets (monkey species) with PCR restriction fragment length polymorphism (RFLP). A fresh primer set was designed to detect ZFX and ZFY. As the fragment length of ZFX (483bp) and ZFY (471bp) was not clearly distinguishable on agarose gel electrophoresis, they were spliced and amplified. Sequencing data of the amplified products from ZFX and ZFY disclosed the recognition sites of Dde1 and Mse1 restriction enzymes, respectively. Further when the products were digested using each enzyme, the different band patterns of the female and male resulted in sex differentiation. Although sex differentiation using ZFX/ZFY loci with RFLP method proved to be sufficient, normal PCR method failed to work with ZFX/ZFY (Takabayashi and Katoh, 2011).

Amelogenin is a major protein (belongs to the extracellular matrix proteins family) that forms an outer protective layer of the tooth (enamel). AMELX/AMELY genes are homologous sequences that are found in the sex chromosomes of mammals. The advantages of using this technique are that amplification of both the amologenins, AMELX/AMELY homologous sequences can done in a single reaction. Tsai et al. (2011) conducted a study on sex differentiation in goat for commercial purposes using this. They first cloned and determined the intron sequences of the goat AMELX/AMELY genes from female and male ear tissues. Their results from the PCR based RFLP and Southern blot hybridization, illustrated that the goat's AMELY gene has more deletions/insertions compared to AMELX gene and also that the intron 5 only shares 48% of the goat's AMELX gene. Moreover, when they amplified the fragments of the X and Y-chromosomes isolated from the agarose gel on a PCR using sex-specific triplex primers, they were able to differentiate the sex with a single blastomere at the blastula stage itself. This method is widely used as they can ease the problems of contamination of the PCR reaction and misdiagnosis, which was detected in the case of SRY gene. Amelogenin sex differentiation methods are also preferred to other PCR-based embryo sexing methods, like ZFY/ZFX gene detection due to its feasibility, sensitivity, accuracy, and reliability and also it can detect even with 0.5ng of template DNA (Shujin et al., 2012).

1.4. Advantages of microsatellites

The first and foremost advantage of using microsatellites for various studies is that it requires very less quantities of template DNA (10-100ng). Availability of microsatellites throughout the genomic DNA and its randomness in occurrence adds a positive criterion. Remarkable level of polymorphism is another considerable feature. Band profiles can be interpreted in terms of loci and alleles co-dominant markers. And the size of allele can be determined to high degree of precision. Microsatellites are well suited to be used in multiplex PCR. Microsatellites are used in number of applications like forensic, diagnosis and identification of human diseases and conservation biology to name a few. And finally the process can be fully automated.

The reliability of the microsatellites in the application of genotyping was assessed by comparing microsatellite genotypes from fox faecal samples using genetic and GIS analyses. The kit fox scat data set were examined for genotyping errors comparing both genetic analysis and Geographic Information system (GIS) analysis, thereby evaluating and concluding that genetic tests indicated low genotyping errors as they were found to be similar with GIS analysis (Smith et al, 2008).

1.5. Microsatellite marker/ primer design

With the help of PRIMER3 software Moore et al., (2010) redesigned 31 microsatellite markers and a sex marker from the existing dog flanking sequences in order to genotype native red foxes in California and Sweden. They also used the MULTIPLX program to group the primers into 5 multiplexes. Similarly, Wandeler and Funk (2005) also established seven microsatellite markers in 3 panels to characterize red foxes. Part of the present study is to use these established markers, redesign the primers and reduce the number of multiplex panels needed to genotype red foxes.

The main goal in designing good primers is to produce efficient and sufficient amplification. Quality of the primers is proportional to the success rate of amplification. The length of the primer is linked with the specificity of the amplification. Primers with reasonable GC content are designed. Care is taken so that primers are not complimentary especially at 3’ prime end (Dieffenbach, Lowe and Dveksler, 2010).

Table 3 Selected markers from Moore et al. (2010) (highlighted region) and Wandeler and Funk (2005).

Loci	Dye	Number of alleles	H_o	H_e
AHT121	VIC	9	0.81	0.82
AHT133	NED	5	0.46	0.65
AHT137	NED	8	0.6	0.76
CPH7	FAM	4	0.35	0.40
FH2054	NED	7	0.82	0.75
FH2328	VIC	9	0.84	0.83
FH2848	FAM	6	0.71	0.74
REN105LO3	FAM	9	0.73	0.84
REN162CO4	VIC	8	0.63	0.75
REN169018	PET	6	0.76	0.79
REN247M23	FAM	7	0.59	0.64
K9AMELO	FAM			
AHT142		10	0.85	0.80
CXX374		6	0.88	0.89
CXX468		7	0.96	0.92

H_o, observed heterozygosity; H_e, expected heterozygosity.

1.6. Common techniques used for genotyping

Genotyping is achieved by many techniques. Restriction fragment length polymorphism (RFLP), random amplified polymorphism (RAPD), amplified fragment length polymorphism (AFLPD), polymerase chain reaction (PCR), single-nucleotide

polymorphism (SNP) are some of the most powerful and efficient techniques used for generating large numbers of anonymous DNA markers for both plants and animals.

SNP genotyping method comprising of the melting curve analysis (MCA) lead to the study of allele's discrimination and also detection of DNA fragments in the presence of dsDNA specific fluorescent dye SYBR Green. As the MCA depends on the melting temperature differences between DNA fragments with different sequences and length, it is suitable for genotyping (Figure 5). This technique is proved to be an accurate means of genotyping the SNP variations but this method is expensive and requires previous knowledge of the sequence (Akey et al., 2001).

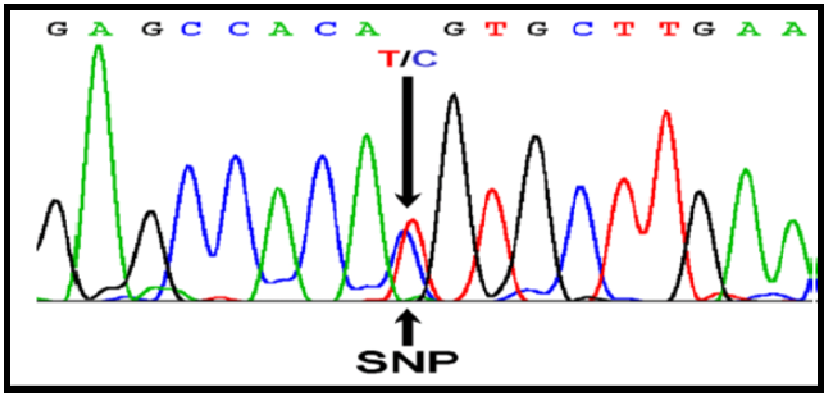


Figure 5: Interpretation of SNP genotyping (Alison, 2009)

AFLPs (Amplified Fragment Length Polymorphisms) have been proved to produce highly replicable markers for genotyping. AFLPs are dominant, multilocus and PCR-based markers commonly used to analyse the genetic diversity of organisms. It can access random genetic deviations among groups of individuals and also lineages that were developed independently. The main advantage of AFLP-PCR is that it can simultaneously screen different DNA regions that are scattered in a genome. AFLP-PCR techniques are widely used for DNA fingerprinting to identify individuals and parentage. AFLP-PCR is time consuming and an expensive technique as well (Ulrich and LaReesa, 1999).

Multiplex PCR is a very powerful and widely used genotyping technique that enables amplification of two or more products in a single reaction. It simultaneously amplifies multiple regions of a DNA template or multiple DNA templates using more than one primer set comprising of forward and reverse primer in single tube (Shen et al, 2010). It is also used for qualitative and semi –quantitative gene expression analysis, DNA testing in research, forensic and diagnostic laboratories on both eukaryotes as well as prokaryotic

sources (Qiagen Multiplex PCR handbook, October 2010). A fluorescent primer based PCR has the ability to genotype animal population carrying normal or mutant alleles.

1.6.1. Applications of Multiplex PCR

- In Animals/human: Analysis of satellite DNA, genotyping of transgenic animals, lineage analysis, detection of pathogens, diet analysis, sex determination, mutation detection and qualitative and semi quantitative analysis of gene expression.
- In Plants: GMO analysis, analysis of satellite DNA, lineage analysis, pathogen detection, qualitative and semi quantitative analysis of gene expression and genotyping of transgenic plants.
- In Bacteria/viruses: Hygiene analysis, diagnostics/pathogen detection and qualitative and semi quantitative analysis of gene expression.

1.7. PCR Inhibitors

Many scientists have faced PCR inhibitors as an obstacle for successful PCR amplification. These inhibitors either directly interact with DNA or with DNA polymerase thereby preventing the amplification completely or reduced product yield. The magnesium that acts as a critical cofactor for the DNA polymerase is found to be the major target for the inhibitors to act on. These inhibitors normally bind to or reduce the magnesium, and thus inhibit the PCR. Commonly these inhibitors are found in blood, fabrics, tissues and soil sources as a result of excess of potassium chloride, sodium chloride and other salts. They also occur in the presence of iso-detergents, phenol, and ethanol and other alcohols (Katcher, 1994).

1.7.1. Methods to overcome Inhibition

The best way to avoid PCR inhibition is by preventing it from being processed along with the samples. Prevention is almost impossible while handling samples like blood or other tissue material. Scientists have approached prevention of PCR inhibition or failure using the following steps:

- Swab-transfer method for sample collection has been proved to avoid inhibitor inhibitor containing materials rather than processing spliced materials. DNA purification methods were performed with specialised kits like the QIAamp DNA Stool Mini kit. This kit contains Inhibitex tablets, which has the property to adsorb the inhibitors in the

purification process while extracting the DNA from the stool. When the tablet adsorbs the inhibitors, ASL buffer is used to remove them. Bacteria and other pathogens are lysed or killed by incubating the stool homogenate at 70°C (QIAamp® DNA Stool Handbook).

- Increasing the DNA polymerase amount in the reaction and also by the use of additives like Bovine Serum Albumin (BSA). BSA is found to provide resistance against inhibitors in blood (Bessetti, 2007).
- Use of Internal Positive Control (IPC) in the multiplex real-time PCR helps to detect the inhibitors by analysing target amplification efficiency (Bessetti, 2007).

From the above references, genotyping was carried out on the DNAs isolated from the muscle tissues, hair and scat of the organisms using multiplex PCR. As the hair and scat can be collected without capturing the animals, these methods show great promise in estimating the population data (Creel et.al, 2003).

1.8. 16-capillary 3130*xl* Genetic Analyzer (Applied Biosystems)

Applied Biosystems has a longterm business to provide instruments, reagents and software for life science department. The custom of initiating and innovation in the genetic field of analysis continues with the introduction of 3130xl Genetic Analyzers. The 16-capillary 3130*xl* Genetic Analyzer has proved to give all the benefits of the Applied Bio-systems suite of fluorescence-based capillary electrophoresis (CE) systems. The versatile 3130xl system delivers significantly high data quality, faster turnaround times, automation and reliability over a range of sequencing, re-sequencing (mutational profiling), genotyping and fragment analysis applications.

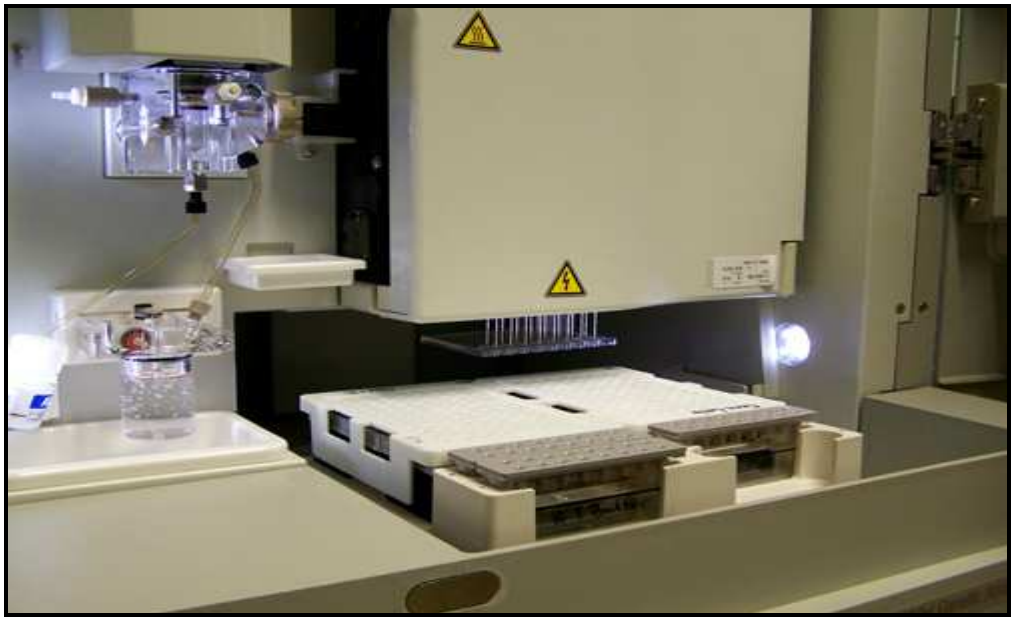


Figure 6: 16-capillary 3130xl Genetic Analyzer (http://www.uidaho.edu).

The system also comprises of many software such as Sequencing analysis software, SeqScape® and Gene Mapper. The Sequencing analysis provides, analyze and display sequencing data, SeqScape® is used for mutation detection and profiling. The GeneMapper is an ideal tool comprising multiple features for genotyping, allele calling, fragment sizing and SNP analysis. Quality Values (QV) is assigned to each fragment analyzed for easy automation and throughput, and questionable fragments are easily identified. Applications include microsatellite analysis (diploid and polyploid), linkage mapping, SNP analysis, AFLP, relative fluorescent quantitation including loss of heterozygosity, and conformational sizing.

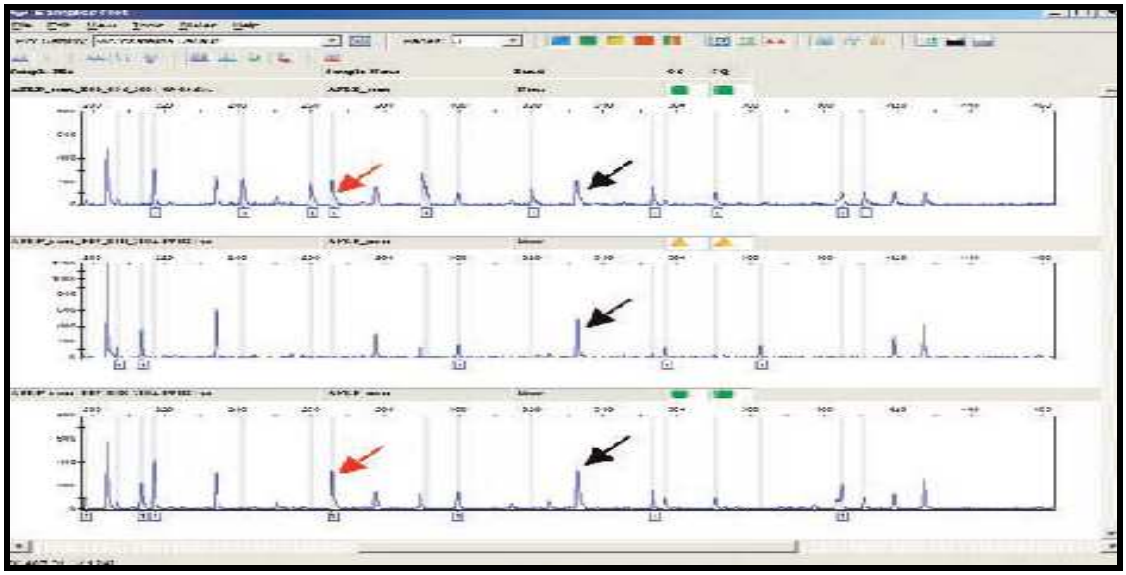


Figure 7: A picture of genotype analysis of data

The genotyping data were used to analyse and study genome mapping, population genetics. Genotyping data with neutral markers led to the determination of underlying mutation, which is proportional to the polymorphism of neutral markers. The rate of mutation helps in estimating the genetic distance and is also observed that it leads to direct transmission of alleles from the parent to offspring. The data is further analysed to determine the linkage and lineage among the individuals. Genotyping data involving microsatellite markers are also used in identification of paternity, forensic studies and linkage disequilibrium mapping studies (Ellegren, 2004).

2. Aim of Study

This study was designed to establish a microsatellite-based genotyping pipeline to determine genetic structures among red foxes in Scandinavia. In order to achieve this, DNA extraction protocols had to be tested on different sample types like tissue, hair and scat. New microsatellite primers had to be designed for the existing markers and redesigned to optimize the panel composition, thereby reducing the number of multiplex panels needed. Finally, the established pipeline had to be tested on capillary electrophoresis for 33 individual fox samples in order to check for marker suitability in population genetic studies. Based on the microsatellite analysis, the genetic distances within the individuals had to be analysed exploiting different genotyping models.

3. Materials and Methods

3.1. Scheme of the study

DNAs were extracted from each tissue, hair and scat samples and quantified in a Nano-drop to check the efficiency of the DNA to perform the analysis on it. Primers were designed, amplified using a single sample (positive control) and tested on an agarose gel. Failed primers were further modified to amplify the DNA template and labelled primers were ordered. Depending upon the sizes and colour of the labelled primers, they were grouped into multiplexes and genotyped in five-coloured laser induced fluorescence capillary electrophoresis system. The result obtained is further studied to analyze the sex, population diversity, expected and observed heterozygosity among the individuals.

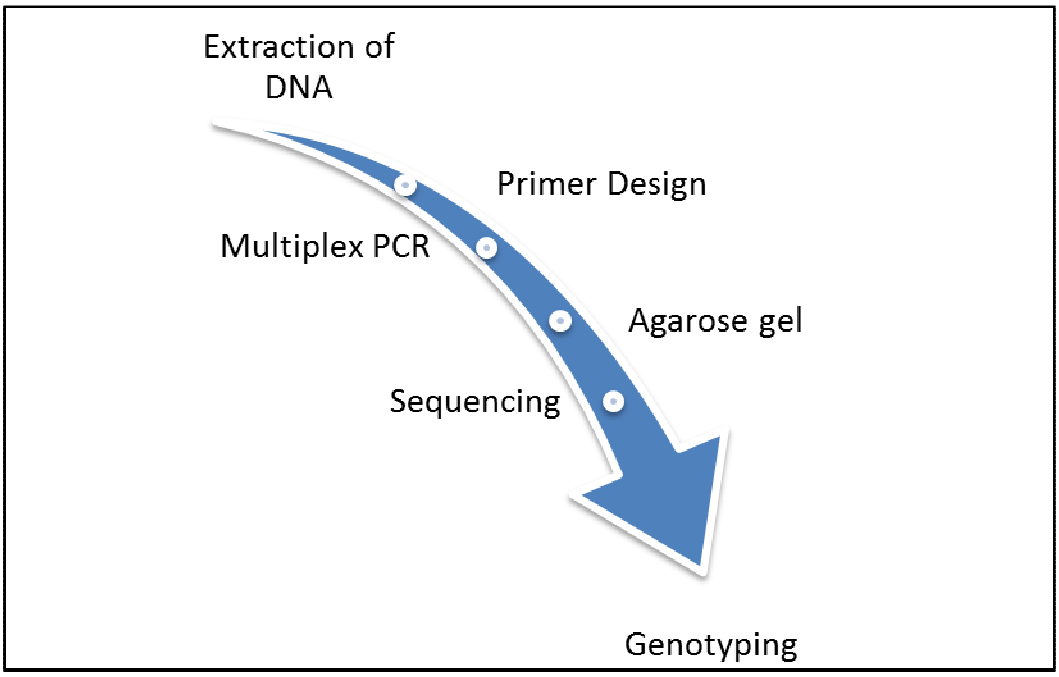


Figure: 8 Schematic of the workflow in this study

3.2. Extraction of DNA

3.2.1. From tissue

A small piece of tissue about <10 mm^3 was sliced and added to 500 µl of set buffer. 13 µl of 20 % SDS solution and 7.5 µl of proteinase K (20 mg/ml) were added to the samples and mixed well. The samples were incubated overnight at 55 °C with regular intervals of mixing. Next day 50 µl of 5 M NaCl (pH=8) was added to the sample mixed well and spun down. 500 µl of phenol chloroform was added and mixed with the help of vortex to form a homogenous mixture and was left to rest in a fume hood. After 60 minute, the mixture was

centrifuged at 11,000 rpm for 15 min. The supernatant was moved to new tubes discarding the top layer of the protein. 50µl of 3M NaOAc (pH 8) was added to the supernatant, to which twice its volume 99 % ethanol was also added and mixed well. The tubes were left in the freezer at -20 °C for 1 hour and again centrifuged at 11,000 rpm for 10 min. Pure DNA pellet was located, dissolved in 200 µl of TE buffer, quantified and stored in the refrigerator, whereas the supernatant was discarded (personal communication, Mikael Åkesson, Grimsö Wildlife Research Station, Swedish University of Agricultural Sciences).

3.2.2. From Hair

A number of hairs containing intact roots were separated from the bunch of hair cut from different foxes. The root part of the hair was only placed in 100 µl of lysis buffer along with 1.5 µl of proteinase K (20 mg/ml). The samples were mixed well and incubated at 56 °C with regular intervals of mixing for 3 hours. After incubation, the samples were spun down for 10 minutes in a centrifuge at 10,000 rpm. The supernatant was moved into new tubes with 10 µl of NaOAc and 220 µl of 99 % ethanol. The samples were mixed well and spun down at 10,000 rpm for 5 minutes. The ethanol was removed and the pellet was kept to dry in a fume-hood overnight. Next day the pellet was dissolved in 200 µl of TE buffer, quantified and stored in the refrigerator (personal communication, Mikael Åkesson, Grimsö Wildlife Research Station, Swedish University of Agricultural Sciences).

3.2.3. From Scat

The DNA was isolated from the scat samples using Qiagen kit. Samples were scraped from the outer layer of the scat sample carefully avoiding the bones and hair present in it. About 180-220 mg of the scat was taken in a 2 ml micro-centrifuge tube and placed on ice. 1.6 ml of ASL buffer was added to the sample and vortexed for 1 min to produce homogenized mixture. The sample was spun down at full speed for 1 min in a centrifuge tube and the pellet was discarded. 1 Inhibitex tablet was added to each of the sample and vortexed and allowed to rest for 1 min for the inhibitors to get adsorbed to the inhibitEX matrix. After vortexing, the sample was centrifuged for 6 min for the inhibitors and the pellet scat particles to bind to the inhibitEX matrix. The supernatant was transferred to a new tube containing 25 µl proteinase K and 600 µl of Buffer AL and mixed well. The lysate was later incubated at 70 °C for 10 minute followed by 56 °C for 10 min. After 20 min 600 µl of 96-100 % ethanol was added to the lysate. 600 µl of the lysate was applied to QIAamp spin

column carefully and spun down at full speed for 1 min in a centrifuge. The tubes containing the filtrate were discarded and the same process was repeated for the remaining lysate. The QIAamp spin column was carefully placed in a new container and wash buffer AW1 was applied to the spin column and centrifuged at full speed for 1 min. The filtrate was discarded and the spin column was washed again with wash buffer AW2 for 3 min in a centrifuge and the centrifugation was repeated by replacing the spin column at 180° to the previous position in order to get rid of ethanol. The QIAamp spin column was transferred into a new micro-centrifuge tube and centrifuged with 100 µl of Buffer AE for 1 min at full speed. The DNA was refrigerated.

The isolated genomic DNAs were quantified and analysed using nanodrop. Depending on the absorbance values the quality of the DNA was analysed and used for this study.

3.3. Multiplex Primer Design

Primers used for multiplex PCR were designed using the software MPprimer and the Vector NTI program. The MPprimer program also provided a scoring matrix for multiplex evaluation on Linux (Ubunto v11.1 to v12.04). It is a valuable tool for designing specific, non-dimerizing primer set constrained amplicon size for multiplex PCR assay (Shen et al., 2010). Considering the existing primers used by Moore et al. (2010) and Wandeler and Funk (2005) intensified the number of primers. The sex marker K9AMELO was used as such. Primers have some generalised characteristics like they usually match to DNA 18-30 base pairs, which is unordered. Melting temperature ranges between 50 and 65 °C, with a very little difference in temperature of the order of 3 °C with a total percentage of Guanine (G) and cytosine (C) nucleotide (GC) content between 40-60 %. Primers (forward primers) were designed, one with unlabeled sequence and second one with M13 labeled sequence,

A commercial service provider, Invitrogen (Carlsbad CA, USA) synthesized all primers that were used in this study. A BLAST search was conducted on primers to ensure that the primers were specific to the target region. Further modifications of the primers depended on the Agarose gel electrophoresis results post PCR amplification using unlabeled primers.

The labeled primer sets were then grouped into their corresponding multiplexes depends on their colours and sizes, without any overlapping.

3.4. Multiplex PCR

Various PCR protocols were used to confirm that the designed primers were actually amplifying the desired target region of the STR gene sequence that is used for genotyping the red foxes. To begin with, the PCR reaction was started with dissimilar unlabeled primers as simplex on a solitary genomic DNA section, RR1025 (positive).

For one reaction, 1.5 µl of 10x B1 buffer, 1.5 µl of 2.5 mM $MgCl_2$, 0.3 µl of 0.2 mM dnTPs, 0.15 µl of 100xBSA, 0.2 µl of Hot fire polymerase were used. 2 µl of 50 ng/µl of DNA template, the H_2O volume was adjusted to the volume of primer sets Multiplex 1 and Multiplex2 (see table) to make a total reaction volume of 15 µl. The PCR conditions were as follows 95 °C 12 min, followed by 35 cycles of 95 °C 15 sec, annealing at 60 °C for 30sec, extension at 72 °C for 1 min, later followed by single final extension of 72 °C 60 min and final hold at 10 °C ∞.

The PCR products were analyzed on a 2.5 % agarose 1x TAE gel for further modification of the primer design, its concentration and the annealing temperature of PCR reaction.

3.5. Agarose Gel Electrophoresis

A 2.5% agarose-gel was prepared by weighing 2g of agarose to 40 mL of 1X TAE buffer. The suspension was melted in the microwave until no floating particles could be observed. The solution was cooled to ~60 °C before the addition of 6 µl ethidium bromide. 3 µl of 6x loading buffer was added to all the PCR amplified samples prior to loading them on the gel along with 100 bp DNA ladder. Electrophoresis was performed for 120 minute at 80 V. Under UV-illumination, the gel was photographed.

3.6. Sequencing of PCR Amplicons

A 15 µl PCR reaction was set up for only one genomic DNA sample (positive control) with the different unlabeled primers and amplified. Prior to sequencing, the PCR amplified products were treated with Exonuclease1/ Exo1 to remove excess primer using 2 µl of PCR reagent, 2 µl of 5x sequencing buffer, 0.2 µl of Exo1 (20 u/µl) and 5.8 µl of H_2O for a total volume of 10 µl. Samples were incubated at 37 °C for 60 min, 85 °C for 15min and then held at 10 °C in a thermal cycler. Sequencing reactions were performed following Platt et al's (2007) St_ep method. 10 µl of the sequenced product was cleaned up using

NaOAc/EDTA/EtOH precipitation method. The precipitated sequencing products were denatured in 10 µl deionized formamide and run in a capillary electrophoresis (ABI PRISM 3130xl Genetic Analyzer (ABI, Foster City, CA) using the appropriate non-BDX run module.

The obtained sequences were blasted in NCBI blastn using the Nucleotide collection database in order to confirm that the microsatellites that has been analyzed matches with the targeted microsatellites that has to be studied.

3.7. Genotyping

To prepare capillary electrophoresis 1 µl PCR reaction was added to 10 µl formamide/GS Rox mix, where for 20 reactions 174.6 µl of Hi Di Formamide and 7.5 µl GS500 Rox (size marker) was used. In capillary electrophoresis system 1 µl sample was injected at 15 kV for 1 sec and separation of PCR fragments was done by a four-five coloured laser induced fluorescence capillary electrophoresis system (Prism 3130xl Genetic Analyzer). The electrophoresis was carried out for 15 minutes in POP7 polymer; 1X TAE buffer and one capillary well of 36 cm x 50 µm dimension at 15 kV. The amplicons separates in the lanes thereby demonstrates its ability to distinguish between homozygous and heterozygous animals. The data obtained from Gene-Mapper (version 4.0) were analyzed for the size of the alleles, checked whether they confirm to the ranges by creating panel for each microsatellite (1 and 2) and bins for every microsatellite marker. Depending upon the heights of each allele the concentration of the primers was adjusted.

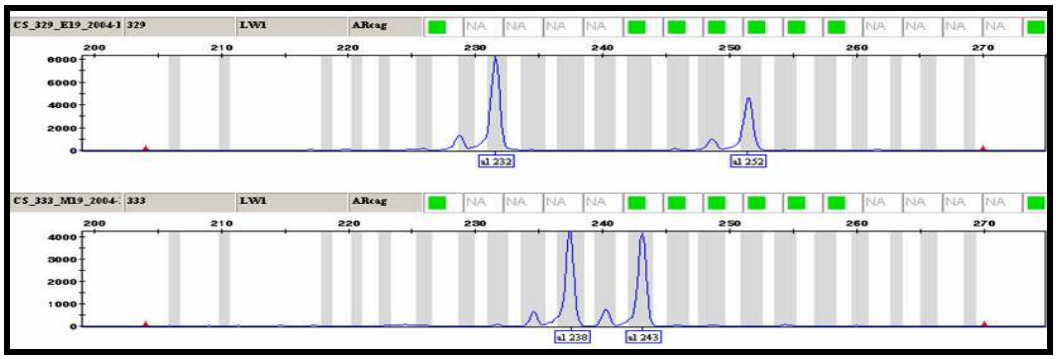


Figure 9 Panel for Microsatellite and their adjustment

3.8. Bioinformatics Analysis

Non-neutral markers were studied to alter the analysis and affect them; therefore 15 microsatellite markers with 1 sex chromosome marker were screened using the Lositan web application. It eliminated the non-neutral markers based on the relationship between F_{st} and H_e (expected heterozygosity. The presence of null alleles and large allele dropouts could bias the genotype analysis; thereby MICRO-CHECKER a windows-based program was used to identify them. The program not only identified the 'null' alleles but also adjusted their value automatically based on the deviations from Hardy-Weinberg equilibrium using a chi-square goodness-of-fit test.

Gene-Pop v.4.10 was used to determine allele frequency, expected and observed heterozygosity for each locus, for all individuals. Estimated the deviation from Hardy-Weinberg equilibrium was done by Markov-chain algorithm and the test of link equilibrium between each locus was done by Fisher's method. Calculating of pairs F_{st} values between all pairs of populations was done by standard AMOVA in Gene-Pop; P value for each population pair was calculated by Distance method. ML-Relate software was used to find the relationship between the foxes.

4. Results

In this study, microsatellite-based genotyping pipeline was established to determine genetic structures among red foxes in Scandinavia. DNAs were isolated successfully from the tissue, hair and scat samples and tested on agarose gel electrophoresis. Microsatellite markers were redesigned to optimize panel composition, which in turn reduced the number of panels. By testing and verification, selected markers were combined together and assembled into two different multiplexes. Both the multiplexes have been amplified with the template and optimized to give equal intensity of the signal at different loci, thereby studying the genotype of the different individuals of the Vulpes group.

4.1. Isolation of DNA

DNAs were isolated successfully from tissue samples of 33 red foxes. The quality of the DNA was tested and quantified using Nano drop measurement.

Table A1 (see appendix) illustrates the Nano drop measurement of different DNAs isolated from their corresponding tissue samples. The table shows the absorbance at 260, 280, 260/280 and 260/230. Samples were selected based on the 260/280 and 260/230-absorbance value for further analysis. DNA samples ranging less the absorbance value 2 for 260/280 and 2 – 2.25 for 260/230 were considered to be pure and good for further analysis. All the samples were observed to be good except for RR1009B sample. The absorbance of the sample for the ratio 260/280 was 2 and for 260/230 was 1.05. This may be because DNA of RR1009B was the second elution of the sample RR1009.

The DNA concentration from the scat samples was found to be too low as shown in table A2. Although the DNA from sample 515 had a concentration of 100.61 ng/µl, the absorbance ratio of 260/230 (1.75) was very low. As the samples were collected during summer, the DNAs would have been degraded. Thereby extraction of DNA from the scat samples were not as efficient as from the tissue samples.

The DNAs from the hair samples were also observed to be low from the Figure s3. The hair follicles for most of the hair in the samples were found to be missing because they were cut from the individuals and not pulled out directly. Hence only small amount of DNA was isolated from the hair.

4.2. Development of new primers for the established markers

Microsatellite primers were successfully designed for the established markers, thereby initiating the microsatellite-based genotype pipeline.

4.2.1. Marker choice

Markers were selected from Moore et al. (2010) (highlighted region) and Wandeler and Funk's (2005) established microsatellite loci for genotyping of red foxes based on the observed heterozygosity and number of alleles (Table 3).

4.2.2. Primer design

Primers were successfully designed by adjusting the melting temperature, GC content, length and also by adjusting the concentration of the monovalent salts (KCl), dNTPs and divalent salts ($MgCl_2$). Twenty microsatellites markers with 1 sex marker were developed, tested and used for genotyping out of which only 15 markers met the following criteria:

1. The primers amplified the fox DNA samples successfully.
2. The selected locus gave rise to PCR products of different sizes (100 bp to 400 bp).
3. The alleles gave good signal intensity during optimization.

In table 4, the highlighted primers (row 1) were redesigned for the microsatellite loci listed in Moore et al. (2010) except CPH7. The primers in row 2 were redesigned for the microsatellite loci listed in Wandeler and Funk (2005). The second version forward (fp2) and reverse primers (rp2) were generated during a secondary round of MPprimer-based primer design in cases where the first round primers failed the PCR amplification tests.

Table 4 Designed Primers

	Primers	DYE	Oligo sequence (5' - 3')	Predicted size range (bp)	Length
1	AHT121 fp2	VIC	AGCATGGCCCTATGGCAGTT	125-151	20
	AHT121 rp2		GGGGCTGGATCCGGATGTTA		20
	AHT133 fp2	NED	ACCCACCCAGGGATAGACAAC	135-149	21
	AHT133 rp2		CCTGCAGGAAGAGGTGCAAT		20
	AHT137 fp2	NED	ACAAGATGAGCTCTCTGCCT	118-144	20
	AHT137 rp2		AGAGCAAGGGCTTGTTGAGA		20
	CPH19 fp2	VIC	TGACCCCATGAAGGAATTTGC	236-250	21
	CPH19 rp2		TCACTCAATCTTCCAGAGCTCCA		23
	CPH7 F	FAM	ACACAACTTTCCATAATACTTCCCA	154-176	25
	CPH7 R		ATCAATGCTCTCCTCCCCAG		20
	FH2054 fp2	NED	TCTGTGTCCGGAAGGCTCAG	205-261	20
	FH2054 rp2		TGCTGAGTTTTGAACTTTCCCT		22
	FH2328 fp2	VIC	TAATAAAGGCTCCCTTCCAGGT	161-197	22
	FH2328 rp2		TGAAGAGGAGGAGACAGTTGTG		22
	FH2848 fp2	FAM	CCCCAAGTCAAAACCAACCCA	230-244	21
	FH2848 rp2		AGTCACAAGGACTTTTCTCCT		21
	K9AMELO fp2	FAM	ACTCCAACCCAACACCACCA	189-204	20
	K9AMELO rp2		AAGCTTCCAGAGGCAGGTCA		20
	REN105L03 fp2	FAM	GCAAAATGGGTGTCAGCAGC	136-152	20
	REN105L03 rp2		AGGCTCACCTACCTCGTTTCT		21
	REN162C04 fp2	VIC	CCGAGCTAAACAAGATCTTCTGCC	137-151	24
	REN162C04 rp2		ATCTTGTTCTCCCTTCCCCCG		21
	REN169018 fp2	FAM	GCACCATTGGCTCTCCCAAG	257-269	20
	REN169O18 rp2		ACCTGTGTTACCAACTCTTCTTCCT		25
	REN247M23 fp2	NED	TGGAGTGACAACACCAAGGC	169-187	20
	REN247M23 rp2		AGACATCAATCCACTCTGGGGA		22
2	AHT142 fp2	FAM	CCACTCCATTGCAGGGATAGGT	122-137	22
	AHT142 rp2		ACACAACACACTCATTCACACG		22
	CXX374 fp2	NED	AGAAAGTGTTTATTAAAACATGTGCG	105-117	26
	CXX374 rp2		GCCTGGCACAGGGATAATGC		20
	CXX468 fp2	FAM	TCAATCTCCCACCCAAATCTCT	79-91	22
	CXX468 rp2		AGACTTTTTAGTCCCGTGAAGA		22

4.2.3. Primer Testing

PCR reactions were set up using the newly designed primers to amplify the target microsatellite loci in the genomic DNA samples. The main goal to perform the agarose gel electrophoresis was to check whether each primer set could amplify the fox DNA under the designed PCR conditions. Successful amplification was confirmed by the presence of bands that were separated by electrophoresis and compared to a 100 bp ladder size standard.

Series of experiments were performed to establish final concentration primers and PCR reaction mix to give successful amplification of the target sequence of the microsatellite genomic samples.

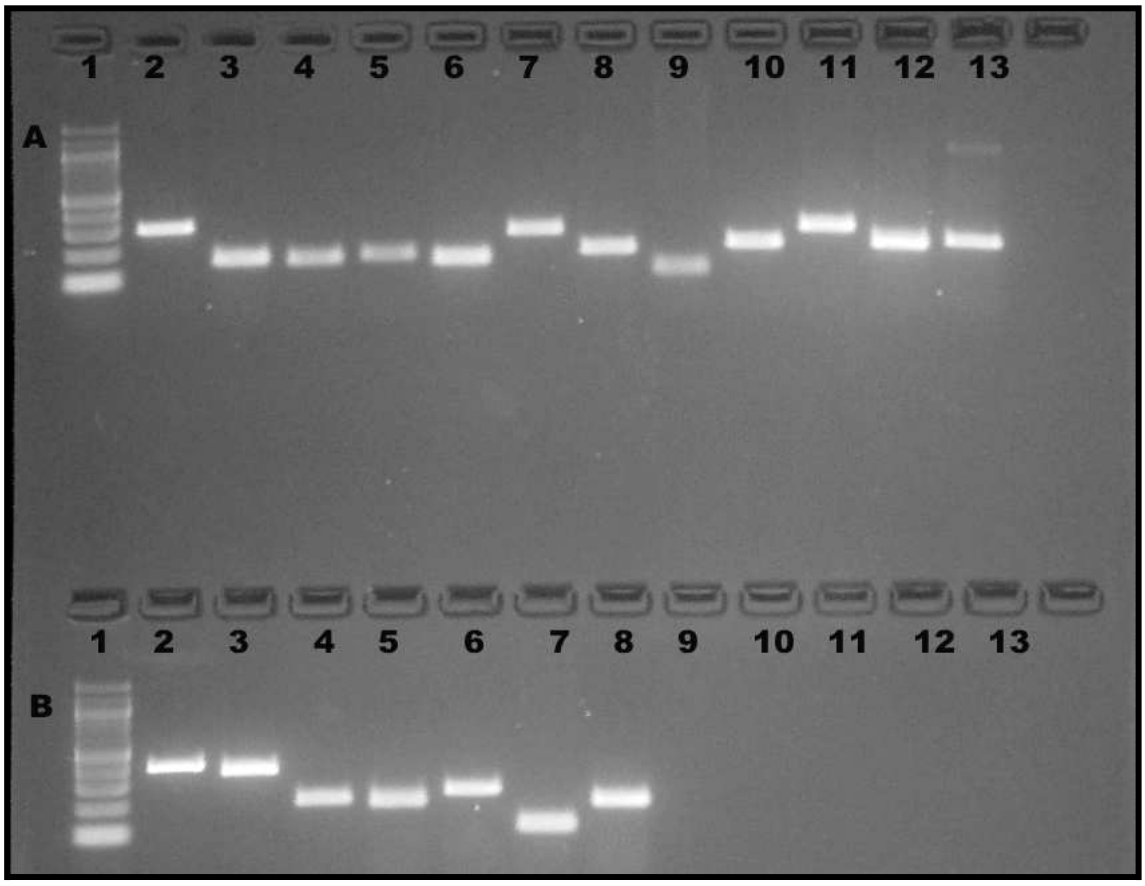


Figure 10: Photo of Agarose gel electrophoresis for PCR products on tissue genomic DNA, RR1025. Panel A: 1: 100 bp LADDER 2: AHT121 fp2+rp2, 3: AHT133 fp2+rp2, 4: AHT137 fp2+rp2, 5: AHT142 fp2+rp2, 6: CXX374 fp2+rp2, 7: CPH19 fp2+rp2, 8: CPH7 F+R, 9: CXX468 fp2+rp2, 10: FH2328 fp2+rp2, 11: FH2848 fp2+rp2, 12: K9AMELO fp2+rp2 and 13: REN105L03 fp2+rp2. Panel B: 1: 100 bp ladder, 2: REN162CO4 fp2+rp2, 3: REN169018 fp2+rp2, 4: REN247M23 fp2+rp2, 5: FH2054 fp2+rp2, 6: K9AMELO F+R, 7: CXX468 fp2+rp2 and 8: K9AMELO fp1+rp1.

From Figure 10 the strong bands confirm that the PCR amplification was successful for DNA from tissue samples and also all the primers amplified the DNA samples. The PCR products were separated on the agarose gel at their corresponding sizes against the 100bp DNA ladder. In the panel A AHT133, AHT137, AHT142 and CXX374 were approximately between 100 bp to 200 bp. CPH7, FH2328, K9AMELO and REN105L03 were between 200

bp to 250 bp. AHT121, CPH19 and FH2848 were at ~300bp. In panel B REN162CO4, REN169018 were at ~400 bp. REN247M23 and FH2054 were at ~250 bp; CXX468 was at 200 bp. K9AMELO F+R and K9AMELO fp1+rp1 were at 300 bp and 250 bp, respectively. REN169018 and K9AMELO F+R were observed to have larger size compared to the predicted size range, 257-269 bp and 202-217 bp respectively.

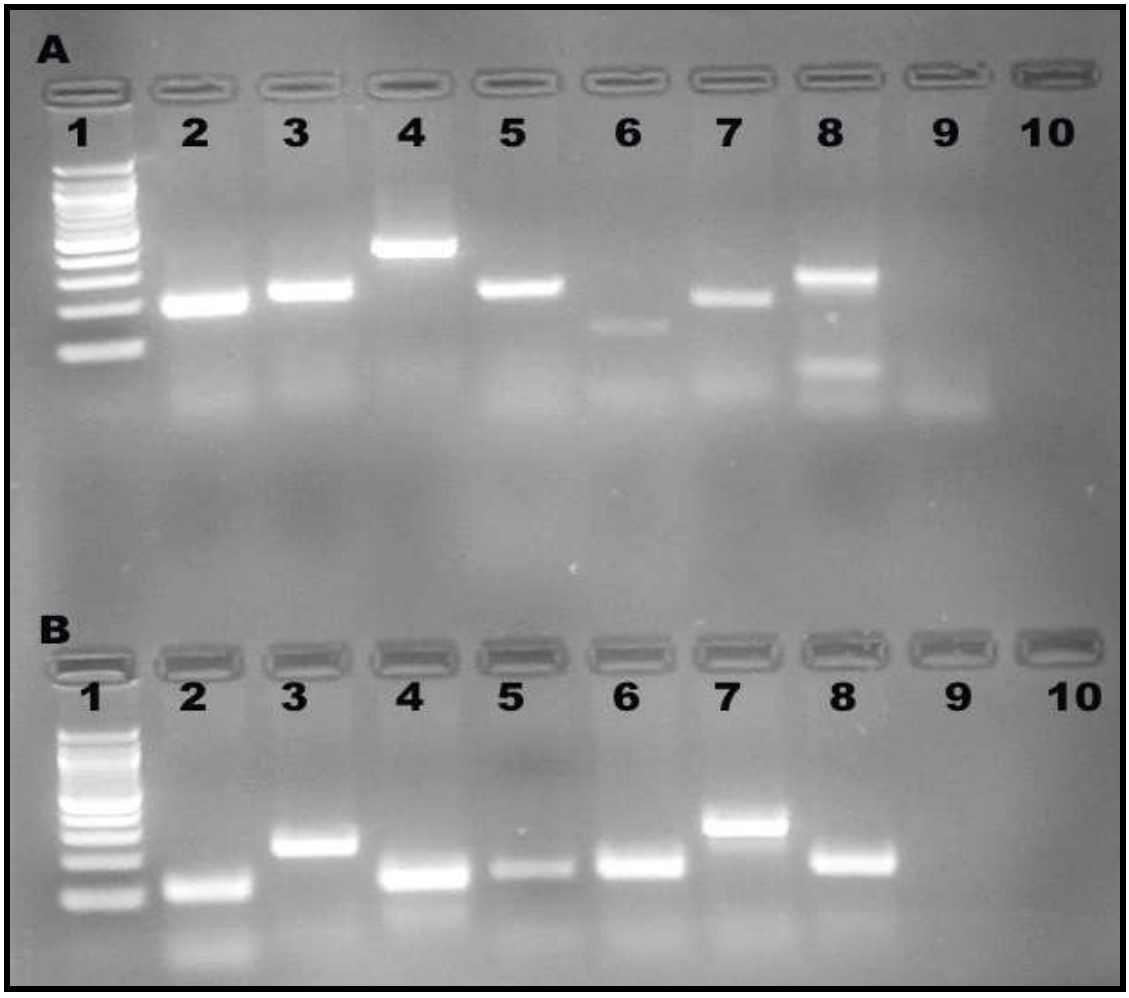


Figure 11: Photo of Agarose gel Electrophoresis for PCR products of fresh scat sample, Sc1042. Panel A: 1: 100 bp LADDER, 2: REN247M23 fp2+rp2, 3: REN169018 fp2+rp2, 4: REN162CO4 fp2+rp2, 5: K9AMELO fp2+rp2, 6: FH2848 fp2+rp2, 7: FH2328 fp2+rp2, 8: FH2054 and 9: CXX468 fp2+rp2. Panel B: 1: 100 bp LADDER, 2: CXX374 fp2+rp2, 3: CPH19 fp2+rp2, 4: AHT142 fp2+rp2, 5: AHT137 fp2+rp2, 6: AHT133 fp2+rp2, 7: AHT121 fp2+rp2, 8: CPH7 F+R and 9: REN105L03 fp2+rp2

Agarose gel from Figure 11 confirms that the primers were able to amplify the DNA from fresh scat samples. The PCR products were separated on the agarose gel at their corresponding sizes against the 100bp DNA ladder. In panel A REN247M23, REN169018, K9AMELO, FH2328 were at ~200 bp, REN162CO4 was between 300-400 bp and FH2054 was at 250 bp. Avery light band ~200 bp has been formed for FH2848 and primer dimer for CXX468. In the panel B Cxx374 has a band at ~110 bp; CPH19, AHT142, AHT137,

AHT133 and CPH7 have bands at less than 200bp. AHT121 has a very strong band at 400bp. REN162CO4 and AHT121 were observed to have larger size compared to the predicted size ranges 137-151 bp and 125-151 bp respectively.

PCR products for dried scat samples were similarly tested by agarose gel electrophoresis (Figure 12).

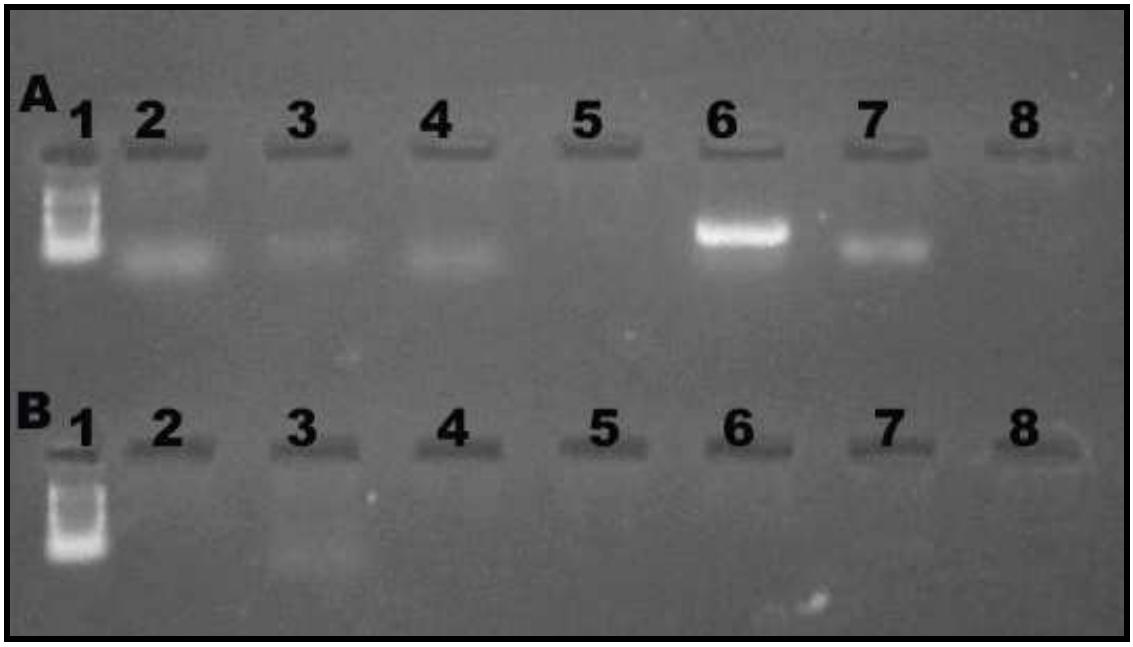


Figure 12: Agarose gel electrophoresis of PCR product of genomic DNA- dried scat, Sc407. Panel A: 1:100bp ladder, 2:AHT121fp2+rp2, 3:AHT133fp2+rp2, 4:AHT137fp2+rp2, 5:AHT142fp2+rp2, 6:k9AMELOfp2+rp2, 7:CPH7F+R, 8:FH2328fp2+rp2. Panel B: 1:100bp ladder, 2: REN162CO4fp2+rp2, 3:RNE169018fp2+rp2, 4:CPH19fp2+rp2, 5:CXX374fp2+rp2, 6:AHT121fp2+rp2, 7:REN247M23fp2+rp2, 8:REN105L03fp2+rp2.

PCR amplification on dried scat samples was not as successful as PCR conducted on tissues and wet scat samples as seen in Figure 10 and 11, respectively. Only K9AMELO has worked, showing a strong band of 200 bp. The bands for AHT121, AHT133, AHT137 and AHT142 are difficult to conclude as the ladder was smeared.

4.2.4. Multiplexes

Multiplex PCR was performed and PCR products were genotyped on genomic tissue DNA sample, RR1025. The primer concentrations were adjusted depending upon the intensity of the fluorescence signal in Gene-Mapper, as in Figure 13 and 14.

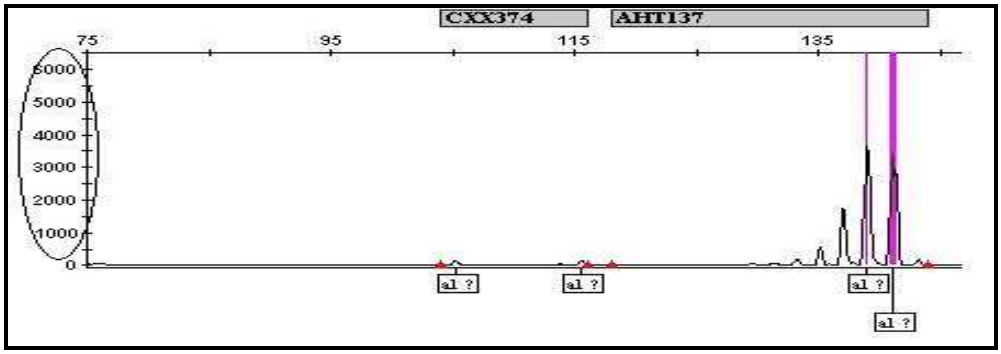


Figure 13: Panel of alleles at their corresponding loci giving high signal intensity around 5000 units for AHT137 and very low signal for CXX374. Tissue sample, RR1025

The Panel in Figure 13 illustrates the signal intensity of CXX374 and AHT137 alleles of multiplex 1 for the sample RR1025. Every allele was framed into 2-3 bp repeat differences. This frame was referred to as 'bins'. The allele sizes were not shown as the bins were yet to be created. AHT137 show high signals and the pink highlighted region could be due to the pull up from the markers in the VIC region (green). CXX374 and AHT137 should have produced signal in yellow colour as they were tagged with NED, but they appeared in black as a result of cross-detection of the colour due to overlapping emission curves. The concentration of the primer was reduced to produce better signal for AHT137. The alleles for CXX374 shows very low signal; thereby its concentration was increased to have a clear view of its tops.

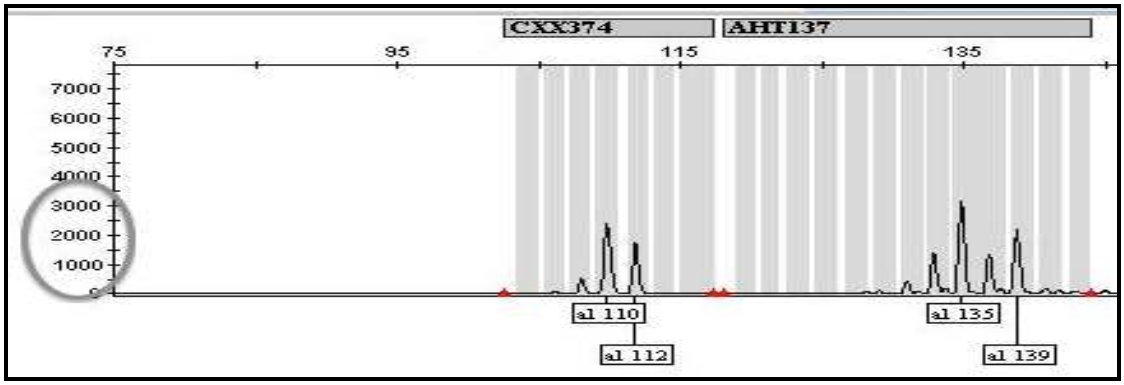


Figure 14: Panel illustrates the alleles giving better signal intensity, around 3000units after the alteration of the primer concentration. Tissue sample, RR1025.

Panel in Figure 14 illustrates alleles giving better signal after alteration of the primer concentration. Primer concentration for CXX374 marker was increased from 0.21 µM to 0.33 µM and primer concentration for AHT137 marker was decreased from 0.42 µM to 0.39 µM. Bins were created to check the allele sizes. CXX374 were found to be heterozygous

showing allele sizes 110 and 112, whereas AHT137 was also found to be heterozygous showing allele sizes 135 and 139.

Table 5 Grouping of primers into Multiplexes

Multiplex Panel	**Primer**	**Size Range (bp)**	**Colour**	**Final Concentration (μM)**
1	AHT121fp2+rp2	261-281	Green	0.033
	AHT133fp2+rp2	135-149	Green	0.033
	AHT137fp2+rp2	118-144	Yellow	0.039
	AHT142fp2+rp2	122-137	Blue	0.026
	CPH19fp2+rp2	236-250	Green	0.072
	CPH7F+R	154-176	Blue	0.05
	K9AMELOfp2+rp2	194-208	Blue	0.023
	CXX468fp2+rp2	76-91	Blue	0.05
	FH2054fp2+rp2	206-262	Yellow	0.033
	CXX374fp2+rp2	104-116	Yellow	0.033
2	FH2848fp2+rp2	230-244	Blue	0.1
	FH2328fp2+rp2	161-197	Blue	0.06
	REN105LO3fp2+rp2	396-412	Blue	0.04
	REN162CO4fp2+rp2	366-380	Green	0.026
	REN169018fp2+rp2	194-208	Blue	0.033
	REN247M23fp2+rp2	169-197	Yellow	0.05

Sixteen labelled microsatellite markers were chosen for multiplex PCR design. The choice of the primers in each multiplexes depended on the colour, size range, so that they didn't overlap each other (table 5).

4.3. Genotyping

Thirty-three individual foxes were genotyped successfully. True alleles for each microsatellite marker for every individual were identified and analysed (see Figure 15 for Multiplex1 and Figure 16 for Multiplex 2).

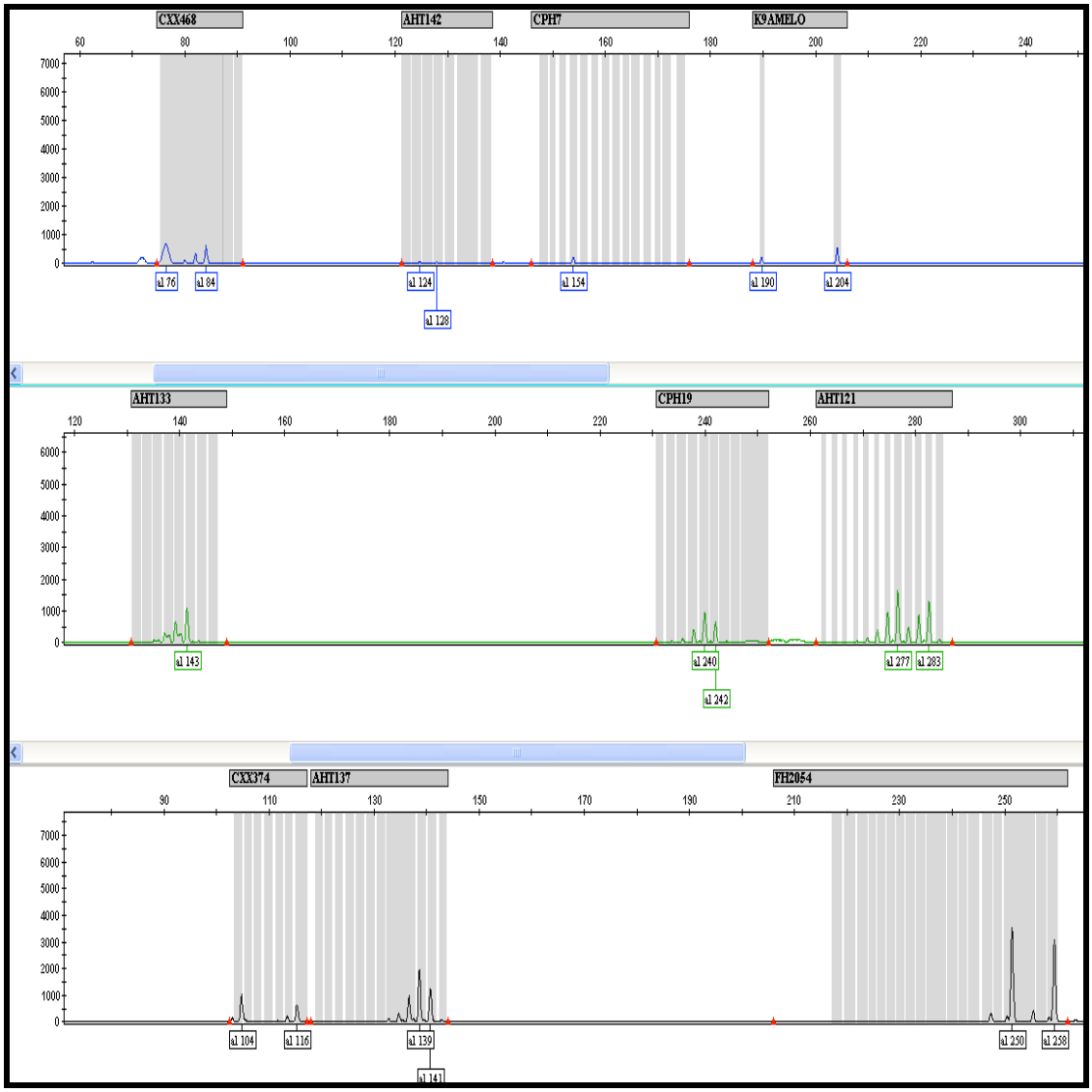


Figure 15: Multiplex1 genotyping panel illustrates the true allele sizes of different markers at their corresponding loci. Tissue sample, RR1025

Figure 15 illustrates how each marker of Multiplex 1 has occupied their corresponding loci according to their size range and also based on their labelled fluorescence. Top panel represents the FAM labelled markers (blue colour). It includes CXX468, which was heterozygous with allele sizes 76 and 84, AHT142 also heterozygous with allele sizes of 124 and 129. K9AMELO was found to be heterozygous with allele sizes 190 and 204 representing a male sample RR1025. The alleles gave very low signal; thereby the primer concentration was increased to give a better signal. Second Panel represents markers labelled with VIC fluorescence (green). It includes AHT133 allele with size 143 (homozygous), allele with sizes 240 and 242 (heterozygous) and also AHT121 with sizes 277 and 283 (heterozygous). Last Panel represents markers labelled with NED fluorescence (yellow). It includes CXX374 allele with sizes 104 and 116 (heterozygous), AHT137 allele

with sizes 139 and 141(heterozygous) and FH2054 allele with sizes 254 and 258 (heterozygous).

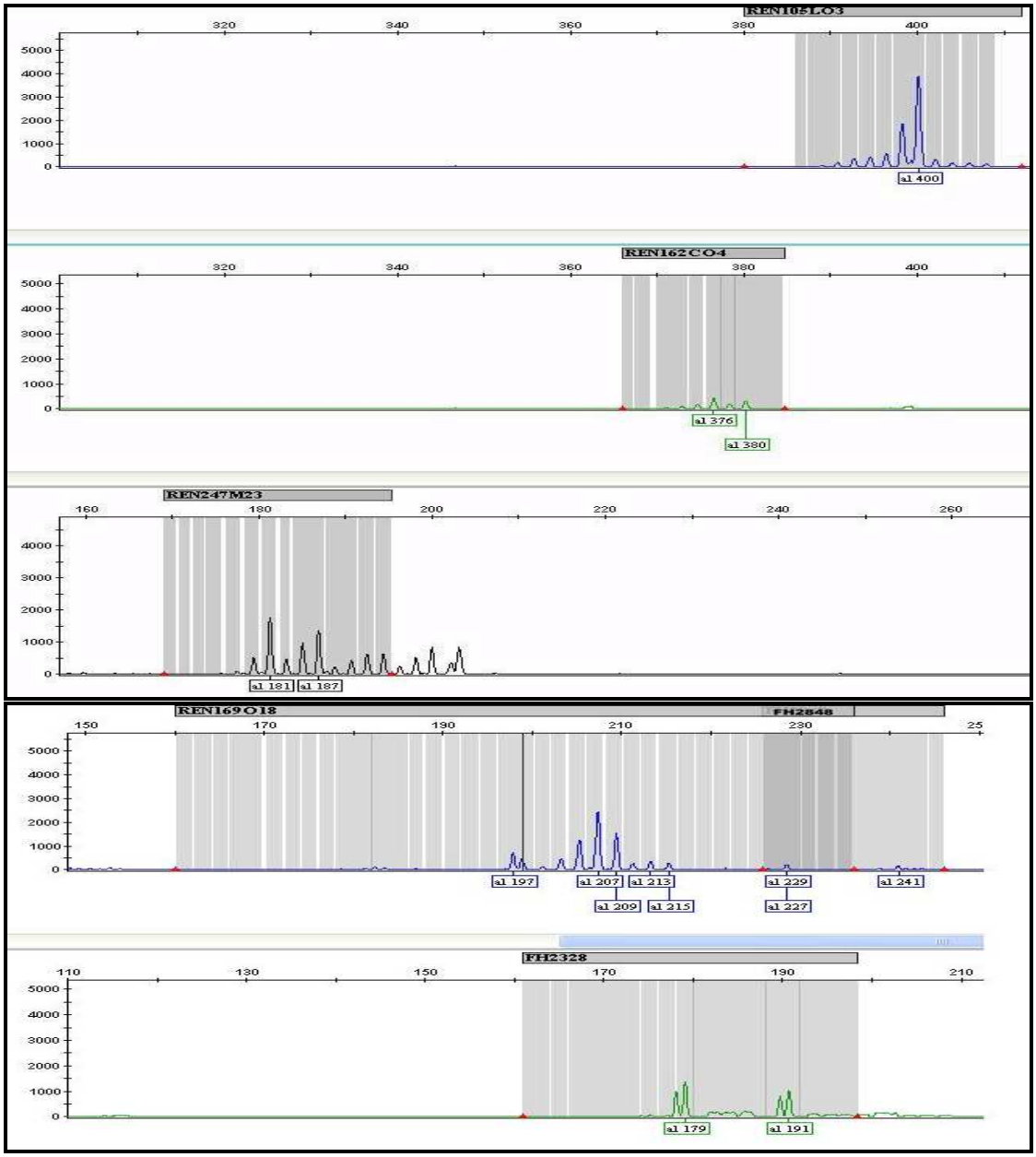


Figure 16: Representation of Multiplex 2 panel illustrating the true allele sizes of each marker at their respective loci. Tissue sample RR1025

Markers in Multiplex 2 covered wide size ranges; hence the panel was split into two (Figure 16) to illustrate all the markers. Top panel represents the FAM labelled markers (blue colour). It includes REN105LO3 marker with allele size 400 (homozygous), FH2848 with allele sizes 229 and 241 (heterozygous). REN169018 marker showed many allele sizes but as alleles in the foxes were known to be diploid the sizes of high peaks were considered for the analysis. The middle panel represents the markers labelled with VIC fluorescence (green). It consists of REN162C04 marker with allele sizes 376 and 380 (heterozygous) and

FH2328 marker with allele sizes 179 and 191 (heterozygous). The last panel represents the markers labelled with NED fluorescence (yellow). It includes REN247M23 marker showing allele sizes of 101 and 107, hence proving heterozygous.

The alleles of two hair samples were compared with the alleles of two tissue samples to confirm that they were collected from the same foxes.

Table 6 Genotyping results

Markers	**Tissue 1025**	**Hair 1025**	**Tissue 1019**	**Hair 1019**
AHT121	277,283	277,283	279	279
AHT133	143	143	141,145	141,145
AHT137	139,141	139,141	141,143	141,143
AHT142	124,128	124,128	128,130	128,130
CPH19	240,242	240,242	234,240	234,240
CPH7	154	154	152	152
CXX374	104,116	104,116	110	110
CXX468	76,84	76,84	**76**,82	82
FH2054	250,258	250,258	250	250
K9AMELO	190,204	190,204	**190**,204	204
FH2328	179	179	191	189
FH2848	231,245	231,245	229	229
REN105L03	404	404	406	406
REN162C04	374,382	374,382	374,380	374
REN169018	187,203	187,203	209,**227**	209
REN247M23	193,195	193,195	181	181

The alleles of different markers matched with each other between the hair and tissue samples but not exactly (table 6). In tissue 1019 sample and hair 1019 samples, CXX468 shows only one allele in hair sample (82) whereas it shows two alleles in tissue sample (76 and 82).

Similarly the alleles of two scat samples (fresh) were compared with the alleles of two tissue samples to prove whether they were collected from the same foxes.

Table 7 Genotyping results of the matched samples

Markers	Tissue 1009	Scat 1009	Tissue 1042	Scat1042
AHT121	277	277	277,279	277,279
AHT133	137,**141**	**135**,137	139,143	139,143
AHT137	139,141	139,141	137,141	137,141
AHT142	128,132	128,132	126,136	126,136
CPH19	234,**250**	234	242,250	242,250
CPH7	152,154	152,154	154	154
CXX374	110,**112**	110,**114**	116	116
CXX468	80,84	80,84	76,78	76,78
FH2054	232,234	232,234	254,258	254,258
K9AMELO	204	204	190,204	190,204
FH2328	191	191,**189**	181,191	181,191
FH2848	229,**241**	229,**243**	229,243	229,243
REN105L03	**400**,402	**398**,402	402,406	402,406
REN162C04	372,380	372,380	382	382
REN169018	207,227	207,227	207,215	207,215
REN247M23	189	**187**,189	193,195	193,195

Like the hair samples, some of the alleles of Sc1009 (Dry Scat sample) did not match exactly with the alleles of the tissue sample (table 7). For example AHT133 showed alleles with sizes 137 and 141 in tissue sample RR1009 but it showed alleles with sizes 135 and 137 in the scat sample Sc1009. In case of scat sample Sc1042 (Wet Scat sample) and tissue sample RR1042 all the alleles matched exactly. This does not mean that Sc1009 and RR1009 were not collected from the same fox; the slight variation in the allele sizes may be due to the quality of the scat DNA.

4.3.1. Sex Differentiation

Sex of each fox was identified using K9AMELO marker. Depending upon the zygosity, the sex of the individuals was accessed. Females are homozygous; Figure 17 and males are heterozygous, Figure 18.

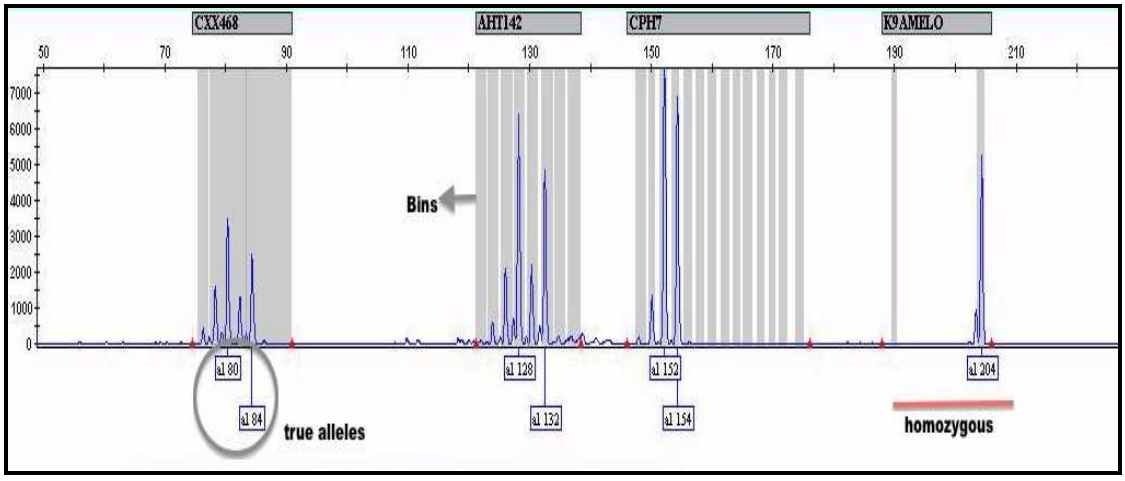


Figure 17: Multiplex panel of a female fox, tissue sample RR1006. K9Amelo shows X allele, homozygous.

Figure 17 illustrates part of Multiplex 1 panel displaying the alleles of a female fox, RR1006. The allele with highest peak was considered as 'true allele'. The panel displays markers with FAM labelled. CXX468 with true allele sizes 80 and 84 (heterozygous), AHT142 with true allele sizes 128 and 132 (heterozygous) and CPH7 with true allele sizes 152 and 154 (heterozygous) were observed along with the sex marker K9AMELO in the panel. Only one true allele representing X chromosome and also proving to be homozygous with consistent size standard 204 was observed under K9AMELO marker; thereby confirming that RR1006 fox was a female fox.

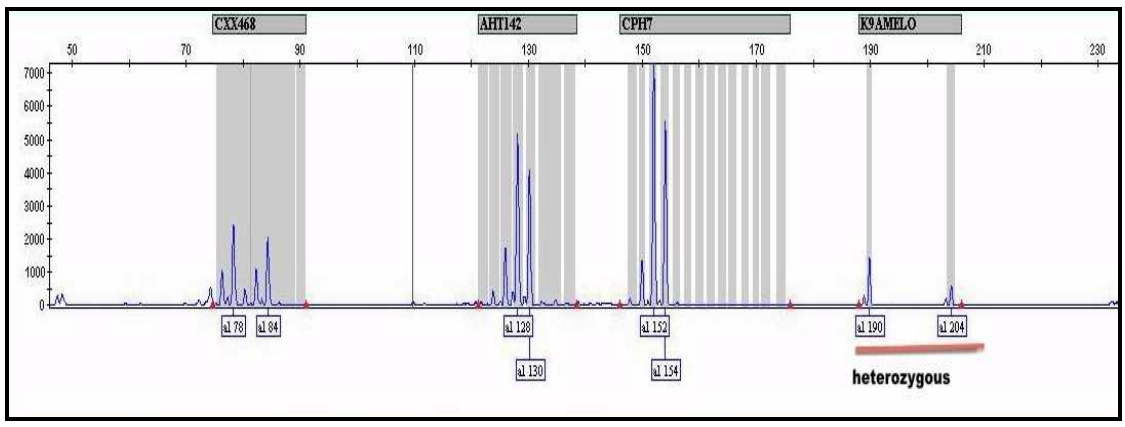


Figure 18: Multiplex panel of a male fox, tissue sample RR1025. K9AMELO shows two true alleles (X and Y), heterozygous.

Figure 18 illustrates part of Multiplex 1 panel displaying the alleles of a male fox, RR1025. The panel displays markers with FAM labelled. CXX468 with true allele sizes 78 and 84 (heterozygous), AHT142 with true allele sizes 128 and 130 (heterozygous) and CPH7 with true allele sizes 152 and 154 (heterozygous) were observed along with the sex marker K9AMELO in the panel. Two true alleles representing X and Y with their sizes 204 and 190 (heterozygous) respectively were observed at K9AMELO locus, hence proving that RR1025 was a male fox.

Table 8 illustrates the sex of the foxes that were genotyped using their genomic DNA isolated from their tissue. Out of 33 foxes, 12 foxes were found to be females and rest of the 21 were males, highlighted row in the table.

Table 8 Sex differentiation between the foxes

Foxes	**Size of the allele K9AMELO**	
RR1000	190	204
RR1004	190	204
RR1005	190	204
RR1007	190	204
RR1018	190	204
RR1019	190	204
RR1021	190	204
RR1022	190	204
RR1023	190	204
RR1024	190	204
RR1025	190	204
RR1027	190	204
RR1029	190	204
RR1031	190	204
RR1033	190	204
RR1034	190	204
RR1035	190	204
RR1037	190	204
RR1042	190	204
RR1046	190	204
RR1048	190	204
RR1006		204

RR1009	204
RR1026	204
RR1028	204
RR1030	204
RR1032	204
RR1038	204
RR1039	204
RR1041	204
RR1044	204
RR1047	204
RR1050	204

From table 8, it was observed that all the male and female foxes show consistent allele size at their respective K9AMELO loci.

4.4. Sequencing Analysis

Direct sequencing was performed for both forward and reverse PCR products to confirm whether the genotyped microsatellite loci were the expected target loci or not. Normally, DNA sequences that included homozygous markers were only sequenced, however as some of the markers like CXX468 were found only as heterozygous they were also sequenced to verify the results.

The best part of the displayed sequencing result from CLC workbench, Figure 19 and 20 was blasted using blastn on NCBI web page.

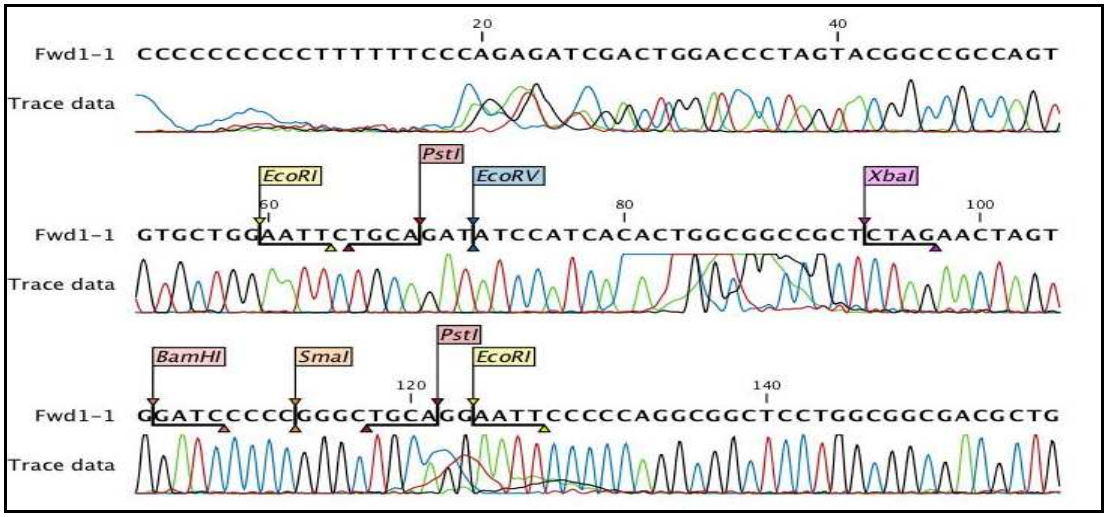


Figure 19: Part of sequencing analysis of FH2328 microsatellite forward reaction (homozygous marker from RR1035 tissue sample) obtained from CLC workbench.

The best part of the sequence was observed from base 36, illustrated in Figure 19. Each colour of the peak represents a nucleotide. Blue for cytosine, green for adenine, red for thiamine and finally black represents guanine. The sequence was later blasted to confirm that it fits FH2328 microsatellite.

The nucleotide blast (Figure A2) gave only one significant hit- FH2328 microsatellite sequence after performing a blast on the best part of the FH2328 sequence (Figure 19); thereby verified the FH2328 loci.

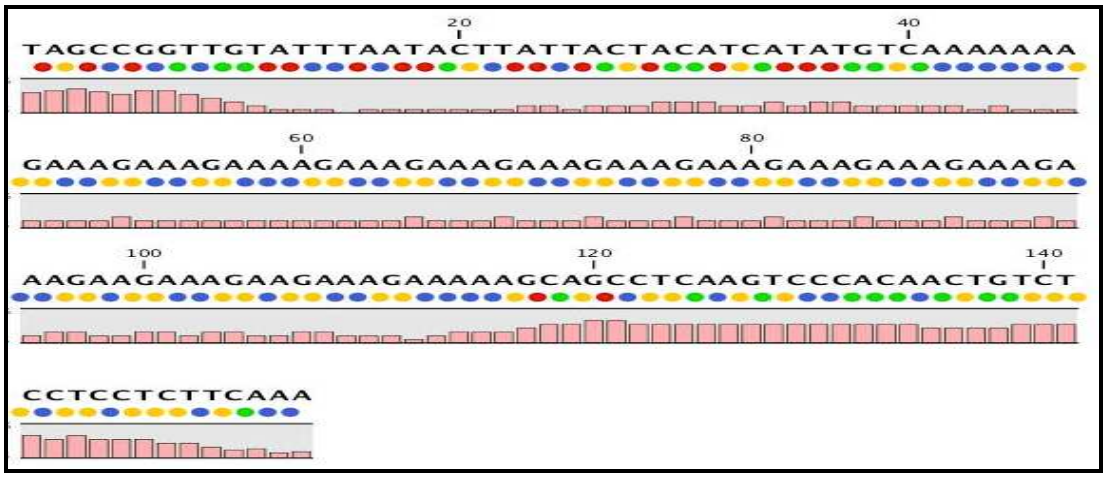


Figure 20: Part of sequencing analysis of K9 Amelo microsatellite (heterozygous condition) obtained from CLC workbench

The sequence illustrated in Figure 20 looked bit complicated as it represents a heterozygous marker K9AMELO from a male fox, RR1025. The sequence from the length 120 was blasted to verify K9AMELO microsatellite sequence in the NCBI.

The Blastn (Figure A3) gave many hits against the sequence from Figure 20 when blasted in the NCBI. The hits were aligned with amelogenin Y (AMELY) gene sequence and amelogenin (AMELY X) gene; thereby confirming the heterozygous K9AMELO microsatellite sequence.

The results illustrated in table A7 (appendix), obtained from the blastn verified that all the other sequences also match their respective targeted microsatellite sequences.

4.5. Bioinformatics Analysis

The genotypic results obtained from the Gene-Mapper were further processed to analyse and estimate the population diversity of the foxes. The data were converted into gene-pop files using Microsatellite tools in the Excel sheet. The values of polymorphic information content (PIC), number of alleles (N_a), expected heterozygosity (He) and observed heterozygosity (Ho) are listed in Table 9.

Table 9 Characterization of 15 microsatellite loci

Locus	**N_a**	**He**	**Ho**	**PIC**
AHT121	7	0.741	0.636	0.701
AHT133	7	0.798	0.606	0.753
AHT137	5	0.781	0.818	0.732
AHT142	7	0.653	0.697	0.605
CPH19	8	0.8	0.879	0.76
CPH7	3	0.38	0.364	0.315
CXX374	6	0.727	0.364	0.671
CXX468	6	0.801	0.879	0.758
FH2054	11	0.893	0.879	0.867
FH2328	12	0.897	0.848	0.873
FH2848	7	0.832	0.636	0.794
REN105LO3	10	0.883	0.879	0.856
REN162O4	8	0.815	0.758	0.776
REN247M23	7	0.62	0.51	0.54
REN169018	15	0.864	0.909	0.837

The highest number of alleles (15) was seen for REN169018 locus within the population. Loci REN105LO3 (10), FH2328 (12), FH2054 (11) show alleles greater than 10 and the remaining alleles show lower number of unique alleles (Table A4). Polymorphic information content allows identifying the marker genotype that is inherited from the parent. Normally PIC value ranges between 0 and 1. Loci FH2328 have the highest PIC

value (0.873) and CPH7 has the lowest PIC value (0.315). The observe heterozygosity differed widely among the analyse loci. Highest observed heterozygosity was observed for REN169018 (0.909) and CPH7, CXX374 have the lowest observed heterozygosity (0.364).

As the analysis depended on diploid microsatellite alleles; K9Amelo (haplodiploid) allele was excluded.

4.5.1. LOSITAN Analysis

The genotyping results were submitted into the LOSITAN web page to identify the markers that are found to be under selection.

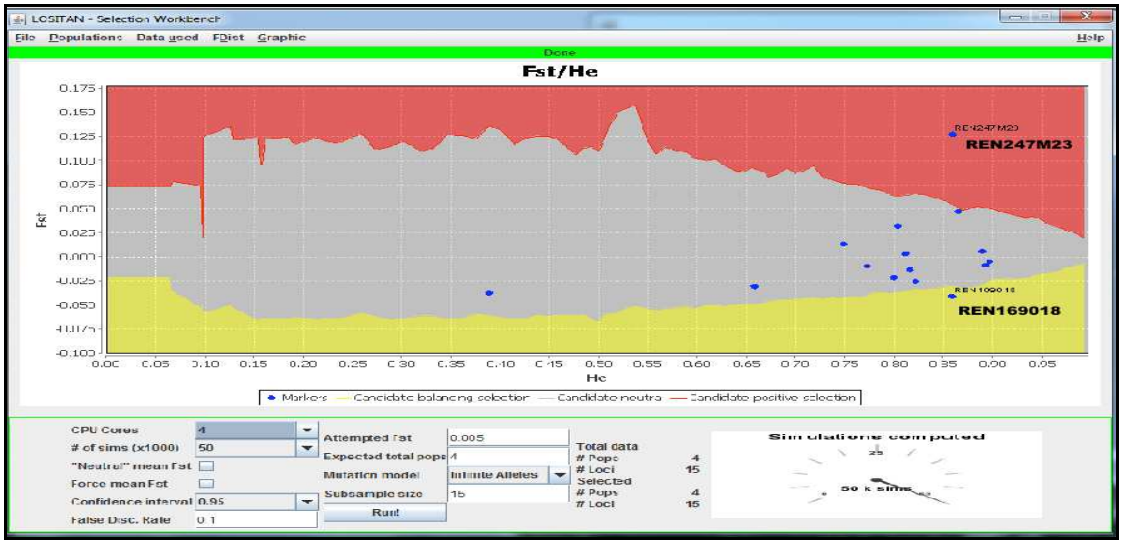


Figure 21: Screen shot of LOSITAN webpage illustrating the neutral loci in the grey region. The locus in red and yellow region is considered under selection.

From the Figure 21 and table A8 (Appendix), it was observed that REN247M23 marker was under direct selection (positive) and REN169018 marker was under balanced selection, therefore they were considered to be non-neutral markers. The neutral markers were displayed in the grey region at their respective loci. Although studies confirmed that it was better to eliminate the non-neutral markers while analysing or performing genetic studies, most researchers included markers under balanced selection along with other neutral markers for their genetic studies (K. Østebye, personal communication). So the genotype data were analysed with and without REN169018 marker in order to understand its effect on the analysis.

4.5.2. MICRO-CHECKER Analysis

MICRO-CHECKER software was used to test the genotyping of diploid population's microsatellites. The program was used to identify the null alleles, large allele drop out and also typographic errors. The most important criteria of this program was to adjust the allele and genotype frequencies of the amplified alleles that can be used for further population genetic analysis, thereby avoiding false interpretation of the genetic analysis (Cock et al., 2004).

Three out of fourteen markers were identified to have null alleles and the following table 10 shows the adjusted alleles of these markers.

Table 10 Adjusted null alleles by MICRO-CHECKER program. '1' refers to the adjusted 'null' allele value

Samples	AHT133		CXX374		FH2848	
RR1000	145	147	112	114	231	1
RR1001	143	145	110	112	229	1
RR1002	143	1	112	112	231	233
RR1003	139	139	114	1	243	243
RR1004	143	145	110	1	229	245
RR1005	137	141	110	1	229	1
RR1006	143	147	112	1	233	241
RR1007	141	145	110	110	229	229
RR1008	139	139	110	1	241	241
RR1009	139	145	112	1	229	245
RR1010	143	147	112	112	241	241
RR1011	133	143	112	112	241	245
RR1012	143	1	112	112	241	241
RR1013	145	145	110	112	243	243
RR1014	143	143	114	114	229	235
RR1015	145	147	114	114	229	241
RR1016	139	139	112	112	229	231
RR1017	139	145	104	116	241	245
RR1018	139	145	110	1	229	229
RR1019	141	141	114	1	229	245

RR1020	143	143	110	1	235	243
RR1021	141	1	112	1	243	243
RR1022	145	1	112	1	231	1
RR1023	141	145	110	1	243	1
RR1024	141	141	110	114	245	245
RR1025	141	143	112	112	229	245
RR1026	141	145	112	112	233	233
RR1027	139	139	110	110	243	1
RR1028	141	145	104	116	229	1
RR1029	141	143	110	110	243	245
RR1030	145	1	110	1	243	243
RR1031	143	1	104	1	231	245
RR1032	139	143	116	116	229	243

There were two choices to approach the null alleles, either to remove them from the analysis or consider them as recessive alleles by adjusting its value to 1 (table 10). Both the approaches were followed to observe their effects on the study.

4.5.3. Population Diversity Analysis on the genotype data

Analysis of genetic variation has been conducted on original genotypic data, the genotyping data that has been adjusted by MICRO-CHECKER (null alleles replaced by 1, considered recessive alleles) and the data without null allele markers. Analysis was all conducted with the addition of non-neutral marker, REN169018 to observe its effect. The gene-pop files were converted to structure format and loaded into STRUCTURE program. The program computed the ratio of the genome of an individual that could be originated from each inferred population by quantitative clustering method. The individuals under their deduced populations were analysed to study the genetic variation among them.

- **Original Genotypic Data** (considering all the alleles as true alleles): Foxes have been assembled into 3 clusters effectively under 1 population based on their predominant value of F_{st} in each cluster, as shown in the Figure 22. Each shade represents a cluster.

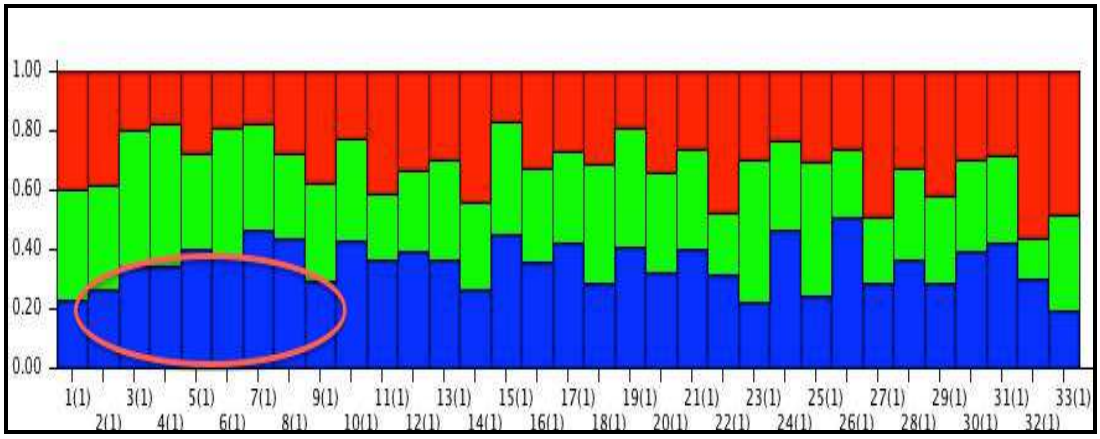


Figure 22: Bar plot representing various populations where each fox shares its proportion in each population. X-axis represents the individuals and Y-axis represents F_{st} values. The bars of the same height, for example the region under the circle are genetically closer compared to the rest.

The X-axis in the bar plots of Figure 22 represents the individuals in the study and the Y-axis represents the F_{st} ranges. Each fox shares its own genetic proportion in every cluster. For example Fox 1 (RR1000) shares equal proportion genetically in cluster 1 (red) and 2 (green) compared to the 3rd cluster. Classification of the foxes was fixed depending on the highest proportion shared by each individual in a group.

Genetic variation among the groups and between the groups were analysed using Arlequin software.

Table 11 Matrix of significant F_{st} (P) values. Where significance level=0.05, using Distance method. '+' Indicates significant and '-' indicates non-significant.

	Group 1	**Group 2**	**Group 3**
Group 1	*		
Group 2	0.07207+-0.0353 (-)	*	
Group 3	0.00000+-0.0000 (+)	0.02703+-0.0194 (+)	*

Table 11 illustrates that individuals of group 3 were observed to vary significantly from group 1 and 2 genetically; thereby this group could be considered as a population. As group 1 and 2 does not vary significantly with each other they could be considered as one group.

Table 12 Amova results from Arlequin software (average over 13 loci) illustrating the genetic variation.

Source of variation	d.f.	Sum of squares	Variance components	Percentage of variation
Among populations	2	16.500	0.14879 Va	2.94
Among individuals within populations	30	156.061	0.29040 Vb	5.74
Within individuals	33	152.500	4.62121 Vc	91.32
Total	65	325.061	5.06041	

Fixation Indices	
FIS :	0.05913
FST :	0.02940
FIT :	0.08679

Percentage of variation from Amova results (table 12) using Arlequin software proved that the genetic variation within all the individuals was predominant to the individuals within the population and among the populations. The percentage of genetic variation within the individuals was 91.32%; whereas among the individuals within populations and among population showed only 5.74% and 2.94% variation respectively. Inbreeding co-efficient (FIS value) was not significant as its value was not less than 0.05. The F_{st} value (0.02) proves that the genetic variation among the individuals were significant.

Table 13 Illustration of Linkage disequilibrium between the loci. Results collected from Arlequin program. '– ' indicates non-significant linkage disequilibrium. Locus 0: AHT121, 1:AHT133, 2:AHT137, 3:AHT142, 4:CPH19, 5:CPH7, 6:CXX374, 7:CXX468, 8:FH2054, 9:FH2328 10:FH2848, 11:REN105LO3, 12:REN162CO4

Histogram of the number of linked loci per locus

Locus:	0	1	2	3	4	5	6	7	8	9	10	11	12
	0	0	0	0	0	0	0	0	0	0	0	0	0

Table of significant linkage disequilibrium (significance level=0.0500):

Locus #	0	1	2	3	4	5	6	7	8	9	10	11	12
0	*	–	–	–	–	–	–	–	–	–	–	–	–
1	–	*	–	–	–	–	–	–	–	–	–	–	–
2	–	–	*	–	–	–	–	–	–	–	–	–	–
3	–	–	–	*	–	–	–	–	–	–	–	–	–
4	–	–	–	–	*	–	–	–	–	–	–	–	–
5	–	–	–	–	–	*	–	–	–	–	–	–	–
6	–	–	–	–	–	–	*	–	–	–	–	–	–
7	–	–	–	–	–	–	–	*	–	–	–	–	–
8	–	–	–	–	–	–	–	–	*	–	–	–	–
9	–	–	–	–	–	–	–	–	–	*	–	–	–
10	–	–	–	–	–	–	–	–	–	–	*	–	–
11	–	–	–	–	–	–	–	–	–	–	–	*	–
12	–	–	–	–	–	–	–	–	–	–	–	–	*

From table 13, the Loci did not deviate from significant Linkage disequilibrium (LD). As this program eliminates minimum LD, if the locus shows slight LD they did not have significant effect on the analysis.

- **Original Genotypic Data** (considering all the alleles as true alleles with REN169018): Foxes have been assembled into 3 clusters effectively under 1 population based on their predominant value of F_{st} in each cluster, as shown in the Figure 23. Each shade represents a cluster.

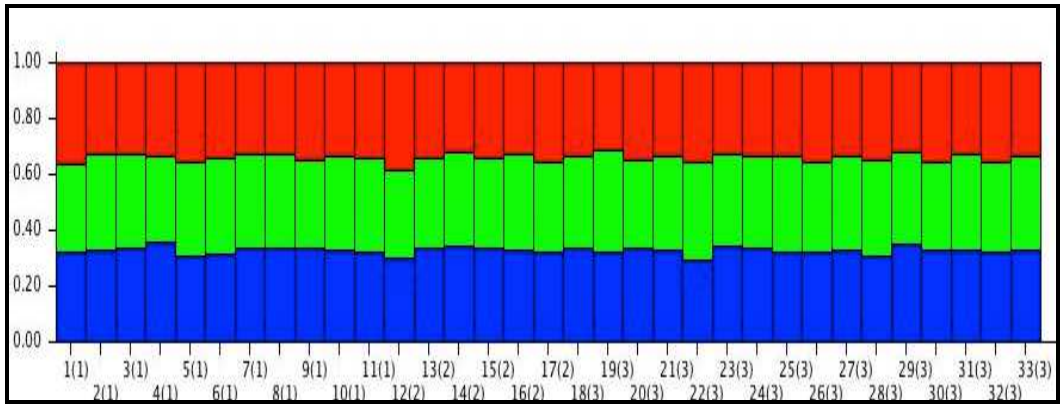


Figure 23: Bar plot representing various populations where each fox shares its proportion in each population. X-axis represents the individuals and Y-axis represents F_{st} values.

Bar plot (Figure 23) from the STRUCTURE program proved to vary slightly from the bar plotted using the original data without the non-neutral marker (REN169018). For example individual, RR1000 on lane 1 has shares high genetic variation in cluster 3 (blue) compared to 1 (red) and 2 (green). Classification of the foxes was fixed depending on the highest proportion shared by each individual in a group. Genetic variation among the groups and between the groups were analysed using Arlequin software.

Table 14 Matrix of significant F_{st} (P) values. Where significance level=0.05, using Distance method. '+' Indicates significant and '-' indicates non-significant.

	Group 1	**Group 2**	**Group 3**
Group 1	*		
Group 2	0.03604+-0.0278 (+)	*	
Group 3	0.00000+-0.0000 (+)	0.01802+-0.0121 (+)	*

Table 14 illustrates that all the three groups were genetically variable to each other; thereby all the three groups were considered as a population.

Table 15 Amova results from Arlequin software (average over 14 loci) illustrating the genetic variation.

Source of variation	d.f.	Sum of squares	Variance components	Percentage of variation
Among populations	2	17.296	0.14822 Va	2.70
Among individuals within populations	30	168.355	0.26804 Vb	4.88
Within individuals	33	167.500	5.07576 Vc	92.42
Total	65	353.152	5.49202	

Fixation Indices

FIS :	0.05016
FST :	0.02699
FIT :	0.07579

Amova results (table 15) from Arlequin software proved that they did not vary from the original data without the marker under balanced selection by illustrating that the genetic variation within all the individuals was predominant to the individuals within the population and among the populations (Percentage of variation). The percentage of genetic variation

within the individuals was 92.42%; whereas among the individuals within populations and among population showed only 4.88% and 2.70% variation respectively. Inbreeding co-efficient (FIS) was not significantly observed. F_{st} value (0.0299) proves that the genetic variation among the individuals were significant.

Table 16 Illustration of significant Linkage disequilibrium between the loci. + Indicates significant linkage disequilibrium and – indicates non-significant linkage disequilibrium. Locus 0: AHT121, 1:AHT133, 2:AHT137, 3:AHT142, 4:CPH19, 5:CPH7, 6:CXX374, 7:CXX468, 8:FH2054, 9:FH2328 10:FH2848, 11:REN105LO3, 12:REN162CO4 and 13:REN169018.

Histogram of the number of linked loci per locus

Locus:	0	1	2	3	4	5	6	7	8	9	10	11	12	13
	0	3	1	1	1	0	1	0	1	2	2	0	0	2

Table of significant linkage disequilibrium (significance level=0.0500):

Locus #	0	1	2	3	4	5	6	7	8	9	10	11	12	13
0	*	–	–	–	–	–	–	–	–	–	–	–	–	–
1	–	*	–	–	–	–	–	–	–	+	+	–	–	+
2	–	–	*	–	–	–	–	–	+	–	–	–	–	–
3	–	–	–	*	+	–	–	–	–	–	–	–	–	–
4	–	–	–	+	*	–	–	–	–	–	–	–	–	–
5	–	–	–	–	–	*	–	–	–	–	–	–	–	–
6	–	–	–	–	–	–	*	–	–	–	+	–	–	–
7	–	–	–	–	–	–	–	*	–	–	–	–	–	–
8	–	–	+	–	–	–	–	–	*	–	–	–	–	–
9	–	+	–	–	–	–	–	–	–	*	–	–	–	+
10	–	+	–	–	–	–	+	–	–	–	*	–	–	–
11	–	–	–	–	–	–	–	–	–	–	–	*	–	–
12	–	–	–	–	–	–	–	–	–	–	–	–	*	–
13	–	+	–	–	–	–	–	–	–	+	–	–	–	*

Addition of the REN169018 (Table 16), which were considered as non-neutral marker proved to have its effect on the analysis by illustrating that loci pairs such as AHT133-FH2328, AHT133-FH2848, AHT133-REN169018, AHT137-FH2054, AHT142-CPH19, CXX374-FH2848, and FH2328-REN169018 exhibited significant linkage disequilibrium.

- **Adjusted Genotypic Data** (where null alleles were considered as Recessive alleles): The Structure program verified that foxes under 3 clusters proved comparatively better variant than in 1 or 2 clusters, as illustrated in Figure 24.

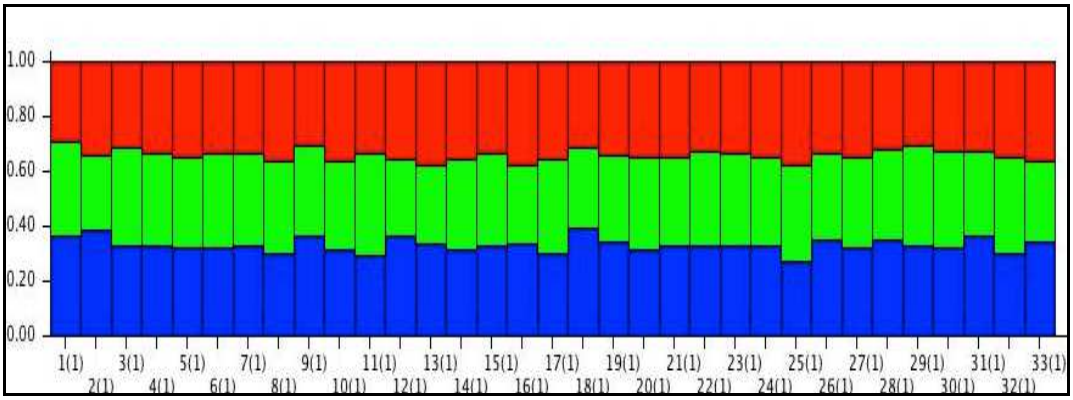


Figure 24: Bar plot representing various populations where each fox has its own share in it.

The X-axis in the bar plots of Figure 24 represents the individuals in the study and the Y-axis represents the F_{st} ranges. Each fox shares its own genetic proportion in every cluster. For example Fox 1 (RR1000)'s genetic proportion appears to be predominant 3rd cluster (blue) rather than in 1st (red) and 2nd cluster (green). Classification of the foxes was fixed depending on the highest proportion shared by each individual in a group.

Genetic variation among the groups and between the groups were analysed using Arlequin software.

Table 17 Matrix of significant F_{st} (P) values. Where significance level=0.05, using Distance method. '+' Indicates significant and '-' indicates non-significant.

	Group 1	**Group 2**	**Group 3**
Group 1	*		
Group 2	1.00000+-0.0000 (-)	*	
Group 3	1.00000+-0.0000 (-)	1.00000+-0.0000 (-)	*

Every group in the table 17 did not prove to vary from each other genetically.

Table 18 Illustration of significant Linkage disequilibrium between the loci. Locus 0: AHT121, 1:AHT133, 2:AHT137, 3:AHT142, 4:CPH19, 5:CPH7, 6:CXX374, 7:CXX468, 8:FH2054, 9:FH2328, 10: FH2848, 11:REN105LO3, 12:REN162CO4.

Histogram of the number of linked loci per locus

Locus:	0	1	2	3	4	5	6	7	8	9	10	11	12
	0	0	2	1	1	0	2	1	0	2	0	0	1

Table of significant linkage disequilibrium (significance level=0.0500):

Locus #	0	1	2	3	4	5	6	7	8	9	10	11	12
0	*	-	-	-	-	-	-	-	-	-	-	-	-
1	-	*	-	-	-	-	-	-	-	-	-	-	-
2	-	-	*	-	-	-	-	+	-	+	-	-	-
3	-	-	-	*	+	-	-	-	-	-	-	-	-
4	-	-	-	+	*	-	-	-	-	-	-	-	-
5	-	-	-	-	-	*	-	-	-	-	-	-	-
6	-	-	-	-	-	-	*	-	-	+	-	-	+
7	-	-	+	-	-	-	-	*	-	-	-	-	-
8	-	-	-	-	-	-	-	-	*	-	-	-	-
9	-	-	+	-	-	-	+	-	-	*	-	-	-
10	-	-	-	-	-	-	-	-	-	-	*	-	-
11	-	-	-	-	-	-	-	-	-	-	-	*	-
12	-	-	-	-	-	-	+	-	-	-	-	-	*

The analysis of genotype data when they were adjusted with the null alleles (table 18) differed widely from the original data (table 13) with respect to LD. Like none of the loci proved to be in LD significantly in the original genotype data, whereas AHT137-CXX468, AHT137-FH2328, AHT142-CPH19, CXX374-FH2328, CXX374-REN162CO4 were the pair of loci in adjusted genotype data that exhibited significant linkage disequilibrium.

- **Adjusted Genotypic Data** (where null alleles were considered as recessive alleles but with REN169018 marker): The Structure program indicates that foxes under 3 clusters proved comparatively better variant than in 1 or 2 clusters, as illustrated in Figure 25.

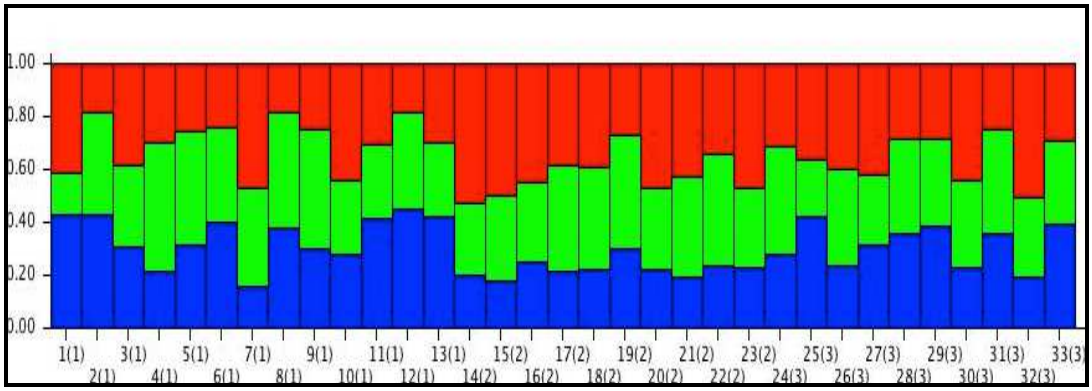


Figure 25: Bar plot representing the two populations where each foxes shares their proportions in the populations

The structure of the individuals (Figure 25) differed widely from the groups with respect to the structure designed for neutral alleles alone. For example RR1000 fox in the lane 1 shares high genetic variation in cluster 3 (blue), then with cluster 1(red) and very low proportion in cluster 2 (green).

Genetic variation among the groups and between the groups were analysed using Arlequin software.

Table 19 Matrix of significant F_{st} (P) values. Where significance level=0.05, using Distance method. '+' Indicates significant and '-' indicates non-significant.

	Group 1	**Group 2**	**Group 3**
Group 1	*		
Group 2	0.15315+-0.0305 (-)	*	
Group 3	0.68468+-0.0594 (-)	0.61261+-0.0485 (-)	*

Genetic variation between the groups (table 19) proved to be non-significant with each other.

Table 20: Illustration of Linkage disequilibrium between the loci. Results collected from Arlequin program. Locus 0: AHT121, 1:AHT133, 2:AHT137, 3:AHT142, 4:CPH19, 5:CPH7, 6:CXX374, 7:CXX468, 8:FH2054, 9:FH2328 10:FH2848, 11:REN105LO3, 12:REN162CO4, 13:REN169018.

Histogram of the number of linked loci per locus

Locus:	0	1	2	3	4	5	6	7	8	9	10	11	12	13
	0	0	2	1	1	1	2	2	0	2	0	0	1	0

Table of significant linkage disequilibrium (significance level=0.0500):

Locus #	0	1	2	3	4	5	6	7	8	9	10	11	12	13
0	*	–	–	–	–	–	–	–	–	–	–	–	–	–
1	–	*	–	–	–	–	–	–	–	–	–	–	–	–
2	–	–	*	–	–	–	–	+	–	+	–	–	–	–
3	–	–	–	*	+	–	–	–	–	–	–	–	–	–
4	–	–	–	+	*	–	–	–	–	–	–	–	–	–
5	–	–	–	–	–	*	–	+	–	–	–	–	–	–
6	–	–	–	–	–	–	*	–	–	+	–	–	+	–
7	–	–	+	–	–	+	–	*	–	–	–	–	–	–
8	–	–	–	–	–	–	–	–	*	–	–	–	–	–
9	–	–	+	–	–	–	+	–	–	*	–	–	–	–
10	–	–	–	–	–	–	–	–	–	–	*	–	–	–
11	–	–	–	–	–	–	–	–	–	–	–	*	–	–
12	–	–	–	–	–	–	+	–	–	–	–	–	*	–
13	–	–	–	–	–	–	–	–	–	–	–	–	–	*

Addition of the non-neutral marker (table 20) appears not to have any effect on Linkage Disequilibrium on the adjusted genotype data (table 18). Loci AHT137-CXX468, AHT137-FH2328, AHT142-CPH19, CPH7-CXX468, CXX374-FH2328, CXX374-REN162CO4, CXX468-AHT142 were the pair of loci that exhibited significant linkage disequilibrium.

- **Genotypic Data without null allele markers**: The Structure program showed effective group of foxes under 2 clusters only, Figure 26.

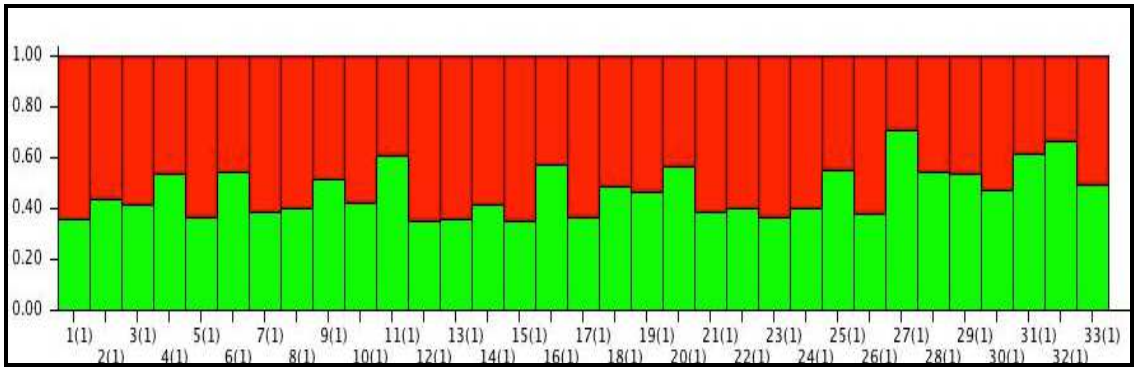


Figure 26: Bar plot representing the two populations where each foxes shares their proportions in the populations

The X-axis in the bar plots of Figure 26 represents the individuals in the study and the Y-axis represents the F_{st} ranges. Each fox shares its own genetic proportion in every cluster.

Like the fox RR1000 in the Lane 1 are predominant in cluster 2 (green rather than cluster 1 (red Classification of the foxes was fixed depending on the highest proportion shared by each individual in a group.

Genetic variation among the groups and between the groups were analysed using Arlequin software.

Table 21 Matrix of significant F_{st} (P) values. Where significance level=0.05, by Distance method. '+' Indicates significant and '-' indicates non-significant.

	Group 1	**Group 2**
Group 1	*	
Group 2	0.03062 (+)	*

Table 22: Illustration of Linkage disequilibrium between the loci. Results collected from Arlequin program. Locus 0: AHT121, 1:AHT137, 2:AHT142, 3:CPH19, 4:CPH7, 5:CXX468, 6:FH2054, 7:FH2328, 8:REN105LO3, 9:REN162CO4.

Histogram of the number of linked loci per locus

Locus:	0	1	2	3	4	5	6	7	8	9
	0	0	0	0	0	0	0	0	0	0

Table of significant linkage disequilibrium (significance level=0.0500):

Locus #	0	1	2	3	4	5	6	7	8	9
0	*	-	-	-	-	-	-	-	-	-
1	-	*	-	-	-	-	-	-	-	-
2	-	-	*	-	-	-	-	-	-	-
3	-	-	-	*	-	-	-	-	-	-
4	-	-	-	-	*	-	-	-	-	-
5	-	-	-	-	-	*	-	-	-	-
6	-	-	-	-	-	-	*	-	-	-
7	-	-	-	-	-	-	-	*	-	-
8	-	-	-	-	-	-	-	-	*	-
9	-	-	-	-	-	-	-	-	-	*

Table 21 and 22 justifies that the there was no effect on the genetic analysis of the data after the elimination of null alleles from the original genotype data.

- **Genotypic Data without null allele markers**: The Structure program showed effective group of foxes under 2 clusters only, Figure 27.

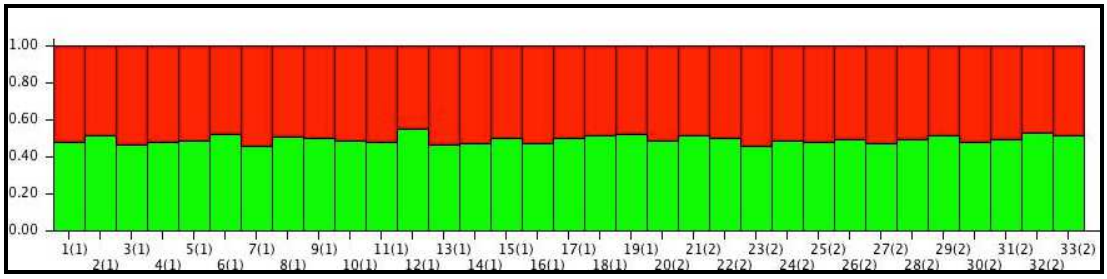


Figure 27: Bar plot representing the two populations where each foxes shares their proportions in the populations

Bar plot shows that the individuals did not shows significant variation in any of the clusters; thereby it was considered that all the individuals belong to a single structure.

4.6. Relationship Analysis

- ML-Relate program structured the relationship between the individuals into a matrix with the help of original genotyping data from the Gene-Mapper version 4.0.

	RR1000	RR1004	RR1024	RR1027	RR1030	RR1035	RR1038	RR1047	RR1050	RR1019	RR1009	RR1005	RR1006	RR1018	RR1021	RR1039	RR1044	RR1007	RR1022	RR1023	RR1026	RR1028	RR1029	RR1031	RR1032	RR1033	RR1034	RR1037	RR1041	RR1046	RR1048	RR1025	RR1042
RR1000	-																																
RR1004	HS	-																															
RR1024	U	U	-																														
RR1027	U	U	HS	-																													
RR1030	HS	HS	U	U	-																												
RR1035	U	U	U	HS	U	-																											
RR1038	U	U	U	U	U	HS	-																										
RR1047	U	U	U	U	U	U	U	-																									
RR1050	U	U	U	U	U	U	U	U	-																								
RR1019	U	U	U	U	U	U	U	FS	HS	-																							
RR1009	U	U	U	U	HS	U	U	U	U	U	-																						
RR1005	U	U	U	U	U	HS	U	U	U	U	U	-																					
RR1006	U	U	U	U	U	U	U	U	U	U	U	U	-																				
RR1018	U	U	U	U	U	U	U	U	U	U	U	HS	U	-																			
RR1021	U	U	U	U	U	U	U	U	U	U	HS	U	U	U	-																		
RR1039	U	U	U	U	U	U	U	U	U	U	U	U	U	U	HS	-																	
RR1044	U	U	HS	U	U	U	U	U	U	U	U	U	U	U	U	U	-																
RR1007	U	U	U	HS	U	U	U	U	U	U	U	U	U	U	U	U	U	-															
RR1022	U	U	U	U	U	U	U	U	U	U	HS	U	U	U	U	U	U	U	-														
RR1023	U	U	U	U	U	U	U	U	U	U	U	U	U	U	U	U	U	U	U	-													
RR1026	U	U	U	HS	U	U	U	U	U	U	U	U	U	U	U	U	U	U	HS	U	U	-											
RR1028	U	U	U	U	U	U	U	U	U	U	U	U	U	U	U	U	U	U	U	U	U	U	-										
RR1029	U	U	U	U	HS	U	U	U	U	U	U	U	U	U	U	U	U	U	U	U	U	U	U	-									
RR1031	HS	U	U	U	U	U	U	U	U	U	U	U	U	U	U	U	U	U	HS	HS	U	HS	U	-									
RR1032	U	U	U	U	U	U	U	U	U	U	U	U	U	U	U	U	U	U	U	U	U	U	U	U	-								
RR1033	U	HS	U	U	U	U	U	U	U	U	U	U	U	U	U	U	U	U	U	U	U	U	U	U	U	-							
RR1034	U	HS	U	U	U	U	U	U	U	U	U	U	U	HS	U	U	U	U	U	U	U	U	U	U	U	U	-						
RR1037	U	U	U	U	U	U	U	U	U	U	U	U	U	U	U	U	U	U	U	HS	U	U	U	HS	U	U	U	-					
RR1041	U	U	U	U	U	U	U	U	U	U	U	U	U	HS	U	U	U	U	U	HS	U	U	U	U	U	U	U	U	-				
RR1046	U	U	U	U	U	U	U	U	U	U	U	U	U	U	U	U	U	U	HS	HS	U	HS	U	U	U	U	U	U	U	-			
RR1048	U	U	U	U	U	U	FS	U	U	U	U	U	U	U	U	U	U	U	U	U	U	U	U	U	U	U	U	U	U	U	-		
RR1025	U	U	U	U	U	U	U	U	U	U	U	HS	U	U	U	U	U	U	U	U	HS	U	U	U	U	U	U	U	U	U	U	-	
RR1042	U	U	U	U	U	U	U	HS	U	U	U	U	U	U	U	U	U	HS	U	U	U	U	U	U	HS	U	U	U	U	U	U	HS	-

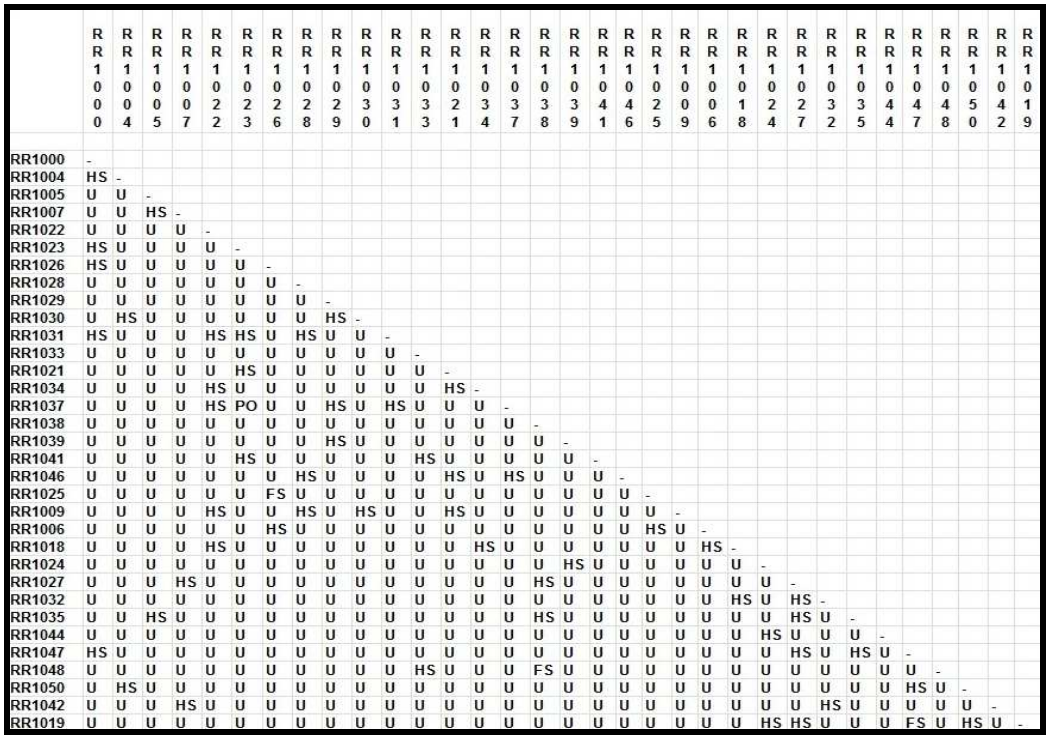
Figure 28: Represents the matrix illustrating the relationship between the individuals using original genotyping data. HS refers to Half-sibling, FS refers to Full sibling, PO refers Parent-offspring relationship and U represents 'no' relationship.

Table 23 Individuals showing half sibling relationship from the matrix, Figure 28

Individuals	HS with
RR1000	RR1004 RR1030 RR1031
RR1004	RR1030 RR1033 RR1034
RR1024	RR1027 RR1044
RR1027	RR1035 RR1007 RR1026
RR1030	RR1009 RR1029
RR1035	RR1038 RR1005
RR1047	RR1042
RR1050	RR1019
RR1009	RR1021 RR1022
RR1005	RR1018 RR1025
RR1018	RR1034
RR1021	RR1039 RR1041
RR1007	RR1026 RR1042
RR1022	RR1031 RR1046
RR1023	RR1031 RR1037 RR1041 RR1046
RR1026	RR1025
RR1028	RR1031 RR1046
RR1031	RR1037
RR1032	RR1042
RR1025	RR1042

From relationship matrix (Figure 28), many foxes appear to be half siblings (table 23) and four foxes have been observed to be in sisters or brother-sister relationship. RR1047 – RR1009 and RR1048 – RR1038 were observed to be full-siblings. These results did not vary with the relationship results obtained using genotyping data with adjusted null allele markers (Figure A8 and table A9).

- ML-Relate program structured the relationship between the individuals into a matrix using genotyping data without null allele markers from the Gene-Mapper version 4.0



	RR1000	RR1004	RR1005	RR1007	RR1022	RR1023	RR1026	RR1028	RR1029	RR1030	RR1031	RR1033	RR1021	RR1034	RR1037	RR1038	RR1039	RR1041	RR1046	RR1025	RR1009	RR1006	RR1018	RR1024	RR1027	RR1032	RR1035	RR1044	RR1047	RR1048	RR1050	RR1042	RR1019
RR1000	-																																
RR1004	HS	-																															
RR1005	U	U	-																														
RR1007	U	U	HS	-																													
RR1022	U	U	U	U	-																												
RR1023	HS	U	U	U	U	-																											
RR1026	HS	U	U	U	U	U	-																										
RR1028	U	U	U	U	U	U	U	-																									
RR1029	U	U	U	U	U	U	U	U	-																								
RR1030	U	HS	U	U	U	U	U	U	HS	-																							
RR1031	HS	U	U	U	HS	HS	U	HS	U	U	-																						
RR1033	U	U	U	U	U	U	U	U	U	U	U	-																					
RR1021	U	U	U	U	U	HS	U	U	U	U	U	U	-																				
RR1034	U	U	U	U	HS	U	U	U	U	U	U	U	HS	-																			
RR1037	U	U	U	U	HS	PO	U	U	HS	U	HS	U	U	U	-																		
RR1038	U	U	U	U	U	U	U	U	U	U	U	U	U	U	U	-																	
RR1039	U	U	U	U	U	U	U	U	HS	U	U	U	U	U	U	U	-																
RR1041	U	U	U	U	U	HS	U	U	U	U	U	HS	U	U	U	U	U	-															
RR1046	U	U	U	U	U	U	U	HS	U	U	U	U	HS	U	HS	U	U	U	-														
RR1025	U	U	U	U	U	U	FS	U	U	U	U	U	U	U	U	U	U	U	U	-													
RR1009	U	U	U	U	HS	U	U	HS	U	HS	U	U	HS	U	U	U	U	U	U	U	-												
RR1006	U	U	U	U	U	U	HS	U	U	U	U	U	U	U	U	U	U	U	U	HS	U	-											
RR1018	U	U	U	U	HS	U	U	U	U	U	U	U	U	HS	U	U	U	U	U	U	U	HS	-										
RR1024	U	U	U	U	U	U	U	U	U	U	U	U	U	U	U	U	HS	U	U	U	U	U	U	-									
RR1027	U	U	U	HS	U	U	U	U	U	U	U	U	U	U	U	HS	U	U	U	U	U	U	U	U	-								
RR1032	U	U	U	U	U	U	U	U	U	U	U	U	U	U	U	U	U	U	U	U	U	U	HS	U	HS	-							
RR1035	U	U	HS	U	U	U	U	U	U	U	U	U	U	U	U	HS	U	U	U	U	U	U	U	U	HS	U	-						
RR1044	U	U	U	U	U	U	U	U	U	U	U	U	U	U	U	U	U	U	U	U	U	U	HS	U	U	U	-						
RR1047	HS	U	U	U	U	U	U	U	U	U	U	U	U	U	U	U	U	U	U	U	U	U	U	HS	U	HS	U	-					
RR1048	U	U	U	U	U	U	U	U	U	U	U	HS	U	U	U	FS	U	U	U	U	U	U	U	U	U	U	U	U	-				
RR1050	U	HS	U	U	U	U	U	U	U	U	U	U	U	U	U	U	U	U	U	U	U	U	U	U	U	U	U	HS	U	-			
RR1042	U	U	U	HS	U	U	U	U	U	U	U	U	U	U	U	U	U	U	U	U	U	U	U	U	U	HS	U	U	U	U	-		
RR1019	U	U	U	U	U	U	U	U	U	U	U	U	U	U	U	U	U	U	U	U	U	U	HS	HS	U	U	U	FS	U	HS	U	-	

Figure 29: Relationship between individuals using genotyping data without null alleles

Table 24 Individuals showing half sibling relationship from the matrix, Figure 30

Individuals	HS with
RR1000	RR1004 RR1023 RR1026 RR1031 RR1047
RR1004	RR1030 RR1050 RR1007 RR1035
RR1005	RR1027 RR1042
RR1007	RR1027 RR1042
RR1022	RR1031 RR1034 RR1037 RR1009 RR1018
RR1023	RR1031 RR1021 RR1041
RR1026	RR1006
RR1028	RR1031 RR1046 RR1009
RR1029	RR1030 RR1037 RR1039
RR1030	RR1009
RR1031	RR1037
RR1033	RR1041 RR1048
RR1021	RR1034 RR1046 RR1009
RR1034	RR1018
RR1037	RR1046
RR1038	RR1027 RR1035

RR1039	RR1024
RR1025	RR1006
RR1006	RR1018
RR1018	RR1032
RR1024	RR1044 RR1019
RR1027	RR1032 RR1035 RR1047 RR1019
RR1032	RR1042
RR1035	RR1047
RR1047	RR1050
RR1050	RR1019

From relationship matrix (Figure 29), many foxes appear to be half siblings (table 24) and four foxes have been observed to be in sisters or brother-sister relationship. RR1047 – RR1009 and RR1048 – RR1038 were observed to be full-siblings. Individuals RR1023-RR1037 exhibit parent-offspring relationship.

5. Discussion

Population genetic study was successful with the established microsatellite-based genotyping pipeline. DNAs were successfully isolated from the tissue, hair and scat samples of the red foxes. Primers were designed for the existing markers and grouped into two multiplexes. The DNA samples were amplified in the multiplex PCR and genotyped. The genotyping data were analysed and studied. The study revealed that although the red foxes were widely prevalent in Scandinavia their genetic deviation among the groups did not meet the expected level of variation. A shallow population structure with low F_{st} value was studied, yet the number of individuals engaged for this study was quite low to justify a significant genetic variation.

5.1. Isolation of DNA

DNAs were isolated successfully from tissues; hair and fresh scat samples of the red fox, but were not same with the silica-dried scat samples. As the scat contains large amount of debris that can inhibit the isolation of DNA, extraction became a challenge. DNA isolation from dried samples could have been improved with the addition of carrier RNA. The carrier RNA reduces the acid-dissociation constant of the silica membrane and also electrostatic repulsion between the DNA and the column, thereby induce the adsorption of DNA to the column (Ram et al., 2006).

5.2. Microsatellite Primers

One of the objectives of this study included the design of microsatellite primers for the established markers. Primers for 23 microsatellite markers with one sex marker (K9AMELO) were designed and tested on tissue sample, RR1042; but only 15 microsatellite markers with one sex marker worked at different levels. By the alteration of primer length, changing the primer concentration, adjusting their melting temperatures and GC concentration newly designed primers were induced to work. The concentration of $MgCl_2$ and BSA was also increased to achieve successful amplification. The Mg^{++} cation in $MgCl_2$ increases the specificity of the target to be amplified, whereas, the BSA (bovine serum albumin) increases the availability of the DNA polymerase for the PCR reaction by getting adsorbing on the surface of the PCR tube. BSA also adds resistance towards the inhibitors (Polymerase Chain Reaction (PCR), 2004). As the annealing temperature ensures the binding properties of the PCR primer pairs, it was adjusted to 60 °C after several trials.

The number of PCR cycles was reduced from 45 to 35 as the concentration of the template DNA was high and also to induce the processivity of the reaction.

Successful amplifications were illustrated on tissue sample, RR1025 and fresh scat sample, Sc1042 in Figure 10 and 11 respectively. The amplified products separated between the 100bp to 400bp against the 100bp DNA ladder according to their respective sizes.

For dried scat sample, the amplification was observed to be weak and very few primers have worked; which could be the result of low DNA concentration and quality as well as presence of PCR inhibitors. The preservation of the scat samples has also been proved to have its impact on the extraction of DNA. Fresh samples preserved in 70% ethanol yielded better quality and quantity of DNA compared to silica-dried samples. This could be the result of DNA degradation caused during high drying pressure. Moreover, since these samples were collected during summer, chances of DNA degradation were very high. Lucchini et al., (2002) also verified that the quality of the faecal DNA extracts were higher in samples that were freshly collected on snow in winter than in samples that were older or collected in summer.

It was also observed that only K9AMELO primer showed successful amplification on the dried-scat DNA sample Sc407 (Figure 12), whereas other primers except K9AMELO showed signals while genotyping the same dried-scat sample using the multiplexes (Figure A6). This could probably relate to different PCR conditions especially different primer concentration applied in both multiplex and simplex PCR.

5.3. Genotype Analysis

Primers were grouped successfully into two different multiplexes. The primer concentrations were adjusted not only to achieve successful genotyping but also to give better intensity of the fluorescent signal. The main purpose to obtain better signal was to induce good clarity of the peak tops, thereby there wont be an issue to check for true alleles.

Thirty-three foxes were successfully genotyped and their alleles sizes for each microsatellite loci were recorded for further analysis. The output depended on not only on the PCR amplification but also on the quality and concentration of the template DNA. As the documents proved that some of the hair/ scat and tissues were obtained from the same fox, the genotype results could not prove their alleles matched exactly. This could be because,

some of the hair strands in the samples lacked follicle, which resulted in low DNA concentration for the analysis and for the scat; extraction was always complicated.

5.4. Sex Differentiation

Awareness of sex of the animals lead to a better understanding of population genetics, as they have an intense influence on the extinction risk of a species as a whole. Genetic variations were studied to arise as a result of an interaction between natural and sexual selection, migration and genetic drift (Ritchie et al., 2007). Likewise, Genetic variation among the foxes also depended upon on the variation among the sexes. The rate of population genetics diversity is proportion to the genetically variable mates. The Arlequin program justified that the male foxes significantly varied from the female foxes.

K9 AMELO proved to be an efficient sex marker from the explicit differentiation of sexes among the foxes. Sex differentiation was implemented on the zygosity of the individuals. The female foxes being homozygous (XX chromosomes) showed X-allele with consistent size, 204, whereas the male foxes illustrated heterozygous alleles X and Y with also consistent sizes 190 and 204. Out of 33 foxes, 12 were female and 21 were observed to be male. The allele sizes differed when Moore et.al. (2010) genotyped the Californian red foxes with K9AMELO, where female homozygous showed 216 bp allele and male heterozygous had 202 bp and 216 bp allele. The difference in the size of allele could have been the result of the adjustment made in the length of K9AMELO to fit into the multiplex.

5.5. Sequencing Analysis

Sequencing results verified the expected microsatellite loci. The resulted hits from the nucleotide blast matched the target loci, thereby proving that the analysis was in the right phase. Sequencing a homozygous template was much easier compared to heterozygous template. The results for a heterozygous template was difficult to interpret, since they were present as mixed trace formed by superimposing of the two allele traits upon each other. The database was not able to extend the information of both direct and reverse mixed sequences of the same template (Dmitry et al., 2008). Heterozygous being a diploid template with insertion or deletion became a challenge for blast to present direct hits for the sequences.

After verification of the results, population genetic diversity was studied based on the genotypic results.

5.6. Population Genetics

15 microsatellite markers were proved to be highly polymorphic averaging 0.722, with He averaging 0.76 (range 0.38-0.87). The number of alleles per locus averaged 7.9 (range = 2-15). These microsatellite markers proved better compared to the established markers of Moore et.al (2010) and Wandeler and Funk (2005) microsatellite markers. Polymorphic average of the established markers were 0.67, with He average 0.72 (range 0.14-0.92) and the number of alleles averaged 5.1 (range= 2-11).

5.6.1. Selection of Markers

Testing for selection formed the basis for the analysis of population genetics data sets in order to avoid non-neutral markers. 'Neutral' genetic markers were the markers for which mutational changes does not change their adaptive fitness within the population whereas 'non-neutral' markers shows adaptive responses to the environmental changes. Thereby it was necessary to use the neutral genetic markers to understand the gene flow in the population. LOSITAN, a web-based was used to eliminate the under selection markers by evaluating the relationship between F_{st} and H_e (expected heterozygosity) in an island model of migration with neutral markers. F_{st} values were not only used for examining genetic basis of adaptation but also for eliminating non-neutral outlier loci from data sets (Tiago et al., 2008).

The LOSITAN model described the distribution of Wright's inbreeding coefficient F_{st} vs. H_e with a pool of neutral markers (grey region). The distribution also identified that Ren247M23 genetic marker (red region) was under positive/direct selection and Ren169018 genetic marker (yellow region) was under balanced selection. As these markers deviated from neutral markers, they were considered to be under selection. Thereby they were removed from data so that they did not affect any further analysis like population genetics, structure and genetic variation among the tested foxes. Moreover as the degree of polymorphism in neutral marker is proportional to the rate of underlying mutation, it adds an important criterion to use neutral markers for the analysis (Ellegren, 2004).

Many researchers also included markers under balanced selection along with neutral markers in their study; thereby REN169018 was included to verify its effect on the analysis (K. Østebye, personal communication).

5.6.2. Identification of Genotyping Errors

Degradation of DNA, primer-site mutations or low DNA concentrations may lead to genotyping errors. Many studies also confirmed that genotyping errors like null alleles, large allele dropouts could mislead or bias the population genetic analysis (Cock et al., 2004). MICRO-CHECKER, windows based software was used to testify the genotype results of the microsatellites. When the genotyping data was loaded in the program, the program illustrated that the observed heterozygosity of the marker was lower than the expected. Three out of fourteen markers proved to have null alleles, as they were not in Hardy-Weinberg equilibrium. The null alleles were automatically adjusted to '0' by the program, which was calculated based on the deviations from Hardy-Weinberg equilibrium using a chi-square goodness-of-fit test. The analysis also used an iterative algorithm on the difference between the observed and expected frequency of the homozygotes to estimate the null allele frequency.

The 'null' alleles under these markers could have existed as 'likely' false homozygous; as it was not sure whether all the alleles under marker were homozygous or only some of them were homozygous. The null allele's frequencies were comparatively weaker than the rest of the alleles; thereby the program adjusted their genotype frequency for the further analysis. The 'null' alleles have been assumed as 'recessive' alleles with their value as '1' in this study.

Occurrence of null alleles could be avoided by increasing the quality of DNA or by designing new primers for the marker. The template sequence could have SNPs at 3' end so adjusting the length of the primers towards 5' end could solve the issue of 'null' alleles.

5.6.3. Genetic Analysis

Population diversity was analyzed on the three different types of data based on various structure designed by the quantitative clustering method in the STRUCTURE program. The program implemented a clustered model with genotype data consisting of unlinked markers. The model assumes that there are K clusters, which were characterized by a set of allele frequencies at each locus. The individuals were assigned to different populations based on genetic variation or admixed property. Within the populations their loci were assumed to be at maximum Hardy-Weinberg equilibrium and minimum linkage disequilibrium.

For the original genotype data, the STRUCTURE program grouped the individual foxes into 3 clusters. Each individual occupied at his or her own genetic proportion in every cluster and they were segregated into their corresponding groups based on their F_{st}. The foxes in each group shared similarity in their loci or distance between their loci on a chromosome. Thereby, further analysis was done to understand the genetic diversity between the clusters with the help of Arlequin software. The results from the program illustrated that the group 3 differs significantly from group 1 and 2. This particular group was considered as a population. The Loci were not significantly in Linkage disequilibrium, but on addition of REN169018 marker in the analysis some of the loci proved Linkage disequilibrium significantly; thereby illustrating the effect of adding non-neutral marker.

For the genotype data that were modified considering the null alleles, the structure program also organized the individuals into 3 clusters with the individuals sharing a different level of proportions compared to the original genotype data. On analysis of these 3 groups on Arlequin software, the groups did not vary significantly with each other. Moreover, the results disclosed that the Loci were significantly in linkage disequilibrium. Addition of the non-neutral marker, REN169018 did not have significant effect on the analysis.

Comparing the original genotype data with the modified data, results from original genotype data looked more promising. Taking Linkage disequilibrium (LD) into account, the results from the original data looked more convincing. Microsatellite markers being highly heterozygous as well as highly mutable with often 15 or more alleles in any given population, regular occurrence of LD could be a question. Studies proved spacing of the markers could illustrate the level of LD across the genome or locus. Therefore in this analysis, as the microsatellite markers were believed to be on different chromosomes the chance of LD to occur was quite low. Even if the alleles were considered to obey the Mendelian inheritance (crossing over principle), over generation occurrence of significant linkage disequilibrium among the loci was a question with respect to mutations.

Observing low observed heterozygosity in the population, the loci showing significant LD could be an artifact. These loci could have significant gametic disequilibrium, which was commonly attributed as LD. Gametic disequilibrium was associated with non-random alleles at different loci and break during recombination, as they have no force to bind them tightly. Inbreeding within the population decreases the heterozygosity, which in turn slows down the decay of gametic disequilibrium. However, the individuals in this population did

not show significant inbreeding co-efficient (FIS=0.06). Thereby considering the loci to have significant gametic disequilibrium could also be a question.

With respect to null alleles, after observing that the adjusted alleles were originally true and could have been morphed as false homozygous, original genotype data analysis appears to be considered over other analysis. Indeed, increasing the number of individuals for the study could provide better clarification of the results.

The analysis on the genotypic data without null alleles were grouped only into 2 clusters, as there were only 10 markers and they proved to vary significantly from each other. On the inclusion of REN169018 marker, the program did not prove significance in grouping of the individuals; thereby all the individuals were considered to be in single population. However, elimination of the null alleles did not have much effect on the analysis.

The study on red foxes in Norway did not vary with the study of farmed and wild red foxes in Poland (Witkowska et al., 2012). The farmed and wild red foxes in Poland exhibited high degree of genetic variability (F_{st} =0.2657) like the red foxes in Norway (F_{st} =0.2699).

5.7. Relationship

Relationship between the foxes was observed using ML-Relate program using original genotype data, adjusted genotype data and genotypic data without null alleles. All the three types of data proved that individuals RR1047 and RR1009, RR1048 and RR1038 were observed to be in sisters and brother-sister relationship, respectively. However, differences were observed among the half-siblings, which could be as a result in the adjustment made in the null allele frequencies. Genotyping data without null allele markers showed that individuals RR1023 and RR1037 were related as parent-offspring. But on observing the genotypic data of these individuals from table A4 (Appendix), it was seen that the alleles did not match with each other proving that they could be wrong. Observing the location from where the samples were collected, for example RR1048 from Stange and RR1038 from Trysil, the foxes may have migrated away from their pack for their generation to survive.

5.8. Future Studies

DNA was successfully isolated from tissue and hair samples of the red foxes but not on silica-dried scat samples. New methods to isolate DNA from dried scat samples could be

designed, for example addition of carrier-RNA to induce DNA concentration. Alternative methods to preserve the collected scat samples can also be deployed. For example, in this study some of the scat samples were preserved in ethanol and some were silica-dried but scat could be preserved by dissolving in DMSO salt solution to prevent from degradation of DNA. Frantzen, et al. (1998) also confirmed that preservation of scat samples in DMSO salt solution proved effective over ethanol and silica-dried preservations. Hair samples yielded less quantity of DNA due to lack of follicles, therefore in the future studies hair samples could be collected by pulling out the hairs directly from the individual rather than cutting them out. 16 Primers were designed to genotype the foxes, the primers could be expanded by designing more. In this study only 33 red foxes were genotyped, for accurate or extensive results more number of foxes could be deployed. For this study all the foxes were used from Norway, it would be more ideal to use other foxes might be from Sweden, Denmark to compare results. The primers showing null alleles could be redesigned to prevent the occurrence of null alleles. Linkage disequilibrium mapping could be focused in the future studies. In this study the DNA samples were extracted from the dead red foxes, in future studies this DNA samples could be extracted from hair or scat of the living foxes instead of from the dead foxes in order to track their genetic transmission of traits over generations and also their migration.

6. Conclusion

The genotypic study on Vulpes *vulpes* was successfully accomplished with the help of 15 microsatellite markers with 1 sex marker. 33 red foxes were analysed and the population diversity, genetic variation and structure were observed, as illustrated in the results. Markers under selection were eliminated and analysis was continued with the genotype data. The markers detected with null alleles were also adjusted and analysed. The individuals were successfully grouped under different clusters. The original genotype data revealed that the individuals differed genetically whereas the adjusted genotype data with null alleles showed no significant variance. Original genotype data looks promising as the loci were not significantly in Linkage Disequilibrium, which can be clarified and verified by increasing the number of individuals for the study.

Bibliography

(n.d.). Retrieved from http://en.wikipedia.org/wiki/Fox

(n.d.). Retrieved from http://www.bbc.co.uk/nature/17270249

(n.d.). Retrieved from http://www.uidaho.edu

(n.d.). Retrieved from http://www.vgl.ucdavis.edu/dogset/help.jsp

Aghanoori, M. R., Vafaei, H., Kavoshi, H., Mohamadi, S., & Goodarzi, H. R. (2012). Sex determination using free fetal DNA at early gestational ages: a comparison between a modified mini-STR genotyping method and real-time PCR. American Journal of Obstetrics and Gynecology , 207 (3), 202 e1-202e8.

Antao, T., Lopes, A., Lopes, R. J., Beja-Pereira, A., & Luikart, G. (2008). LOSITAN: A workbench to detect molecular adaptation based on a Fst -outlier method. BMC Bioinformatics , 9 (323), 1-5.

Bachrach, M. (1953). Fur: a practical treatise.

Bessetti, J. (2007). An Introduction to PCR Inhibitors. Profiles in DNA, 9-10.

Blamire, P. J. (1997-2000). Science at a Distance©.

C, C., & R, L.-B. (2002). SRY and Sex determination: How lazy can it be?. Acta Biologica Columbiana , 17 (1).

Chaboissier, M.-C., Kobayashi, A., Videl, V. I., Lutzkendorf, S., Kant, H. J., Wegner, M., et al. (2004). Functional Analysis of Sox8 and Sox9 during sex determination in the mouse. Development , 131, 1891-1901.

D.N, I., A.L., S., T.T., F., M.L., E., S., H. S., & N.C., P. (2003). Analysis of Genetic Variation in 28 Dog Breed Populations with 100 Microsatellite Markers. Journal of Heredity , 94 (1), 81-87.

Dale, T. F. (1858-1923). Foxes; Hunting.

Davidson, R., Øines, Ø., & Norström, M. (2009). The surveillance and control programme for Echinococcus multilocularis in red foxes (Vulpes vulpes) in Norway. Annual Reports , 1-5.

Dieffenbach, C. W., Lowe, T. M., & Dveksler, G. S. (2010). General concepts for PCR primer design. Genome Res. , S30-S37.

Eenennaam, A. V. (2009). Basics of DNA markers and genotyping .

Ellegren, H. (2004). Microsatellites: Simple sequences with complex evolution. Nat Rev Genet. , 5 (6), 435-445.

Ennis, S., & Gallagher, T. F. (1994). A PCR-based sex-determination assay in cattle based on the bovine amelogenin locus. Animal Genetics , 25, 425-427.

Fabbri, E., Miquel, C., Lucchini, V., Santini, A., Caniglia, R., Duchamp, C., et al. (2007). From the Appenines to the Alps: Colonization genetics of the naturally exxpanding Italian wolf (Canis lupus) population. Mol Ecol. , 16 (8), 1661-1671.

Fedriani, J. M., Fuller, T. K., Sauvajot, R. M., & York, E. C. (2000). Competition and intraguild predation among three sympatric carnivores. Oecologia , 125, 258-270.

Feldhamer, G. A., Thompson, B. C., & Chapman, J. A. (2003). Wild mammals of North America: biology, management, and conservation.

Frantzen, M., Silk, J., Ferguson, J., R.K.Wayne, & M.H.Kohn. (1998). Empirical evolution of preservation methods for faecal DNA. Molecular Ecology , 7, 1423-1428.

Fredholm, M., & Wintero, A. (1994). Variation of short tandem repeats within and between species belonging to the Canidae family. Mammalian Genome , 6, 11-18.

Garzon, F. A., Camacho, N. E., & Sanchez, A. (2012). Sex-Determination systems and their evolution: Mammals. Acta biol. Columb , 17 (1), 3-18.

Handbook, Q. M. (2010). For fast and efficient multiplex PCR without optimization.

Harris. (2008). Mammals of the British Isles.

HEGGLIN, D., BONTADINA, F., CONTESSE, P., GLOOR, S., & DEPLAZES, P. (2007). Plasticity of predation behaviour as a putative driving force for parasite life-cycle dynamics: the case of urban foxes and Echinococcus multilocularis tapeworm. Functional Ecology , 21 (3), 552-560.

Heptner, V. G., & Naumov, N. P. (1998). Mammals of the Soviet Union Vol.II Part 1a, SIRENIA AND CARNIVORA (Sea cows; Wolves and Bears).

Hildebrand, M. (1952). The Integument in Canidae. Mammal , 33, 419-428.

Iwase, M., Satta, Y., Hirai, Y., Hirai, H., Imai, H., & Takahata, N. (2003). The amelogenin loci span an ancient pseudoautosomal boundary in diverse mammalian species. PNAS , 100 (9), 5258-5263.

L, F., M, T., E, P., H.C, H., T.J, R., & V, R. (2008). ZFX and ZFY Gene sequences: Use for Molecular sexing European rabbits, European Broen Hares and Mountain Hares and Perspectives for Sex Determination of other Leporid Species. Genetics , 73-77.

Li, S., Feng, T., Fu, L., Li, Z., Lou, C., Zhang, X., et al. (2012). Pyrosequencing of a short fragment of the amelogenin gene for gender identification. Mol Bio Rep , 39, 6949-6957.

Lloyd, H. G. (1981). The Red Fox.

M, B. L., M, S.-C., C, C.-D., & G, B. (2012). Genetic Diversity of Rabbit Populations in Tunasia Using Microsatellites Markers. World Rabbit Science Association, 31-35.

MacDonald, D. J. (2007). "Red fox (Vulpes vulpes)".

Manunza, A., Zidi, A., Yeghoyan, S., Balteanu, V. A., Carsai3, T. C., Oleg Scherbakov2, O. R., et al. (2013). A High Throughput Genotyping Approach Reveals Distinctive Autosomal Genetic Signatures for European and Near Eastern Wild Boar .

Marcelle Moore, S. K. (n.d.). Red fox microsatellites. Molecular ecology Resources , 2-10.

Mburu, D., & Hannote, O. (2005). A practical approach to microsatellite genotyping with special reference to livestock population genetics.

Merkel, A., & Gemmell, N. J. (2008). Detecting Microsatellites in Genome Data: Variance in Definitions and Bioinformatic Approaches Cause Systematic Bias. Evolutionary Bioinformatics , 1-6.

Min, L., Wang, X. M., Bin, J. W., Yu, H. P., & Huan, W. Z. (2011). Isolation and Characterization of fifteen microsatelite loci in the Tibetan fox (Vulpes ferrilata).

Mueller, U. G., & Wolfenbarger, L. L. (1999). AFLP genotyping and fingerprinting. Trends Ecol Evol. , 14 (10), 389-394.

Nagamine, C. M., Chan, K., Hake, L. E., & Lau, Y.-F. C. (1989). The two candidate testis-determining Y genes (Zfy-1 and Zfy-2) are differentially expressed in fetal and adult mouse tissues. Genes Dev , 4, 63-74.

Nagaoka, T., & Ogihara, Y. (1997). Applicability of inter-simple sequence repeat polymorphisms in wheat for use as DNA markers in comparison to RFLP and RAPD markers. Theor Appl Genet , 94, 597-602.

Nef, S., & Paraada, L. F. (2000). Hormones in male sexual development. Genes Dev. , 14, 3075-3086.

Oishi, T., Uragichi, K., Takahashi, K., & Masuda, R. (2010). Population Structures of the Red Fox (Vulpes vulpes) on the Hokkaido Island, Japan, Revealed by Microsatellite Analysis. Hered , 102 (1), 38-46.

Oosterhout, C. V., Hutchinson, W. F., Wills, D. P., & Shipley, P. (2004). MICROCHECKER: Software for identifying and correcting genotyping errors in microsatellite data. Molecular Ecology Notes , 4, 535-538.

Pailhoux, E., Vigier, B., Schibler, L., Cribiu, E. P., Cotinot, C., & Vaiman, D. (2004). Positional cloning of the PIS mutation in goats and its impact on understanding mammalian sex-differentiation. Genet. Sel.Evol , 37 (1), S55-S64.

Palmer, M. S., Berta, P., Sinclair, A. H., & Goodfellow, B. P. (1990). Comparison of human ZFY and ZFX transcripts. Proc. Natl. Acad. Sci , 87, 1681-1685.

Polymerase Chain Reaction (PCR) . (2004, December). Retrieved from http://www.eeescience.utoledo.edu/Faculty/Sigler/Von_Sigler/LEPR_Protocols_files/PCR.pdf

Rodrigues, K., & S., V. K. (2006). Isolation of 43 Microsatellite Loci from Paphiopedilum rothschildianum, An Endangered Species of Slipper Orchid.

Seiko KA, T. R. (2006). Microsatellites for ecologists: a practical guide to using and evaluating microsatellite markers. Ecology Letters , 9 (5), 615-629.

Selvamani, M. J., Degnan, S. M., & Degnan*, B. M. (2001). Microsatellite Genotyping of Individual Abalone Larvae: Parentage Assignment in Aquaculture. Marine Biotechnology , 3, 478-485.

Shen, Z., Qu, W., Wang, W., Lu, Y., Wu, Y., Li, Z., et al. (2010). MPprimer: a program for reliable multiplex PCR primer design. BMC Bioinformatic , 11:143, 2-7.

Smith, D. A., K.Ralls, Hurt, A., Adams, B., Parker, M., & Maldonado, J. E. (2006). Assessing reliability of microsatellite genotypes from kit fox faecal samples using genetic and GIS analysis. 15 (2), 387-406.

Sun, D. Q., H.Y.Li, Xu, T., & Wang, R. X. (2011). Development and characterization of miicrosatellite markers for the walking goby (Scartelaos viridis; Gobiidae). Genet. Mol. Res. , 10 (1), 203-207.

Takabayashi, S., & Katoh, H. (2011). Sex Identification Using the ZFX and ZFY Genes in Common Marmosets (Callithrix jacchus). Exp. Anim. , 60 (4), 417-420.

Tevosian, S. G., Albrecht, K. H., Crispino, J. D., Fujiwara, Y., Eicher, E. M., & Okin, S. H. (2002). Gonadal differentiation, sex determination and normal Sry expression in mice require direct interaction between transcription partners GATA4 and FOG2. Development , 129, 4627-4634.

Tsai, T. C., Wu, S. H., Chen, H. L., Tung, Y. T., Cheng, W. T., & Chen, J. C. (2011). Identification of sex-specific polymorphic sequences in the goat amelogenin gene for embryo sexing. Journal of Animal Science , 2407-2414.

V., L., E., F., F., M., S., R., L., B., & Randi. (2002). Noninvasive molecular tracking of colonizing wolf (Canis lupus) packs in the western Italian Alps. Molecular Ecology , 11, 857-868.

Vignal, A., Milan, D., Sancristobal, M., & Eggen, A. (2002). A review on SNP and other types of molecular markers and their use in animal genetics. Genet Sel Evol. , 34 (33), 275-305.

Wandeler, P., & Funk, S. M. (2005). Short microsatellite DNA markers for red fox (Vulpes vulpes). Molecular Ecology Notes , 6, 98-100.

Weissenberger, M., Reichert, W., & Mattern, R. (2010). A Multiplex PCR assay to differentiate between dog and red fox. Forensic Science International: Genetics , 5 (5), 411-414.

Williams, J. G., Kubelik, A. R., Livak, K. J., Rafalski, J. A., & Tingey, S. V. (1990). DNA polymorphisms amplified by arbitary primers as genetic markers. Nucleic Acid Research , 18 (22), 6531-6535.

Witkowska, G. J., Horeka, B., Jakubczak, A., Kasperek, K., Slaska, B., Poniewwierska, M. B., et al. (2012). Genetic Variability of Farmed and Free living Populations of Red Foxes (Vulpes vulpes). Ann. Anim. Sci , 12 (4), 501-512.

Appendix

Table A 1 Nano drop measurement of DNA from tissue samples

Sample ID	ng/ul	A260	A280	260/280	260/230
RR1000	366,20	7,324	3,718	1,97	1,70
RR1004	1239,92	24,798	14,626	1,70	1,94
RR1005	514,88	10,298	5,265	1,96	2,05
RR1006	1065,87	21,317	12,077	1,77	2,04
RR1007	491,71	9,834	5,031	1,95	2,05
RR1009	176,80	3,536	1,854	1,91	1,75
RR1009B	45,53	0,911	0,454	2,01	1,05
RR1018	1090,37	21,807	11,589	1,88	1,89
RR1019	568,12	11,362	5,796	1,96	1,95
RR1021	1204,38	24,088	13,910	1,73	1,88
RR1022	471,53	9,431	4,994	1,89	1,95
RR1023	806,97	16,139	9,439	1,71	1,85
RR1024	1684,39	33,688	19,380	1,74	1,98
RR1025	1086,19	21,724	12,663	1,72	1,80
RR1026	943,85	18,877	10,029	1,88	1,94
RR1027	1340,68	26,814	14,540	1,84	1,82
RR1028	937,42	18,748	10,058	1,86	1,77
RR1029	1575,06	31,501	16,311	1,93	1,88
RR1030	1738,19	34,764	19,424	1,79	1,93
RR1031	1490,57	29,811	17,323	1,72	1,61
RR1032	2000,98	40,020	23,538	1,70	1,82
RR1033	1011,63	20,233	11,864	1,71	1,91
RR1034	1236,55	24,731	14,680	1,68	1,78
RR1035	1373,89	27,478	16,521	1,66	1,80
RR1037	1045,44	20,909	11,980	1,75	1,80
RR1038	1438,64	28,773	16,089	1,79	1,91
RR1039	1159,51	23,190	13,402	1,73	1,92
RR1041	4958,12	99,162	58,553	1,69	1,71
RR1042	39,46	0,789	0,414	1,91	0,64
RR1044	1439,77	28,795	15,894	1,81	1,96
RR1046	1207,44	24,149	13,363	1,81	1,86
RR1047	3742,94	74,859	41,055	1,82	2,08
RR1048	1168,65	23,373	13,517	1,73	1,90
RR1050	1466,17	29,323	15,738	1,86	1,89
RR1025Post	717,22	14,344	8,274	1,73	1,86

Table A 2 Nano drop measurement of DNA from the dried scat samples

Sample ID	ng/ul	A260	A280	260/280	260/230
021sc	19,01	0,380	0,230	1,66	-13,32
038sc	10,81	0,216	0,112	1,92	-1,47
212sc	24,56	0,491	0,288	1,70	2,61
216sc	6,34	0,127	0,033	3,84	0,45
222sc	320,74	6,415	4,198	1,53	0,78
226sc	51,72	1,034	0,528	1,96	1,27
244sc	26,31	0,526	0,249	2,11	3,17
252sc	10,53	0,211	0,045	4,72	9,27
255sc	13,58	0,272	0,088	3,09	4,00
405sc	-2,29	-0,046	-0,073	0,62	0,11
407sc	63,25	1,265	0,573	2,21	1,70
418sc	85,84	1,717	0,944	1,82	0,63
515sc	100,61	2,012	1,131	1,78	1,75
612sc	10,33	0,207	0,139	1,49	-5,23
614sc	19,22	0,384	0,305	1,26	0,68
615sc	57,37	1,147	0,676	1,70	1,21
621sc	24,18	0,484	0,311	1,56	2,00

Table A 3 Nano drop measurement of the DNA of hair samples

Sample ID	ng/ul	A260	A280	260/280	260/230
h1018	0,98	0,020	0,075	0,26	0,03
h1021	-0,35	-0,007	-0,010	0,69	-0,08
h1025	6,43	0,129	0,048	2,66	0,64
h1019	12,20	0,244	0,322	0,76	0,15

Table A 4 Genotype data for different tissue samples obtained from Gene-Mapper

MARKERS	MULITIPLEX	RR1000		RR1004		RR1005		RR1006		RR1007		RR1018	
AHT121	1	277		277	281	275	283	275	283	277	279	277	283
AHT133	1	145	147	143	145	143		139		143	145	143	147
AHT137	1	139	143	135	139	137	141	137	139	139	141	139	141
AHT142	1	126	130	128	130	128	132	128		128	136	126	128
CPH19	1	240	242	234		240	242	240	250	234	242	234	244
CPH7	1	154		152	154	154		154		152	154	152	154
CXX374	1	112	114	110	112	112		114		110	112	112	
CXX468	1	84		78	84	80	84	82	84	78	80	82	84
FH2054	1	250	254	254	258	220	234	250	254	246	254	234	254
K9AMELO	1	190	204	190	204	190	204	204		190	204	190	204
FH2328	3	183	185	183	185	193	197	175	181	193	197	163	191
FH2848	3	231	241	229	235	231	233	243		229	245	233	241
REN105L03	3	400	406	400		394	406	404	408	398	404	394	404
REN162C04	3	368	380	384		374	380	374	382	380	382	382	
REN169018	3	209	231	211	227	219	233	207	211	211	227	193	205
REN247M23	3	189	195	193		195		193		171	175	193	

MARKERS	MULITIPLEX	RR1022		RR1023		RR1024		RR1026		RR1027		RR1028	
AHT121	1	277	283	277		275	279	277		275	279	277	281
AHT133	1	139		139	145	143	147	133	143	143		145	145
AHT137	1	137	139	137	143	137	143	137	139	141	143	137	139
AHT142	1	128		126	136	128		128		128		128	132
CPH19	1	240	244	234	248	238	244	242	250	234	246	240	244
CPH7	1	154		154		154		154		154		152	154
CXX374	1	110		112		112		112		112		110	112
CXX468	1	80	84	80	84	82	84	76	84	76	78	84	86
FH2054	1	234	242	234	250	236	250	250	254	240	250	240	244
K9AMELO	1	190	204	190	204	190	204	204		190	204	204	
FH2328	3	183	189	183	195	189	191	179	197	191,181	197	187	191
FH2848	3	241		229	245	241		241	245	241		243	
REN105L03	3	392	394	402	406	398	406	398	400	402	404	396	402
REN162C04	3	372	382	378	382	380	384	374	378	374	382	380	382
REN169018	3	207	211	207	215	195	207	211	217	215	217	211	215
REN247M23	3	181	183	193	195	181	193	179	183	181		93	

MARKERS	MULITIPLEX	RR1029		RR1030		RR1031		RR1032		RR1033		RR1021	
AHT121	1	273	277	277	281	277		273	277	273	285	275	277
AHT133	1	143		145	147	139		139	145	139	145	141	
AHT137	1	135	139	135	137	143		137		137	141	141	143
AHT142	1	126	128	130	132	126	128	128		126	128	128	
CPH19	1	238	244	234		240	242	234	244	234	240	234	
CPH7	1	154		154		154		152	154	154		152	154
CXX374	1	114		114		112		104	116	110		114	
CXX468	1	76	78	76	80	84	86	78		76	80	80	84
FH2054	1	232	236	232	244	236	240	218	254	254	254	234	246
K9AMELO	1	190	204	204		190	204	204		190	204	190	204
FH2328	3	181	187	179	191	183	191	191		185	187	175	179
FH2848	3	229	235	229	241	229	231	241	245	229		229	245
REN105L03	3	392	408	400		392	402	402	404	390	400	406	
REN162C04	3	376	380	376	380	378	382	378	382	378	382	372	378
REN169018	3	207	235	209	215	207	231	207	209	215	227	209	215
REN247M23	3	187	195	181	187	191	195	181	195	187	195	185	193

MARKERS	MULITIPLEX	RR1034		RR1035		RR1037		RR1038		RR1039		RR1041	
AHT121	1	281		275	279	277		277	281	277	279	275	279
AHT133	1	143		141	143	145		141	145	141		141	143
AHT137	1	139	143	141		135	143	137	141	137		137	143
AHT142	1	126	128	126	128	126	128	126	128	126	128	126	128
CPH19	1	244		234	240	234	238	234	246	234	244	234	240
CPH7	1	152	154	154		154		154		154	156	152	154
CXX374	1	110	112	112	118	112		110	116	110	114	112	
CXX468	1	80	84	78	80	76	80	76	84	82		80	
FH2054	1	246	254	236	240	234	254	240	244	232	236	244	
K9AMELO	1	190	204	190	204	190	204	204		204		204	
FH2328	3	179	183	181		175	195	175	183	175	179	189	195
FH2848	3	235	243	243		231	241	243		245		229	245
REN105L03	3	394	408	394	398	392	402	398	404	394	408	390	406
REN162C04	3	378	382	374		382		374		372	380	374	382
REN169018	3	197	211	207	215	205	207	205	211	207	215	207	
REN247M23	3	193		185	193	195		191	193	187	195	181	187

MARKERS	MULITIPLEX	RR1044		RR1046		RR1047		RR1048		RR1050			
AHT121	1	273	281	277		279		273	281	277			
AHT133	1	141	145	139		141	145	141	143	145	147		
AHT137	1	137		143		137	141	137	141	135	137		
AHT142	1	128	130	128		130		128		122	130		
CPH19	1	240	244	234	244	234	240	234	246	234	244		
CPH7	1	154		152	154	152	154	154		152			
CXX374	1	112		110		104	116	110		110			
CXX468	1	80	84	76	80	82	84	76	82	82	84		
FH2054	1	236		234	244	234	240	240	244	234	250		
K9AMELO	1	204		190	204	204		190	204	204			
FH2328	3	175	179	179	187	181	197	163	183,193	183	191		
FH2848	3	233		243	245	229	241	243	245	243			
REN105L03	3	404	406	402	404	402	406	398	406	406	408		
REN162C04	3	382	384	380	382	374	380	378		374			
REN169018	3	215	217	207		209	217	207	211	207			
REN247M23	3	193		195		187	195	193	195	187	195		

Table A 5 Genotype data for different tissue samples obtained from Gene-Mapper

MARKERS	**TISSUE-1025**		**TISSUE-1042**	
AHT121	277	283	277	279
AHT133	143		139	143
AHT137	139	141	137	141
AHT142	124	128	126	136
CPH19	240	242	242	250
CPH7	154		154	
CXX374	104	116	116	
CXX468	76	84	76	78
FH2054	250	258	254	258
K9AMELO	190	204	190	204
FH2328	179		181	191
FH2848	231	245	229	243
REN105L03	404		402	406
REN162C04	374	382	382	
REN169018	187	203	207	215
REN247M23	193	195	193	195
MARKERS	**TISSUE-1019**		**TISSUE-1009**	
AHT121	279		277	
AHT133	141	145	137	141
AHT137	141	143	139	141
AHT142	128	130	128	132
CPH19	234	240	234	250
CPH7	152		152	154
CXX374	110		110	112
CXX468	76	82	80	84
FH2054	250		232	234
K9AMELO	190	204	204	
FH2328	191		191	
FH2848	229		229	241
REN105L03	406		400	402
REN162C04	374	380	372	380
REN169018	209	227	207	227
REN247M23	181		189	

Table A 6 Genotype data for different silica-dried scat samples obtained from Gene-Mapper

MARKERS	MULITIPLEX	scat226		Scat244		Scat407	
AHT121	1						
AHT133	1						
AHT137	1						
AHT142	1	128	130	126	128	128	
CPH19	1						
CPH7	1	152	154				
CXX374	1			114			
CXX468	1	76	88	76	90	76	
FH2054	1						
K9AMELO	1						
FH2328	3	187	193	187	193	171	185
FH2848	3	243		229			
REN105L03	3	394	406	394	406	394	408
REN162C04	3	370	372				
REN169018	3	167	169	197	207	207	213
REN247M23	3	187	195	187	195	187	195

MARKERS	MULITIPLEX	scat418		Scat515		SCAT212	
AHT121	1						
AHT133	1						
AHT137	1						
AHT142	1	126		128		128	
CPH19	1						
CPH7	1						
CXX374	1						
CXX468	1	76	84	76		80	
FH2054	1					246	
K9AMELO	1						
FH2328	3	187	197	187	193		
FH2848	3						
REN105L03	3	394	408	394	406		
REN162C04	3	380		378	380		
REN169018	3	197	207	207	215		
REN247M23	3	187	191	187	193		

Table A 7 Sequencing results after performing blastn

Sequences used for Blastn	**E Value**	**Hits**
ACTACATCATATGTCAAAAAAAGAAAGAAAGAAAAGAAAGAAAG AAAGAAAGAAAGAAAGAAAGAAAGAAAGAAGAAAGAAGAAAGA AAAAGCAGCCTCAAGTCCCACAACTGTCTCCTCCTCTTCAA	4e-40	FH2328 microsatellite
CTTCACACCAAAGGCAGTTGACTCGAGATTATCAGTTAGGGGTCA AAACAAAAACAAACCAAACCAGGCCAAACTTTCCTTGCCACGTTAT TGCGAATGTCCTGCTTTCCACACACACACACCCCCCCCCCCC	1e-51	AHT121 microsatellite
GAGCACTTGCTTGCCTTACTCATTGCAGTTAGGGTTGTAATAAAAG CAGAAACATTGGAGCTATTATCTAATCTAATCTATTAATTATCTATC TATCTATCTATCTATCTATCTATCTATC	4e-50	FH2054 microsatellite
AAGGCTTATGCCTTAAATGTAATAATATGTTCATTTAAACCCAATT TTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTG	1e-40	FH2848 microsatellite
CACACACACACACAGTGTTTATGGCACCCCCATTACTTCCTGTGTG TAGCAATCTCTGTGCCAGGGGCGA	2e-53	CXX374 microsatellite
CCTTCCAGCCACAGCCTCACCAGCCCATTCAGCCACAGCCACCTAT GCACCCTATCCAGCCCCTGCTGCCAGAGCCACCTCTACCTCCGATG TTCCCCATACAGCCCCTTCCCCCCATGCTTCCTGACCTGCCTCTGGA AGCTTAACTCTGGAAGCTTT	1e-51	Canis K9AMELO microsatellite (homoxygous)
TGTATGGGGACATCGGAGGTAGAGGTGGCTCTGGCAGCAGGGG CTGGATAGGGTGCATAGGTGGCTGTGGCTGAATGGGCTGGTGAG GCTGTGGCTGGAAGGGCTGCTGTGCAGGCAGAGGGAGGTTTGGC TGGTGGTGTTGGGTTGGAGTG	3e-51	Canis K9AMELO microsatellite (heterozygous)
GTGTGTGTGTGTGTGTGTGGTGTCTGCGACACAGAAAAGAGCTCT GCAAAAAACAAAAAGGGGTGGGGGGTATA	2e-56	Ren105L03 microsatellite
TTCTACCCCTTTCGATTCTCTTCCTATTTCTCTTTTGTTAGCGTCTAC GCTTGCGCCGATGCTCACGTTTGATTAATAGGTAGACAAAGCAAA GGAGTAGAAGTTCTTAGGTGTTCCTGGGAGATGACAGCATTGCA GGGTATCTGCGTGTGAATGAGTGTGTTGTG	3e-42	CPH19 microsatellite
CTTTGGCACATGATGAACAGTTTCCCTTTCGGGGTTTCAAGAATAA AAAATGAGAGGCTTTAAATCTCGCTAACTTGTGCAAGAGTGTGTG TGGTGTGTGTGTG	6e-43	REN162C04 microsatellite
TATGTTTCTCTCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTG TAAAAGAGAATAGTATAAGGGATTGGCTCACACAGATACAGAAA CTGACGGGGCCTGAGATTTACAAGATGAGGTGGCGAGCTGGAAA CTTGAGAGAGCCAATGGCAC	2e-61	REN247M23 microsatellite
TCTGTAACTGTGTGAGCCAATCCCTTATACTAAATCTCTTTTACACA CACACACACACACACACACACACACACAGAGAGAGAAACATATTA	5e-55	REN169018 microsatellite

CTATGCTGATGGCAATGATCAAGAGGAAGAAGAGTTGGTAACAC AGATA		
GCGCGTGTGTGTGTGTGTGTGTGTGTTACTACGTTAGCCCTCTGG GGAATTATTTCTGTGTTTATGATGGGAAGTATTATGGAAAGTTGT ATA	6e-37	CPH7 microsatellite
GGACCCCACACACACACACACACACACACACTCTCCTCACGACACT AAAATATATA	2e-37	AHT137 microsatellite
GGGTTGGTGGTGTGTGTGTGTCTCTCTCAAGAGATTTGGGAGAG AGAA	4e-50	CXX468 Microsatellite (heterozygous)
AGTATACCTTATTCCTTGCCTACAGTACTTTATCAACTCTTTCCCTA AACACACACACACACACACACACACACACACACACACACACACAAAA TGACAAAGAGGTGAAAACAAGGCTTTTTTTTATATTTTAATTAAAT GCCTAAGGTTCTTCTCACCCAACTTCTAGAATATCTGTGAAATGAG AAAAATGGGGTGGTTGTGTCTTTGGGGAAAAGGGAAAAAAAGTT CTGTGCGCCTGTATGAGTTTATCTATATAAATAAATAGGAGAAAG GGCCGTGTGTCTATGTGAATAA	1e-24	FH2848 microsatellite (heterozygous)
CAGACACACACACACACACACACACACACACACACACACACACAACA AACGCTCAGATAATCTCAACAAGCCCTTTATATAACCCTTGGGGG G	3e-61	AHT142

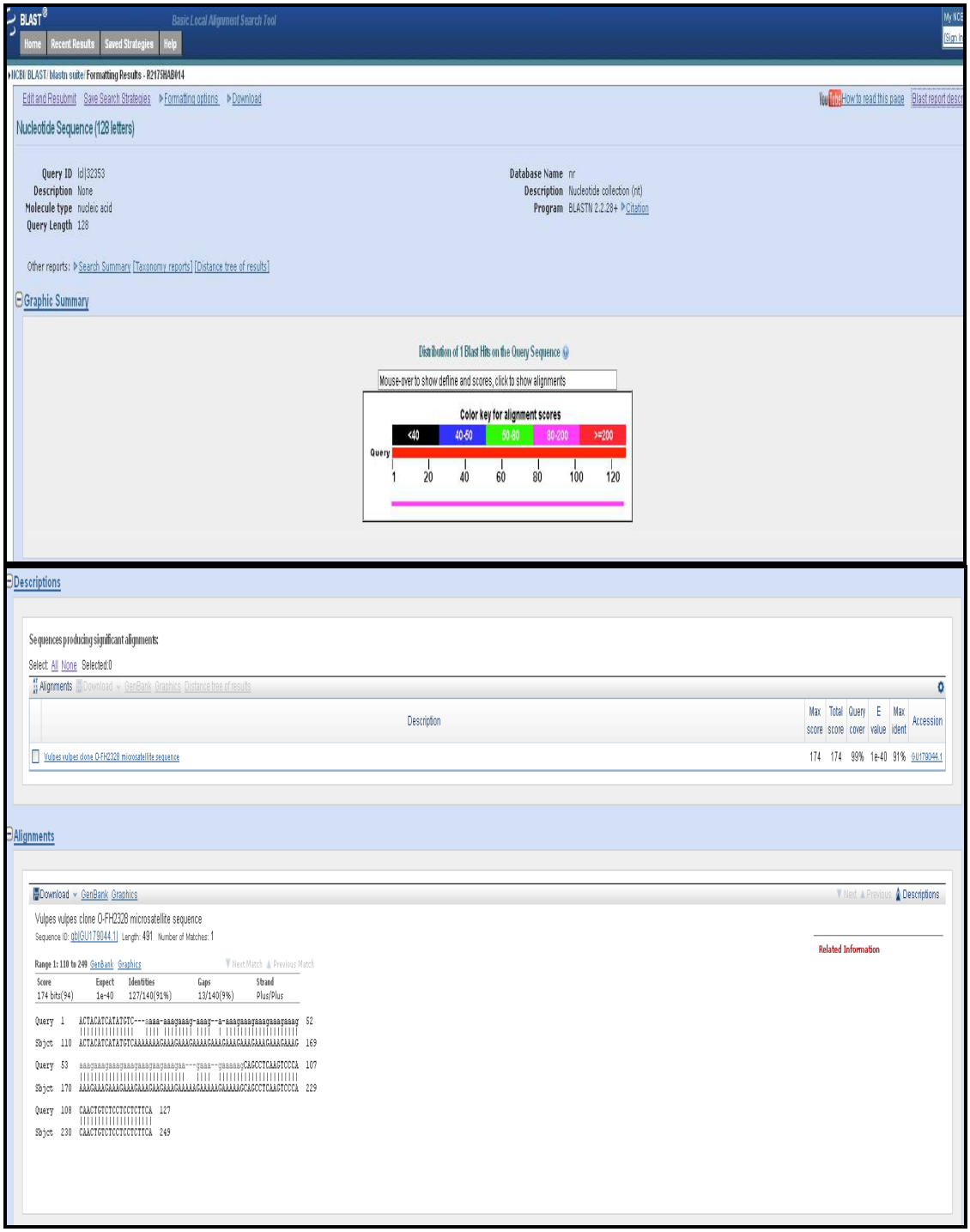


Figure A 2: Output of Blastn from NCBI for the sequence of homozygous marker FH2328 (Figure 19)

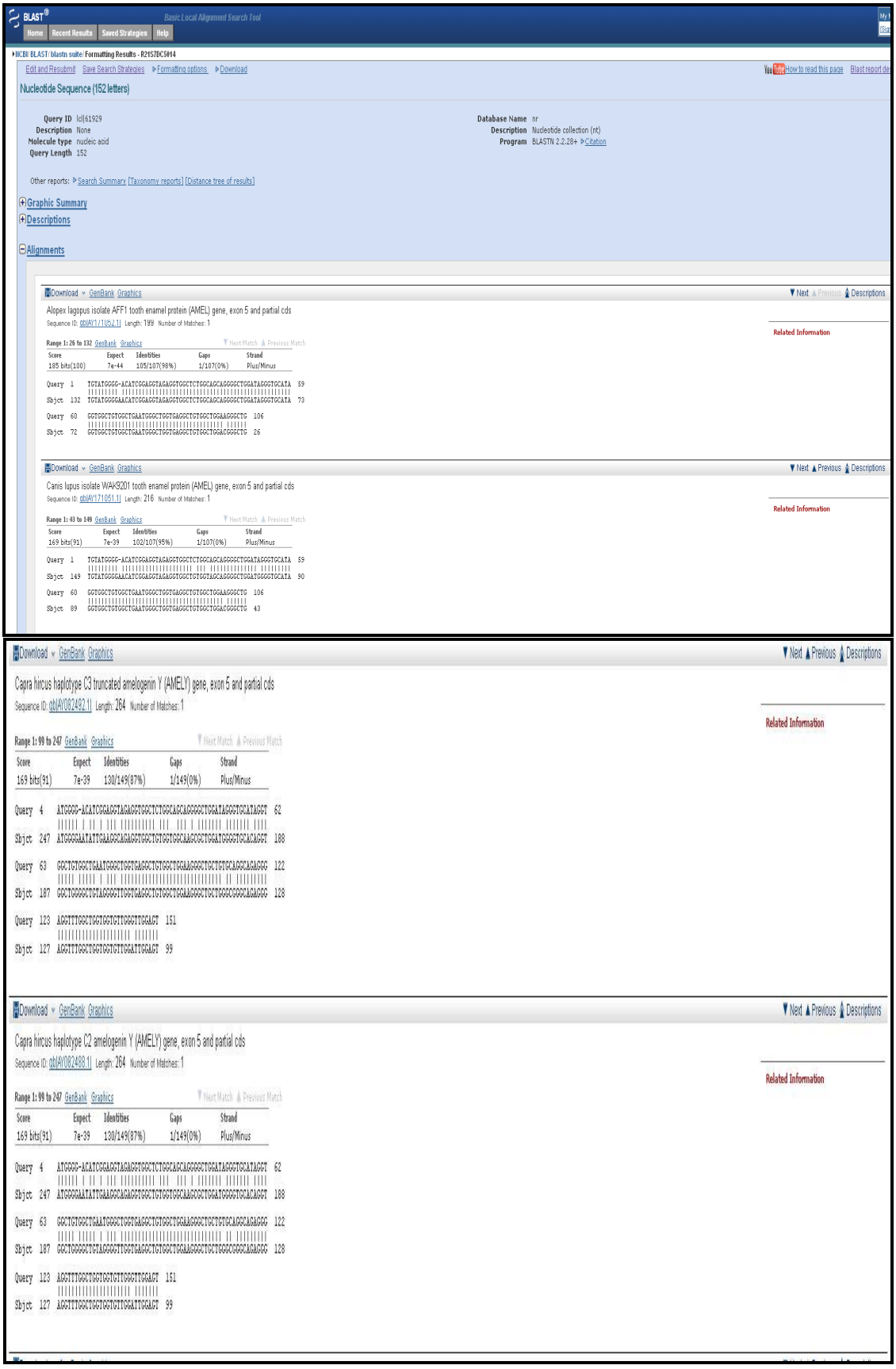


Figure A 3: Output of Blastn from NCBI for heterozygous sequence K9AMELO (Figure 20)

Table A 8 Results from LOSITAN program showing neutral and non-neutral markers. Marker highlighted in yellow falls under balancing selection and marker highlighted in red falls under directed selection (positive).

Locus	Heterozygosity	F_{st}	P(Simulated F_{st} <sample F_{st})
AHT121	0.748264	0.013477	0.671399
AHT133	0.821181	-0.025305	0.104471
AHT137	0.772184	-0.009344	0.394134
AHT142	0.657986	-0.030461	0.20947
CPH19	0.79919	-0.021445	0.195665
CPH7	0.387731	-0.037398	0.257097
CXX374	0.803048	0.031922	0.838042
CXX468	0.810571	0.00352	0.523169
FH2054	0.89159	-0.008594	0.319971
FH2328	0.895833	-0.004827	0.399801
FH2848	0.864005	0.047291	0.955573
REN105L03	0.88831	0.006008	0.595378
REN162C04	0.815008	-0.013208	0.287909
REN169018	0.857832	-0.040857	0.000638
REN247M23	0.858603	0.127273	0.999929

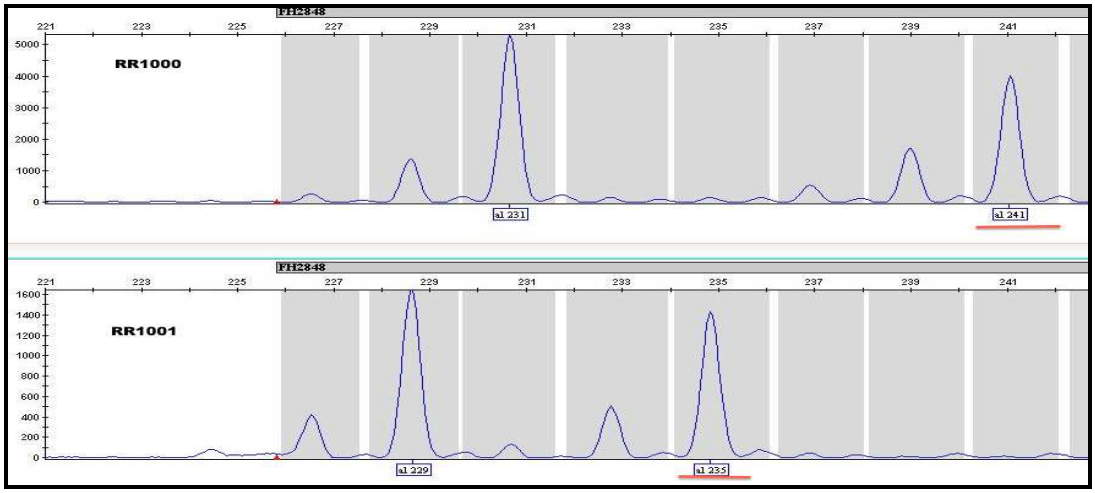
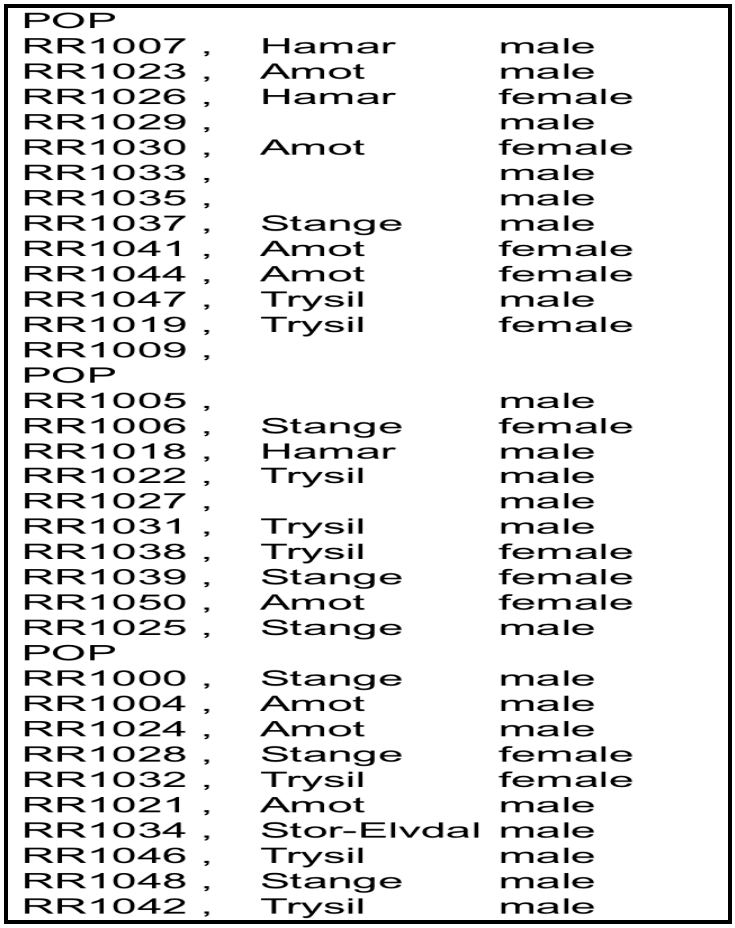


Figure A4: Electropherogram illustrating the alleles of FH2848 Loci of different samples that were considered as null allele by MICRO-CHECKER program



POP		
RR1007 ,	Hamar	male
RR1023 ,	Amot	male
RR1026 ,	Hamar	female
RR1029 ,		male
RR1030 ,	Amot	female
RR1033 ,		male
RR1035 ,		male
RR1037 ,	Stange	male
RR1041 ,	Amot	female
RR1044 ,	Amot	female
RR1047 ,	Trysil	male
RR1019 ,	Trysil	female
RR1009 ,		
POP		
RR1005 ,		male
RR1006 ,	Stange	female
RR1018 ,	Hamar	male
RR1022 ,	Trysil	male
RR1027 ,		male
RR1031 ,	Trysil	male
RR1038 ,	Trysil	female
RR1039 ,	Stange	female
RR1050 ,	Amot	female
RR1025 ,	Stange	male
POP		
RR1000 ,	Stange	male
RR1004 ,	Amot	male
RR1024 ,	Amot	male
RR1028 ,	Stange	female
RR1032 ,	Trysil	female
RR1021 ,	Amot	male
RR1034 ,	Stor-Elvdal	male
RR1046 ,	Trysil	male
RR1048 ,	Stange	male
RR1042 ,	Trysil	male

Figure A 5: Individuals segregated under different clusters in STRUCTURE program using data modified with null-allele

Title line:"foxorig"
AHT121
AHT133
AHT137
AHT142
CPH19
CPH7
CXX374
CXX468
FH2054
FH2328
FH2848
REN105L03
REN162C04
POP1

RR1000,	277277	145147	139143	126130	240242	154154	112114	084084	250254	183185	231241	400406	368380	Stange	male
RR1004,	277281	143145	135139	128130	234234	152154	110112	078084	254258	183185	229235	400400	384384	Amot	male
RR1024,	275279	143147	137143	128128	238244	154154	112112	082084	236250	189191	241241	398406	380384	Amot	male
RR1027,	275279	143143	141143	128128	234246	154154	112112	076078	240250	191197	241241	402404	374382		male
RR1030,	277281	145147	135137	130132	234234	154154	114114	076080	232244	179191	229241	400400	376380	Stange	female
RR1035,	275279	141143	141141	126128	234240	154154	112118	078080	236240	181181	243243	394398	374374		male
RR1038,	277281	141145	137141	126128	234246	154154	110116	076084	240244	175183	243243	398404	374374	Trysil	female
RR1047,	279279	141145	137141	130130	234240	152154	104116	082084	250254	181197	229241	402406	374380	Trysil	female
RR1050,	277277	145147	135137	122130	234244	152152	110110	082084	234250	183191	243243	406408	374374	Amot	female
RR1019,	279279	141145	141143	128130	234240	152152	110110	076082	250250	191191	229229	406406	374380	Amot	male
RR1009,	277277	137141	139141	128132	234250	152154	110112	080084	232234	191191	229241	400402	372380		female

POP2

RR1005,	275283	143143	137141	128132	240242	154154	112112	080084	220234	193197	231233	394406	374380		male
RR1006,	275283	139139	137139	128128	240250	154154	114114	082084	250254	175181	243243	404408	374382	Stange	female
RR1018,	277283	143147	139141	126128	234244	152154	112112	082084	234254	163191	233241	394404	382382	Hamar	male
RR1021,	275277	141141	141143	128128	234234	152154	114114	080084	234246	175179	229245	400406	372378	Amot	male
RR1039,	277279	141141	137137	126128	234244	154156	110114	082082	232236	175179	245245	394408	372380	Stange	female
RR1044,	273281	141145	137137	128130	240244	154154	112112	080084	236236	175179	233233	404406	382384	Amot	female

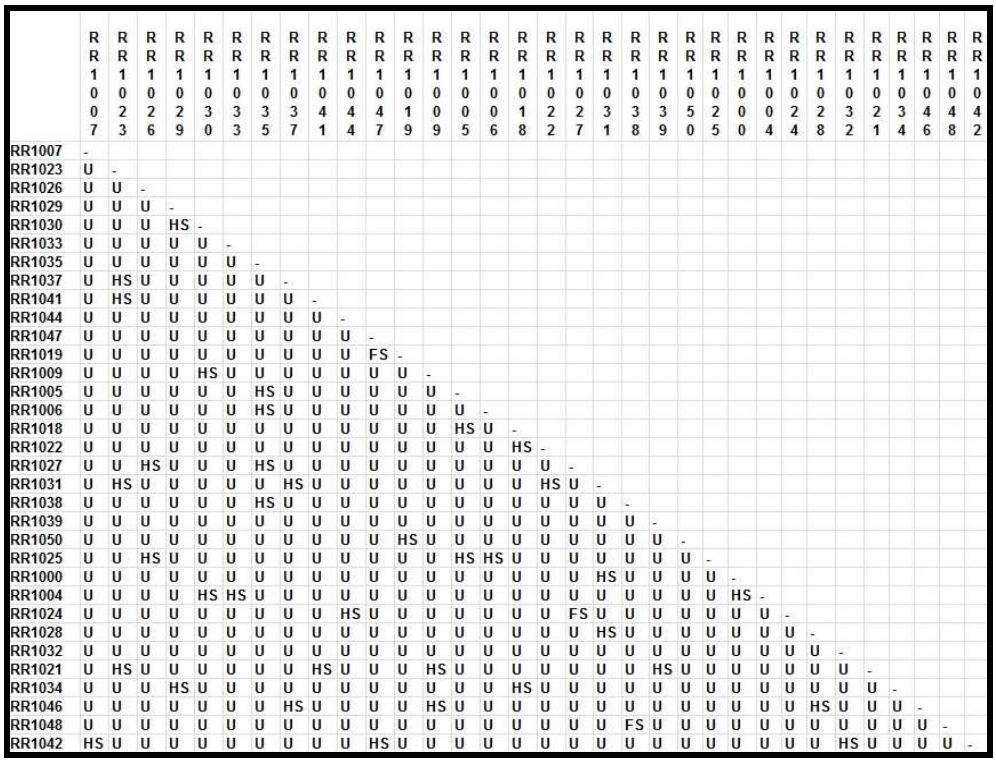
POP3

RR1007,	277279	143145	139141	128136	234242	152154	110112	078080	246254	193197	229245	398404	380382	Hamar	male
RR1022,	277277	139139	137139	128128	240244	154154	110110	080084	234242	183189	241241	392394	372382	Trysil	male
RR1023,	277277	139145	137143	126136	234248	154154	112112	080084	234250	183195	229245	402406	378382	Amot	male
RR1026,	277277	133143	137139	128128	242250	154154	112112	076084	250258	179197	241245	398400	374378	Hamar	female
RR1028,	277281	145145	137139	128132	240244	152154	110112	084086	240244	187191	243243	396402	380382	Stange	female
RR1029,	273277	143143	135139	126128	238244	154154	114114	076078	232236	181187	229235	392408	376380		male
RR1031,	277277	139139	143143	126128	240242	154154	112112	084086	236240	183191	229231	392402	378382	Trysil	male
RR1032,	273277	139145	137137	128128	234244	152154	104116	078078	220254	191191	241245	402404	378382	Trysil	female
RR1033,	273285	139145	137141	126128	234240	154154	110110	076080	254254	185187	229229	390400	378382		male
RR1034,	281281	143143	139143	126128	244244	152154	110112	080084	246254	179183	235243	394408	378382	Stor-Elvdal	male
RR1037,	277277	145145	135143	126128	234238	154154	112112	076080	234254	175195	231241	392402	382382	Stange	male
RR1041,	275279	141143	137143	126128	234240	152154	112112	080080	244244	189195	229245	390406	374382	Amot	female
RR1046,	277277	139139	143143	128128	234244	152154	110110	076080	234244	179187	243245	402404	380382	Trysil	male
RR1048,	273281	141143	137141	128128	234246	154154	110110	076082	240244	163183	243245	398406	378378	Stange	male
RR1025,	277283	143143	139141	124128	240242	154154	104116	076084	250258	179179	231245	404404	374382	Stange	male
RR1042,	277279	139143	137141	126136	242250	154154	116116	076078	254258	181191	229243	402406	382382	Trysil	male

Figure A 6: Representation of gene-pop files with the individuals segregated into their corresponding clusters from STRUCTURE program using original genotype data

```
POP
RR1000 ,   Stange      male
RR1004 ,   Amot        male
RR1005 ,               male
RR1007 ,   Hamar       male
RR1022 ,   Trysil      male
RR1023 ,   Amot        male
RR1026 ,   Hamar       female
RR1028 ,   Stange      female
RR1029 ,               male
RR1030 ,   Amot        female
RR1031 ,   Trysil      male
RR1033 ,               male
RR1021 ,   Amot        male
RR1034 ,   Stor-Elvda  male
RR1037 ,   Stange      male
RR1038 ,   Trysil      female
RR1039 ,   Stange      female
RR1041 ,   Amot        female
RR1046 ,   Trysil      male
RR1025 ,   Stange      male
RR1009 ,               female
POP
RR1006 ,   Stange      female
RR1018 ,   Hamar       male
RR1024 ,   Amot        male
RR1027 ,               male
RR1032 ,   Trysil      female
RR1035 ,               male
RR1044 ,   Amot        female
RR1047 ,   Trysil      female
RR1048 ,   Stange      male
RR1050 ,   Amot        female
RR1042 ,   Trysil      male
RR1019 ,   Trysil      male
```

Figure A 7: Representation of different groups where individuals are segregated using STRUCTURE program with the help of genotype data without null alleles



	RR1007	RR1023	RR1026	RR1029	RR1030	RR1033	RR1035	RR1037	RR1041	RR1044	RR1047	RR1019	RR1009	RR1005	RR1006	RR1018	RR1022	RR1027	RR1031	RR1038	RR1039	RR1050	RR1025	RR1000	RR1004	RR1024	RR1028	RR1032	RR1021	RR1034	RR1046	RR1048	RR1042
RR1007	-																																
RR1023	U	-																															
RR1026	U	U	-																														
RR1029	U	U	U	-																													
RR1030	U	U	U	HS	-																												
RR1033	U	U	U	U	U	-																											
RR1035	U	U	U	U	U	U	-																										
RR1037	U	HS	U	U	U	U	U	-																									
RR1041	U	HS	U	U	U	U	U	U	-																								
RR1044	U	U	U	U	U	U	U	U	U	-																							
RR1047	U	U	U	U	U	U	U	U	U	U	-																						
RR1019	U	U	U	U	U	U	U	U	U	U	FS	-																					
RR1009	U	U	U	U	HS	U	U	U	U	U	U	U	-																				
RR1005	U	U	U	U	U	U	HS	U	U	U	U	U	U	-																			
RR1006	U	U	U	U	U	U	HS	U	U	U	U	U	U	U	-																		
RR1018	U	U	U	U	U	U	U	U	U	U	U	U	U	HS	U	-																	
RR1022	U	U	U	U	U	U	U	U	U	U	U	U	U	U	U	HS	-																
RR1027	U	U	HS	U	U	U	HS	U	U	U	U	U	U	U	U	U	U	-															
RR1031	U	HS	U	U	U	U	U	HS	U	U	U	U	U	U	U	U	HS	U	-														
RR1038	U	U	U	U	U	U	HS	U	U	U	U	U	U	U	U	U	U	U	U	-													
RR1039	U	U	U	U	U	U	U	U	U	U	U	U	U	U	U	U	U	U	U	U	-												
RR1050	U	U	U	U	U	U	U	U	U	U	U	HS	U	U	U	U	U	U	U	U	U	-											
RR1025	U	U	HS	U	U	U	U	U	U	U	U	U	U	HS	HS	U	U	U	U	U	U	U	-										
RR1000	U	U	U	U	U	U	U	U	U	U	U	U	U	U	U	U	U	U	HS	U	U	U	U	-									
RR1004	U	U	U	U	HS	HS	U	U	U	U	U	U	U	U	U	U	U	U	U	U	U	U	U	HS	-								
RR1024	U	U	U	U	U	U	U	U	U	HS	U	U	U	U	U	U	U	FS	U	U	U	U	U	U	U	-							
RR1028	U	U	U	U	U	U	U	U	U	U	U	U	U	U	U	U	U	U	HS	U	U	U	U	U	U	U	-						
RR1032	U	U	U	U	U	U	U	U	U	U	U	U	U	U	U	U	U	U	U	U	U	U	U	U	U	U	U	-					
RR1021	U	HS	U	U	U	U	U	U	HS	U	U	U	HS	U	U	U	U	U	U	U	HS	U	U	U	U	U	U	U	-				
RR1034	U	U	U	HS	U	U	U	U	U	U	U	U	U	U	U	HS	U	U	U	U	U	U	U	U	U	U	U	U	U	-			
RR1046	U	U	U	U	U	U	U	HS	U	U	U	U	HS	U	U	U	U	U	U	U	U	U	U	U	U	U	HS	U	U	U	-		
RR1048	U	U	U	U	U	U	U	U	U	U	U	U	U	U	U	U	U	U	U	FS	U	U	U	U	U	U	U	U	U	U	U	-	
RR1042	HS	U	U	U	U	U	U	U	U	U	HS	U	U	U	U	U	U	U	U	U	U	U	U	U	U	U	HS	U	U	U	U	-	

Figure A 8: Relationship between the individuals using adjusted genotyping data

Table A 9 Individuals showing half sibling relationship from the matrix, Figure A 8

Individuals	HS with
RR1007	RR1042
RR1023	RR1037 RR1041 RR1031 RR1021
RR1026	RR1027 RR1025
RR1029	RR1030 RR1034
RR1030	RR1009 RR1004
RR1033	RR1004
RR1035	RR1005 RR1006 RR1027 RR1038
RR1037	RR1031 RR1046
RR1041	RR1021
RR1044	RR1024
RR1047	RR1042
RR1019	RR1050
RR1009	RR1021 RR1046
RR1005	RR1018 RR1025
RR1006	RR1025
RR1018	RR1022 RR1034
RR1022	RR1031

RR1031	RR1000 RR1028
RR1039	RR1021
RR1000	RR1004
RR1028	RR1046
RR1032	RR1042