Population Genetics of the Tibetan wild ass (*Equus kiang*) and Evolution of the Equids

Rita Mafalda Soares Monteiro

Mestrado em Biodiversidade Genética e Evolução
Departamento de Biologia,
2013

Orientador
Albano Beja-Pereira,

Investigador Auxiliar
CIBIO - FCUP

Todas as correções determinadas pelo júri, e só essas, foram efetuadas.

O Presidente do Júri,

Porto, _____/_____/_____

This work was fund by

FEDER funds through the Programa Operacional Factores de Competitividade – COMPETE and the national funds from FCT – Fundação para a Ciência e a Tecnologia in the frame work of the project with the codes PTDC/BIA-BIC/118107/2010 and FCOMP-01-0124-FEDER-019757.

ACKNOWLEDGEMENTS

The achievement of my objectives would probably be impossible without the support of several people.

First of all, I would like to thank to my supervisor, Albano Beja Pereira for the enormous support and for the trust in my work and Shanyuan Chen for the help with sequences analysis software.

I am very grateful to Vânia Costa for all the care, for teaching me all the laboratory proceedings, for always be there to help me patiently even from so many kilometers away. To Sónia Rosenbom, for the motivation and the precious help during the most stressful step of my thesis, and for the moral support when dealing with so difficult samples.

I also want to thank Liliana for support and help when I was freaking out with my data analysis.

A great thanks to all my friends and colleagues for making this year a little less difficult and for never refuse some help or advise, despite all the stress and difficulties that affected the most part of us during this time.

For last but not least, I really want to thank to my family for the great comprehension and patience, which have supported me throughout my life and for been responsible for what I am today.

I´m also very grateful to Tiago, for been always on my side encouraging me.

RESUMO

Durante o Pleistoceno, todas as espécies de equídeos apresentavam uma distribuição bastante abrangente, mas com o passar do tempo, algumas espécies acabaram por desaparecer, e as que sobreviveram sofreram uma redução muito significativa das suas distribuições.

Equus kiang é uma das espécies cujo declínio se supõe que seja causado maioritariamente por pressões antropogénicas. *Equus kiang*, também designado por Burro selvagem do Tibete, habita o planalto Tibetano e foi considerado uma das espécies de mamíferos de grande porte menos estudadas em todo o mundo, sobre o qual não existe nenhuma informação a nível populacional.

O principal objectivo deste trabalho foi: (i) inferir os parâmetros da genética populacional para esta espécie, recorrendo a uma metodologia de amostragem não-invasiva para a obtenção de DNA. Informação obtida através de um conjunto de microssatélites foi utilizada para inferir acerca da diversidade genética, estruturação populacional e deriva génica, que poderá ter resultado da fragmentação populacional e (ii) para comparar a diversidade e evolução dos genes candidatos para a coloração da pelagem (*ASIP*, *MC1R*, *SLC45A2*, *SILV* e *SLC36A1*) entre espécies de equídeos selvagens e domésticas.

Os nossos resultados não demonstraram qualquer evidência de fragmentação entre as cinco populações estudadas apesar das informações prévias. No entanto, todas as populações apresentaram um elevado défice de heterozigotía (F_{IS}), o qual pode ser preocupante, no entanto a elevada taxa de alelos nulos observada pode ser responsável pelos resultados obtidos.

A análise das sequências dos genes de coloração da pelagem revelou que apesar da grande divergência fenotípica entre selvagens e domésticos, ambos apresentam vários alelos em comum. Isto demonstra que uma grande parte da variação observada em espécies domésticas ocorreu antes da domesticação, no entanto a forte selecção artificial conduzida pelo Homem conduziu ao aumento da frequência alélica, bem como a fixação desses mesmos alelos.

Em geral, este trabalho contribuiu para aumentar o conhecimento a nível genético acerca da diversidade e evolução das espécies de equídeos, a qual á particularmente importante para o desenvolvimento de planos de conservação adequados.

Palavras-Chave: Kiang, genética populacional, microssatélites, conservação, cor da pelagem, selvagem, domesticação

ABSTRACT

During the Pleistocene, the equid species were widely distributed across the old world. However since that time the genus equus has been losing species and, those that survived suffered significant reduction on their ranges. The *Equus kiang,* also known as the Tibetan wild ass*,* inhabits high altitude ranges across the Tibetan plateau and, perhaps for this reason it has been one of the least threatened equid species. It is also the least studied equid species, and one of the least studied large mammalian species in the world, about which there are no genetics information at population level.

This work was divided in two main objectives: (i) to estimate the population genetics health using microsatellite markers data and, (ii) to study the evolutionary patterns of five coat color candidate genes (*ASIP*, *MC1R*, *SLC45A2*, *SILV* and *SLC36A1**)*** in the equid species. For this we have obtained samples using a noninvasive approach and genotyped

Our results showed no evidence of population sub structuring due to fragmentation on the five studied populations, despite the previous indication. However, all populations displayed a high heterozygous deficit (F_{IS}), which could be worrying, although the high null allele rates observed could be responsible for the obtained results. The coat color genes sequence analysis revealed that, despite the high phenotypic divergence between wild and domestic, they present several alleles in common. This showing that a large part of the variation observed on domestic species occurred before domestication, although the strong human-mediated artificial selection conducted to the frequency increase and fixation of those alleles.

In general our work have contributed to increase the genetic knowledge about diversity and evolution of the equid species, which is particular important to design appropriate conservation plans.

Keywords: Kiang, population genetics, microsatellites, non-invasive, conservation, coat color, wild, domestication

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

DNA – Deoxiribonucleic acid

mtDNA – Mithocondrial DNA

nDNA – nuclear DNA

rRNA – ribosomal RNA

aDNA – ancient DNA

SNP – Single nucleotide polymorphism

PCoA - Principal Coordinate analysis

AMOVA - Analysis of Molecular variance

HWE - Hardy-Weinberg Equilibrium

LD - Linkage Disequilibrium

SE - Standard Error

CITES - Convention on International Trade in Endangered Species of Wild Fauna and Flora

PCR – Polymerase chain reaction

IUCN - International Union for Conservation of Nature

MC1R - Melanocortin 1 receptor gene

ASIP - Agouti signaling protein

SLC36A1 - Solute carrier family 36 member 1 gene

SLC45A2 - Solute carrier family 45, member 2 gene

PMEL17 - Melanocyte protein 17 precursor

CHAPTER 1

Introduction

1.1 The species

1.1.1. A brief characterization

Presently, there are two species of Asiatic wild ass (*E. kiang* and *E. hemionus*), which were the most abundant large herbivore species rooming across the entire Eurasia, during the Pleistocene (Moehlman 2002). However, its actual distribution is confined to marginal areas (Moehlman 2002), mainly due to the increase of human density on their distribution area (St-louis & Côte 2009) and habitat fragmentation. From these two species, the Tibetan wild ass (*Equus kiang*) (Fig. 1) is one of the least studied large mammal species of the world (Sharma *et al.* 2004).

Although, the distinction between the two Asiatic wild ass species is yet controversial, as until 1990, kiang was considered a subspecies of *E. hemionus*. However, the results from cytogenetic studies by Ryder & Chemnick (1990), have promoted the identification of *E.kiang* as separate species. Another argument in favor of this splitting come from the fact that both species are allopatric (distribution ranges never cross), and a slightly morphologically distinctintiveness (Geigl & Grange 2012), such as the kiang being the larger of the two wild asses (Schaller 1998).

1.1.2. Taxonomy

The Kiang (*Equus kiang,* Moorcroft, 1841) belongs to the Mammalia class, Perissodactyla order, Equidae family, Equus genus (Sharma *et al.* 2004) and is divided in three subspecies with geographically distinct populations and morphologic characters : *E. k. kiang* (western Tibet)*, E. k. polyodon* (southern Tibet), *E. k. holdereri* (easthern Tibet), according to Groves (1974).

In terms of morphology, this species is characterized by a relatively short body, large head, short snout and convex nose, with long limbs and broad hooves, and a short, erect and tufted tail (Bennett, 1980; Groves 1974). On average it has a chestnut-brown coat color covering the upper part of the body, such as thighs, back, upper flanks, and dorsal part of the neck, cheeks, and forehead. It has a white coloration on rostrum, throat, inside of the ears and legs. Its mane, edges and ear tips are black. It possesses a broad dorsal stripe from the mane to the end of the tail (Groves 1974; Schaller 1998). The coat color can vary from darker brown during winter to paler reddish on summer and in the winter the coat is ticker and longer than summer, ranging, respectively between 35–46 mm and 14–16 mm and long (Groves & Mazák 1962; Groves 1974).

Figure 1- Equus kiang female from Xinjiang (China)

Among the three kiang subspecies, is possible to find several morphological distinctive characters, like body size, coat color patterns, and cranial and dental features (Groves & Mazák, 1967). But this division is not consensual, as referred by Wang (2003), which just validated the existence of two subspecies (*E. k. kiang* and *E. k. holdereri*;) and by Schaller (1998), which claimed the absence of significant morphological differences between subspecies to separate them. Nonetheless, the intra-specific genetic variation had never been assessed (Ryder & Chemnick 1990).

1.1.3. Distribution

E. kiang is endemic to the Tibetan Plateau (Central Asia) at altitudes ranging between 2.700 and 5.300m above sea level (Schaller 1998; Shah 2002). It is often observed in small patches of population across a large range, occurring mainly in China, but also in adjacent areas like India, Nepal, Pakistan and probably also in Bhutan (Shah 2002). All the three subspecies of Kiang are found in China, one of them, the eastern kiang (*E. k. holdereri*) is exclusive of Chinese territory while the other two, southern (*E. k. polyodon) and westhern (E. k. kiang)* are found on border areas of the previously referred countries (Moehlman 2002), as shown on Fig. 2 . However, despite this broad range, the Kiang geographical distribution is becoming highly fragmented and currently, a great part of the populations inhabit protected areas or areas under jurisdiction (Shah 2002).

Figure 2 - Geographic distribution of the E. kiang (in orange) across the Tibetan Plateau

It is estimated that the total number of individuals is around the 60,000 and 70,000 (Schaller 1998). About 90% of kiang population is found in China, on a number between 56,500-68,500 individuals (Shah 2002). From this, the most part is found on Tibet (37,000-48,000 individuals) (Schaller 1998), and the remain are located in Xinjiang (4,500-5,500 individuals) (Shah and Huibin, 2000), Qinghai and Gansu (15,000 individuals) (Schaller 1998).

In India, the species is mostly found in Ladakh region (1500- 2000 individuals) (Fox 1991), but also in Sikkim, even although in smaller number (104-120 individuals) (Shah 2002). In Nepal, there are records of this species in Doplo and Mustang areas (37-500 individuals)

(Shah 2002; Sharma *et al.* 2004), while in Pakistan only an isolate population of about 25 individuals restricted to Khunjerab National Park had been reported (Rasool 1992). Finally, in Bhutan, the species have never been signed, although Shah (1997) defended that it is very likely that species occur on north and northwest of the territory.

1.1.4. Habitat, ecology and diet habits

The Tibetan plateau is considered one of the highest and most extensive mountain region of the world, occupying around 25% of the Chinese territory (Foggin 2012) and presenting a continental climate, whereas most part of steppe is arid or semi-arid. The area is so wide that is possible to find a great variety of cold deserts, alpine and semi-arid steppe, as well as alpine meadows (Sheehy *et al.* 2006).

From this vast territory, *Equus kiang* inhabits preferentially alpine steppes and alpine meadows (Schaller 1998), and according to Harris & Miller (1995) they are mostly found in south-facing xeric basins, feeding preferentially on open terrains, since it allows a more easily observation of potential hazards and thus more easily escape.

Equus kiang is classified as a variable grazer, with a grass-dominant diet, showing geographical and seasonal variation on their diet, as found in other ungulate species (Gagnon & Chew 2000). This variation was studied by Harris & Miller (1995), which showed that during summer, *Stipa* sp constitutes almost 95% of Kiangs' diet. Since they inhabit an arid environment like Tibetan plateau, the water availability is scarce, so they pick it mainly through vegetation and snow (Schaller 1998).

Despite the kiang does not migrate regularly (Schaller 1998), they often do several seasonal movements, alone or in herd (St-Louis & Côte, 2009), probably searching for more available and better quality forage. This nomadic behavior, essential to survival, can be compromised due to the increasing fragmentation and climate changes observed.

In terms of social structure, the only permanent association is observed between the females and fouls, whereas young males tend to aggregate and adult males are solitary and territorial (St-Louis and Côte, 2009).

1.1.5 Conservation

1.1.5.1 *Status*

Currently, *E. kiang* is listed as “least concerned” (LC) according to the IUCN Red List Categories, and is included on Appendix II of CITES (Convention on International Trade in Endangered Species of Wild Fauna and Flora). However, this classification could not be correctly assigned, since it is only based on information about one of the three subspecies, the eastern kiang (*E. k. holdereri*, which is classified as LC. The other two subspecies, western (*E. k. kiang*) and southern kiang (*E. k. polyodon*) are unclassified and have been marked as “data deficient“(DD),.

As referred previously, habitat fragmentation, mainly driven by the increase of human activities, is the major threat affecting this species. In fact, fragmentation is often a direct consequence of the human population increase (Wilcox & Murphy 1985), and have been described as one of the major cause for species threat and extinction. These effects have been already described as threatened causes on several species, like on a bear (*Tremarctos ornatos)* (Kattan *et al.* 2004)*,* Siberian tiger (*Panthera tigris altaica*) (Xiaofeng *et al.* 2011) and in Sumatran and Bornean orangutans (*Pongo abelii*, and *P. pygmaeus*) (Goossens *et al.* 2006), but concerning the impact of fragmentation in high-altitude ecosystems very little or nothing is known.

The habitat fragmentation *per si*, in the Tibetan plateau, is mostly due to the increasing number of pastoralist that become settled and increased their livestock numbers (Harris & Miller 1995; Sheehy *et al.* 2006). This demographic changes leads to an increasing need of space and resources, which consequently lead to a change on land uses, intensification of agricultural practices, transformations on pastoral livestock systems (Sheehy *et al.* 2006) and infrastructure construction (Baofa *et al.* 2006). Some regional conflicts with livestock has also been reported, for example in India (Ladakh) (Bhatnagar *et al.* 2006), along with some illegal poaching practices (Schaller 1998). Habitat fragmentation conducts to habitat loss and isolation of remnant patches. This isolation decreases connectivity between populations, which directly affects dispersal and gene flow, leading to loss of genetic diversity and increase of differentiation among populations due to genetic drift (Frankham 2005), which are known for being the main contributors for the increasing of extinction risk. Looking at this scenario becomes urgent to infer about the genetic condition of *E. kiang* populations, in order to verify whether some populations have become isolated and start drifting.

Although there is no consensus on the intraspecific partition in three kiang subspecies (Groves Mazák,1967; Schaller, 1998; Wang ,2003), it is nonetheless consensual that each of the disputed subspecies are indeed three different populations, with slightly distinct phenotypic traits that might have resulted from translating environmental adaptions.
Therefore it is urgent to genetically characterize these three subspecies, to verify their level of divergence and clarify inherent taxonomic questions.

1.1.6 Genetics characterization

A study developed by Ryder & Chemnick (1990) showed that at chromosomal level there are two different diploid chromosomal number among kiang individuals: 2n=51 and 2n= 52; however both present the same number of autosomal arms=92. This polymorphism was detected both in males and females, and may have resulted from a Robertsonian translocation. In kiangs the X chromosome presents a sub-metacentric pattern, whereas Y is acrocentric (Ryder & Chemnick 1990).

Besides this latter studies, very few molecular studies focusing *Equus kiang* was published. Moreover, at the population genetics level there is nothing published about this species. The only studies besides these, where phylogenetic assessments of genus Equus, which includes the subspecies *E. k. holdereri* using microsatellite loci (Krüger *et al.* 2005) and a study applying DNA sequences of 12S rRNA (Oakenfull *et al.* 2001). Both studies were conducted on captive animals, which probably do not represent the genetic diversity range of this species. Another piece of information come from a couple of mtDNA sequences of this species, which resulted from a study to infer the origins of domestic donkey, (Beja-Pereira *et al.* 2004).

1.2 Non-Invasive sampling techniques

The approaches used to obtain DNA from a biological sample of a wild animal can be divided in three main groups: destructive, which requires the animal sacrifice to obtain the source material for DNA extraction (e.g., internal organ); Nondestructive sampling, which frequently requires the animal capture in order to collect the sample, using an invasive way (e.g., blood, skin); Non-invasive, does not requires the capture of the animal and consists on the collection

of some biological material containing DNA (e.g., feces, urine, saliva) left behind by the animal (Taberlet *et al.* 1999).

Frequently, in the literature, the nondestructive sampling is wrongly considered as noninvasive (Taberlet *et al.* 1999). However, trapping and manipulate an animal in order to pluck some hairs or feathers, or a buccal swab (Broquet, Berset-Braendli, *et al.* 2006), should not be described as noninvasive. In fact, this goes against the noninvasive assumptions, since this methodology does not require the capture or animal disturbance to samples collection (Taberlet *et al.* 1999).

Noninvasive sampling techniques allow to study natural populations' genetics without the necessity of capture, handle or even observance of the individuals (Piggott & Taylor 2003a). For scientists working on free ranging, rare or sensitive species, this could be the most adequate sampling approach (Taberlet *et al.* 1997), since trapping and manipulating (to collect the biological sample) could result in animals' damage or even in death (Greenwood,1996), and are always a source of additional stress, which effects the species dynamics on an unpredictably way (Harrison et al. 1991). Therefore, the invasive sampling approaches will always influence somehow the individuals behavior, and could be problematic, for example, on ethological studies (Morin *et al.* 1994).

The noninvasive sampling is widely applied on several fields of the biology as molecular ecology, conservation genetics, wildlife forensics, epidemiology, behavioral ecology, among others (Beja-Pereira *et al.* 2009), as it is probably the only way of sampling when dealing with free range species from remote places (Morin & Woodruff 1992), like the case of our subject species, the kiang. Noninvasive approaches made possible the population genetic studies for conservation purposes of several threatened species, which until then had been impossible to access by invasive methods (Morin & Woodruff. 1996).

The fact of a great number of species are facing declining and extinction, becomes urgent designing management conservation plans, for which genetic information is essential for demographic and life story inferences. Due to the species declining, its direct observation becomes more difficult, and therefore, noninvasive technology is the most appropriated way to achieve this information (Kohn & Wayne 1997).

Noninvasive genetic sampling started to be used on wildlife research more than twenty years ago, mainly driven by the PCR discovery and application (Mullis *et al.* 1986; Taberlet & Luikart 1999). In 1992, noninvasive genetic sampling methods were firstly applied on the study of a threatened European brown bear population (*Ursos arctus*) to access the population genetics parameters using fecal and hair samples (Hoss *et al.* 1992; Taberlet & Bouvet 1992) and to infer about the social structure of chimpanzees (*Pan troglodytes*) (Morin & Woodruff 1992).

Today there is a very wide list of different biological materials from which DNA was successfully extracted. From these feces (Hoss *et al.* 1992; Kohn *et al.* 1995), shed hair (Taberlet *et al.* 1997), shed feathers (Rudnick *et al.* 2007), urine (Valiere & Taberlet 2000), blood on substratum (Scandura 2005), eggshell and foot mucus from invertebrates (Kawai *et al.* 2004), water samples (in which usually the species inhabits) (Ficetola *et al.* 2008), and shedded skin are among the most widely used (Waits & Paetkau 2004).

From this list of noninvasive samples, feces are definitively the most frequently used in studies involving wild ungulates. The main advantages are the fact these samples can be collected totally on a noninvasive way, and a single portion of stool can contain a great amount of DNA from the excretory individual (Albaugh *et al.* 1992). Besides this, scats are easiest to find on field and are the most informative since they can also provide information on the animals' diet and immunity status, among others (Kohn & Wayne 1997). The correct molecular analysis of fecal samples can provide information not only at the species but also can be used for individual identification, sex, pathogenic organisms, physiological hormones and food habits (Kohn & Wayne 1997; Luikart *et al.* 2008).

Fecal samples can provide a wide range of information on very different biological fields: (1) behavioral biology, (2) census population size, (3) home range and territory size, (4) effective population size, (5) population genetic variation, (6) phylogeography, (7) diet and, (8) diseases, (Kohn & Wayne 1997; Beja-Pereira *et al.* 2009). Another advantage of fecal samples is that it does not require permission from international organizations such as CITES (Luikart *et al.* 2008), since often the high number of permits required for the capture and transport of invasive samples from endangered species incapacitates the study (Kohn & Wayne 1997). Although noninvasive sampling presents several advantages, it also has some downsides such as the high rate of genotyping errors, caused by the low-quantity and low-quality of extracted DNA, as well as the presence of PCR inhibitors and cross-contaminations during extraction and amplification (Taberlet *et al.* 1999).

Since feces were the type of noninvasive samples used to achieve the objectives of the current work, and in general are the most commonly noninvasive samples used in ungulates, the rest of this section is focused on stool noninvasive sampling.

On fecal samples, the DNA is only present on intestinal epithelial cells which covers its otter surface (Waits & Paetkau 2004). During the DNA extraction process, not only DNA from the targeted individual is obtained, but also other undesirable material, as DNA from other different sources, and several PCR inhibitor compounds (Kohn & Wayne 1997) are obtained. Allied to this, is the fact that sometimes samples have been exposed to the natural elements

such as the ultra-violet waves from sunlight, which are known for damaging of the DNA chains, affecting significantly the quality and quantity of amplifiable DNA *(*Reddy *et al.* 2012).

Despite these handicaps, the amplification and genotyping success of fecal material can be improved on several ways, like the collection (if possible) of fresh samples (Taberlet *et al.* 1999), the selection of the most suitable storage method, capable of ensure the sample integrity for further analysis, from the field collection up to the the laboratory processing (Frantzen *et al.* 1998; Murphy *et al.* 2002). The selection of an appropriate extraction method, capable to recovering DNA with sufficient quality and quantity (Piggott & Taylor 2003b) is another important variant that affects the quality and quantity of the extracted DNA. Despite the large amount of storage and extraction methods already published and the different methodological comparisons among them (Frantzen *et al.* 1998; Murphy *et al.* 2002; Piggott & Taylor 2003b; Soto-Calderón *et al.* 2009), it is very important to take into account particularities of the species under study to avoid compromise the remaining work (Wasser *et al.* 1997).

At amplification level, the use of PCR primers that amplify short DNA fragments, and are specific (instead of universal), can overtake the problem of DNA degradation (Taberlet *et al.* 1999), and avoid the amplification of non-target DNA (Broquet, *et al.* 2006). However, genotyping errors cannot always be avoided by laboratory procedures improvement (Broquet & Petit 2004).

There are three types of genotypic errors often observed: (1) allelic dropout (detects a false homozygote), (2) false allele (i.e., false heterozygous) and (3) human error (Beja-Pereira *et al.* 2009), which can be source of miscalculations of genetic diversity, individual identification and populations structure and assignment, compromising the reliability of the studies (Taberlet *et al.* 1999). However, these errors can be minimized by using a “Multiple tubes” approach. This approach consists on repeating amplifications for the same samples several times, and obtain a consensus genotype by analyzing the entire set of repeats. This method is indicated when dealing with very low amounts of DNA (Taberlet *et al.* 1996). To avoid contaminations, it is recommended that PCR’s preparation place is isolated from the DNA extraction bench, and the use of non-invasive samples dedicated pipetting devices. The contaminations can easily be motorized by adding at each PCR reaction, one negative control sample in which the DNA template is replaced by sterile water. For the detection of genotyping errors, rates calculation and even evaluation of its effects can be assessed by several software programs that are publically available (Pompanon *et al.* 2005).

1.3. Molecular markers

The last two decades, have been quite productive on the identification of different molecular markers with different applications, which consequently permitted answering to many biological questions. The first molecular markers applied were allozymes, used to accessing genetic variation on an indirect way by protein analysis (Schlötterer 2004). After this and essentially due to the possibility of DNA amplification by PCR, the numbers of available markers increased largely, and now we can assess genetic variation by many different ways, and according to the purpose, choose the most suitable marker according to their characteristics. Microsatellites, mitochondrial and nuclear DNA are the most frequently used nowadays (Selkoe & Toonen 2006).

For this work were used two different markers: microsatellites and nuclear DNA sequences polymorphisms.

1.3.1. Microsatellites

Microsatellites were until recent the most applied genetic markers on population genetics (Vial *et al.* 2003) and despite the success of the NGS techniques they are still widely used in wildlife molecular ecology studies. They are characterized for being neutral, co-dominant, widely distributed among the genome, highly polymorphic, and capable of being amplified by PCR even with poor quality samples, collected using non-invasive sampling approaches (Beja-Pereira *et al.* 2009). Another advantage of microsatellites is the fact that their flanking regions are frequently conserved across close related species and families, particularly on ungulates (Vial *et al.* 2003), which share an high degree of homology at DNA sequence level (Krüger et al., 2005). This characteristic enables cross amplification, which consists on applying the same microsatellites in studies with close related taxa, like was used in a recent study by Rosenbom et al., (2011) with African wild ass. This method permits to save considerable time and funds, since it avoids the laborious step of cloning and design specific microsatellite primers (Hammond *et al.* 1997), as well as can be used in comparative studies between closely related species (Vial *et al.* 2003).

The high mutation rate observed on microsatellites, compel it to accumulate mutations in a small time scale. This characteristic is fundamental when the aims are the access of genetic structuring of small and endangered populations, recent historic and demographic events, as

it permits to evaluate the variation of genetic diversity among time and space (Beaumont 1999). On population genetics, since it deals frequently with close related taxa (like variation between and among populations) is necessary to resort to a fast evolving marker, in order to enlighten the differences among them (Oakenfull & Ryder 2002).

In order to access a population structure, as propose in the present work (on *E.kiang*), several demographic parameters need to be achieved, like estimation of effective population size (Ne), migration, and demographic fluctuations (Morin *et al.* 2012). Once this study is based on noninvasive sampling, the microsatellites are the most suitable molecular markers to be used, as it present short allele sizes, which increases the amplification success on degraded samples.

1.3.2. Nuclear DNA sequence polymorphisms

DNA sequencing consists on obtaining the nucleotide sequence of a given part of the genome from several individuals, by PCR amplification with region specific primers for the desired region (Mills 2007). It presents a great level of resolution since it has the power of provide total information about the sequenced region which prevents its use of being affected by ascertainment bias (Schlötterer 2004). By PCR, is possible to obtain any specific part of the genome. Despite on early times this was considered expensive and time-consumer, currently, the advances among sequencing technology allowed to invert this reality.

Briefly, nuclear DNA is constituted by two main kinds of regions: coding-regions, possible to be translated in to amino acid chains (proteins) in which are included exons; and by non-coding regions, which includes introns (found between coding-regions on eukaryotes), promoter regions (located upstream the exons and presenting regulatory function), and all extra genic regions, where none amino acid translation is observed (Zhang & Hewitt 2003; Calonje *et al.* 2008).

Mutation rate varies among nuclear genome, been higher on non-coding regions, whereas coding-regions are more conserved, presenting a lower mutation rate (Zhang & Hewitt 2003). The highest level of diversity found on nuclear genome is mainly caused by polymorphisms occurrence on nonfunctional regions or by synonym mutations occurring on coding-regions (which does not change the amino acid conformation) while the lowest diversity if found among functional regions, like coding or regulatory regions (Frankham *et al.* 2002). This is also related with the fact of only a really small part of the genome corresponds to functional regions (e.g. almost 1.5% in Humans), involved in protein codifying (Goodwin *et al.* 2007). Indeed, among nuclear genome is possible to find two different kinds of loci: those evolving

on neutrality, present on non-coding and non-functional regions, without fitness effects, and loci evolving under selection, present on functional and coding-regions, showing fitness effects (Luikart *et al.* 2003). These characteristics allow us to apply this molecular marker on our coat color variation study.

Nuclear loci evolving under neutrality, non-linked to selected locus, are suitable for population demography (Pluzhnikov *et al.* 2002), phylogenetic inferences (Hare 2001; Luikart *et al.* 2003) and also for the establishment of molecular clocks (Hasegawa & Yano 1989), due to the relatively high mutation rate observed on this regions. On the other hand, loci evolving under selection (selected loci), occurring on more conserved regions (Frankham *et al.* 2002), are ideal for the detection of traits under selection (e.g., coat color), with effects on fitness and animals adaptation since variation among this regions is often reflected on individuals phenotype, like observed among coat color genes (e.g. (Mariat *et al.* 2003; Brunberg *et al.* 2006) (see Manuscript 2).

Evolutionary process as domestication can be accessed by genetically analysis of a selected phenotypic characteristic, as coat color (Manuscript 2).

1.4 The importance of population genetics to inform conservation plans

According to IUCN most recent data, about 18,788 species of the 52,017 assessed until now are in risk of extinction, and from the total of 5,490 mammals species, 78 are considered as extinct (or extinct in the Wild), 188 are critically endangered, 540 endangered and 492 vulnerable (http://www.iucn.org/search.cfm?uSearchTerm=cite-, at 30/08/2013). The main threats to biodiversity are habitat loss and fragmentation, invasive species, over-exploitation and pollution, associated with the increasing population density or human activities such as mining (Wilcox & Murphy 1985; Frankham *et al.* 2002); however some stochastic factors (demographic, environmental, catastrophic and genetic) can also contribute for population extinction (Frankham 2005). These factors conduct largely to the decrease of isolation and reduction of population size, which necessarily led to the occurrence of inbreeding, reduction of gene flow, genetic drift, and loss of genetic diversity (Frankham *et al.* 2002; Frankham 2005).

Inbreeding affects largely the reproduction potential and survival (fitness), since the mate between related individuals conduces to loss of heterozygosity and fixation of deleterious alleles, reducing the effective population size (number of individuals with reproduction potential) and conducting to the loss of genetic diversity, which is mainly responsible for the decrease of the adaptive capacity to environmental changes or pressures (Frankham *et al.* 2002).

Looking into this worrying scenario, it became urgent the demand for most accurate and efficient methods capable to provide relevant information at species and populations level, in order to inform and implement conservation plans (Bertorelle *et al.* 2009). This search, carried out for years, led into the population genetics field, which is currently on the vanguard of the analysis on conservation biology and molecular ecology (Pool *et al.* 2010), because of the increasingly power and accuracy of the molecular markers and genetic statistical analysis methods (Sunnucks 2000).

The main proposes of population genetics is to infer about demographic processes or selection to study adaptive genetic variation within and between populations, on time and space (Chikhi & Bruford.M 2005; Laird & Lange 2011), and in fact, on a more generic definition, population genetics consists on the study of gene lineages within a micro-temporal scale (Lewontin 1985). With this tool became possible to infer several parameters capable to inform about evolutionary history and genetic health of the populations.

A great improvement is also the possibility that a great part of these studies can now be based on DNA obtained from noninvasive sampling, avoiding the introduction of additional stress on already threatened population (e.g. (Kohn *et al.* 1995; Quéméré *et al.* 2009). This combination was labeled by Broquet, *et al.* (2006) as “noninvasive population genetics”. Moreover, another emergent field, landscape genetics, which consists on the conjugation between population genetics and landscape ecology, in order to define how landscape influences micro-evolutionary processes (as gene flow, genetic drift and adaptation) and population structure (Manel *et al.* 2003), can provide more powerful and efficient tools for management and conservation purposes. Finally, population genetics, as referred previously, can also be useful to increase the knowledge about historic and demographic evolutionary processes level, like domestication (Larson & Burger 2013).

1.5 - Objectives

In the current work, two main questions were assessed and are represented at two different manuscripts:

Manuscript 1 - "Population Genetics Assessment of the *Equus kiang*"

Manuscript 2 - "A Coat Color Assessment on Domestic and Wild Equids"

The main objective of Manuscript 1, was the assessment if the genetic status of several Tibetan wild ass (*Equus kiang*) groups/populations distributed across the Tibetan plateau. For this we have used samples collected noninvasivelly from five regions on Tibetan plateau (China), to estimate diversity and demographic parameters, like, demographic dynamics (e.g., expansion retraction), degree of inbreeding, detect social and spatial structuring, degree of inter-population gene flow, and differentiation. These parameters will improve our knowledge about the effects of the fragmentation suffered in recent years by these populations.

Manuscript 2 consists of an exploratory study on the evolution of coat color genes from wild and domestic Equids. For this we have used samples from the Asiatic wild ass – *E. kiang* and *E. hemionus -,* and two African wild Equids*: E. grevy* and *E. africanus*. Our aim was to compare the evolutionary history between domestic and wild species, at the level of the coat color genes.

The goals of this study is the comparison of exonic regions from five coat color responsible genes: *MC1R*, *ASIP*, *SLC36A1*, *SLC45A2* (or *MATP*), and *PMEL17* (or *SILV),* in order to search for nucleotide polymorphisms among wild Equids, and check if they are coincident with those already found on domestic ones, and responsible for the huge coat color diversity observed. This work will allow the inference of several questions about domestication processes like how the observed phenotypically differences at coat color behave at genetic level, see how divergent are domestic from wild Equids at coat color genes level, evaluating the effects of artificial selection against natural selection.

CHAPTER 2

Manuscript 1

Manuscript 1

Population Genetics Of The *Equus kiang* (Moorcroft, 1841)

Rita Monteiro[1], Shanyuan Chen[1], Chamba Yangzon[2], Albano Beja-Pereira[1]

[1] CIBIO-INBIO, Universidade do Porto, Campus Agrário de Vairão, Vairão, Portugal

[2] College of Agriculture and Animal Husbandry, Tibet University, LINZHI 860000, Tibet, China

Abstract

The Tibetan wild ass, *E. kiang*, is one of least studied large ungulate species in the world. It inhibits one of the most extreme environments, characterized by its high-altitude location and continental weather. For this reason the survival of this species as not being much affected but recent anthropogenic pressures (pastoralism and mining activities) may start to impact their habitat. In this study is pretended to assess the population genetics "health" of this species. For this propose we have collected samples from five distant geographic points distributed across the Tibetan Plateau using a noninvasive approach. A set of 12 microsatellite loci was genotyped for 42 individuals representing those five sampling points. Our results showed a high heterozygosity deficit, which can be related with high levels of null alleles, we did not find any sign of population fragmentation as population sub structuring was considerably low. However the low heterozigosity observed among these populations can be a matter of concern, nevertheless the null allele influence should be taken into consideration.

Keywords: *Equus kiang*; Tibetan wild ass; conservation; noninvasive; microsatellites

Introduction

During Pleistocene, all the equid species presented a widely distribution, along Eurasia, Africa, and the American continent, but, over the years, some of species disappeared and those that survive suffered for a great reduction on their ranges (Waring 2003).

This is the case of Tibetan wild ass (*Equus kiang*), which is considered as one of the least studied species in the world (Sharma *et al.* 2004), and, along with another wild Equid species (*Equus hemionus),* was once one of the most abundant large herbivore of Eurasia (Shah 2002). Presently Kiang´s distribution is highly fragmented mainly due to increase of human pressure and changes in the land use (St-louis & Côte 2009).

Equus kiang (Moorcroft, 1841) is classified in Perisodactyl, from the Equidae family, which is composed of a single genus, the Equus (Sharma *et al.* 2004). This species is subdivided in three morphologically and geographically distinct subspecies: *E. k. kiang* (western Tibet)*, E. k. polyodon* (southern Tibet), *E. k. holdereri* (eastern Tibet), following Groves (1974). However, this division is controversial and is not clarified yet within scientific community (Groves 1974; Schaller 1998; Wang 2003). It inhabits the Tibetan plateau at altitudes ranging from 2.700 and 5.300 m above sea level (Schaller 1998; Shah 2002), and despite the most part be found on Chinese territory (at least 90% of population), it can also be found on border areas of Nepal, India, Pakistan, and probably in Bhutan (Shah 2002).

According to the IUCN Red List Categories, Kiang is listed as least concerned (LC), however, this classification is not very accurate, since it is only based on data from the eastern kiang (*E. k. holdereri*); the other two subspecies western (*E. k. kiang)* and southern kiang (*E. k. polyodon)* are classified as data deficient (DD), which emphasize the great need for studies about this species. Despite this species not being considered as endangered, the fact that has been subject to extensive habitat fragmentation (St-louis & Côte 2009), is challenging its population genetic health. Until now, this species have never been analyzed at population genetics level, and there is only a few number of molecular studies conducted. The first consisted on the inference of the chromosomal structure, which led to the taxonomic separation of *E.kiang* and *E.hemionus* (Ryder & Chemnick 1990). So far the only genetic studies on this species where on the frame work of the phylogeny of the Equus genus and only have used captive individuals (Oakenfull *et al.* 2001; Beja-Pereira *et al.* 2004; Krüger *et al.* 2005).

Since this species inhabits a very human-hostile environment, it is quite difficult capture individuals. Since noninvasive sampling does not requires the capture or even visualization of the animals (Piggott & Taylor 2003) the use of such sampling approach will permit to gather enough samples to conduct population genetic studies on this species. As the noninvasive samples often provide limited amounts of DNA, the use of molecular markers that can be easily amplified is crucial. The microsatellite loci are still the most widely used markers on noninvasive population genetics studies of non-model species. The main reason for this is that they display high levels of polymorphism, have a nearly neutral behavior and can be easily amplified on degraded samples like the ones collected non-invasively (Beja-Pereira *et al.* 2009).

Therefore, it is urgent access this species at genetic level since the little available information is based just on species observation and occurrence, which is insufficient to improve appropriate management plans.

Materials and Methods

Sampling

Where collected 114 fecal samples from five geographical points, in China: Nyalam, Tingri, Ngamring, Gyirong and Nagqu, (Fig.3)

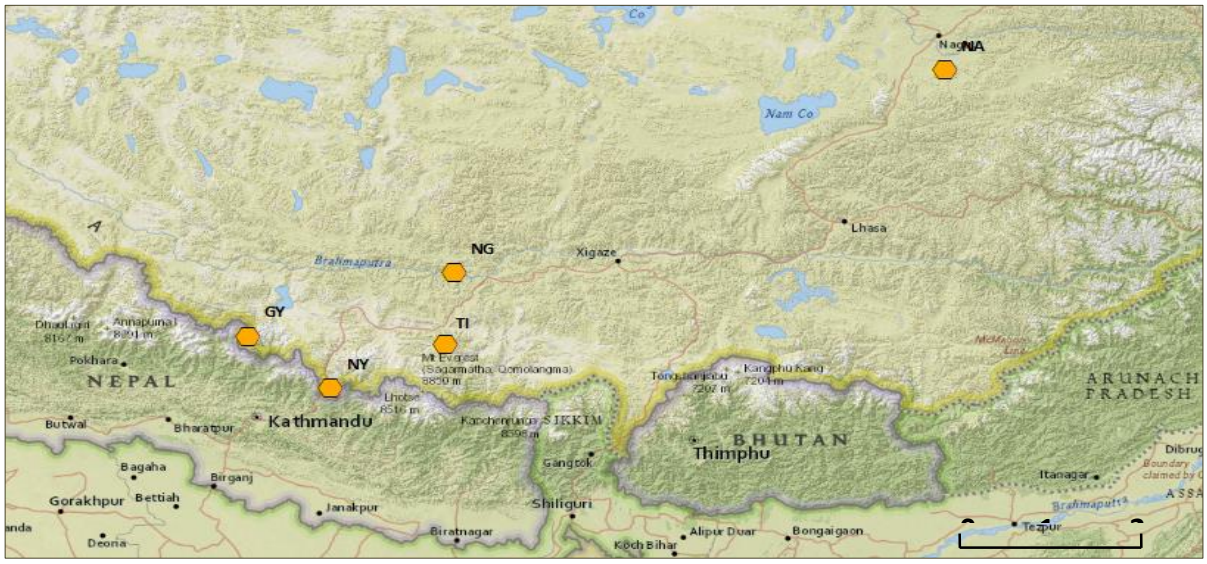

Figure 3 - Map where are represented the 5 sampling points (orange marks): **NY** - Nyalam, **TI**- Tingri, **NG**-Ngamring, **GY**-Gyirong and NA-Nagqu.

The sampling procedure consisted in the collection of feces (noninvasive sampling) from previously observed animals, which were stored in bags with silica for a faster desiccation of the samples to avoid DNA deterioration. At the arrival to the laboratory facilities, the samples were further desiccated in an oven. Finally the samples were stored at room temperature until the extraction.

DNA extraction

Genomic DNA extraction was performed using JETQUICK Tissue DNA Spin Kit (Genomed ®), following a slightly modified manufacture (Costa et al. in prep). After extraction samples were tested on a 0.8% agarose gel, stained with GelRed™ (Biotum), in order to evaluate extraction quality and determine dilution factors for each sample.

Microsatellite Amplification

Initially we have tested a set of 36 microsatellite loci, previously isolated from horse (*Equus caballus*) – AHT4, AHT5, COR20, COR58, COR90, HMS6, HMS7, HMS20, HTG6, LEX68, LEX74, NVHEQ18, UM11, VHL20, ASB17, HMS3, ASB2, ASB23, COR32, AHT17, HTG02, HTG05, HTG07, HTG10, HTG14, HTG15, LEX46, LEX61, NVHEQ11, NVHEQ29, TKY16, UM04, UM10, UMNE61, VHL150, VHL209. From those we have selected 12 based on the best amplification success on degree of polymorphism and easiness of allele scoring (Table 1).

The markers were all amplified using GeneAmp® PCR System 9700 Dual 96-Well (Applied Biosystems®) thermocycler. The forward primers were modified to hybridize with a tail labelled with a fluorescent dyes (6-FAM™, VIC®, NED™, PET®) at the 5′ end during PCR. Individual PCR amplifications were performed on 10 µl reactions containing either 1µl of DNA, water, primers (0.6 µM of primer reverse and 0.06 µM of primer forward), fluorescent tails, bovine serum albumin (BSA) and Platinum® Multiplex PCR Master Mix, 2✕ (Applied Biosystems®).

Samples were amplified with the following conditions: initial denaturation at 94ºC for 15 min, 45 cycles of 30 s at 94ºC, 45 s at a range of temperatures between 52 ºC and 61ºC (depending of the locus), and 45 s at 72ºC.This step was followed with 10 cycles of 30 s at 94ºC, 45 s at 53ºC and 45s at 72 ºC, essential to the annealing of the fluorescent tail, a final extension of 72ºC during 20 min, and stored at 12ºC.

The amplification product was tested on a 2% agarose gel and stained with GelRed™ (Biotum).

The amplification products were separated by capillary electrophoresis using a 3130XLGenetic Analyzer® (AppliedBiosystems™) sequencer; the microsatellite scoring were done using GeneMapper® Software v4.0 (Applied Biosystems™) and Peak Scanner™ Software v1.0 (Applied Biosystems™).

In order to ensure the quality control, and thereby the reliability of our data set, a multiple tube approach was performed, which is considered the most accurate method to track allelic dropout, false alleles and sporadic contaminations on non-invasive samples (Taberlet *et al.* 1996). According to this method, genotypes were validated only after three consensual scores for heterozygous, and four for homozygous individuals. At the end we have obtained genotype scores of 42 samples were representing all the five sampling regions.

Data analysis

To estimate the data quality and verify errors rate, null allele frequency was calculated for all the 12 loci on Cervus 3.0 software (Kalinowski *et al.* 2007). GENALEX v 6.5 (Peakall & Smouse 2006, 2012) was used to calculate several parameters such as allele frequency, number of alleles (Na), observed heterozygosity (Ho) expected heterozygosity (He), mean number of alleles and F_{IS}. Allele frequency and its patterns over populations were also obtained with this software.

Deviations from Hardy-Weinberg equilibrium proportions were performed for each population–locus combination, with exact probability tests, using GENEPOP v.4.1.4 (Rousset, 2008), with Markov chain parameters set to 5000 batches and 1000 iterations. Using this same software, we the gametic disequilibrium (LD) was used to test for significantly non-random associations between alleles of different loci, with Markov chain parameters set to 5000 batches and 1000 iterations. A hierarchical analysis was performed to examine the population differentiation using analysis of molecular variance (AMOVA) using Arlequin v3.5 software (Excoffier & Lischer 2010). GENALEX v6.5 (Peakall & Smouse 2006, 2012) was used to calculate a Principal Coordinate Analysis (PcoA), in order to provide a spatial distribution of the genetic distances. To summarize population structure F-statistics were computed (Wright's F coefficients) using Arlequin v3.5 software (Excoffier & Lischer 2010).

Table 1 - Characteristics of the 12 microsatellite polymorphic loci accessed – name, primer sequences, dye label, size range (base pairs), annealing temperature (TA) and respective reference.

Locus	Primer sequences	Dye Label	Size range (bp)	TA (ºC)	Reference
AHT17	**F** - TGTAAAACGACGGCCAGTCCCCATAACCACAAGTG **R** - GAAGTGGGAGAGTCGGTAAGG	FAM	111-135	58 ºC	Swinburne et al. 1997
HTG05	**F -** GATAACAATTTCACACAGGTGCTAAGCCTCAG **R -** TGGAAATAAGGTTAGCAGGGATGC	PET	75–95	54 ºC	Ellegren et al. 1992
HTG10	**F -** GATAACAATTTCACACAGGCAATTCCCGCCCCACC **R -** TTTTTATTCTGATCTGTCACATTT	PET	94–110	58 ºC	Lindgren et al. 1998
HTG14	**F -** TAATACGACTCACTATAGGGCCAGTCTAAGTTTGTT **R -** CAAAGGTGAGTGATGGATGGAAGC	VIC	129–137	60 ºC	Lindgren et al.1998
HTG15	**F -** TAATACGACTCACTATAGGGTCTTGATGGCAGAGC **R -** AATGTCACCATGCGGCACATGACT	VIC	128–140	58 ºC	Lindgren et al.1998
LEX046	**F -** TTTCCCAGTCACGACGTTGATAAGCCAATCCACTTT **R -** ATTACCACCCCATTTCCTT	NED	114-120	57 ºC	Coogle et al. 1997
NVHEQ29	**F -** GATAACAATTTCACACAGGGAGATTTTGCCCCAAA **R -** CTCTTCTTTCTTCCCCAGGTCT	PET	91–103	61 ºC	Roed et al. 1998
TKY16	**F -** TAATACGACTCACTATAGGGGGTTATGGTTTGGTAT **R -** AAAACAATGGCTTCCTGGTCA	VIC	114–130	52 ºC	Kakoi et al. 1999
UM004	**F -** TTTCCCAGTCACGACGTTGAGGTCAGGTTCACTTTT **R -** AGGTCACTGTGCCTAGTTG	NED	110–122	53 ºC	Meyer et al. 1997
UM007	**F -** TGTAAAACGACGGCCAGTGGGAATAGAGAAAGGT **R -** TGTAAAACGACGGCCAGTGGGAATAGAGAAAGGT	FAM	110–148	60 ºC	Meyer et al. 1997
UM010	**F -** TTTCCCAGTCACGACGTTGTACAGCCATTGGAAATC **R -** CACCATTACATTTTCCCAG	NED	108–122	60 ºC	Meyer et al. 1997
VHL150	**F -** TTTCCCAGTCACGACGTTGCTCCTATGGGTGTCAGA **R -** TGTAATCTGTGGGGGAGACT	NED	92-104	60 ºC	van Haeringen et al. 1998

To infer the level of population genetic structure a Bayesian clustering method implemented on STRUCTURE v2.3.3 (Pritchard *et al.* 2000) was used. This method assigns individuals into genetic clusters and can reveal hidden population structure. To infer the number of genetic clusters, 20 independent runs of K between 1 and 5, were performed at 1,000,000 Markov Chain Monte Carlo (MCMC) repetitions with a 100,000 burn-in period. Results generated by STRUCTURE were accessed on STRUCTURE HARVESTER (Earl & vonHoldt 2011), which allows the access of likelihood scores in order to determine the most suitable level of population subdivision (best K value) performing, if possible, the Evanno correction (Evanno *et al.* 2005).

An assignment test was also performed to verify the population origin and it was carried on GENALEX v 6.5 (Peakall & Smouse 2006, 2012), which will also contribute to infer about structuration levels.

Results

A total of 42 individuals representing five different regions (Fig. 3) provided reliable genotypes for the 12 markers. A total of 128 different alleles were identified. The total number of alleles per locus ranged from five (HTG14) to 16 (AHT17). Locus HTG14 shows to be monomorphic on two populations (Gyrong and Nagqu), and for Ngamring population, five of the loci show also be monomorphic (VHL150, TKY16, NVHEQ29, HTG15 and UM004). In total, were detected 48 putative private alleles (observed only in one population) distributed in all the populations across all loci.

Table 2 - Summary statistics (mean values and SE): **N**-Mean number of alleles, **He** – Expected heterozygosity, **uHe -** unbiased expected heterozygosity**, Ho -** Observed heterozygosity

Population		N	He	uHe	Ho	F_{is}
Gyirong	Mean	7.500	0.713	0.768	0.152	0.788
	SE	0.485	0.067	0.072	0.046	0.057
Nagqu	Mean	5.583	0.620	0.684	0.133	0.804
	SE	0.417	0.071	0.079	0.058	0.081
Ngamring	Mean	1.500	0.323	0.431	0.125	0.657
	SE	0.230	0.084	0.111	0.065	0.123
Nyalam	Mean	5.000	0.693	0.776	0.101	0.853
	SE	0.369	0.033	0.036	0.032	0.049
Tingri	Mean	8.583	0.787	0.840	0.155	0.814
	SE	0.679	0.020	0.021	0.047	0.056

Different parameters of diversity showed a large variation across populations with the highest values often observed on Tingri, and the lower on Ngamring, which is expectable, due to the differences observed at expected heterozygosity level, (Table 2). The mean and standard deviation values for each population are present on Table S1.

The highest value of expected heterozygosity was found on Tingri population (0.787± 0.020), and the minimum on Ngamring (0.323± 0.084), whereas the observed heterozygosity presented values significantly lower, ranging between 0.155 ± 0.047 on Tingri and 0.101± 0.032 on Nyalam. The population which presents lower heterozygosity and mean number of alleles is Ngamring, which can be justified by the low number of sampled individuals.

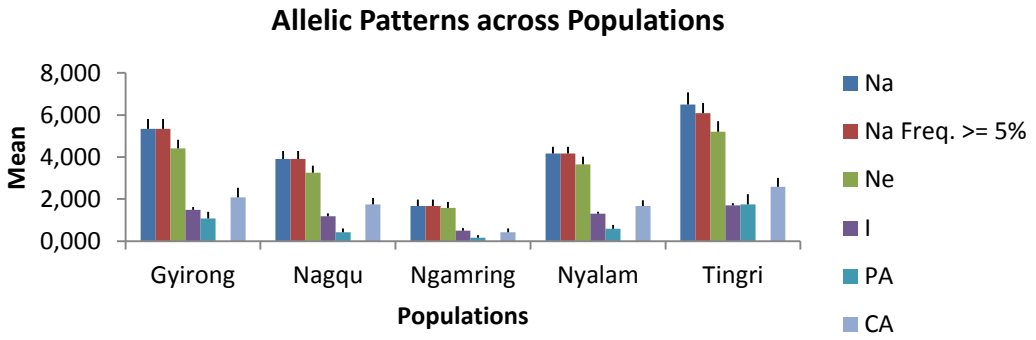

Figure 4- Allelic patterns observed across the 5 populations. **Na** - number of different alleles, **Na (Freq >= 5%)** - number of different Alleles with a frequency ≥5%, **Ne** - number of effective Alleles, **I** - Shannon's Information Index , **PA** - Number of private alleles (number of alleles unique to a single population), **CA** – Number of common alleles observed among 50% or fewer populations

The general F_{IS} values obtained were high (mean F_{IS} = 0.794 ± 0.032) ranging from 0.657 ± 0.123 on Ngamring and 0.853± 0.049 on Nyalam. The high level of F_{IS} is in agreement with the deficit of heterozygosity observed among all the samples. These values are probably influenced by the high rates of null alleles observed for all the loci, but can also be congruent with the occurrence of non-random mating.

Figure 5 shows how the different alleles are distributed among the populations, and is possible to observe that Ngamring is the populations that presented less number of alleles, due probably to sample size.

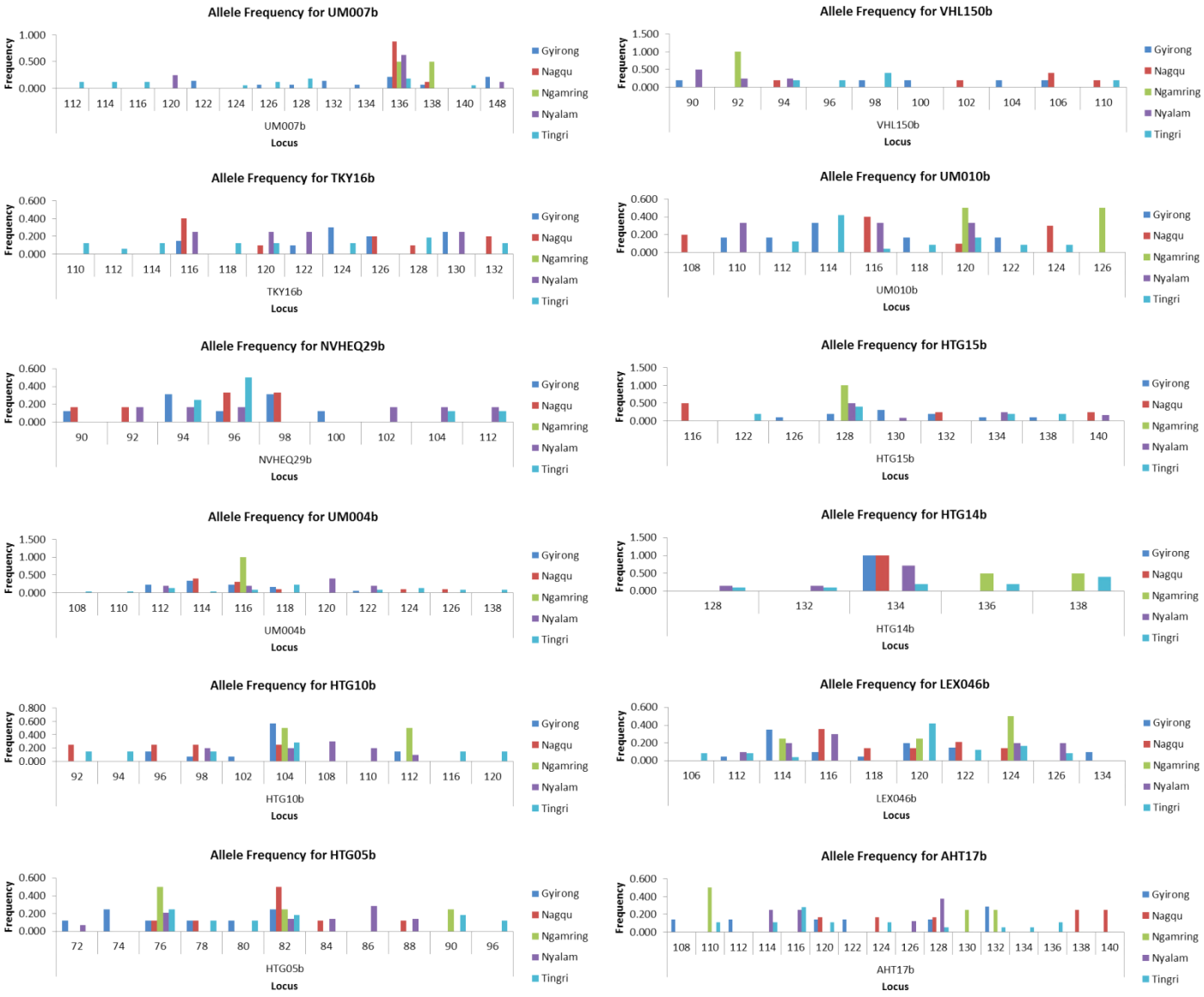

Figure 5 - Representation of the allelic frequency distributed for each loci and population. The colored bars correspond to the five sampling points/population (as indicated in the legend). Y axis indicates the frequency values, whereas X axis represents each allele of the locus.

The genetic structuration of the populations was accessed with STRUCTURE v2.3.3 software (Pritchard *et al.* 2000), and the obtained results were analyzed on STRUCTURE HARVESTER (Earl & vonHoldt 2011), which identified the maximum likelihood value for a K = 3, which means that theoretically the most suitable level of population subdivision corresponds to three clusters (Fig. 6). However, the analysis of clustering assignment shows that in general and

comparing among all the K values, the population structure level is not significant, and the most noteworthy differentiation occurred at individual's level.

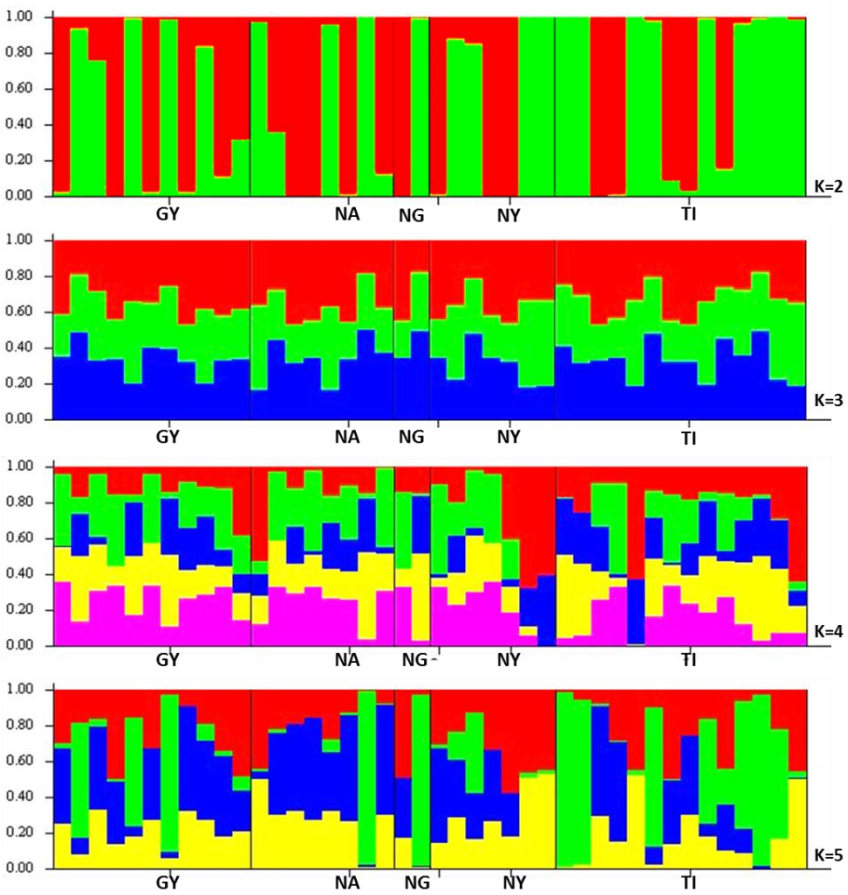

Figure 6 - Clustering outcomes for all the samples for the different K values (from K = 2 to K = 5). Each bar corresponds to an individual and each color represents one cluster (K). The amount of color in each vertical segment shows the level of association between the individual and the different clusters. Black lines separate the individual from the five different populations. Populations are labeled on the bottom of each graphic, **GY** - Girong, **NA** - Nagqu , **NG** - Ngamring , **NY** - Nyalam and **TI** - Tingri.

Global F_{ST} (measure of population subdivision) across all five populations was 0.111 .Since F_{ST} values ranges between zero (populations with no differentiation) and one (total differentiation), this value could be considered low, showing reduced differentiation within populations.

According to the pairwise F_{ST} values between populations (level of differentiation between pairs of populations) presented in Table 3, Ngamring (NG) is the most differentiated population (F_{ST}= 0.250), whereas the other four show similar F_{ST} values. According to pairwise values, the highest level of differentiation is found between Ngamring and Gyrong (F_{st} = 0.328), and the most close related populations are Nagqu and Nyalam (F_{ST} = 0.028). Apparently this

divergence could not be explained by geographic distance, since the farther and apparently most isolated population is Nagqu (see Fig. 7).

Table 3 - Population pairwise FSTs values

Population	Gyirong	Nagqu	Ngamring	Nyalam	Tingri
Gyirong	0.00000				
Nagqu	0.03744	0.00000			
Ngamring	0.32824	0.25377	0.00000		
Nyalam	0.03914	0.02769	0.26523	0.00000	
Tingri	0.14931	0.08167	0.15319	0.10651	0.00000
Mean	0.139	0.100	0.250	0.110	0.123

On Figure 7 are represented all the five sampled populations, and the respective genetic distances between the populations on a geographic view. It is possible to observe that despite the geographic distance of NA (Nagqu) population, the genetic distance between this and the other populations is in general very low.

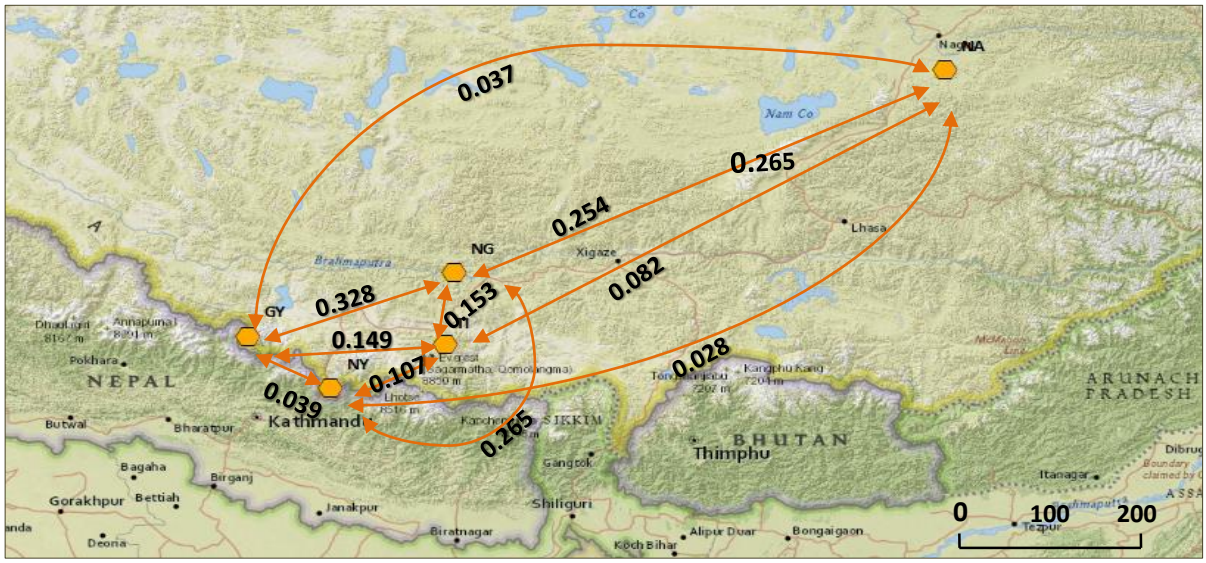

Figure. 7 - Geographical representation of pairwise Fst distance between all the five populations. **NY** - Nyalam, **TI**- Tingri, **NG**-Ngamrin, **GY**-Gyirong and NA-Nagqu.

The Principal Coordinate Analysis (PCoA) obtained through analysis of genetic distance, despite presenting low variation explained by the three axis, is capable to show some evident separation on Tingri (Axis 1) and Nagqu (Axis 3) (Figure 8 – A).

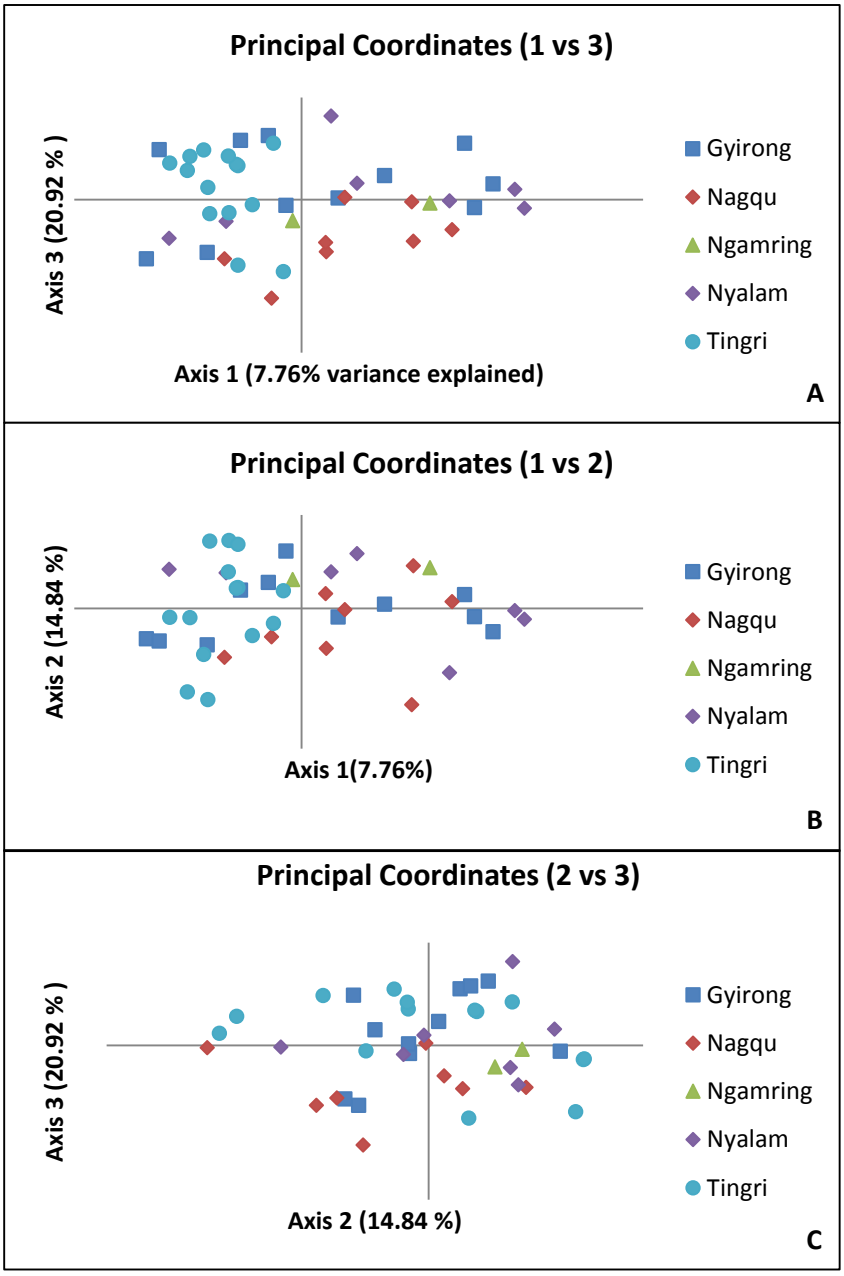

Figure 8 – Plotting of the first three Principal Coordinate analysis (PCoA). The variation explained by PC 1, PC 2 and PC 3, is, respectively 7.8%, 14.8% and 20.9%.

The population assignment tests showed that 98% of the individuals were correctly assigned, and only one individual was incorrectly assigned to the population of origin. This individual (EK012) belongs to Tingri but was wrongly allocated on Nagqu population (see Table 4). The complete result of population assignment test is presented on Table S2.

Table 4 – Summary of Population Assignment

Population	Self-Pop	Other Pop
Gyirong	11	
Nagqu	8	
Ngamring	2	
Nyalam	7	
Tingri	13	1
Total	41	1
Percent	98%	2%

Analysis of Molecular variance (AMOVA) performed on Arlequin v3.5 software (Excoffier & Lischer 2010), showed that approximately 89% of the total variance is observed within populations, and the extant 11% of the variance is observed among populations (Table 5). Which means that the great part of the genetic divergence observed comes from differences between individuals within populations instead of differences among populations. This can explain the low population structuration observed on the previous tests.

Table 5 - Analysis of Molecular variance (AMOVA) for all the five populations.

Source of variation	d.f	Sum of squares	Variance components	Percentage variation
Among populations	4	10.824	0.11337	11.07
Within populations	79	71.973	0.91105	88.93
Total		83	1.02442	

5. Discussion

On general the main pattern depicted across all five sampling points/populations was a relatively high heterozygosity deficit, with a large number of the loci significantly deviating from the proportions of HWE. One of the reasons for this is the high rate of null alleles observed across all loci, which ranged from 0.5062 (UM004) and 1 (VHL150). These kind of error can result from (1) the amplification of only one of the two alleles from an heterozygote individual, which originates genotyping errors by masking the occurrence of heterozygotes (Morin *et al.* 1994), (2) due to the low quality and amount of DNA and (3) the ascertainment bias of the amplification range caused by the use of primers designed for the horse (Chapuis & Estoup 2007), . However, the second hypothesis seems to be very plausible as the samples were only dehydrated upon arrival to the lab around two months after its collection, which probably compromised the DNA quality and the subsequent analysis. Another factor contributing to DNA degradation was probably the samples exposure to UV levels on field, as the Tibetan plateau presents the highest levels of ultraviolet radiation in the world (Norsang *et al.* 2011). The influence of UV lights was previously proved by Broquet *et al.*(2006) has been an influent factor of DNA degradation, constraining the amplification success. The animal diet could also influence DNA extraction and amplification success rates (Murphy *et al.* 2003). Since Kiang is an herbivore species, the presence of secondary compounds (eg. phenols) from plants can also be source for amplification inhibition (Wehausen *et al.* 2004). Seasonal variations on vegetation and diet quality can also present differential effects over amplification (Maudet *et al.* 2004), which means that samples collected on different seasons could present different amplification rates.
According to literature, the null allele values should not be above 0.2, and loci which present values higher than this limit must be excluded from the subsequent analysis (Handley *et al.* 2007).

Besides null allele, heterozygosity deficit and deviation for HWE proportions can also be caused by the occurrence of non-random mating (inbreeding). In fact, when looking to F_{IS} values, in general, they are substantially high for all the populations. These values can also be related to problems associated with field sampling, because since this species defecate in dung piles (Denzau & Denzau, 1999), it is probable that in some situations occurs the accidental collection of two samples from the same individual, which contributed for the increase of inbreeding coefficient values. Usually this kind of situations are detectable during the subsequent analysis, and the individual is eliminated from the dataset, unless the number of replications is not enough to validate the genotype, and be sure that it is not from the same individual. This could be difficult given the high levels of missing data within our dataset.

Since on this species the only permanent aggregation is observed between females and respective fouls (Schaller 1998, Denzau & Denzau,1999) it is also possible that in much cases the samples are collected from the female and from the foul, which are very close related, contributing to the increase of F_{IS} values.

The relatively moderate F_{ST} values, confirm the existence of population differentiation, although considerably lower, as observed on cluster analysis and PCoA, where the separation observed is weakly supported. Low levels of population structuration was also observed among plains zebra (*E. quagga*) (Lorenzen *et al.* 2008), whereas the levels of genetic diversity was considerably higher than those obtained for Kiang. On the other hand, the levels of genetic differentiation observed on Kiang, despite low ere considerable higher than those obtained for *E. hemionus* populations from Mongolian Gobi (Kaczensky *et al.* 2011).

The pairwise differentiation between sampling points did not correlate with the geographic distance among these regions (Fig. 7). As Nagqu is the most geographically distant population, it would be expected that this population would display the highest level of genetic differentiation, but that was not verified. Looking for this scenario, the most plausible explanation could be the fact of the effect of geographic distance over the gene flow has less impact than other barriers like mountain chains or other topographic elements that function like dispersal barrier, as observed on Mongolian Gobi wild ass populations (*E. hemionus*) (Kaczensky *et al.* 2011). Therefore, this could be one explanation for the low differentiation of Nagqu population compared with the other four: despite this population is the most geographically distant, the dispersal is probably not compromised and the individuals disperse more easily. This can be also applied the cases where geographically close populations show higher genetic differentiation. In fact, while carefully looking to the satellite images from the five populations locations, it can be inferred that Nagqu population inhabits a plateau, whereas the other populations, despite more close, inhabits extremely roughest terrains, surrounded by several mountain chains which probably difficult the dispersal.

Despite no migratory movements have been described on Kiang (Schaller 1998), it is possible to detect several seasonal movements mainly driven by food availability on different habitats. During summer they tend to move in small aggregations among more hilly terrain, whereas in winter they prefer to concentrate in more flat terrains (St-louis & Côte, 2009), which is probably related with the fact of during winter the species movements are highly conditioned by the climacteric conditions, since during this season, a great part of the terrain is covered by snow (Schaller *et al.* 2007). Therefore it is plausible to conclude that probably these populations are isolated during the harsh winter months, but during summer this connectivity is restored, matching with Kiang mating season, which occurs between July and August (Schaller 1998).

It is likely to conclude that apparently, and based on the results obtained for these samples, habitat fragmentation is not affecting these populations since the results evidences the occurrence of inter-population breeding. In fact, this fragmentation probably occurs, but just seasonally and due to natural causes like isolation due to climacteric conditions.

According to previous information about Kiang subspecies distribution (Schaller 1998; Shah 2002), we can observe that the five studied populations are approximately located on the range described for *E. k. polyodon (*Gyirong, Ngamring, Nyalam and Tingri*)* and *E. k. holdereri* (Nagqu) and it is possible that the sampled individuals belong to this subspecies; however, according to the results obtained in this study, the level of differentiation is very low. Nonetheless, further analysis and sampling increase are required to thoroughly test the validity of this subdivision.

In conclusion, in this work we did not find any sign of population fragmentation in E*. kiang,* but the high deficit of heterozigotes obtained for these populations can be a matter for concern, although is necessary to take in to account that the obtained results can be highly biased by the high rate of null alleles.

6. References

Beja-Pereira A, England PR, Ferrand N *et al.* (2004) African Origins of the Domestic Donkey. *Science*, **304**, 1781.

Beja-Pereira A, Oliveira R, Alves PC, Schwartz MK, Luikart G (2009) Advancing ecological understandings through technological transformations in noninvasive genetics. *Molecular Ecology Resources*, **9**, 1279–301.

Chapuis M-P, Estoup A (2007) Microsatellite null alleles and estimation of population differentiation. *Molecular Biology and Evolution*, **24**, 621–31.

Coogle L, Bailey E (1997) Equine dinucleotide repeat loci LEX049-LEX063. *Animal Genetics*, **28**, 378.

Costa et al, A new approach to improve DNA extraction from noninvasive sources, in preparation.

Earl DA, vonHoldt BM (2011) STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genetics Resources*, **4**, 359–361.

Ellegren H, Johansson M, Sandberg K, Andersson L (1992) Cloning of highly polymorphic microsatellites in the horse. *Animal Genetics*, **23**, 133–42.

Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Molecular Ecology*, **14**, 2611–20.

Excoffier L, Lischer HEL (2010) Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Molecular Ecology Resources*, **10**, 564–7.

Groves, CP,(1974) *Horses, asses and zebras in the wild*. David and Charles, London, United Kingdom.

Groves S, Handley L, Byrne K, Santucci F *et al.* (2007) Genetic structure of European sheep breeds. *Heredity*, **99**, 620–31.

Handley LJ, Byrne K, Santucci F*, et al.* (2007) Genetic structure of European sheep breeds. *Heredity (Edinb)* **99**, 620-631.

Kaczensky P, Kuehn R, Lhagvasuren B *et al.* (2011) Connectivity of the Asiatic wild ass population in the Mongolian Gobi. *Biological Conservation*, **144**, 920–929.

Kalinowski ST, Taper ML, Marshall TC (2007) Revising how the computer program CERVUS accommodates genotyping error increases success in paternity assignment. *Molecular Ecology*, **16**, 1099–106.

Krüger K, Gaillard C, Stranzinger G, Rieder S (2005) Phylogenetic analysis and species allocation of individual equids using microsatellite data. *Zeitschrift für Tierzüchtung und Züchtungsbiologie*, **122**, 78–86.

Lindgren, G., Sandberg, K., Persson et al., (1998) A primary male autosomal linkage map of the horse genome. *Genome Research* **8**, 951–966.

Lorenzen ED, Arctander P, Siegismund HR (2008) High variation and very low differentiation in wide ranging plains zebra (Equus quagga): insights from mtDNA and microsatellites. *Molecular Ecology*, **17**, 2812–24.

Maudet C, Luikart G, Dubray D, Von Hardenberg a., Taberlet P (2004) Low genotyping error rates in wild ungulate faeces sampled in winter. *Molecular Ecology Notes*, **4**, 772–775.

Meyer AH, Valberg SJ, Hillers KR et al., 1997 Sixteen new polymorphic equine mi- crosatellites. *Animal Genetics*, **28**, 58–71.

Moehlman PD, Kebede F, Yohannes H (1998) The African wild ass *(Equus africanus):* conservation status in the horn of Africa. *Applied Animal Behaviour Science,* **60**, 115–124.

Morin PA, Wallis J, Moore JJ, Woodruff DS (1994) Paternity exclusion in a community of wild chimpanzees using hypervariable simple sequence repeats. *Molecular ecology*, **3**, 469–77.

Murphy MA, Waits LP, Kendall KC (2003) The influence of diet on faecal DNA amplification and sex identification in brown bears (Ursus arctos). *Molecular Ecology*, **12**, 2261–2265.

Norsang G, Kocbach L, Stamnes J, Tsoja W, Pincuo N (2011) Spatial Distribution and Temporal Variation of Solar UV Radiation over the Tibetan Plateau. *Applied Physics Research*, **3**, 37–46.

Oakenfull EA, Lim HN, Ryder OA (2000) A survey of equid mitochondrial DNA : Implications for the evolution , genetic diversity and conservation of *Equus*. *Conservation Genetics*,**1**, 341–355.

Peakall R, Smouse PE (2006) genalex 6: genetic analysis in Excel. Population genetic software for teaching and research. *Molecular Ecology Notes*, **6**, 288–295.

Peakall R, Smouse PE (2012) GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research--an update. *Bioinformatics (Oxford, England)*, **28**, 2537–9.

Piggott MP, Taylor AC (2003) Remote collection of animal DNA and its applications in conservation management and understanding the population biology of rare and cryptic species. *Wildlife Research*, **30**, 1.

Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. *Genetics*, **155**, 945–59.

Rosenbom S, Costa V, Steck B et al., (2011) Cross-species genetic markers : a useful tool to study the world ' s most threatened wild equid — *Equus africanus*.*European Journal of Wildlife Research,* **3,** 609–613**.**

Rousset F (2008) genepop'007: a complete re-implementation of the genepop software for Windows and Linux. *Molecular ecology resources*, **8**, 103–106.

Ryder OA, Chemnick LG (1990) Chromosomal and molecular evolution in Asiatic wild asses. *Genetica*, **83**, 67–72.

Schaller GB (1998) *Wildlife of the Tibetan steppe*. University of Chicago Press, Chicago.

Schaller GB, Aili K, Tashi - Dorjie H, Ping C (2007) A winter wildlife survey in the northern Qiangtang of Tibet Autonomous Region and Qinghai Province , China. *Acta Theriologica Sinica*, **27**, 309–316.

Shah N (2002) Status and action plan for the kiang (Equus kiang). In: *Equids: zebras, asses and horses. Status survey and conservation action plan* (ed P. D. Moehlman), pp. 72–81. International Union for Conservation of Nature and Natural Resources/Species Survival Commission, Equid Specialist Group, Gland, Switzerland.

Sharma BD, Clevers J, Graaf R De, Chapagain NR (2004) Mapping Equus kiang (Tibetan Wild Ass) Habitat in Surkhang, Upper Mustang, Nepal. *Mountain Research and Development*, **24**, 149–156.

St-louis A, Côte SD (2009) Equus kiang (Perissodactyla: Equidae). *American Society of Mammalogists*, **835**, 1–11.

Swinburne E, Marti E, Breen M *et al.* (1997) Characterization of twelve new horse microsatellite loci : AHT12-AHT23. *Animal Genetics*, **28**, 453–454.

Taberlet P, Griffin S, Goossens B *et al.* (1996) Reliable genotyping of samples with very low DNA quantities using PCR. *Nucleic acids research*, **24**, 3189–94.

Taberlet P, Luikart G (1999) Non-invasive genetic sampling and individual identification. *Biological Journal of the Linnean Society*, **68**, 41–55.

van Haeringen WA, van de Goor LHP, van der Hout N et al.,(1998) Characterization of 24 equine microsatellite loci. *Animal Genetics*, **29**, 150–160.

Wang Y (2003) *A Complete Checklist of Mammal Species and Subspecies in China: A Taxonomic and Geographic Reference*. China Forestry Publishing House, Beijing.

Waring GH (2003) *Horse Behaviour*. Noyes Publications/William Andrew Publishing, New York.

Wehausen JD, Ramey RR, Epps CW (2004) Experiments in DNA extraction and PCR amplification from bighorn sheep feces: the importance of DNA extraction method. *The Journal of heredity*, **95**, 503–9.

7. Supplementary material

Supplementary tables

Table S1 - Summary Statistics with values for population and locus. **Na**- number of alleles, **He** – Expected heterozygosity, **Ho –** Observed heterozygosity and **Fis** – inbreeding coefficient

Population		UM007	VHL150	TKY16	UM010	NVHEQ29	HTG15	UM004	HTG14	HTG10	LEX046	HTG05	AHT17	Mean	SD
Gyirong	Ho	0.429	0.000	0.300	0.000	0.125	0.400	0.222	0.000	0.143	0.200	0.000	0.000	0.152	0.046
	He	0.847	0.800	0.775	0.778	0.758	0.800	0.759	0.000	0.622	0.790	0.813	0.816	0.713	0.067
	na	8	5	5	5	5	6	5	1	5	7	6	6	5.333	0.482
	Fis	0.494	1.000	0.613	1.000	0.835	0.500	0.707	#N/D	0.770	0.747	1.000	1.000	0.788	0.057
Nagqu	Ho	0.000	0.000	0.200	0.200	0.000	0.000	0.600	0.000	0.000	0.429	0.000	0.167	0.133	0.058
	He	0.219	0.720	0.740	0.700	0.722	0.625	0.720	0.000	0.750	0.765	0.688	0.792	0.620	0.071
	Na	2	4	5	4	4	3	5	1	4	5	5	5	3.917	0.379
	Fis	1.000	1.000	0.730	0.714	1.000	1.000	0.167	#N/D	1.000	0.440	1.000	0.789	0.804	0.081
Ngamring	Ho	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.500	0.500	0.500	0.125	0.065
	He	0.500	0.000	0.000	0.500	0.000	0.000	0.000	0.500	0.500	0.625	0.625	0.625	0.323	0.084
	Na	2	1	NA	2	NA	1	1	2	2	3	3	3	1.667	0.310
	Fis	1.000	#N/D		1.000		#N/D	#N/D	1.000	1.000	0.200	0.200	0.200	0.657	0.123
Nyalam	Ho	0.250	0.000	0.000	0.000	0.000	0.167	0.000	0.000	0.200	0.200	0.143	0.250	0.101	0.032
	He	0.531	0.625	0.750	0.667	0.833	0.653	0.720	0.449	0.780	0.780	0.806	0.719	0.693	0.033
	Na	3	3	4	3	6	4	4	3	5	5	6	4	4.167	0.322
	Fis	0.529	1.000	1.000	1.000	1.000	0.745	1.000	1.000	0.744	0.744	0.823	0.652	0.853	0.049
Tingri	Ho	0.375	0.000	0.250	0.333	0.000	0.000	0.364	0.000	0.000	0.083	0.125	0.333	0.155	0.047
	He	0.859	0.720	0.867	0.760	0.656	0.720	0.872	0.740	0.816	0.760	0.820	0.852	0.787	0.020
	Na	8	4	8	7	4	4	10	5	6	7	6	9	6.500	0.584
	Fis	0.564	1.000	0.712	0.562	1.000	1.000	0.583	1.000	1.000	0.890	0.848	0.609	0.814	0.056
F(Null allele)		0.5245	1	0.599	0.6547	0.9174	0.7086	0.5062	0.9999	0.8239	0.5873	0.8068	0.6197		

Table S2 - Results of population assignment test obtained with Genalex

Sample	Pop	Gyirong	Nagqu	Ngamring	Nyalam	Tingri	Assigned Pop	
EK029	Gyirong	-8.997	-20.796	-25.204	-18.894	-21.639	1	Gyirong
EK030	Gyirong	-12.683	-22.750	-30.000	-21.737	-26.195	1	Gyirong
EK031	Gyirong	-8.486	-16.408	-23.010	-20.778	-22.700	1	Gyirong
EK032	Gyirong	-14.058	-30.797	-30.107	-23.834	-25.488	1	Gyirong
EK033	Gyirong	-13.768	-24.213	-33.204	-31.477	-24.273	1	Gyirong
EK034	Gyirong	-11.613	-19.815	-26.107	-18.433	-23.783	1	Gyirong
EK060	Gyirong	-8.063	-14.644	-24.000	-18.852	-17.251	1	Gyirong
EK061	Gyirong	-9.276	-18.216	-22.903	-19.107	-18.836	1	Gyirong
EK062	Gyirong	-11.278	-22.857	-34.000	-26.945	-20.364	1	Gyirong
EK063	Gyirong	-5.496	-14.542	-20.602	-12.644	-9.033	1	Gyirong
EK064	Gyirong	-7.445	-19.866	-20.903	-17.590	-16.868	1	Gyirong
EK035	Nagqu	-19.303	-8.391	-24.107	-19.743	-21.334	2	Nagqu
EK036	Nagqu	-22.375	-9.177	-28.903	-18.910	-25.604	2	Nagqu
EK037	Nagqu	-16.005	-7.248	-24.301	-14.242	-20.373	2	Nagqu
EK038	Nagqu	-18.945	-8.737	-24.903	-20.187	-24.570	2	Nagqu
EK039	Nagqu	-22.348	-9.311	-33.505	-24.692	-21.015	2	Nagqu
EK040	Nagqu	-15.243	-7.674	-23.505	-20.260	-18.202	2	Nagqu
EK067	Nagqu	-20.558	-10.150	-25.505	-25.548	-19.468	2	Nagqu
EK069	Nagqu	-14.894	-5.699	-21.505	-15.797	-17.159	2	Nagqu
EK022	Ngamring	-20.578	-19.561	-4.515	-15.195	-14.090	3	Ngamring
EK026	Ngamring	-24.444	-27.396	-4.816	-25.544	-23.400	3	Ngamring
EK001	Nyalam	-23.908	-27.806	-28.602	-10.299	-20.400	4	Nyalam
EK002	Nyalam	-13.690	-13.156	-19.699	-6.408	-12.571	4	Nyalam
EK003	Nyalam	-18.901	-21.808	-30.408	-10.929	-22.304	4	Nyalam
EK004	Nyalam	-15.255	-19.118	-26.903	-10.799	-26.722	4	Nyalam
EK042	Nyalam	-27.947	-29.965	-29.806	-12.253	-31.002	4	Nyalam
EK043	Nyalam	-23.554	-26.477	-27.699	-8.761	-20.317	4	Nyalam
EK044	Nyalam	-24.000	-18.602	-24.000	-7.891	-17.204	4	Nyalam
EK010	Tingri	-20.253	-23.139	-28.301	-24.334	-10.390	5	Tingri
EK011	Tingri	-26.760	-32.954	-33.204	-28.760	-12.186	5	Tingri
EK012	Tingri	-11.487	-9.399	-22.000	-18.468	-9.586	2	Nagqu
EK013	Tingri	-17.681	-13.593	-19.806	-17.363	-10.086	5	Tingri
EK015	Tingri	-30.352	-26.292	-33.204	-22.859	-13.936	5	Tingri
EK016	Tingri	-23.601	-28.708	-22.903	-29.097	-11.859	5	Tingri
EK018	Tingri	-14.919	-20.961	-19.204	-15.210	-10.215	5	Tingri
EK019	Tingri	-13.849	-15.806	-13.204	-14.408	-9.058	5	Tingri
EK020	Tingri	-18.801	-28.544	-28.903	-31.352	-12.767	5	Tingri
EK051	Tingri	-14.104	-17.973	-16.408	-22.389	-10.498	5	Tingri
EK053	Tingri	-21.505	-17.699	-20.301	-21.255	-9.746	5	Tingri
EK056	Tingri	-20.261	-28.000	-21.204	-22.602	-10.090	5	Tingri
EK058	Tingri	-19.137	-15.681	-17.204	-15.641	-9.090	5	Tingri
EK079	Tingri	-19.748	-16.769	-17.505	-14.903	-10.701	5	Tingri

Table S3 - Results obtained for deviations from Hardy-Weinberg Equilibrium (HWE), across all the samples and loci. **ns** - Not significant, * P<0.05, ** P<0.01, *** P<0.001

Pop	Locus	DF	ChiSq	Prob	Signif
Gyirong	UM007b	28	42.778	0.037	*
Gyirong	VHL150b	10	20.000	0.029	*
Gyirong	TKY16b	10	20.733	0.023	*
Gyirong	UM010b	10	24.000	0.008	**
Gyirong	NVHEQ29b	10	26.880	0.003	**
Gyirong	HTG15b	15	20.556	0.152	ns
Gyirong	UM004b	10	21.000	0.021	*
Gyirong	HTG14b	Monomorphic			
Gyirong	HTG10b	10	28.000	0.002	**
Gyirong	LEX046b	21	41.315	0.005	**
Gyirong	HTG05b	15	40.000	0.000	***
Gyirong	AHT17b	15	35.000	0.002	**
Nagqu	UM007b	1	8.000	0.005	**
Nagqu	VHL150b	6	15.000	0.020	*
Nagqu	TKY16b	10	20.000	0.029	*
Nagqu	UM010b	6	10.556	0.103	ns
Nagqu	NVHEQ29b	6	18.000	0.006	**
Nagqu	HTG15b	3	8.000	0.046	*
Nagqu	UM004b	10	6.806	0.744	ns
Nagqu	HTG14b	Monomorphic			
Nagqu	HTG10b	6	12.000	0.062	ns
Nagqu	LEX046b	10	17.764	0.059	ns
Nagqu	HTG05b	10	32.000	0.000	***
Nagqu	AHT17b	10	18.667	0.045	*
Ngamring	UM007b	1	2.000	0.157	ns
Ngamring	VHL150b	Monomorphic			
Ngamring	TKY16b	Monomorphic			
Ngamring	UM010b	1	2.000	0.157	ns
Ngamring	NVHEQ29b	Monomorphic			
Ngamring	HTG15b	Monomorphic			
Ngamring	UM004b	Monomorphic			
Ngamring	HTG14b	1	2.000	0.157	ns
Ngamring	HTG10b	1	2.000	0.157	ns
Ngamring	LEX046b	3	4.000	0.261	ns
Ngamring	HTG05b	3	4.000	0.261	ns
Ngamring	AHT17b	3	4.000	0.261	ns
Nyalam	UM007b	3	4.160	0.245	ns
Nyalam	VHL150b	3	8.000	0.046	*
Nyalam	TKY16b	6	12.000	0.062	ns
Nyalam	UM010b	3	6.000	0.112	ns
Nyalam	NVHEQ29b	15	30.000	0.012	*
Nyalam	HTG15b	6	12.667	0.049	*
Nyalam	UM004b	6	15.000	0.020	*
Nyalam	HTG14b	3	14.000	0.003	**
Nyalam	HTG10b	10	15.556	0.113	ns
Nyalam	LEX046b	10	15.556	0.113	ns
Nyalam	HTG05b	15	28.778	0.017	*
Nyalam	AHT17b	6	8.444	0.207	ns
Tingri	UM007b	28	39.111	0.079	ns
Tingri	VHL150b	6	15.000	0.020	*
Tingri	TKY16b	28	46.222	0.017	*
Tingri	UM010b	21	45.880	0.001	**
Tingri	NVHEQ29b	6	24.000	0.001	***
Tingri	HTG15b	6	15.000	0.020	*
Tingri	UM004b	45	76.951	0.002	**
Tingri	HTG14b	10	40.000	0.000	***
Tingri	HTG10b	15	35.000	0.002	**
Tingri	LEX046b	21	61.333	0.000	***
Tingri	HTG05b	15	32.889	0.005	**
Tingri	AHT17b	36	61.560	0.005	**

Table S4 - Results obtained for Linkage Disequilibrium across all the populations for each pair of loci.

Pop	Locus#1	Locus#2	P-Value	S.E.	Switches
Gyirong	UM007	VHL150	N.Info		
Gyirong	UM007	TKY16	N.Info		
Gyirong	VHL150	TKY16	N.Info		
Gyirong	UM007	UM010	N.Info		
Gyirong	VHL150	UM010	N.Info		
Gyirong	TKY16	UM010	N.Info		
Gyirong	UM007	NVHEQ29	N.Info		
Gyirong	VHL150	NVHEQ29	N.Info		
Gyirong	TKY16	NVHEQ29	1	0	4082
Gyirong	UM010	NVHEQ29	0.10412	0.00641	10274
Gyirong	UM007	HTG15	N.Info		
Gyirong	VHL150	HTG15	N.Info		
Gyirong	TKY16	HTG15	N.Info		
Gyirong	UM010	HTG15	N.Info		
Gyirong	NVHEQ29	HTG15	N.Info		
Gyirong	UM007	UM004	N.Info		
Gyirong	VHL150	UM004	N.Info		
Gyirong	TKY16	UM004	1.000.000	0	3868
Gyirong	UM010	UM004	N.Info		
Gyirong	NVHEQ29	UM004	1.000.000	0	7649
Gyirong	HTG15	UM004	N.Infon		
Gyirong	UM007	HTG14	No. cont. tale		
Gyirong	VHL150	HTG14	No. cont. tale		
Gyirong	TKY16	HTG14	No. cont. tale		
Gyirong	UM010	HTG14	No. cont. tale		
Gyirong	NVHEQ29	HTG14	No. cont. tale		
Gyirong	HTG15	HTG14	No. cont. tale		
Gyirong	UM004	HTG14	No. cont. tale		
Gyirong	UM007	HTG10	N.Infon		
Gyirong	VHL150	HTG10	N.Infon		
Gyirong	TKY16	HTG10	N.Infon		
Gyirong	UM010	HTG10	N.Infon		
Gyirong	NVHEQ29	HTG10	1.000.000	0	10351
Gyirong	HTG15	HTG10	N.Infon		
Gyirong	UM004	HTG10	1.000.000	0	8392
Gyirong	HTG14	HTG10	No. cont. tale		
Gyirong	UM007	LEX046	N.Infon		
Gyirong	VHL150	LEX046	N.Infon		
Gyirong	TKY16	LEX046	1.000.000	0	4407
Gyirong	UM010	LEX046	1.000.000	0	6267
Gyirong	NVHEQ29	LEX046	1.000.000	0	5672
Gyirong	HTG15	LEX046	N.Info		
Gyirong	UM004	LEX046	1.000.000	0	3796
Gyirong	HTG14	LEX046	No. cont. tale		
Gyirong	HTG10	LEX046	0.39503	0.00984	11171
Gyirong	UM007	HTG05	N.Info		
Gyirong	VHL150	HTG05	N.Info		
Gyirong	TKY16	HTG05	1.000.000	0	5804
Gyirong	UM010	HTG05	N.Info		
Gyirong	NVHEQ29	HTG05	N.Info		
Gyirong	HTG15	HTG05	N.Info		
Gyirong	UM004	HTG05	1.000.000	0	5889
Gyirong	HTG14	HTG05	No. cont. tale		
Gyirong	HTG10	HTG05	1.000.000	0	14279
Gyirong	LEX046	HTG05	1.000.000	0	7206
Gyirong	UM007	AHT17	N.Info		
Gyirong	VHL150	AHT17	N.Info		
Gyirong	TKY16	AHT17	N.Info		
Gyirong	UM010	AHT17	N.Info		
Gyirong	NVHEQ29	AHT17	N.Info		
Gyirong	HTG15	AHT17	N.Info		
Gyirong	UM004	AHT17	1.000.000	0	6333
Gyirong	HTG14	AHT17	No. cont. tale		
Gyirong	HTG10	AHT17	N.Info		
Gyirong	LEX046	AHT17	1.000.000	0	6221
Gyirong	HTG05	AHT17	1.000.000	0	10601
Nagqu	UM007	VHL150	0.59902	0.00446	19933
Nagqu	UM007	TKY16	0.60516	0.00419	20270
Nagqu	VHL150	TKY16	N.Info		
Nagqu	UM007	UM010	No. cont. tale		
Nagqu	VHL150	UM010	1.000.000	0	19366
Nagqu	TKY16	UM010	No. cont. tale		
Nagqu	UM007	NVHEQ29	1.000.000	0	19370
Nagqu	VHL150	NVHEQ29	N.Info		
Nagqu	TKY16	NVHEQ29	1.000.000	0	19615
Nagqu	UM010	NVHEQ29	N.Info		
Nagqu	UM007	HTG15	No. cont. tale		
Nagqu	VHL150	HTG15	N.Info		
Nagqu	TKY16	HTG15	N.Info		
Nagqu	UM010	HTG15	N.Info		
Nagqu	NVHEQ29	HTG15	N.Info		
Nagqu	UM007	UM004	N.Info		
Nagqu	VHL150	UM004	N.Info		
Nagqu	TKY16	UM004	N.Info		
Nagqu	UM010	UM004	N.Info		
Nagqu	NVHEQ29	UM004	N.Info		
Nagqu	HTG15	UM004	N.Info		
Nagqu	UM007	HTG14	No. cont. tale		
Nagqu	VHL150	HTG14	No. cont. tale		
Nagqu	TKY16	HTG14	No. cont. tale		
Nagqu	UM010	HTG14	No. cont. tale		
Nagqu	NVHEQ29	HTG14	No. cont. tale		
Nagqu	HTG15	HTG14	No. cont. tale		
Nagqu	UM004	HTG14	No. cont. tale		
Nagqu	UM007	HTG10	No. cont. tale		
Nagqu	VHL150	HTG10	N.Info		
Nagqu	TKY16	HTG10	N.Info		
Nagqu	UM010	HTG10	N.Info		

Pop	Locus#1	Locus#2	P-Value	S.E.	Switches	Pop	Locus#1	Locus#2	P-Value	S.E.	Switches
Nagqu	NVHEQ29	HTG10	N.Info			Ngamring	UM007	UM004	No. cont. tale		
Nagqu	HTG15	HTG10	N.Info			Ngamring	VHL150	UM004	No. cont. tale		
Nagqu	UM004	HTG10	N.Info			Ngamring	TKY16	UM004	No data		
Nagqu	HTG14	HTG10	No. cont. tale			Ngamring	UM010	UM004	No. cont. tale		
Nagqu	UM007	LEX046	0.71353	0.00554	14202	Ngamring	NVHEQ29	UM004	No data		
Nagqu	VHL150	LEX046	N.Info			Ngamring	HTG15	UM004	No. cont. tale		
Nagqu	TKY16	LEX046	N.Info			Ngamring	UM007	HTG14	N.Info		
Nagqu	UM010	LEX046	N.Info			Ngamring	VHL150	HTG14	No. cont. tale		
Nagqu	NVHEQ29	LEX046	1.000.000	0	10393	Ngamring	TKY16	HTG14	No data		
Nagqu	HTG15	LEX046	N.Info			Ngamring	UM010	HTG14	N.Info		
Nagqu	UM004	LEX046	N.Info			Ngamring	NVHEQ29	HTG14	No data		
Nagqu	HTG14	LEX046	No. cont. tale			Ngamring	HTG15	HTG14	No. cont. tale		
Nagqu	HTG10	LEX046	N.Info			Ngamring	UM004	HTG14	No. cont. tale		
Nagqu	UM007	HTG05	1.000.000	0	12464	Ngamring	UM007	HTG10	N.Info		
Nagqu	VHL150	HTG05	1.000.000	0	14243	Ngamring	VHL150	HTG10	No. cont. tale		
Nagqu	TKY16	HTG05	1.000.000	0	14151	Ngamring	TKY16	HTG10	No data		
Nagqu	UM010	HTG05	1.000.000	0	10414	Ngamring	UM010	HTG10	N.Info		
Nagqu	NVHEQ29	HTG05	1.000.000	0	9238	Ngamring	NVHEQ29	HTG10	No data		
Nagqu	HTG15	HTG05	1.000.000	0	19514	Ngamring	HTG15	HTG10	No. cont. tale		
Nagqu	UM004	HTG05	N.Info			Ngamring	UM004	HTG10	No. cont. tale		
Nagqu	HTG14	HTG05	No. cont. tale			Ngamring	HTG14	HTG10	N.Info		
Nagqu	HTG10	HTG05	N.Info			Ngamring	UM007	LEX046	N.Info		
Nagqu	LEX046	HTG05	1.000.000	0	7070	Ngamring	VHL150	LEX046	No. cont. tale		
Nagqu	UM007	AHT17	N.Info			Ngamring	TKY16	LEX046	No data		
Nagqu	VHL150	AHT17	N.Info			Ngamring	UM010	LEX046	N.Info		
Nagqu	TKY16	AHT17	N.Info			Ngamring	NVHEQ29	LEX046	No data		
Nagqu	UM010	AHT17	N.Info			Ngamring	HTG15	LEX046	No. cont. tale		
Nagqu	NVHEQ29	AHT17	N.Info			Ngamring	UM004	LEX046	No. cont. tale		
Nagqu	HTG15	AHT17	N.Info			Ngamring	HTG14	LEX046	N.Info		
Nagqu	UM004	AHT17	N.Info			Ngamring	HTG10	LEX046	N.Info		
Nagqu	HTG14	AHT17	No. cont. tale			Ngamring	UM007	HTG05	N.Info		
Nagqu	HTG10	AHT17	N.Info			Ngamring	VHL150	HTG05	No. cont. tale		
Nagqu	LEX046	AHT17	N.Info			Ngamring	TKY16	HTG05	No data		
Nagqu	HTG05	AHT17	N.Info			Ngamring	UM010	HTG05	N.Info		
Ngamring	UM007	VHL150	No. cont. tale			Ngamring	NVHEQ29	HTG05	No data		
Ngamring	UM007	TKY16	No data			Ngamring	HTG15	HTG05	No. cont. tale		
Ngamring	VHL150	TKY16	No data			Ngamring	UM004	HTG05	No. cont. tale		
Ngamring	UM007	UM010	N.Info			Ngamring	HTG14	HTG05	N.Info		
Ngamring	VHL150	UM010	No. cont. tale			Ngamring	HTG10	HTG05	N.Info		
Ngamring	TKY16	UM010	No data			Ngamring	LEX046	HTG05	N.Info		
Ngamring	UM007	NVHEQ29	No data			Ngamring	UM007	AHT17	N.Info		
Ngamring	VHL150	NVHEQ29	No data			Ngamring	VHL150	AHT17	No. cont. tale		
Ngamring	TKY16	NVHEQ29	No data			Ngamring	TKY16	AHT17	No data		
Ngamring	UM010	NVHEQ29	No data			Ngamring	UM010	AHT17	N.Info		
Ngamring	UM007	HTG15	No. cont. tale			Ngamring	NVHEQ29	AHT17	No data		
Ngamring	VHL150	HTG15	No data			Ngamring	HTG15	AHT17	No. cont. tale		
Ngamring	TKY16	HTG15	No data			Ngamring	UM004	AHT17	No. cont. tale		
Ngamring	UM010	HTG15	No. cont. tale			Ngamring	HTG14	AHT17	N.Info		
Ngamring	NVHEQ29	HTG15	No data			Ngamring	HTG10	AHT17	N.Info		

Pop	Locus#1	Locus#2	P-Value	S.E.	Switches	Pop	Locus#1	Locus#2	P-Value	S.E.	Switches
Ngamring	LEX046	AHT17	N.Info			Nyalam	TKY16	HTG05	N.Info		
Ngamring	HTG05	AHT17	N.Info			Nyalam	UM010	HTG05	N.Info		
Nyalam	UM007	VHL150	1.000.000	0	33513	Nyalam	NVHEQ29	HTG05	N.Info		
Nyalam	UM007	TKY16	N.Info			Nyalam	HTG15	HTG05	N.Info		
Nyalam	VHL150	TKY16	N.Info			Nyalam	UM004	HTG05	0.09908	0.00559	10488
Nyalam	UM007	UM010	N.Info			Nyalam	HTG14	HTG05	1.000.000	0	9376
Nyalam	VHL150	UM010	N.Info			Nyalam	HTG10	HTG05	N.Info		
Nyalam	TKY16	UM010	N.Info			Nyalam	LEX046	HTG05	N.Info		
Nyalam	UM007	NVHEQ29	N.Info			Nyalam	UM007	AHT17	N.Info		
Nyalam	VHL150	NVHEQ29	N.Info			Nyalam	VHL150	AHT17	No. cont. tale		
Nyalam	TKY16	NVHEQ29	N.Info			Nyalam	TKY16	AHT17	N.Info		
Nyalam	UM010	NVHEQ29	N.Info			Nyalam	UM010	AHT17	N.Info		
Nyalam	UM007	HTG15	N.Info			Nyalam	NVHEQ29	AHT17	N.Info		
Nyalam	VHL150	HTG15	N.Info			Nyalam	HTG15	AHT17	No. cont. tale		
Nyalam	TKY16	HTG15	N.Info			Nyalam	UM004	AHT17	N.Info		
Nyalam	UM010	HTG15	N.Info			Nyalam	HTG14	AHT17	N.Info		
Nyalam	NVHEQ29	HTG15	N.Info			Nyalam	HTG10	AHT17	N.Info		
Nyalam	UM007	UM004	N.Info			Nyalam	LEX046	AHT17	N.Info		
Nyalam	VHL150	UM004	N.Info			Nyalam	HTG05	AHT17	N.Info		
Nyalam	TKY16	UM004	N.Info			Tingri	UM007	VHL150	N.Info		
Nyalam	UM010	UM004	N.Info			Tingri	UM007	TKY16	N.Info		
Nyalam	NVHEQ29	UM004	N.Info			Tingri	VHL150	TKY16	N.Info		
Nyalam	HTG15	UM004	N.Info			Tingri	UM007	UM010	N.Info		
Nyalam	UM007	HTG14	0.50358	0.00379	25298	Tingri	VHL150	UM010	1.000.000	0	19648
Nyalam	VHL150	HTG14	0.492	0.00306	24792	Tingri	TKY16	UM010	N.Info		
Nyalam	TKY16	HTG14	No. cont. tale			Tingri	UM007	NVHEQ29	N.Info		
Nyalam	UM010	HTG14	No. cont. tale			Tingri	VHL150	NVHEQ29	N.Info		
Nyalam	NVHEQ29	HTG14	N.Info			Tingri	TKY16	NVHEQ29	N.Info		
Nyalam	HTG15	HTG14	0.50426	0.00539	16675	Tingri	UM010	NVHEQ29	1.000.000	0	8066
Nyalam	UM004	HTG14	1.000.000	0	14163	Tingri	UM007	HTG15	N.Info		
Nyalam	UM007	HTG10	N.Info			Tingri	VHL150	HTG15	No. cont. tale		
Nyalam	VHL150	HTG10	N.Info			Tingri	TKY16	HTG15	N.Info		
Nyalam	TKY16	HTG10	N.Info			Tingri	UM010	HTG15	1.000.000	0	10356
Nyalam	UM010	HTG10	N.Info			Tingri	NVHEQ29	HTG15	N.Info		
Nyalam	NVHEQ29	HTG10	N.Info			Tingri	UM007	UM004	N.Info		
Nyalam	HTG15	HTG10	N.Info			Tingri	VHL150	UM004	N.Info		
Nyalam	UM004	HTG10	N.Info			Tingri	TKY16	UM004	N.Info		
Nyalam	HTG14	HTG10	No. cont. tale			Tingri	UM010	UM004	1.000.000	0	2748
Nyalam	UM007	LEX046	N.Info			Tingri	NVHEQ29	UM004	N.Info		
Nyalam	VHL150	LEX046	N.Info			Tingri	HTG15	UM004	N.Info		
Nyalam	TKY16	LEX046	N.Info			Tingri	UM007	HTG14	N.Info		
Nyalam	UM010	LEX046	N.Info			Tingri	VHL150	HTG14	N.Info		
Nyalam	NVHEQ29	LEX046	N.Info			Tingri	TKY16	HTG14	N.Info		
Nyalam	HTG15	LEX046	N.Info			Tingri	UM010	HTG14	0.58461	0.00884	11657
Nyalam	UM004	LEX046	N.Info			Tingri	NVHEQ29	HTG14	1.000.000	0	10382
Nyalam	HTG14	LEX046	N.Info			Tingri	HTG15	HTG14	N.Info		
Nyalam	HTG10	LEX046	N.Info			Tingri	UM004	HTG14	N.Info		
Nyalam	UM007	HTG05	N.Info			Tingri	UM007	HTG10	N.Info		
Nyalam	VHL150	HTG05	N.Info			Tingri	VHL150	HTG10	N.Info		

Pop	Locus#1	Locus#2	P-Value	S.E.	Switches
Tingri	TKY16	HTG10	N.Info		
Tingri	UM010	HTG10	1.000.000	0	8312
Tingri	NVHEQ29	HTG10	1.000.000	0	14295
Tingri	HTG15	HTG10	No. cont. tale		
Tingri	UM004	HTG10	N.Info		
Tingri	HTG14	HTG10	N.Info		
Tingri	UM007	LEX046	N.Info		
Tingri	VHL150	LEX046	1.000.000	0	10424
Tingri	TKY16	LEX046	N.Info		
Tingri	UM010	LEX046	1.000.000	0	3965
Tingri	NVHEQ29	LEX046	0.81736	0.007	10804
Tingri	HTG15	LEX046	N.Info		
Tingri	UM004	LEX046	1.000.000	0	3484
Tingri	HTG14	LEX046	0.68345	0.01237	6957
Tingri	HTG10	LEX046	1.000.000	0	8381
Tingri	UM007	HTG05	N.Info		
Tingri	VHL150	HTG05	N.Info		
Tingri	TKY16	HTG05	N.Info		
Tingri	UM010	HTG05	1.000.000	0	6197
Tingri	NVHEQ29	HTG05	1.000.000	0	24845
Tingri	HTG15	HTG05	No. cont. tale		
Tingri	UM004	HTG05	1.000.000	0	6201
Tingri	HTG14	HTG05	1.000.000	0	16769
Tingri	HTG10	HTG05	N.Info		
Tingri	LEX046	HTG05	0.16567	0.01294	5139
Tingri	UM007	AHT17	N.Info		
Tingri	VHL150	AHT17	N.Info		
Tingri	TKY16	AHT17	N.Info		
Tingri	UM010	AHT17	1.000.000	0	6908
Tingri	NVHEQ29	AHT17	0.16322	0.0049	19467
Tingri	HTG15	AHT17	N.Info		
Tingri	UM004	AHT17	N.Info		
Tingri	HTG14	AHT17	N.Info		
Tingri	HTG10	AHT17	N.Info		
Tingri	LEX046	AHT17	1.000.000	0	3661
Tingri	HTG05	AHT17	N.Info		

CHAPTER 3

Manuscript 2

Manuscript 2

A COAT COLOR ASSESSMENT ON DOMESTIC AND WILD EQUIDS

Rita Monteiro[1], Shanyuan Chen[1], Albano Beja-Pereira[1]

[1] CIBIO-INBIO, Universidade do Porto, Rua Padre Armando Quintas, Campus Agrário de Vairão, Vairão, Portugal

1. Abstract

Since the very first moments of domestication, herders have been selecting their animals to favor a wide range of phenotypic traits. Among those, the coat color is probably the most celebrate phenotypic trait. It is clear that when comparing domestic with their wild ancestral populations the first have a much broad range of colors whereas the wild ones tend to present a great homogeneity. Equids are is an excellent genus to study the evolutionary history of the candidate color genes, as is the only genus having to species domesticated. Here we assessed the levels of molecular diversity at the five main coat color candidate genes (*ASIP*, *MC1R*, *SLC45A2*, *SILV* and *SLC36A1*) for several domestic and wild equid species. We were able to observe that despite the phenotypic differences, most of the mutations are shared among wild and domestic species. This showing that the majority of these mutations arose before domestication and that human-mediated selection or/and selection pressure relaxation have just acted to increase the frequency of those mutations in the domestic species.

Keywords: coat color; domestication; wild; phenotypic diversity; genetic diversity; polymorphisms; natural selection; artificial selection

2. Introduction

The domestication process and the subsequent breeding of the domestic species are mostly characterized by the human selection of their animals based in some desired phenotypes. Among these, coat color is one of the most noticeable ones as it is controlled by the expression and interaction of relatively few genes and thus easy to impose selection directions (Svensson *et al.* 2012). The first records of coat color on domestic animals date back from at least 5000 years B.P, during the UR III dynasty in Mesopotamia, when administrative officers annotated the livestock coat color (Zeder et al.,1994).

As coat color genes, which mostly follow a Mendelian inheritance and can be easily recorded and recognizable among generations (Hauswirth *et al.* 2012), opened the possibility of being one of the very first traits to be analyzed at molecular level (Rieder et al., 2009), and becoming a widely used model to study gene action and interaction (Reissmannet al., 2007). Furthermore, coat color also proved to be an interesting tool to reconstruct the domestication processes. In pigs, studies focused on two coat color genes, MC1R (Fang et al.,2009; Jeon et al.,2000) and TYRP1 (Jeon *et al.* 2000) allowed to infer about pig domestication origins and in 2009. Ludwig et al. (2009) published a study comparing the occurrence and frequency of coat color responsible alleles along the horse domestication process, from ancient wild horses from the late Pleistocene until Iron Age horses. In these studies, it is shown that color variation have largely increased during domestication, probably due to human selection (Fang *et al.* 2009; Ludwig *et al.* 2009). Indeed the majority of the domestic species show a high level of coat color variation, contrary to the uniform color phenotypes observed on wild populations (Cieslak et al.,2011).

The large phenotypic differences between domestic and wild species, are an interesting fact since, most likely, the current domestic populations were founded by a relatively small number of wild individuals and subjected to a strong selection pressures to differentiate from their ancestral forms. The studies on coat color variation among domestic and wild equids can contribute to answer two old questions: (*i*) how could all coat color phenotypes have been originated from a uniform ancestor? And, (*ii*) how such a shallow time scale since the domestication of equids was capable to produce the current variation, when compared with the deep time period that took the speciation in equids?

The fact that the Equus genus is composed by several species showing distinct, but intra-specifically uniform, phenotypes is very interesting. The study of its molecular basis can shed some light on the evolution of coat color genes. Moreover, this is the only animal genus from which two independent species (horse and donkey) were domesticated, and this can serve as a natural replica in any coat color evolution study.

Coat color is a process triggered by a considerable number of genes in interaction, and also influenced by several epistatic effects (Cieslak et al., 2011). MC1R (E) gene along with ASIP (A) gene are responsible for the three basic coat colors observed in horses: black (EE/Ee, aa), chestnut (ee, AA/Aa/aa) and bay (EE/Ee, AA/ Aa). Both loci are dominantly inherited [MC1R (E) is inherited dominantly to chestnut (e), and Bay (A) is inherited dominantly to black (a)] and chestnut is recessively epistatic to ASIP. Two mutations on MC1R gene (E) have been reported as responsible for the two known chestnut alleles: (e) C to T missense mutation at

codon 83, resulting in the replacement of a serine for a phenylalanine (83Ser → Phe) (Marklund et al., 1996); (e^a) G to A missense mutation at codon 84 resulting in the replacement of an aspartic acid for an asparagine (84 Asp→Asn) = (84 GAC→AAC) (Wagner & Reissmann 2000). In ASIP gene (A), an 11 base pair deletion on exon 2, which extends the termination signal by 210 bp to 612 bp, was reported by Rieder et al. (2001) as being associated with horse recessive black coat color black allele (A^a / A^a).

The SLC45A2/MATP, SLC36A1 and PMEL17/SILV, are three of the genes responsible for the dilution of the basic coat colors. SLC45A2/MATP (C^{CR}), is associated to the cream dilution and inherited by incomplete dominance. According to the background basic color (chestnut, bay and black, respectively), on heterozygosity this dilution originates palomino (ee , AA/Aa/aa, C^{CR}/C), buckskin (EE/Ee, AA/Aa) and smoky seal brown (EE/Ee, aa , C^{CR}C). In dominant homozygosity all the basic colors are diluted to a very light coat color and light blue eyes, Perlino/ Cremello (C^{CR}/C^{CR}). The mutation associated with C^{CR} allele was described by Mariat et al.,(2003) as a single base substitution in exon 2 (G153A), leading to substitution of an aspartic acid for an asparagine.

SLC36A1 gene is associated with champagne dilution and follows an autosomal dominant mode of inheritance. The produced phenotypes are very similar to ones resulting of the cream dilution (Cook et al.,2008). According to the basic color (chestnut, bay and black respectively), the dilution results on a Gold champagne (ee, AA/Aa/aa, CH/Ch), Amber champagne (EE/Ee, AA/Aa, CH/ch) and Classical champagne color (EE/Ee, aa, CH/Ch) (Thiruvenkadan et al., 2008). This dilution is linked to nucleotide substitution in exon 2, resulting on the change of a threonine to an arginine (T63R) (Cook et al.,2008).

PMEL17/SILV (Z) gene is responsible for the silver color, almost only diluting black pigment present on bay and black horses, but a few dilution could also be observed on chestnut horses, following a dominant mode of inheritance. This dilution is more evident on the tail and main hairs, diluting it to almost white. A bay horse with silver dilution is defined as silver bay/ silver maine - chestnut (EE/Ee, AA/Aa, ZZ/Zz), and a black is diluted to an black-chocolate color (EE/Ee, aa, ZZ/Zz). Brunberg et al.,(2006) detected primarily a missense mutation, consisting on a C to T transition at position five of the Exon 11, that conduces to an amino acid change from arginine to cysteine (Arg618Cys), and later, Reissmann et al. (2007) detected more two polymorphism on PMEL17 gene associated to silver color: the 697A>T and 1457C>T. In fact, they found that black and bay horses with silver dilution carry the [g.697T; g.1457T] haplotype and the chestnut horses carry the [g.697T; g.1457T] haplotype despite few or none phenotypic effects caused by this dilution.

In equids, there is a dilution factor, identified as Dun (D) which is characterized by appearing along with several primitive color marks, as dorsal stripe and zebra stripes (Rieder, 2009). This is inherited as a dominant trait, however, since none gene association have been described yet (Bricker et al., 2003) it is impossible to genetically distinguish between a dun (Dd/DD) and a non dun horse (dd), but is very likely that this dilution occurred on pre-domestic horses, as in modern Przewalski horses, and is also present on wild equids, like wild ass, showing a bay-dun coloration (Pruvost *et al.* 2011).

In this work, we have partially sequenced five coat color associated genes - MC1R, ASIP, SLC45A2/MATP, PMEL17/SILV, and SLC36A1 - to compare levels of diversity between wild and domestic equid species.

3. Materials and Methods

Sampling

Noninvasive fecal samples were used to obtain DNA from asiatic wild ass (*E. kiang* [N=1], E. *hemionus* [N=1]*) and Grevy zebra (E. gravy* [*N*=1]*)*. Blood samples were used as DNA source from domestic horses (*E. caballus* [*N*=7]), domestic donkey (*E.asinus* [*N*=3]), and African wild ass (*E. africanus* [*N*=3]). With this sampling we intended to collect the widest number of coat colors as possible (Table 6).

Genomic DNA extraction was performed using JETQUICK Tissue DNA Spin Kit (Genomed ®), using adapted protocol developed by Costa et al (in prep) for the fecal samples, and standard protocols, according to the manufacturer to the blood samples.

Primers were designed using the *PrimerBlast* NCBI tool, based on previously published gene sequences from horse (*Equus caballus*). All these sequences were obtained on Genbank, and the Accession numbers are given on Table 5S.

All primer pairs were previously optimized for an invasive and a noninvasive sample as the conditions can be influenced by quality and quantity of the template. These optimizations were carried out with two different polymerase enzymes, AmpliTaq Gold® 360 DNA Polymerase (Applied Biosystems®) and Platinum® Taq DNA Polymerase ((0.3 U); Invitrogen™) with two different amount of $MgCl_2$ concentration on the reaction: 1.5 and 2mM. By this way it was possible to use the condition which most increases the amplification success and quality for each primer pair.

Table 6 - List of extracted samples, individuals coat color and source of DNA

Species and code (N = 14)	Color	Sample Type
Equus kiang (EK)	Wild (dun)	Feces
Equus grevy(GZ)	Wild(strips)	Feces
Equus hemionus(WAI)	Wild(dun)	Feces
Equus caballus (1)	Grey	blood
Equus caballus (2)	Grey	blood
Equus caballus (3)	Brown	blood
Equus caballus (4)	Brown	blood
Equus caballus (5)	Possibly black	blood
Equus africanus (HOI)	Dun	blood
Equus caballus (Garrano)	Dun	blood
Equus caballus(Sorraia)	Dun	blood
Equus asinus(Oman)	White	blood
Equus asinus (Somalia)	Grey	blood
Equus asinus (Miranda)	Brown	blood

PCR amplifications were performed on 20µl volume reactions containing 2µl of DNA, ddH_2O, both primers (between 0.6µM and 1 µM according to the primer), 0.2 µg/µLl of bovine serum albumin (BSA), (50mM) $MgCl_2$, and one of the two different polymerase enzymes, according to the fragment: in some cases the reaction was performed using AmpliTaq Gold® 360 DNA Polymerase (Applied Biosystems®), and in others, with Platinum® Taq DNA Polymerase ((0.3 U); Invitrogen™). The magnesium amount is dependent of the Taq used. All the PCR conditions as well as primer sequences are present on Table 5S.

Samples were amplified on a Veriti® Dx 96-Well Thermal Cycler (Applied Biosystems®) thermocycler, with the following conditions: initial denaturation at 94ºC for 15 min, 45 cycles of 30 s at 94ºC, 45 s at a range of temperatures between 52ºC and 61ºC (according to the loci), and 45 s at 72ºC.This step was followed by a final extension of 72ºC during 20 min, and stored at 12ºC. A 2µl aliquot of the amplified product were tested on a 2% agarose gel and stained with GelRed™ (Biotum).

As some reactions presented secondary/unspecific bands, the band with the expected size was excised from the agarose gel and purified using the JETQUICK Gel Extraction Spin Kit (Genomed®). The amplicons were then sequenced on an ABI3730XL capillary sequencer (Applied Biosystems®).

Data analysis

Sequence editing was performed on DNASTAR (DNASTAR Inc., Madison) and the alignment on MEGA 5.2. (Tamura *et al.* 2011). Due to the very limiting quality and quantity of the genomic DNA it was not possible to amplify or obtain sequences for all exons of each gene from each of the tested individuals. All sequences were aligned using the same reference sequences that were previously used for primer design – *ASIP* (Accession number: NC_009165.2), *MC1R* (Accession number: NC_009146), *SILV* (Accession number: NC_009149.2), *SLC36A1* (Accession number: NC_009157) and *SLC45A2* (Accession number: NC_009164.2). All the polymorphic sites were annotated in MEGA software, using one sequence representative of each of the described coat color phenotypes (Table 7S).

All the obtained sequences were phased using DNAsp (Librado & Rozas 2009), and then the phylogenetic relationships between haplotypes were calculated by Median-Joining network method contained in the NETWORK 4.6.1.1 software (Bandelt et al., 1999). In this analysis all unidentified nucleotides (N) were trimmed and sequences were all of the same size. Therefore the number of represented mutations (Table S6) is higher than those observed in the haplotype networks.

4. Results

A total of 145 polymorphisms have been detected. The most part of these SNPs were observed in coding-regions; 43 of them are synonymous, not leading to amino acid changes, and 61 are non-synonymous, conducting to functional changes. On non-coding regions were observed 41 mutations.

In the *ASIP*, we have observed three polymorphisms distributed in both domestic and wild equids. Two were located in non-coding-regions and one non-synonym mutation on a coding region. An 11 bp deletion on exon 2 of two domestic horses was also observed In fact, an 11 base pair deletion have been previously described as responsible for the black coat color phenotype in horses Rieder et al.(2001), however, the deletion obtained with this work was observed on a different position than the previously annotated.

In the *MC1R* we have observed 15 polymorphisms, all located in coding regions (this gene presents a single exon). From this, four were synonymous (silent) and 11 were non-synonymous. Two of these polymorphic sites were exclusive of previously published sequences (Table S1).The two previous described mutation for this gene, responsible for the chestnut color (Marklund *et al.* 1996; Wagner & Reissmann 2000), were observed only in domestic horses and *E. przewalskii*, and were not observed in any other species.

In the *SILV* (or *PMEL17*) we have observed 48 SNPs, 23 non-synonymous, 9 synonymous and 16 in non-coding regions. We did not found the SNP that was previously associated to the silver dilution on horse (Brunberg *et al.* 2006; Reissmann *et al.* 2007).

From the total 40 SNPs detected in *SLC36A1*, 13 were silent, 9 functional and 18 occurred on non-coding regions. Only a horse sequence collected from NCBI showed the annotated mutation associated with this gene (champagne dilution).

In SLC45A2, were detected a total of 36 polymorphisms, from which 15 were silent, 17 functional and 4 occurred on non-coding regions. The mutation that that have been previously associated with cream dilution on horse (Mariat et al. 2003) was only found on an horse (E.caballus(5)).

Table 7 – Characterization of the mutations observed by gene. N – Number of mutations, SY – Synonymous, NSY – Non-synonymous, NC – located in non-coding regions.

GENE	N	SY	NSY	NC
ASIP	**6**	**2**	**1**	**3**
MC1R	**15**	**4**	**11**	**-**
SILV	**48**	**9**	**23**	**16**
SLC36A1	**40**	**13**	**9**	**18**
SLC45A2	**36**	**15**	**17**	**4**
Total	**145**	**43**	**61**	**41**

The Median-joining network (Fig. 9) depicts the genetic relationship between the haplotypes in each of the five genes. In some cases it was not possible to represent all the exons or all the samples due to amplification failure as well as those which only have shown two haplotypes. However, all information on these cases is given in Table 7.

In the *ASIP*, three haplotypes were reconstructed and H1 and H2 were only present in domestic horses, whereas the H3 is shared between the donkey and its wild ancestor, *E. africanus,* and the Asiatic wild ass *E. hemionus*.

In the *MC1R*, 17 haplotypes were obtained from all the four amplified regions. The *E. caballus* displayed considerably higher number of haplotypes than all the wild equid species. However this diversity can be biased by the fact that the number of available sequences for domestic horse is much higher than from wild equids or from donkey. From this representation is possible to observe the evolutionary proximity of *E. prezewalskii* and horse by looking to since this species share always the same haplotypes by observing H1 (*MC1R_e1a*), H1 and H5 (*MC1R_e1b*), H1 (*MC1R_e1c*). The closest relationship between the donkey and the African wild ass is also highlighted in this gene by the H2 (*MC1R_e1a* and, *MC1R_E1c*). All the extant equids tend to share the most part of haplotypes; however, it is possible to observe some haplotypes that are exclusive of some species (Fig. 9).

In the *SILV,* a total of 39 haplotypes were observed. Despite the horse being the most diverse, it is possible to observe several shared haplotypes with other wild species. On the other hand, the domestic donkey displayed four exclusive haplotypes in this gene (exon 2, 5, and 6A, 11), as well as *E. grevyi* (exon 2, 4, 11) and was also observed that the two species of Asiatic wild ass (*E. kiang and E. hemionus)* shared several haplotypes, on exon 2, 4, 9 and 10.

In the *SLC36A1*, we have observed a total of 29 haplotypes. Here, the horse rarely shares alleles with other species (this was just observed on exon 1 with kiang and zebra), showing high diversity within this gene with a high number of haplotypes distributed among horses. The extant equids, mostly tend to share the same haplotypes (Fig. 9). On this gene it observed some haplotypes exclusive of Kiang (on Exon 1 and 9) and Grevy zebra (exon 1, 2 and 6).

Finally, in the *SLC45A1 gene* a total of 32 haplotypes were observed. Once more the genetic proximity between domestic donkey and *E. africanus* is highlighted. Despite the haplotype diversity observed among horse, it was also observed some diversity on domestic donkeys with several being shared wild species (exon 2, 3b, and 5).

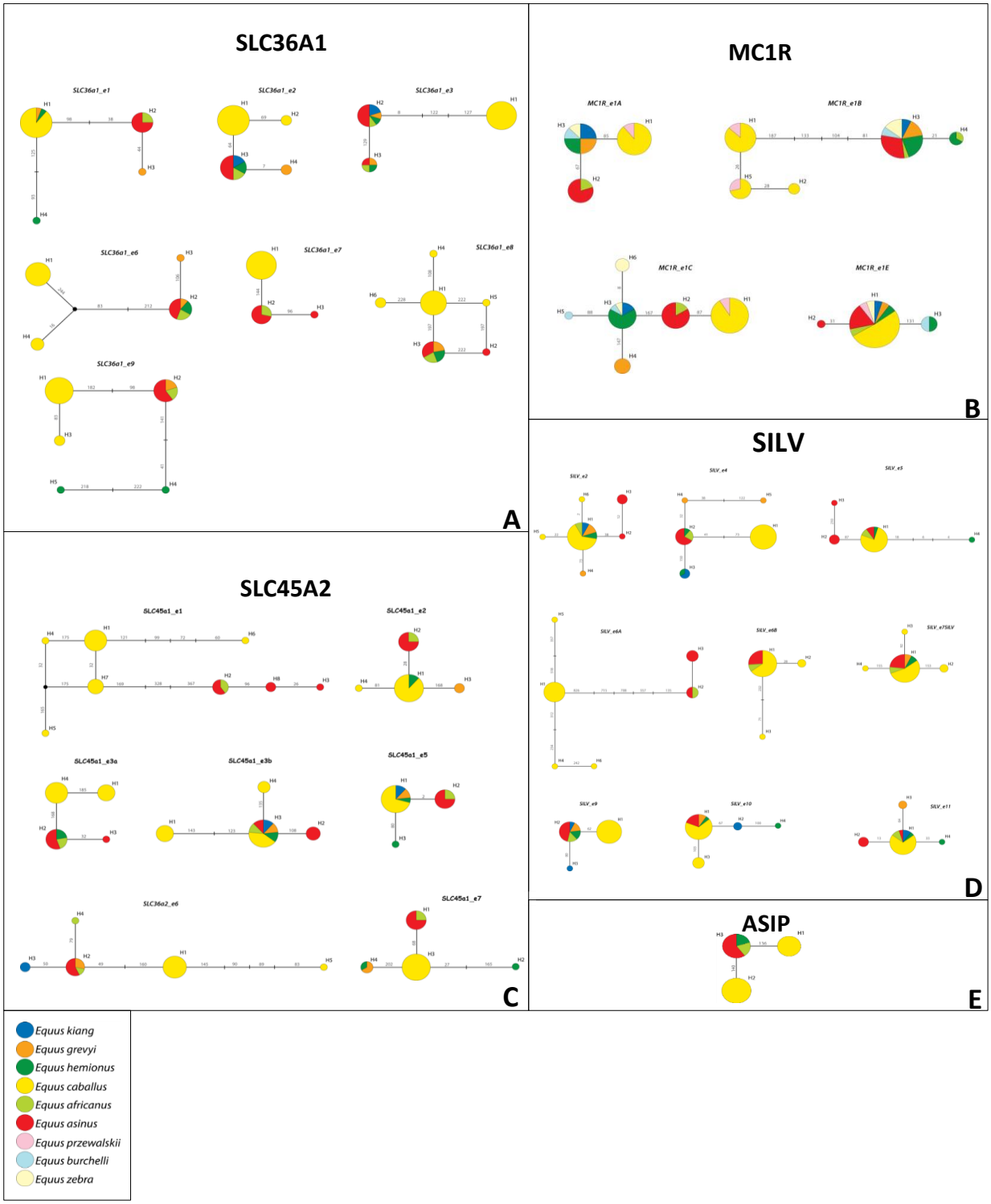

Fig.9 – Median-joining network for all the five assessed genes, *SLC36A1*A), *MC1R*(B), *ASIP*(E), *SLC45A2*(C), *SILV*(D). The size of each haplotype is proportional to the number of allocated individuals, and the branch size is proportional to the number of mutations. The legend on the right represents the correspondence between the colors of each represented species.

Table 8 – Characterization of all five coat color studied genes haplotypes found in the analyzed samples. Hap – haplotype number in each exon or fragment. N - number of sampled individual allocated in each haplotype. Species codes: EC - *Equus caballus*, **EA** - *Equus asinus*, **HOI** - *Equus africanus*, **GZ** - *Equus grevyi*, **EK** - *Equus kiang*, **EH** - *Equus hemionus* , **EP** - *Equus przewalskii*, **EB** - *Equus burchelli*, **EZ** - *Equus zebra*.

ASIP

Hap	N	Species allocated
1	7	7 EC
2	10	6 EA, 2 HOI, 2EH
3	11	11 EC

MC1R

E1_A

Hap	N	Species allocated
1	18	16 EC, 2 EP
2	10	8 EA, 2 HOI
3	16	4 EK, 4GZ, 2 EB, 4 EH, 2 EZ

E1_B

Hap	N	Species allocated
1	15	13 EC, 2 EP
2	2	2 EC
3	27	8 EA, 1 HOI, 4 GZ, 4 EZ, 2 EB, 6 EH, 2 EK
4	3	2 EA, 1 HOI
5	7	5 EC, 2 EP

E1_C

Hap	N	Species allocated
1	24	20 EC, 2 EP
2	12	10 EA, 2 HOI
3	12	8 EH, 1 EB, 2 EK, 1 EZ
4	4	4 GZ
5	1	1 EB
6	3	3 EZ

E1_D

Hap	N	Species allocated
1	54	22 EC, 8 EA, 2 HOI, 4 EP, 4 GZ, 4 EZ, 6EH, 2 EK, 2EB
2	2	2 EA

E1_E

Hap	N	Species allocated
1	39	20EC, 7EA, 2GZ, 2 EH, 2 EK, 2 EP, 2 EZ, 2 HOI
2	1	1 EA
3	4	2 EB, 2 EH

SILV

E1

Hap	N	Species allocated
1	22	16 EC, 2 EK, 2 GZ, 2 EH
2	8	6 EA, 2 HOI

E2

Hap	N	Species allocated
1	25	16 EC, 2 EK, 2 GZ, 2 EH, 2 HOI, 1 GZ
2	1	1 EA
3	3	3 EA
4	1	1 GZ
5	1	1 EC
6	1	1 EC

E4

Hap	N	Species allocated
1	20	20 EC
2	9	6 EA, 2 HOI, 1 EH
3	3	2 EK, 1 EH
4	1	1 GZ
5	1	1 GZ

E5

Hap	N	Species allocated
1	21	16 EC, 2 EA, 2 HOI, 1 EH
2	3	3 EA
3	1	1 EA
4	1	1 EH

E6_A

Hap	N	Species allocated
1	13	13 EC
2	4	2 EA, 2 HOI
3	4	4 EA
4	1	1EC
5	1	1EC
6	1	1EC

E6_B

Hap	N	Species allocated
1	23	15 EC, 6 EA, 2 HOI
2	2	2EC
3	1	1EC

E6_C

Hap	N	Species allocated
1	18	18 EC
2	14	6 EA, 2 EK, 2 GZ, 2 HOI, 2 EH

E7

Hap	N	Species allocated
1	26	14 EC, 6 EA, 2 HOI, 2 GZ, 2 EH
2	2	2 EC
3	1	1 EC
4	1	1 EC

E8

Hap	N	Species allocated
1	24	18 EC, 4 EA, 2 HOI
2	2	EA

E9

Hap	N	Species allocated
1	18	18 EC
2	13	6 EA, 1 EK, 2 GZ, 2 HOI, 2 WAI
3	1	1 EK

E10

Hap	N	Species allocated
1	21	14 EC, 4 EA, 2 GZ, 1 EH
2	2	2 EK,
3	4	4 EC
4	1	1 EH

E11

Hap	N	Species allocated
1	20	14 EC, 2 EK, 2 HOI, 1 EA, 1 EH
2	3	3 EA
3	2	2 GZ
4	1	1 EH

SLC36A1

E1

Hap	N	Species allocated
1	18	16 EC, 1 EH, 1 GZ
2	8	6 EA, 2 HOI
3	1	1 GZ
4	1	1 EH

E2

Hap	N	Species allocated
1	18	18 EC
2	2	2 EC
3	12	6 EA, 2 EK, 2 HOI, 2 EH
4	2	2 GZ

E3

Hap	N	Species allocated
1	16	16 EC
2	10	5 EA, 2 EK, 1 GZ, 1 EH, 1 HOI
3	4	1 GZ, 1 HOI, 1 EH, 1 EA

E4

Hap	N	Species allocated
1	16	16 EC
2	10	6 EA, 2 HOI, 2 EH

E6

Hap	N	Species allocated
1	11	11 EC
2	9	4 EA, 1 GZ, 2 HOI, 2 EH
3	1	1 GZ
4	3	3 EC

E7

Hap	N	Species allocated
1	16	16 EC
2	7	5 EA, 2 HOI
3	1	1 EA

E8

Hap	N	Species allocated
1	12	12 EC
2	1	1 EA
3	9	3 EA, 2 GZ, 2 HOI, 2 EH
4	1	1 EC
5	1	1 EC
6	2	2 EC

E9

Hap	N	Species allocated
1	14	14 EC
2	10	6 EA, 2 GZ, 2 HOI
3	2	2 EC
4	1	1 EH
5	1	1 EH

E10

Hap	N	Species allocated
1	14	14 EC
2	8	6 EA, 2 HOI

SLC45A2

E1

Hap	N	Species allocated
1	10	10 EC
2	5	3 EA, 2 HOI
3	1	1EA
4	1	1 EC
5	1	1 EC
6	1	1 EC
7	5	5 EC
8	2	2 EA

E2

Hap	N	Species allocated
1	17	15 EC, 2 EH
2	8	6 EA, 2 HOI
3	2	2 GZ
4	1	1 EC

E3_A

Hap	N	Species allocated
1	6	6 EC
2	9	5 EA, 2 HOI, 2 EH
3	1	1 EA
4	10	10 EC

E3_B

Hap	N	Species allocated
1	6	6 EC
2	4	4 EA
3	17	2 EK, 2 GZ, 2 HOI, 7 EC, 2 EH, 2 EA
4	3	3 EC

E4

Hap	N	Species allocated
1	28	16 EC, 6 EA, 2 GZ, 2 HOI, 2 EH
2	2	2 EC

E5

Hap	N	Species allocated
1	17	12 EC, 2 EK, 2 GZ, 1 EH
2	8	6 EA, 2 HOI
3	1	1 EH

E6

Hap	N	Species allocated
1	11	11 EC
2	7	4 EA, 1 HOI, 2 GZ
3	2	2 EK
4	1	1 HOI
5	1	1 EC

E7

Hap	N	Species allocated
1	16	16 EC
2	8	6 EA, 2 HOI
3	1	1 EH
4	1	1 EH

5. Discussion

Despite the high phenotypically differences between the equid species included on this study, it is possible to observe that in most cases, several alleles are shared by all wild equid species. This fact evidences that coat color is a very complex trait, controlled by the interaction of several genes simultaneously, which led to so divergent phenotypes (e.g. a zebra and a Kiang are very divergent at phonotypic level, but share several mutations) (Cieslak et al., 2011).

Among all the genes, domestic horse always shows a higher level of diversity, when compared with the remaining species, even with domestic donkey. This fact highlights the differential directional selection for this trait on horse and donkey. Despite both species have been mainly selected for transportation capacity during domestication process, donkey is considered one of the less managed domestic animals, and consequently less subjected to directional selection (Marshall & Weissbrod 2011a), namely at coat color level. On the other hand, during the horse domestication process, coat color is supposed to have been one of the first phenotypic traits subjected to artificial selection, which led in to the huge phenotypic and genotypic diversity observed among the modern domestic horses (Ludwig *et al.* 2009).

As referred previously, it was possible to observe an evident relation between domestic horse and *E.* przewalskii, highlighting the evolutionary proximity between both species, despite the origin of domestic horse still not fully clarified (Jansen *et al.* 2002; Goto *et al.* 2011). This study also demonstrated a great closeness between the domestic donkey (*E. asinus*) and the African wild Ass *E. africanus*, the observation of several alleles, among the five genes that are merely shared by these two species, supports the hypothesis that the African wild ass was the ancestral of the domestic donkey (Beja-pereira *et al.* 2004a; Kimura *et al.* 2010).

In SILV and SLC36A1 gene is also possible to observe a several differentiation of *E. grevyi* from the remaining species, but in general this species shares a great part of alleles with the other equids, despites the high phenotypic divergence.

From these results we can conclude that the highest level of genotypic variation was observed in the domestic species, coinciding with the highest level of phenotypic differences observed among them. This been in agreement with the idea that wild species tend to present less variation than domestic relatives (Cieslak et al.,2011).

Despite the domestication processes of horse and donkey started only at approximately between 7000-6000 yBP (Beja-pereira *et al.* 2004b; Marshall & Weissbrod 2011a; Marshall & Weissbrod 2011b), which is a short time period when compared with evolution time of wild

equids, it was enough to accumulate a high number of mutations underlining a huge phenotypic diversity in the domestic species (mainly on horses).

The shared alleles between domestic and wild equids, shows that its origin predates the domestication process. However, due to natural selection forces acting on wild equid populations, those mutations leading to more exuberant phenotypic traits, tend to be eliminated, since changes in the camouflage patterns could compromise individuals survival and fitness in their natural environment (Linderholm & Larson 2013). On the other hand, since domestication implies the separation of animals from their natural environment, some of the observed phenotypic variation may have resulted from the relaxation of natural selection against those more exuberant phenotypes. This and the human-mediated selection, which prized and positively selected the some coat color phenotypes have played a major role in the fixation of those new mutations in the domestic species (Andersson 2013). Finally, the accumulation of those fixed mutations will increase the genetic differentiation between wild and domestic relatives (Linderholm & Larson 2013)

The idea of a large pre-domestication variation in equids given by the cave paintings existing in Pech-Merle, France (24.000 yBP), conjugated with ancient DNA analysis, revealed the existence of at least three coat color variations among pre-domestic horse populations: black, bay and leopard spotting (Pruvost et al., 2011). Ludwig et al.(2009) studied the variation observed in five genes, on which four are common to those of this work (MC1R, SLC45A2, ASIP) and concluded that on pre-domestic horses there were only black and bay phenotypes, and that mutations related with diluted phenotypes, as cream (SCL45A2) and silver (SILV), appeared later on after domestication process. The only gene that apparently encountered this previous results was SLC36A2 (champagne dilution), in which almost neither species shared the same alleles as domestic horse.

An interesting fact that can also validate the early emergence of coat color variation is the fact that most part of domesticated species, as donkey (Raudseppet al., 1999) dog (Schmutz & Berryere 2007), sheep (Fontanesi et al., 2010; Gratten et al., 2012), or even hamster (Alizadeh *et al.* 2009), the genes responsible for the color phenotypes are almost the same, independently of the pathway that the species follows after domestication process (Linderholm & Larson 2013).

However, despite the high number of shared alleles observed between domestic and wild species, where also evidenced some exclusive mutations among domestic equids. These mutations probably appeared after the domestication process mainly driven by genetic drift, inbreeding, cross between similar phenotypes or even by the posterior hybridization between

domestic females and wild males. The correct understanding of the evolutionary processes inherent to domestication and those effects over the species are not only important to increase the knowledge surrounding it but also to more deeply assess the origins of modern human society (Diamond 2002).

6. References

Alizadeh A, Hong LZ, Kaelin CB *et al.* (2009) Genetics of Sex-linked yellow in the Syrian hamster. *Genetics*, **181**, 1427–36.

Andersson L (2013) Molecular consequences of animal breeding. *Current opinion in genetics & development*, **23**, 295–301.

Bandelt H-J, Forster P, Rohl A (1999) Median-Joining Networks for Inferring Intraspecific Phylogenies. *Molecular biology and evolution*, **16**, 37–48.

Beja-pereira A, England PR, Ferrand N *et al.* (2004) African Origins of the Domestic Donkey. *Science*, **304**, 1781.

Bruford MW, Bradley DG, Luikart G (2003) DNA markers reveal the complexity of livestock domestication. *Nature reviews. Genetics*, **4**, 900–10.

Brunberg E, Andersson L, Cothran G *et al.* (2006) A missense mutation in PMEL17 is associated with the Silver coat color in the horse. *BMC genetics*, **7**, 46.

Cieslak M, Reissmann M, Hofreiter M, Ludwig A (2011) Colours of domestication. *Biological reviews of the Cambridge Philosophical Society*, **86**, 885–99.

Cook D, Brooks S, Bellone R, Bailey E (2008) Missense mutation in exon 2 of SLC36A1 responsible for champagne dilution in horses. *PLoS genetics*, **4**, e1000195.

Costa et al, A new approach to improve DNA extraction from noninvasive sources, in preparation

Diamond J (2002) Evolution, consequences and future of plant and animal domestication. *Nature*, **418**, 700–7.

Fang M, Larson G, Ribeiro HS, Li N, Andersson L (2009) Contrasting mode of evolution at a coat color locus in wild and domestic pigs. *PLoS genetics*, **5**, e1000341.

Fontanesi L, Beretti F, Riggio V *et al.* (2010) Sequence characterization of the melanocortin 1 receptor (MC1R) gene in sheep with different coat colours and identification of the putative allele at the ovine Extension locus. *Small Ruminant Research*, **91**, 200–207.

Goto H, Ryder O a, Fisher AR *et al.* (2011) A massively parallel sequencing approach uncovers ancient origins and high genetic variability of endangered Przewalski's horses. *Genome biology and evolution*, **3**, 1096–106.

Gratten J, Pilkington JG, Brown EA *et al.* (2012) Selection and microevolution of coat pattern are cryptic in a wild population of sheep. *Molecular Ecology*, **21**, 2977–2990.

Hauswirth R, Haase B, Blatter M *et al.* (2012) Mutations in MITF and PAX3 cause “splashed white” and other white spotting phenotypes in horses. *PLoS genetics*, **8**, e1002653.

Jansen T, Forster P, Levine M a *et al.* (2002) Mitochondrial DNA and the origins of the domestic horse. *Proceedings of the National Academy of Sciences of the United States of America*, **99**, 10905–10.

Jeon J, Andersson L, Giuffra E, Kijas JMH, Amarger V (2000) The Origin of the Domestic Pig : Independent Domestication and Subsequent Introgression. *Time*, **1791**, 1785–1791.

Kimura B, Marshall FB, Chen S *et al.* (2010) Ancient DNA from Nubian and Somali wild ass provides insights into donkey ancestry and domestication. *Proceedings. Biological sciences / The Royal Society*, **278**, 50–7.

Librado P, Rozas J (2009) DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics*, **25**, 1451–1452.

Linderholm A, Larson G (2013) The role of humans in facilitating and sustaining coat colour variation in domestic animals. *Seminars in cell & developmental biology*, **24**, 587–93.

Ludwig A, Pruvost M, Reissmann M *et al.* (2009) Coat color variation at the beginning of horse domestication. *Science*, **324**, 485.

Mariat DM, Aourit ST, Uérin GG (2003) A mutation in the MATP gene causes the cream coat colour in the horse. *Genetics* Selection *Evolution* , **35**, 119–133.

Marklund L, Moller MJ, Sandberg K, Andersson L (1996) A missense mutation in the gene for melanocyte-stimulating hormone receptor (MC1R) is associated with the chestnut coat color in horses. *Mammalian genome*, **7**, 895–9.

Marshall F, Weissbrod L (2011) Domestication Processes and Morphological Change. *Current Anthropology*, **52**, 397–413.

Pruvost M, Bellone R, Benecke N *et al.* (2011) Genotypes of predomestic horses match phenotypes painted in Paleolithic works of cave art. *Proceedings of the National Academy of Sciences of the United States of America*, **108**, 18626–18630.

Raudsepp T, Kijas J, Godard S (1999) Comparison of horse Chromosome 3 with donkey and human chromosomes by cross-species painting and heterologous FISH mapping. *Mammalian Genome*, **282**, 277–282.

Reissmann M, Bierwolf J, Brockmann G a (2007) Two SNPs in the SILV gene are associated with silver coat colour in ponies. *Animal genetics*, **38**, 1–6.

Rieder S (2009) Molecular tests for coat colours in horses. *Zeitschrift für Tierzüchtung und Züchtungsbiologie*, **126**, 415–24.

Rieder S, Taourit S, Mariat D, Langlois B, Guérin G (2001) Mutations in the agouti (ASIP), the extension (MC1R), and the brown (TYRP1) loci and their association to coat color phenotypes in horses (Equus caballus). *Mammalian genome*, **12**, 450–5.

Schmutz SM, Berryere TG (2007) Genes affecting coat colour and pattern in domestic dogs: a review. *Animal genetics*, **38**, 539–49.

Svensson EM, Telldahl Y, Sjöling E *et al.* (2012) Coat colour and sex identification in horses from Iron Age Sweden. *Anatomischer Anzeiger ,* **194**, 82–7.

Tamura K, Peterson D, Peterson N *et al.* (2011) MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Molecular biology and evolution*, **28**, 2731–9.

Thiruvenkadan AK, Kandasamy N, Panneerselvam S (2008) Coat colour inheritance in horses. *Livestock Science*, **117**, 109–129.

Vilà C, Jennifer L a., Beja-pereira A (2006) Genetic Documentation of Horse and Donkey Domestication. In: Documenting Domestication, University of California Press.

Wagner HJ, Reissmann M (2000) New polymorphism detected in the horse MC1R gene. *Animal genetics*, **31**, 289–90.

Zeder MA (1994) *Of kings and shepherds: specialized animal economy in Ur III Mesopotamia*; Stein G, Rothman M, eds. Madison, WI: Prehistory Press.

7. Supplementary material

Supplementary tables

Table S1 - List of sequences downloaded from NCBI. **A.N** - Accession Number

Gene	Species	A.N	Reference
ASIP	E.caballus	NC_009165.2	Wade et al.,2009
MC1R	E.caballus	NC_009146.2	Wade et al.,2009
MC1R	E.caballus	AF288357	Rieder et al.,2001
MC1R	E. asinus	JF718952_1	Steiner & Ryder, 2011
MC1R	E.grevi	JF718958.1	Steiner & Ryder, 2012
MC1R	E.kiang	JF718953.1	Steiner & Ryder, 2013
MC1R	E.burchelli	JF718956.1	Steiner & Ryder, 2014
MC1R	E.hemionus	JF718954.1	Steiner & Ryder, 2015
MC1R	E.zebra	JF718957.1	Steiner & Ryder, 2016
MC1R	E.przewalskii	JF718951.1	Steiner & Ryder, 2017
MC1R	E.h.onager	JF718955.1	Steiner & Ryder, 2018
MC1R	E.h.kulan	JF718954.1	Steiner & Ryder, 2019
MC1R	E.caballus	AF252541.1	Wagner & Reissmann, 2000
MC1R	E.hemionus	AF141365.1	Pistacchi & Ryder, 1999
MC1R	E.asinus	AF141364.1	Pistacchi & Ryder, 2000
MC1R	E.grevy	AF141363.1	Pistacchi & Ryder, 2001
MC1R	E.zebra	AF097749.1	Pistacchi, 1998
MC1R	E.przewalskii	AF039388.1	Wagner & Reissmann, 1998
SILV	E.caballus	NC_009149.2	Wade et al.,2009
SILV	E.caballus	DQ855465.1	Brunberg et al., 2006
SLC36A1	E.caballus	NC_009157.2	Wade et al.,2009
SLC36A2	E.caballus	EU432177.1	Cook et al., 2008
SLC36A3	E.caballus	EU432176.1	Cook et al., 2008
SLC45A2	E.caballus	NC_009164.2	Wade et al.,2009
SLC45A2	E.caballus	AH012460.2	Mariat et al., 2003
SLC45A2	E.caballus	EU272794.1	Del Valle & Penedo, 2007

Table S2 - Characteristics of Primers used on amplification of Coat color genes

Gene	Chr.	Gene(bp)	Exon	Exon(bp)	Primer sequence	Fragment (bp)	TA (º C)	Taq	MgCl2 (mM	Gene AN
ASIP	**22**	3995	Exon1	160	F - CCTGACTGCCTTCCCTGC R - GGCTGGAGGCCAGAATTGAG	238	52ºC	TaqPlatinum	1.5	NC_009165.2
			Exon2	65	F - ACCAGAAACATCTGGCTTTGC R - TTAAGTCTGGTCCCACCCCT	150	59ºC	AmpliTaqGold	1.5	
			Exon3	177	F - AAAGAGCTCCCAGGAGGGG R - CTCTAGGGTCTCCAGGGTCC	305	57ºC	AmpliTaqGold	1.5	
MC1R	**3**	954	Exon1 - A	954	F - CTGGAGAGTGAGGACCCCTT R - GCCAGGCAGCAGATGAAGTA	281	56ºC	AmpliTaqGold	2	NC_009146.2
			Exon1 - B		F - GCTGGTGAGCCTAGTGGAAA R - TGTAGCGGTCTACGGCAATG	284	54ºC	AmpliTaqGold	2	
			Exon1 - C		F - TTGCAGCAGCTGGACAACAT R - ACGTACAGCACTGCCATGAG	287	54ºC	AmpliTaqGold	2	
			Exon1 - D		F - TCTAGCACCCTCTTCATCGC R - GGTGTTGAGGGCAGAGGATA	289	57ºC	AmpliTaqGold	2	
			Exon1 - E		F - CACCCTCACCATCCTGCTG R - TGATGTCACCACCTCCCTCT	290	52ºC	AmpliTaqGold	2	
SILV	**6**	7822	Exon1	76	F - CAGTCTCCCAGCGCCTT R - TGCAACCCCAAATTCACACT	160	52ºC	TaqPlatinum	1.5	NC_009149.2
			Exon2	111	F - CTGAGTGAGGATCCTAGGGC R - CCCATTCCCACCATACCCTC	194	57ºC	TaqPlatinum	1.5	
			Exon3	147	F - GGAAGCCAAGGAGTCCCAC R - TTGGAGCCCTTAGAAGAGGG	228	52ºC	TaqPlatinum	1.5	
			Exon4	135	F -GGAATACCAGGGTTTGAGGGT R - AGGAAGAGGAGAATCCAACTGT	230	57ºC	TaqPlatinum	1.5	
			Exon5	162	F - TTGAACAGTAACCCCCACCC R - CCTGCCCTGCAACTGATCT	250	52ºC	TaqPlatinum	1.5	
			Exon6 - A	720	F - CCTGGCCTGGTTCTCACCT R - CTGTGGTGCCTGGAACTGG	338	58ºC	AmpliTaqGold	1.5	
			Exon6 - B		F - AGGGCACTTGTGGTCACTC R - GGAACTCTGAGGTTGACACC	340	61ºC	TaqPlatinum	1.5	
			Exon6 - B		F - GTGCCAACCACAGAGGTCAT R - AATGCCCCACCTCAGTGAAC	340	59ºC	AmpliTaqGold	1.5	
			Exon7	117	F - GCCTCTGTTTCCAGGTGAGT R - AAACACCTCCCTCCACCAGC	220	52ºC	TaqPlatinum	1.5	
			Exon8	85	F - GGAGGCTGGGGAAAGGAATC R - GGAATACTAGACTCCCAAGAGCA	178	57ºC	AmpliTaqGold	1.5	
			Exon9	206	F -TCCTTCCTAGACCCCCTGAC R - CAGATAGCCTCTGCCCACTC	299	52ºC	TaqPlatinum	1.5	
			Exon10	88	F- GCAGACACTGACTGAAGCTG R - GGTGGTAATAAGGGATAAGGGGG	178	52ºC	TaqPlatinum	2	
			Exon11	142	F - CCCTTGTCCATTGCTTACCA R - GCTGTCCACACCTGGGTAAA	230	52ºC	TaqPlatinum	1.5	
SLC36A1	**14**	28177	Exon1	143	F - CCCCTGGAGGTCTTAGTGCT R - ACTTAGGGACACGAATCCCGC	234	52ºC	TaqPlatinum	1.5	NC_009157.2
			Exon2	91	F - TCATTCTTGTTAACACGCTC R - TTGTCACAATACTTCCCTCC	173	52ºC	TaqPlatinum	1.5	
			Exon3	89	F - CTGCCTGTCTCTCCCTCTCT R - GGGCTGCTAAAGGAACCCAA	191	54ºC	TaqPlatinum	1.5	
			Exon4	96	F - TCACTGTCTAACCCCCACCT R - CGTCTCACCCCACTACCTGA	220	52ºC	TaqPlatinum	1.5	
			Exon5	85	F - CTTCCTTCCCCACGTTCAGG R - AAGTCCCCCATCTCTTACCA	162	54ºC	TaqPlatinum	1.5	
			Exon6	219	F - GTCTGAGAGGCTTTCTCTTTCT R - CGCTCAGTGTTTTGTTGATGA	315	52ºC	TaqPlatinum	1.5	
			Exon7	99	F - CTGTCTGCTGTTGGGGTTCA R - CAGACCTCGAGGACACCGA	219	52ºC	TaqPlatinum	1.5	
			Exon8	167	F - GGGAAATTCTGGGGATAATGAAGT R - AAAATATCCAGCACTCCCAGG	320	52ºC	TaqPlatinum	1.5	
			Exon9	170	F - TGCAGAGCAGGCTTTTGAGT R - ACACGCTTGCTCTCAGGAAA	293	52ºC	TaqPlatinum	1.5	
			Exon10	272	F - CAGCTCTGGGCAGTCTTGAA R - CAGGTGCTGGGGAACATACA	374	52ºC	TaqPlatinum	1.5	
SLC45A2	**21**	28849	Exon1-A	385	F - TGAGGCACAAACCCGAGG R - GGACTGAGCAGCCACACC	300	54ºC	TaqPlatinum	1.5	NC_009164.2
			Exon1-B		F - GGAATTTTGCTACGCGGTGG R - ATGATGGGACCCTTTTCCCC	299	54ºC	TaqPlatinum	2	
			Exon2	176	F - ACGTGGGTCATTCTAAACCAGG R - TCCCTGTCTGGGGAGCTAAA	279	56ºC	AmpliTaqGold	1.5	
			Exon3-A	322	F - GGAGAGAGTCCTTCTAAATATGCTA R - GGGTCTTGGGAGTCTTGC	260	59ºC	TaqPlatinum	1.5	
			Exon3-B		F - TGCTTACTTTGTGTGTTGTT R - CAAGAATATTTTTCCCAAGTAAAG	258	52ºC	TaqPlatinum	1.5	
			Exon4	143	F - TGCATGTCTGCTCTGAAGGT R - GACAGGTGCTGAATGAGGGA	258	52ºC	TaqPlatinum	1.5	
			Exon5	123	F - GGATCACACTGCATTTGTAGC R - GACCAGAGCTCTTTAGAAGACA	215	52ºC	TaqPlatinum	2	
			Exon6	214	F - TGTGGTGAGACACAGTCAGC R - GGAGTCCAGATGTTACCCAG	305	53ºC	TaqPlatinum	1.5	
			Exon7	224	F - GGATAGTTCGTCTAGATCAAATGC R - TTCTTCCATATGCTCCACTGAG	319	52ºC	TaqPlatinum	1.5	

Table S3 - Polymorphisms observed among the five study coat color genes. The sequences collected from NCBI presents the Accession numbers close to sequence name. The codes are according to the haplotypes information on Table 8 and colors are according to the Medium-joining Network legend for each species on Figure 9 .

GENE	Exon	REF.Sequence	Sequence name	Position	Region	Aminoacid change	Mutation	Code
ASIP	1	NC_009165.2	E. africanus (HOI)	163G>A	Non-coding	-	-	HOI
ASIP	1	NC_009165.2	E. asinos (Miranda)	163G>A	Non-coding	-	-	EA
ASIP	1	NC_009165.2	E. asinus (Somalia)	163G>A	Non-coding	-	-	EA
ASIP	1	NC_009165.2	E. caballus (2)	163G>A	Non-coding	-	-	EC
ASIP	1	NC_009165.2	E. caballus (5)	163G>A	Non-coding	-	-	EC
ASIP	1	NC_009165.2	E. caballus (Garrano)	163G>A	Non-coding	-	-	EC
ASIP	1	NC_009165.2	E. caballus(Sorraia)	163G>A	Non-coding	-	-	EC
ASIP	1	NC_009165.2	E. hemionus(WAI)	163G>A	Non-coding	-	-	EH
ASIP	1	NC_009165.2	E. asinus (Oman)	163G>A	Non-coding	-	-	EA
ASIP	1	NC_009165.2	E. caballus (1)	163G>R	Non-coding	-	-	EC
ASIP	1	NC_009165.2	E. caballus (3)	163G>R	Non-coding	-	-	EC
ASIP	1	NC_009165.2	E. caballus (4)	163G>R	Non-coding	-	-	EC
ASIP	1	NC_009165.2	E. africanus (HOI)	168G>A	Non-coding	-	-	HOI
ASIP	1	NC_009165.2	E. asinos (Miranda)	168G>A	Non-coding	-	-	EA
ASIP	1	NC_009165.2	E. asinus (Oman)	168G>A	Non-coding	-	-	EA
ASIP	1	NC_009165.2	E. asinus (Somalia)	168G>A	Non-coding	-	-	EA
ASIP	1	NC_009165.2	E. hemionus(WAI)	168G>A	Non-coding	-	-	EH
ASIP	2	NC_009165.2	E. caballus (Garrano)	1500 (11 bp deletion)>1511	Coding	?	?	EC
ASIP	2	NC_009165.2	E. caballus(Sorraia)	1500 (11 bp deletion)>1511	Coding	?	?	EC
ASIP	2	NC_009165.2	E. caballus (Garrano)	1451C>A	Non-coding	-	-	EC
ASIP	2	NC_009165.2	E. asinus (Somalia)	1451C>G	Non-coding	-	-	EA
ASIP	2	NC_009165.2	E. caballus (Garrano)	1488G>T	Coding	SER14ILE (14 AGC > ATG)	Functional	EC
MC1R	1	NC_009146.2	E. africanus (HOI)	111A>G	Coding	GLU37GLU(37GAA>GAG)	Silent	HOI
MC1R	1	NC_009146.2	E. asinos (Miranda)	111A>G	Coding	GLU37GLU(37GAA>GAG)	Silent	EA
MC1R	1	NC_009146.2	E. asinos (Miranda)	111A>G	Coding	GLU37GLU(37GAA>GAG)	Silent	EA
MC1R	1	NC_009146.2	E. asinus (Somalia)	111A>G	Coding	GLU37GLU(37GAA>GAG)	Silent	EA
MC1R	1	NC_009146.2	E. grevyi(GZ)	111A>G	Coding	GLU37GLU(37GAA>GAG)	Silent	GZ
MC1R	1	NC_009146.2	E.asinos (JF718952.1)	111A>G	Coding	GLU37GLU(37GAA>GAG)	Silent	EA
MC1R	1	NC_009146.2	E.burchelli(JF718956.1)	111A>G	Coding	GLU37GLU(37GAA>GAG)	Silent	EB
MC1R	1	NC_009146.2	E.grevyi(JF718958.1)	111A>G	Coding	GLU37GLU(37GAA>GAG)	Silent	GZ
MC1R	1	NC_009146.2	E.h.kulan(JF718954.1)	111A>G	Coding	GLU37GLU(37GAA>GAG)	Silent	EH
MC1R	1	NC_009146.2	E.h.onager(JF718955.1)	111A>G	Coding	GLU37GLU(37GAA>GAG)	Silent	EH
MC1R	1	NC_009146.2	E.kiang(EK)	111A>G	Coding	GLU37GLU(37GAA>GAG)	Silent	EK
MC1R	1	NC_009146.2	E.kiang(JF718953.1)	111A>G	Coding	GLU37GLU(37GAA>GAG)	Silent	EK
MC1R	1	NC_009146.2	E.zebra(JF718957.1)	111A>G	Coding	GLU37GLU(37GAA>GAG)	Silent	EZ
MC1R	1	NC_009146.2	E. africanus (HOI)	200G>C	Coding	ARG67ARG(67GGC>GCC)	Silent	HOI
MC1R	1	NC_009146.2	E. africanus (HOI)	243C>T	Coding	ALA81ALA(GCC<GCT)	Silent	HOI
MC1R	1	NC_009146.2	E. asinos (Miranda)	243C>T	Coding	ALA81ALA(GCC<GCT)	Silent	EA
MC1R	1	NC_009146.2	E. asinus (Somalia)	243C>T	Coding	ALA81ALA(GCC<GCT)	Silent	EA
MC1R	1	NC_009146.2	E. caballus (4)	248C>T	Coding	SER83PHE(83TCC>TTC)	Functional	EC
MC1R	1	NC_009146.2	E. caballus (5)	248C>T	Coding	SER83PHE(83TCC>TTC)	Functional	EC
MC1R	1	NC_009146.2	E.przewalskii(JF718951.1)	248C>T	Coding	SER83PHE(83TCC>TTC)	Functional	EP
MC1R	1	NC_009146.2	E.caballus(AF252541.1)	250G>A	Coding	ASP84ASN(84GAC>AAC)	Functional	EC
MC1R	1	NC_009146.2	E. africanus (HOI)	303>G>A	Coding	LEU101LEU(101CTG>CTA)	Functional	HOI
MC1R	1	NC_009146.2	E. asinos (Miranda)	303>G>A	Coding	LEU101LEU(101CTG>CTA)	Functional	EA
MC1R	1	NC_009146.2	E. asinus (Oman)	303>G>A	Coding	LEU101LEU(101CTG>CTA)	Functional	EA
MC1R	1	NC_009146.2	E. asinus (Somalia)	303>G>A	Coding	LEU101LEU(101CTG>CTA)	Functional	EA
MC1R	1	NC_009146.2	E.asinos (JF718952.1)	303>G>A	Coding	LEU101LEU(101CTG>CTA)	Functional	EA
MC1R	1	NC_009146.2	E.asinus(AF141364.1)	303>G>A	Coding	LEU101LEU(101CTG>CTA)	Functional	EA
MC1R	1	NC_009146.2	E.burchelli(JF718956.1)	303>G>A	Coding	LEU101LEU(101CTG>CTA)	Functional	EB
MC1R	1	NC_009146.2	E.grevyi(AF141363.1)	303>G>A	Coding	LEU101LEU(101CTG>CTA)	Functional	GZ
MC1R	1	NC_009146.2	E.grevyi(JF718958.1)	303>G>A	Coding	LEU101LEU(101CTG>CTA)	Functional	GZ
MC1R	1	NC_009146.2	E.h.kulan(JF718954.1)	303>G>A	Coding	LEU101LEU(101CTG>CTA)	Functional	EH
MC1R	1	NC_009146.2	E.h.onager(JF718955.1)	303>G>A	Coding	LEU101LEU(101CTG>CTA)	Functional	EH
MC1R	1	NC_009146.2	E.hemionus(AF141365.1)	303>G>A	Coding	LEU101LEU(101CTG>CTA)	Functional	EH
MC1R	1	NC_009146.2	E.kiang(JF718953.1)	303>G>A	Coding	LEU101LEU(101CTG>CTA)	Functional	EK

GENE	Exon	REF.Sequence	Sequence name	Position	Region	Aminoacid change	Mutation	Code
MC1R	1	NC_009146.2	E.zebra(AF097749.1)	303>G>A	Coding	LEU101LEU(101CTG>CTA)	Functional	EZ
MC1R	1	NC_009146.2	E.zebra(JF718957.1)	303>G>A	Coding	LEU101LEU(101CTG>CTA)	Functional	EZ
MC1R	1	NC_009146.2	E. africanus (HOI)	326A>G	Coding	GLN109ARG(109CAG>CGG)	Functional	HOI
MC1R	1	NC_009146.2	E. asinos (Miranda)	326A>G	Coding	GLN109ARG(109CAG>CGG)	Functional	EA
MC1R	1	NC_009146.2	E. asinus (Oman)	326A>G	Coding	GLN109ARG(109CAG>CGG)	Functional	EA
MC1R	1	NC_009146.2	E. asinus (Somalia)	326A>G	Coding	GLN109ARG(109CAG>CGG)	Functional	EA
MC1R	1	NC_009146.2	E.asinos (JF718952.1)	326A>G	Coding	GLN109ARG(109CAG>CGG)	Functional	EA
MC1R	1	NC_009146.2	E.asinus(AF141364.1)	326A>G	Coding	GLN109ARG(109CAG>CGG)	Functional	EA
MC1R	1	NC_009146.2	E.burchelli(JF718956.1)	326A>G	Coding	GLN109ARG(109CAG>CGG)	Functional	EB
MC1R	1	NC_009146.2	E.grevyi(AF141363.1)	326A>G	Coding	GLN109ARG(109CAG>CGG)	Functional	GZ
MC1R	1	NC_009146.2	E.grevyi(JF718958.1)	326A>G	Coding	GLN109ARG(109CAG>CGG)	Functional	GZ
MC1R	1	NC_009146.2	E.h.kulan(JF718954.1)	326A>G	Coding	GLN109ARG(109CAG>CGG)	Functional	EH
MC1R	1	NC_009146.2	E.hemionus(AF141365.1)	326A>G	Coding	GLN109ARG(109CAG>CGG)	Functional	EH
MC1R	1	NC_009146.2	E.kiang(JF718953.1)	326A>G	Coding	GLN109ARG(109CAG>CGG)	Functional	EK
MC1R	1	NC_009146.2	E.zebra(AF097749.1)	326A>G	Coding	GLN109ARG(109CAG>CGG)	Functional	EZ
MC1R	1	NC_009146.2	E.zebra(JF718957.1)	326A>G	Coding	GLN109ARG(109CAG>CGG)	Functional	EZ
MC1R	1	NC_009146.2	E. asinos (Miranda)	355A>G	Coding	ILE119VAL(119ATC>GTG)	Functional	EA
MC1R	1	NC_009146.2	E. asinus (Oman)	355A>G	Coding	ILE119VAL(119ATC>GTG)	Functional	EA
MC1R	1	NC_009146.2	E.asinos (JF718952.1)	355A>G	Coding	ILE119VAL(119ATC>GTG)	Functional	EA
MC1R	1	NC_009146.2	E.asinus(AF141364.1)	355A>G	Coding	ILE119VAL(119ATC>GTG)	Functional	EA
MC1R	1	NC_009146.2	E.burchelli(JF718956.1)	355A>G	Coding	ILE119VAL(119ATC>GTG)	Functional	EB
MC1R	1	NC_009146.2	E.grevyi(AF141363.1)	355A>G	Coding	ILE119VAL(119ATC>GTG)	Functional	GZ
MC1R	1	NC_009146.2	E.grevyi(JF718958.1)	355A>G	Coding	ILE119VAL(119ATC>GTG)	Functional	GZ
MC1R	1	NC_009146.2	E.h.kulan(JF718954.1)	355A>G	Coding	ILE119VAL(119ATC>GTG)	Functional	EH
MC1R	1	NC_009146.2	E.h.onager(JF718955.1)	355A>G	Coding	ILE119VAL(119ATC>GTG)	Functional	EH
MC1R	1	NC_009146.2	E.hemionus(AF141365.1)	355A>G	Coding	ILE119VAL(119ATC>GTG)	Functional	EH
MC1R	1	NC_009146.2	E.kiang(JF718953.1)	355A>G	Coding	ILE119VAL(119ATC>GTG)	Functional	EK
MC1R	1	NC_009146.2	E.zebra(AF097749.1)	355A>G	Coding	ILE119VAL(119ATC>GTG)	Functional	EZ
MC1R	1	NC_009146.2	E.zebra(JF718957.1)	355A>G	Coding	ILE119VAL(119ATC>GTG)	Functional	EZ
MC1R	1	NC_009146.2	E.zebra(AF097749.1)	418G>A	Coding	VAL140ILE(140GTA>ATA)	Functional	EZ
MC1R	1	NC_009146.2	E.zebra(JF718957.1)	418G>A	Coding	VAL140ILE(140GTA>ATA)	Functional	EZ
MC1R	1	NC_009146.2	E. africanus (HOI)	497T>C	Coding	VAL166>ALA(166GTG>GCG)	Functional	HOI
MC1R	1	NC_009146.2	E. asinos (Miranda)	497T>C	Coding	VAL166>ALA(166GTG>GCG)	Functional	EA
MC1R	1	NC_009146.2	E. asinus (Oman)	497T>C	Coding	VAL166>ALA(166GTG>GCG)	Functional	EA
MC1R	1	NC_009146.2	E. asinus (Somalia)	497T>C	Coding	VAL166>ALA(166GTG>GCG)	Functional	EA
MC1R	1	NC_009146.2	E. hemionus(WAI)	497T>C	Coding	VAL166>ALA(166GTG>GCG)	Functional	EH
MC1R	1	NC_009146.2	E.asinos (JF718952.1)	497T>C	Coding	VAL166>ALA(166GTG>GCG)	Functional	EA
MC1R	1	NC_009146.2	E.asinus(AF141364.1)	497T>C	Coding	VAL166>ALA(166GTG>GCG)	Functional	EA
MC1R	1	NC_009146.2	E.burchelli(JF718956.1)	497T>C	Coding	VAL166>ALA(166GTG>GCG)	Functional	EB
MC1R	1	NC_009146.2	E.grevyi(AF141363.1)	497T>C	Coding	VAL166>ALA(166GTG>GCG)	Functional	GZ
MC1R	1	NC_009146.2	E.grevyi(JF718958.1)	497T>C	Coding	VAL166>ALA(166GTG>GCG)	Functional	GZ
MC1R	1	NC_009146.2	E.h.kulan(JF718954.1)	497T>C	Coding	VAL166>ALA(166GTG>GCG)	Functional	EH
MC1R	1	NC_009146.2	E.h.onager(JF718955.1)	497T>C	Coding	VAL166>ALA(166GTG>GCG)	Functional	EH
MC1R	1	NC_009146.2	E.hemionus(AF141365.1)	497T>C	Coding	VAL166>ALA(166GTG>GCG)	Functional	EH
MC1R	1	NC_009146.2	E.kiang(JF718953.1)	497T>C	Coding	VAL166>ALA(166GTG>GCG)	Functional	EK
MC1R	1	NC_009146.2	E.zebra(JF718957.1)	497T>C	Coding	VAL166>ALA(166GTG>GCG)	Functional	EZ
MC1R	1	NC_009146.2	E.grevyi(AF141363.1)	557G>A	Coding	THR186>MET(186ACG>ATG)	Functional	GZ
MC1R	1	NC_009146.2	E.grevyi(JF718958.1)	557G>A	Coding	THR186>MET(186ACG>ATG)	Functional	GZ
MC1R	1	NC_009146.2	E. hemionus(WAI)	577G>A	Coding	VAL193>ILE(193GTC>ATC)	Functional	EH
MC1R	1	NC_009146.2	E.burchelli(JF718956.1)	577G>A	Coding	VAL193>ILE(193GTC>ATC)	Functional	EB
MC1R	1	NC_009146.2	E.grevyi(AF141363.1)	577G>A	Coding	VAL193>ILE(193GTC>ATC)	Functional	GZ
MC1R	1	NC_009146.2	E.grevyi(JF718958.1)	577G>A	Coding	VAL193>ILE(193GTC>ATC)	Functional	GZ
MC1R	1	NC_009146.2	E.h.kulan(JF718954.1)	577G>A	Coding	VAL193>ILE(193GTC>ATC)	Functional	EH
MC1R	1	NC_009146.2	E.h.onager(JF718955.1)	577G>A	Coding	VAL193>ILE(193GTC>ATC)	Functional	EH
MC1R	1	NC_009146.2	E.hemionus(AF141365.1)	577G>A	Coding	VAL193>ILE(193GTC>ATC)	Functional	EH
MC1R	1	NC_009146.2	E.kiang(JF718953.1)	577G>A	Coding	VAL193>ILE(193GTC>ATC)	Functional	EK

GENE	Exon	REF.Sequence	Sequence name	Position	Region	Aminoacid change	Mutation	Code
MC1R	1	NC_009146.2	E.zebra(AF097749.1)	577G>A	Coding	VAL193>ILE(193GTC>ATC)	Functional	EZ
MC1R	1	NC_009146.2	E.zebra(JF718957.1)	577G>A	Coding	VAL193>ILE(193GTC>ATC)	Functional	EZ
MC1R	1	NC_009146.2	E. asinus (Oman)	629T>C	Coding	Met210THR(210ATG>ACG)	Functional	EA
MC1R	1	NC_009146.2	E. asinos (Miranda)	822C>A	Coding	GLY271GLY(121GGC>GGA)	Functional	EA
MC1R	1	NC_009146.2	E. africanus (HOI)	93G>A	Coding	THR31THR(31ACG>ACA)	Silent	HOI
MC1R	1	NC_009146.2	E. asinos (Miranda)	93G>A	Coding	THR31THR(31ACG>ACA)	Silent	EA
MC1R	1	NC_009146.2	E. asinus (Oman)	93G>A	Coding	THR31THR(31ACG>ACA)	Silent	EA
MC1R	1	NC_009146.2	E. asinus (Somalia)	93G>A	Coding	THR31THR(31ACG>ACA)	Silent	EA
MC1R	1	NC_009146.2	E.asinos (JF718952.1)	93G>A	Coding	THR31THR(31ACG>ACA)	Silent	EA
MC1R	1	NC_009146.2	E. asinos (Miranda)	940C>G	Coding	LEU314VAL(314CTG>GTG)	Functional	EA
SILV	1	NC_009149.2	E. africanus (HOI)	44T>C	Coding	ALA13ALA(13GCT>GCC)	Silent	HOI
SILV	1	NC_009149.2	E. asinos (Miranda)	44T>C	Coding	ALA13ALA(13GCT>GCC)	Silent	EA
SILV	1	NC_009149.2	E. asinus (Oman)	44T>C	Coding	ALA13ALA(13GCT>GCC)	Silent	EA
SILV	1	NC_009149.2	E. asinus (Somalia)	44T>C	Coding	ALA13ALA(13GCT>GCC)	Silent	EA
SILV	2	NC_009149.2	E. caballus(Sorraia)	2041C>A	Coding	ASP28GLU(28GAC>GAA)	Functional	EC
SILV	2	NC_009149.2	E. asinus (Oman)	2051G>A	Coding	GLY32ARG(32GGA>AGA)	Functional	EA
SILV	2	NC_009149.2	E. asinus (Somalia)	2051G>A	Coding	GLY32ARG(32GGA>AGA)	Functional	EA
SILV	2	NC_009149.2	E. caballus (2)	2061C>T	Coding	SER35LEU(TCG>TTG)	Functional	EC
SILV	2	NC_009149.2	E. asinos (Miranda)	2077C>T	Coding	SER40SER(40AGC>AGT)	Silent	EA
SILV	2	NC_009149.2	E. asinus (Oman)	2077C>T	Coding	SER40SER(40AGC>AGT)	Silent	EA
SILV	2	NC_009149.2	E. asinus (Somalia)	2077C>T	Coding	SER40SER(40AGC>AGT)	Silent	EA
SILV	2	NC_009149.2	E. grevyi(GZ)	2112A>C	Coding	GLN52PRO(52CAG>CCG)	Functional	GZ
SILV	4	NC_009149.2	E. grevyi(GZ)	3259G>A	Coding	VAL115VAL(GTG>GTA)	Silent	GZ
SILV	4	NC_009149.2	E. africanus (HOI)	3268T>A	Coding	GLY118GLY(118GGT>GGA)	Silent	HOI
SILV	4	NC_009149.2	E. asinos (Miranda)	3268T>A	Coding	GLY118GLY(118GGT>GGA)	Silent	EA
SILV	4	NC_009149.2	E. asinus (Oman)	3268T>A	Coding	GLY118GLY(118GGT>GGA)	Silent	EA
SILV	4	NC_009149.2	E. asinus (Somalia)	3268T>A	Coding	GLY118GLY(118GGT>GGA)	Silent	EA
SILV	4	NC_009149.2	E. grevyi(GZ)	3268T>A	Coding	GLY118GLY(118GGT>GGA)	Silent	GZ
SILV	4	NC_009149.2	E. hemionus(WAI)	3268T>A	Coding	GLY118GLY(118GGT>GGA)	Silent	EH
SILV	4	NC_009149.2	E.kiang(EK)	3268T>A	Coding	GLY118GLY(118GGT>GGA)	Silent	EK
SILV	4	NC_009149.2	E. africanus (HOI)	3330G>C	Coding	SER129THR(129AGC>ACC)	Functional	HOI
SILV	4	NC_009149.2	E. asinos (Miranda)	3330G>C	Coding	SER129THR(129AGC>ACC)	Functional	EA
SILV	4	NC_009149.2	E. asinus (Oman)	3330G>C	Coding	SER129THR(129AGC>ACC)	Functional	EA
SILV	4	NC_009149.2	E. asinus (Somalia)	3330G>C	Coding	SER129THR(129AGC>ACC)	Functional	EA
SILV	4	NC_009149.2	E. grevyi(GZ)	3330G>C	Coding	SER129THR(129AGC>ACC)	Functional	GZ
SILV	4	NC_009149.2	E. hemionus(WAI)	3330G>C	Coding	SER129THR(129AGC>ACC)	Functional	EH
SILV	4	NC_009149.2	E.kiang(EK)	3330G>C	Coding	SER129THR(129AGC>ACC)	Functional	EK
SILV	4	NC_009149.2	E. grevyi(GZ)	3349G>C	Coding	GLN145HIS(145CAG>CAC)	Functional	GZ
SILV	4	NC_009149.2	E. grevyi(GZ)	3357G>A	Coding	SER148ASN(148AGC>AAC)	Functional	GZ
SILV	4	NC_009149.2	E. hemionus(WAI)	3377A>G	Coding	THR155ALA(155ACC>GCC)	Functional	EH
SILV	4	NC_009149.2	E.kiang(EK)	3377A>G	Coding	THR155ALA(155ACC>GCC)	Functional	EK
SILV	5	NC_009149.2	E. hemionus(WAI)	3761A>M	Non-coding	-	-	EH
SILV	5	NC_009149.2	E. hemionus(WAI)	3773G>C	Non-coding	-	-	EH
SILV	5	NC_009149.2	E. asinos (Miranda)	3842A>G	Coding	THR176THR(176ACA>ACG)	Silent	EA
SILV	5	NC_009149.2	E. asinus (Oman)	3842A>G	Coding	THR176THR(176ACA>ACG)	Silent	EA
SILV	5	NC_009149.2	E. asinus (Somalia)	3842A>G	Coding	THR176THR(176ACA>ACG)	Silent	EA
SILV	5	NC_009149.2	E. asinus (Somalia)	3965C>A	Non-coding	-	-	EA
SILV	6	NC_009149.2	E. africanus (HOI)	4319C>T	Coding	ALA254VAL(254GCT>GTT)	Functional	HOI
SILV	6	NC_009149.2	E. asinos (Miranda)	4319C>T	Coding	ALA254VAL(254GCT>GTT)	Functional	EA
SILV	6	NC_009149.2	E. asinus (Oman)	4319C>T	Coding	ALA254VAL(254GCT>GTT)	Functional	EA
SILV	6	NC_009149.2	E. asinus (Somalia)	4319C>T	Coding	ALA254VAL(254GCT>GTT)	Functional	EA
SILV	6	NC_009149.2	E. asinos (Miranda)	4356C>T	Coding	THR287ILE(287ACT>ATT)	Functional	EA
SILV	6	NC_009149.2	E. asinus (Oman)	4356C>T	Coding	THR287ILE(287ACT>ATT)	Functional	EA
SILV	6	NC_009149.2	E. asinus (Somalia)	4356C>T	Coding	THR287ILE(287ACT>ATT)	Functional	EA
SILV	6	NC_009149.2	E. caballus (1)	4418C>T	Coding	THR287ILE(287ACT>ATT)	Functional	EC
SILV	6	NC_009149.2	E. caballus (5)	4418C>T	Coding	THR287ILE(287ACT>ATT)	Functional	EC
SILV	6	NC_009149.2	E. caballus (5)	4426G>C	Coding	VAL290LEU(290GTG>CTG)	Functional	EC

GENE	Exon	REF.Sequence	Sequence name	Position	Region	Aminoacid change	Mutation	Code
SILV	6	NC_009149.2	E. caballus (3)	4622C>T	Coding	THR355ILE(355ACC>ATC)	Functional	EC
SILV	6	NC_009149.2	E. africanus (HOI)	4761G>C	Coding	THR401THR(401ACG>ACC)	Silent	HOI
SILV	6	NC_009149.2	E. asinos (Miranda)	4761G>C	Coding	THR401THR(401ACG>ACC)	Silent	EA
SILV	6	NC_009149.2	E. asinus (Oman)	4761G>C	Coding	THR401THR(401ACG>ACC)	Silent	EA
SILV	6	NC_009149.2	E. asinus (Somalia)	4761G>C	Coding	THR401THR(401ACG>ACC)	Silent	EA
SILV	6	NC_009149.2	E. grevyi(GZ)	4761G>C	Coding	THR401THR(401ACG>ACC)	Silent	GZ
SILV	6	NC_009149.2	E. hemionus(WAI)	4761G>C	Coding	THR401THR(401ACG>ACC)	Silent	EH
SILV	6	NC_009149.2	E.kiang(EK)	4761G>C	Coding	THR401THR(401ACG>ACC)	Silent	EK
SILV	6	NC_009149.2	E. africanus (HOI)	4812G>A	Coding	THR418THR(418ACG>ACA)	Silent	HOI
SILV	6	NC_009149.2	E. asinos (Miranda)	4812G>A	Coding	THR418THR(418ACG>ACA)	Silent	EA
SILV	6	NC_009149.2	E. asinus (Oman)	4812G>A	Coding	THR418THR(418ACG>ACA)	Silent	EA
SILV	6	NC_009149.2	E. asinus (Somalia)	4812G>A	Coding	THR418THR(418ACG>ACA)	Silent	EA
SILV	6	NC_009149.2	E. grevyi(GZ)	4812G>A	Coding	THR418THR(418ACG>ACA)	Silent	GZ
SILV	6	NC_009149.2	E. hemionus(WAI)	4812G>A	Coding	THR418THR(418ACG>ACA)	Silent	EH
SILV	6	NC_009149.2	E.kiang(EK)	4812G>A	Coding	THR418THR(418ACG>ACA)	Silent	EK
SILV	6	NC_009149.2	E. africanus (HOI)	4819G>A	Coding	GLU421LYS(GAG>AAG)	Functional	HOI
SILV	6	NC_009149.2	E. asinos (Miranda)	4819G>A	Coding	GLU421LYS(GAG>AAG)	Functional	EA
SILV	6	NC_009149.2	E. asinus (Oman)	4819G>A	Coding	GLU421LYS(GAG>AAG)	Functional	EA
SILV	6	NC_009149.2	E. asinus (Somalia)	4819G>A	Coding	GLU421LYS(GAG>AAG)	Functional	EA
SILV	6	NC_009149.2	E. grevyi(GZ)	4819G>A	Coding	GLU421LYS(GAG>AAG)	Functional	GZ
SILV	6	NC_009149.2	E. hemionus(WAI)	4819G>A	Coding	GLU421LYS(GAG>AAG)	Functional	EH
SILV	6	NC_009149.2	E.kiang(EK)	4819G>A	Coding	GLU421LYS(GAG>AAG)	Functional	EK
SILV	6	NC_009149.2	E. africanus (HOI)	4929C>T	Non-coding	-	-	HOI
SILV	6	NC_009149.2	E. asinos (Miranda)	4929C>T	Non-coding	-	-	EA
SILV	6	NC_009149.2	E. asinus (Oman)	4929C>T	Non-coding	-	-	EA
SILV	6	NC_009149.2	E. asinus (Somalia)	4929C>T	Non-coding	-	-	EA
SILV	6	NC_009149.2	E. grevyi(GZ)	4929C>T	Non-coding	-	-	GZ
SILV	6	NC_009149.2	E. hemionus(WAI)	4929C>T	Non-coding	-	-	EH
SILV	6	NC_009149.2	E.kiang(EK)	4929C>T	Non-coding	-	-	EK
SILV	7	NC_009149.2	E. caballus (5)	6196G>A	Coding	ARG468>GLN(468>CGA>CAA)	Functional	EC
SILV	7	NC_009149.2	E. caballus (2)	6257C>T	Coding	ILE488ILE(488ATC>ATT)	Silent	EC
SILV	7	NC_009149.2	E. caballus(Sorraia)	6259T>A	Coding	VAL489ASP(489GTC>GAC)	Functional	EC
SILV	8	NC_009149.2	E. asinus (Oman)	6544C>T	Coding	SER504SER(504AGC>AGT)	Silent	EA
SILV	9	NC_009149.2	E. africanus (HOI)	6745A>G	Coding	THR544ALA(ACA544>GCA)	Functional	HOI
SILV	9	NC_009149.2	E. asinos (Miranda)	6745A>G	Coding	THR544ALA(ACA544>GCA)	Functional	EA
SILV	9	NC_009149.2	E. asinus (Oman)	6745A>G	Coding	THR544ALA(ACA544>GCA)	Functional	EA
SILV	9	NC_009149.2	E. asinus (Somalia)	6745A>G	Coding	THR544ALA(ACA544>GCA)	Functional	EA
SILV	9	NC_009149.2	E. grevyi(GZ)	6745A>G	Coding	THR544ALA(ACA544>GCA)	Functional	GZ
SILV	9	NC_009149.2	E. hemionus(WAI)	6745A>G	Coding	THR544ALA(ACA544>GCA)	Functional	EH
SILV	9	NC_009149.2	E.kiang(EK)	6745A>G	Coding	THR544ALA(ACA544>GCA)	Functional	EK
SILV	9	NC_009149.2	E.kiang(EK)	6763C>A	Coding	GLN550LYS(550CAG>AAG)	Functional	EK
SILV	9	NC_009149.2	E. asinos (Miranda)	6893T>G	Non-coding	-	-	EA
SILV	9	NC_009149.2	E. caballus (1)	6893T>G	Non-coding	-	-	EC
SILV	9	NC_009149.2	E. caballus (2)	6893T>G	Non-coding	-	-	EC
SILV	9	NC_009149.2	E. caballus (3)	6893T>G	Non-coding	-	-	EC
SILV	9	NC_009149.2	E. caballus (4)	6893T>G	Non-coding	-	-	EC
SILV	9	NC_009149.2	E. caballus (5)	6893T>G	Non-coding	-	-	EC
SILV	9	NC_009149.2	E. caballus (Garrano)	6893T>G	Non-coding	-	-	EC
SILV	9	NC_009149.2	E. caballus(Sorraia)	6893T>G	Non-coding	-	-	EC
SILV	10	NC_009149.2	E. caballus (Garrano)	6954T>A	Non-coding	-	-	EC
SILV	10	NC_009149.2	E. caballus (Garrano)	6965C>T	Non-coding	-	-	EC
SILV	10	NC_009149.2	E. caballus (Garrano)	6968C>T	Non-coding	-	-	EC
SILV	10	NC_009149.2	E. hemionus(WAI)	7037A>G	Coding	MET608VAL(608ATG>GTG)	Functional	EH
SILV	10	NC_009149.2	E.kiang(EK)	7037A>G	Coding	MET608VAL(608ATG>GTG)	Functional	EK
SILV	10	NC_009149.2	E. hemionus(WAI)	7080C>A	Non-coding	-	-	EH
SILV	10	NC_009149.2	E. caballus (2)	7083C>T	Non-coding	-	-	EC
SILV	10	NC_009149.2	E. caballus (3)	7083C>T	Non-coding	-	-	EC

GENE	Exon	REF.Sequence	Sequence name	Position	Region	Aminoacid change	Mutation	Code
SILV	10	NC_009149.2	E. caballus (Garrano)	7083C>T	Non-coding	-	-	EC
SILV	11	NC_009149.2	E. asinos (Miranda)	7657T>A	Non-coding	-	-	EA
SILV	11	NC_009149.2	E. asinus (Oman)	7658T>G	Non-coding	-	-	EA
SILV	11	NC_009149.2	E. asinus (Oman)	7659C>T	Non-coding	-	-	EA
SILV	11	NC_009149.2	E. asinos (Miranda)	7660C>T	Non-coding	-	-	EA
SILV	11	NC_009149.2	E. asinus (Oman)	7660C>T	Non-coding	-	-	EA
SILV	11	NC_009149.2	E. asinos (Miranda)	7662T>C	Non-coding	-	-	EA
SILV	11	NC_009149.2	E. asinus (Oman)	7662T>C	Non-coding	-	-	EA
SILV	11	NC_009149.2	E. asinos (Miranda)	7676A>T	Non-coding	-	-	EA
SILV	11	NC_009149.2	E. hemionus(WAI)	7676A>T	Non-coding	-	-	EH
SILV	11	NC_009149.2	E. asinos (Miranda)	7701G>C	Coding	GLY623ALA(623GGC>GCC)	Functional	EA
SILV	11	NC_009149.2	E. asinus (Oman)	7701G>C	Coding	GLY623ALA(623GGC>GCC)	Functional	EA
SILV	11	NC_009149.2	E. grevyi(GZ)	7752G>A	Coding	ARG640HIS(640CGT>CAT)	Functional	GZ
SILV	11	NC_009149.2	E. caballus (2)	7814C>A	Coding	GLN661LYS(661CAG>AAG)	Functional	EC
SILV	11	NC_009149.2	E. hemionus(WAI)	7844C>A	Coding	GLN630LYS(630CAG>AAG)	Functional	EH
SLC36A1	1	NC_009157	E. hemionus(WAI)	115C>A	Coding	GLN39LYS(39CAG>AAG)	Functional	EH
SLC36A1	1	NC_009157	E. africanus (HOI)	120G>A	Coding	ARG40ARG(40CGG>CGA)	Silent	HOI
SLC36A1	1	NC_009157	E. asinos (Miranda)	120G>A	Coding	ARG40ARG(40CGG>CGA)	Silent	EA
SLC36A1	1	NC_009157	E. asinus (Oman)	120G>A	Coding	ARG40ARG(40CGG>CGA)	Silent	EA
SLC36A1	1	NC_009157	E. asinus (Somalia)	120G>A	Coding	ARG40ARG(40CGG>CGA)	Silent	EA
SLC36A1	1	NC_009157	E. grevyi(GZ)	120G>A	Coding	ARG40ARG(40CGG>CGA)	Silent	GZ
SLC36A1	1	NC_009157	E. hemionus(WAI)	147A>C	Non-coding	-	-	EH
SLC36A1	1	NC_009157	E. africanus (HOI)	60C>T	Coding	SER20SER(20AGC>AGT)	Silent	HOI
SLC36A1	1	NC_009157	E. asinos (Miranda)	60C>T	Coding	SER20SER(20AGC>AGT)	Silent	EA
SLC36A1	1	NC_009157	E. asinus (Oman)	60C>T	Coding	SER20SER(20AGC>AGT)	Silent	EA
SLC36A1	1	NC_009157	E. asinus (Somalia)	60C>T	Coding	SER20SER(20AGC>AGT)	Silent	EA
SLC36A1	1	NC_009157	E. grevyi(GZ)	60C>T	Coding	SER20SER(20AGC>AGT)	Silent	GZ
SLC36A1	1	NC_009157	E. grevyi(GZ)	66C>T	Coding	ASP22ASP(22GAC>GAC)	Silent	GZ
SLC36A1	2	NC_009157	E. grevyi(GZ)	5614A>G	Non-coding	-	-	GZ
SLC36A1	2	NC_009157	E. africanus (HOI)	5671T>C	Coding	ILE61ILE(61ATT>ATC)	Silent	HOI
SLC36A1	2	NC_009157	E. asinos (Miranda)	5671T>C	Coding	ILE61ILE(61ATT>ATC)	Silent	EA
SLC36A1	2	NC_009157	E. asinus (Oman)	5671T>C	Coding	ILE61ILE(61ATT>ATC)	Silent	EA
SLC36A1	2	NC_009157	E. asinus (Somalia)	5671T>C	Coding	ILE61ILE(61ATT>ATC)	Silent	EA
SLC36A1	2	NC_009157	E. grevyi(GZ)	5671T>C	Coding	ILE61ILE(61ATT>ATC)	Silent	GZ
SLC36A1	2	NC_009157	E. hemionus(WAI)	5671T>C	Coding	ILE61ILE(61ATT>ATC)	Silent	EH
SLC36A1	2	NC_009157	E.kiang(EK)	5671T>C	Coding	ILE61ILE(61ATT>ATC)	Silent	EK
SLC36A1	2	NC_009157	E.caballus(EU432177.1)	5676C>G	Coding	THR63ARG(63ACA>AGA)	Functional	EC
SLC36A1	3	NC_009157	E. africanus (HOI)	6598A>G	Non-coding	-	-	HOI
SLC36A1	3	NC_009157	E. asinos (Miranda)	6598A>G	Non-coding	-	-	EA
SLC36A1	3	NC_009157	E. asinus (Oman)	6598A>G	Non-coding	-	-	EA
SLC36A1	3	NC_009157	E. asinus (Somalia)	6598A>G	Non-coding	-	-	EA
SLC36A1	3	NC_009157	E. grevyi(GZ)	6598A>G	Non-coding	-	-	GZ
SLC36A1	3	NC_009157	E. hemionus(WAI)	6598A>G	Non-coding	-	-	EH
SLC36A1	3	NC_009157	E.kiang(EK)	6598A>G	Non-coding	-	-	EK
SLC36A1	3	NC_009157	E. africanus (HOI)	6712A>G	Non-coding	-	-	HOI
SLC36A1	3	NC_009157	E. asinos (Miranda)	6712A>G	Non-coding	-	-	EA
SLC36A1	3	NC_009157	E. asinus (Oman)	6712A>G	Non-coding	-	-	EA
SLC36A1	3	NC_009157	E. asinus (Somalia)	6712A>G	Non-coding	-	-	EA
SLC36A1	3	NC_009157	E. grevyi(GZ)	6712A>G	Non-coding	-	-	GZ
SLC36A1	3	NC_009157	E. hemionus(WAI)	6712A>G	Non-coding	-	-	EH
SLC36A1	3	NC_009157	E.kiang(EK)	6712A>G	Non-coding	-	-	EK
SLC36A1	3	NC_009157	E. africanus (HOI)	6717A>G	Non-coding	-	-	HOI
SLC36A1	3	NC_009157	E. asinos (Miranda)	6717A>G	Non-coding	-	-	EA
SLC36A1	3	NC_009157	E. asinus (Oman)	6717A>G	Non-coding	-	-	EA
SLC36A1	3	NC_009157	E. asinus (Somalia)	6717A>G	Non-coding	-	-	EA
SLC36A1	3	NC_009157	E. grevyi(GZ)	6717A>G	Non-coding	-	-	GZ
SLC36A1	3	NC_009157	E. hemionus(WAI)	6717A>G	Non-coding	-	-	EH

GENE	Exon	REF.Sequence	Sequence name	Position	Region	Aminoacid change	Mutation	Code
SLC36A1	3	NC_009157	E.kiang(EK)	6717A>G	Non-coding	-	-	EK
SLC36A1	3	NC_009157	E. africanus (HOI)	6719C>A	Non-coding	-	-	HOI
SLC36A1	3	NC_009157	E. asinus (Somalia)	6719C>A	Non-coding	-	-	EA
SLC36A1	3	NC_009157	E. grevyi(GZ)	6719C>A	Non-coding	-	-	GZ
SLC36A1	3	NC_009157	E. hemionus(WAI)	6719C>A	Non-coding	-	-	EH
SLC36A1	4	NC_009157	E. africanus (HOI)	7141A>C	Non-coding	-	-	HOI
SLC36A1	4	NC_009157	E. asinos (Miranda)	7141A>C	Non-coding	-	-	EA
SLC36A1	4	NC_009157	E. asinus (Oman)	7141A>C	Non-coding	-	-	EA
SLC36A1	4	NC_009157	E. asinus (Somalia)	7141A>C	Non-coding	-	-	EA
SLC36A1	4	NC_009157	E. hemionus(WAI)	7141A>C	Non-coding	-	-	EH
SLC36A1	4	NC_009157	E. africanus (HOI)	7170C>T	Non-coding	-	-	HOI
SLC36A1	4	NC_009157	E. asinos (Miranda)	7170C>T	Non-coding	-	-	EA
SLC36A1	4	NC_009157	E. asinus (Oman)	7170C>T	Non-coding	-	-	EA
SLC36A1	4	NC_009157	E. asinus (Somalia)	7170C>T	Non-coding	-	-	EA
SLC36A1	4	NC_009157	E. hemionus(WAI)	7170C>T	Non-coding	-	-	EH
SLC36A1	4	NC_009157	E. africanus (HOI)	7248G>A	Coding	ARG133ARG(133CGG>CGA)	Silent	HOI
SLC36A1	4	NC_009157	E. asinos (Miranda)	7248G>A	Coding	ARG133ARG(133CGG>CGA)	Silent	EA
SLC36A1	4	NC_009157	E. asinus (Oman)	7248G>A	Coding	ARG133ARG(133CGG>CGA)	Silent	EA
SLC36A1	4	NC_009157	E. asinus (Somalia)	7248G>A	Coding	ARG133ARG(133CGG>CGA)	Silent	EA
SLC36A1	4	NC_009157	E. hemionus(WAI)	7248G>A	Coding	ARG133ARG(133CGG>CGA)	Silent	EH
SLC36A1	6	NC_009157	E. caballus (1)	9750C>T	Coding	ALA172VAL(172GCG>GTG)	Functional	EC
SLC36A1	6	NC_009157	E. caballus (2)	9750C>T	Coding	ALA172VAL(172GCG>GTG)	Functional	EC
SLC36A1	6	NC_009157	E. caballus (5)	9750C>T	Coding	ALA172VAL(172GCG>GTG)	Functional	EC
SLC36A1	6	NC_009157	E. caballus (Garrano)	9750C>T	Coding	ALA172VAL(172GCG>GTG)	Functional	EC
SLC36A1	6	NC_009157	E. caballus (Garrano)	9790G>A	Coding	THR185THR(172ACG>ACA)	Functional	EC
SLC36A1	6	NC_009157	E. africanus (HOI)	9817G>A	Coding	SER194SER(194TCG>TCA)	Silent	HOI
SLC36A1	6	NC_009157	E. asinus (Oman)	9817G>A	Coding	SER194SER(194TCG>TCA)	Silent	EA
SLC36A1	6	NC_009157	E. asinus (Somalia)	9817G>A	Coding	SER194SER(194TCG>TCA)	Silent	EA
SLC36A1	6	NC_009157	E. grevyi(GZ)	9817G>A	Coding	SER194SER(194TCG>TCA)	Silent	GZ
SLC36A1	6	NC_009157	E. hemionus(WAI)	9817G>A	Coding	SER194SER(194TCG>TCA)	Silent	EH
SLC36A1	6	NC_009157	E. grevyi(GZ)	9840T>C	Coding	LEU202PRO(202CTG>CCG)	Functional	GZ
SLC36A1	6	NC_009157	E. africanus (HOI)	9946A>G	Coding	GLN237GLN(237CAA>CAG)	Silent	HOI
SLC36A1	6	NC_009157	E. asinus (Oman)	9946A>G	Coding	GLN237GLN(237CAA>CAG)	Silent	EA
SLC36A1	6	NC_009157	E. asinus (Somalia)	9946A>G	Coding	GLN237GLN(237CAA>CAG)	Silent	EA
SLC36A1	6	NC_009157	E. grevyi(GZ)	9946A>G	Coding	GLN237GLN(237CAA>CAG)	Silent	GZ
SLC36A1	6	NC_009157	E. hemionus(WAI)	9946A>G	Coding	GLN237GLN(237CAA>CAG)	Silent	EH
SLC36A1	6	NC_009157	E. asinus (Oman)	9978G>A	Non-coding	-	-	EA
SLC36A1	6	NC_009157	E. caballus (1)	9978G>A	Non-coding	-	-	EC
SLC36A1	6	NC_009157	E. caballus (2)	9978G>A	Non-coding	-	-	EC
SLC36A1	6	NC_009157	E. caballus (3)	9978G>A	Non-coding	-	-	EC
SLC36A1	6	NC_009157	E. caballus (4)	9978G>A	Non-coding	-	-	EC
SLC36A1	6	NC_009157	E. caballus (5)	9978G>A	Non-coding	-	-	EC
SLC36A1	6	NC_009157	E. caballus(Sorraia)	9978G>A	Non-coding	-	-	EC
SLC36A1	6	NC_009157	E. grevyi(GZ)	9978G>A	Non-coding	-	-	GZ
SLC36A1	7	NC_009157	E. asinus (Somalia)	13429G>A	Coding	LIS256LIS(256AAG>AAA)	Silent	EA
SLC36A1	7	NC_009157	E. africanus (HOI)	13477C>T	Coding	ILE272ILE(272ATC>ATT)	Silent	HOI
SLC36A1	7	NC_009157	E. asinos (Miranda)	13477C>T	Coding	ILE272ILE(272ATC>ATT)	Silent	EA
SLC36A1	7	NC_009157	E. asinus (Oman)	13477C>T	Coding	ILE272ILE(272ATC>ATT)	Silent	EA
SLC36A1	7	NC_009157	E. asinus (Somalia)	13477C>T	Coding	ILE272ILE(272ATC>ATT)	Silent	EA
SLC36A1	8	NC_009157	E. caballus (1)	16162G>A	Coding	GLY295GLU(295GGA>GAA)	Functional	EC
SLC36A1	8	NC_009157	E. africanus (HOI)	16253T>C	Coding	ASN325ASN(AAT>AAC)	Silent	HOI
SLC36A1	8	NC_009157	E. asinus (Oman)	16253T>C	Coding	ASN325ASN(AAT>AAC)	Silent	EA
SLC36A1	8	NC_009157	E. asinus (Somalia)	16253T>C	Coding	ASN325ASN(AAT>AAC)	Silent	EA
SLC36A1	8	NC_009157	E. grevyi(GZ)	16253T>C	Coding	ASN325ASN(AAT>AAC)	Silent	GZ
SLC36A1	8	NC_009157	E. hemionus(WAI)	16253T>C	Coding	ASN325ASN(AAT>AAC)	Silent	EH
SLC36A1	8	NC_009157	E. asinus (Oman)	16278G>A	Non-coding	-	-	EA
SLC36A1	8	NC_009157	E. caballus (4)	16278G>A	Non-coding	-	-	EC

GENE	Exon	REF.Sequence	Sequence name	Position	Region	Aminoacid change	Mutation	Code
SLC36A1	8	NC_009157	E. caballus(Sorraia)	16284G>A	Non-coding	-	-	EC
SLC36A1	9	NC_009157	E. africanus (HOI)	19095C>T	Non-coding	-	-	HOI
SLC36A1	9	NC_009157	E. hemionus(WAI)	19154C>T	Coding	TYR347TYR(347TAC>TAT)	Silent	EH
SLC36A1	9	NC_009157	E. caballus (3)	19196C>A	Coding	PHE361LEU(361TTC>TTA)	Functional	EC
SLC36A1	9	NC_009157	E. caballus(Sorraia)	19196C>A	Coding	PHE361LEU(361TTC>TTA)	Functional	EC
SLC36A1	9	NC_009157	E. africanus (HOI)	19211A>G	Coding	VAL366VAL(366GTA>GTG)	Silent	HOI
SLC36A1	9	NC_009157	E. asinos (Miranda)	19211A>G	Coding	VAL366VAL(366GTA>GTG)	Silent	EA
SLC36A1	9	NC_009157	E. asinus (Oman)	19211A>G	Coding	VAL366VAL(366GTA>GTG)	Silent	EA
SLC36A1	9	NC_009157	E. asinus (Somalia)	19211A>G	Coding	VAL366VAL(366GTA>GTG)	Silent	EA
SLC36A1	9	NC_009157	E. grevyi(GZ)	19211A>G	Coding	VAL366VAL(366GTA>GTG)	Silent	GZ
SLC36A1	9	NC_009157	E. hemionus(WAI)	19211A>G	Coding	VAL366VAL(366GTA>GTG)	Silent	EH
SLC36A1	9	NC_009157	E. hemionus(WAI)	19254A>G	Coding	MET381VAL(381ATG>GTG)	Functional	EH
SLC36A1	9	NC_009157	E. africanus (HOI)	19295C>T	Non-coding	-	-	HOI
SLC36A1	9	NC_009157	E. asinos (Miranda)	19295C>T	Non-coding	-	-	EA
SLC36A1	9	NC_009157	E. asinus (Oman)	19295C>T	Non-coding	-	-	EA
SLC36A1	9	NC_009157	E. asinus (Somalia)	19295C>T	Non-coding	-	-	EA
SLC36A1	9	NC_009157	E. grevyi(GZ)	19295C>T	Non-coding	-	-	GZ
SLC36A1	9	NC_009157	E. hemionus(WAI)	19295C>T	Non-coding	-	-	EH
SLC36A1	9	NC_009157	E. hemionus(WAI)	19331G>C	Non-coding	-	-	EH
SLC36A1	9	NC_009157	E. hemionus(WAI)	19335C>G	Non-coding	-	-	EH
SLC36A1	10	NC_009157	E. africanus (HOI)	27901T>C	Non-coding	-	-	HOI
SLC36A1	10	NC_009157	E. asinos (Miranda)	27901T>C	Non-coding	-	-	EA
SLC36A1	10	NC_009157	E. asinus (Oman)	27901T>C	Non-coding	-	-	EA
SLC36A1	10	NC_009157	E. asinus (Somalia)	27901T>C	Non-coding	-	-	EA
SLC36A1	10	NC_009157	E. africanus (HOI)	27902G>A	Non-coding	-	-	HOI
SLC36A1	10	NC_009157	E. asinos (Miranda)	27902G>A	Non-coding	-	-	EA
SLC36A1	10	NC_009157	E. asinus (Oman)	27902G>A	Non-coding	-	-	EA
SLC36A1	10	NC_009157	E. asinus (Somalia)	27902G>A	Non-coding	-	-	EA
SLC36A1	10	NC_009157	E. africanus (HOI)	27916T>C	Coding	ALA390ALA(390GCT>GCC)	Silent	HOI
SLC36A1	10	NC_009157	E. asinos (Miranda)	27916T>C	Coding	ALA390ALA(390GCT>GCC)	Silent	EA
SLC36A1	10	NC_009157	E. asinus (Oman)	27916T>C	Coding	ALA390ALA(390GCT>GCC)	Silent	EA
SLC36A1	10	NC_009157	E. asinus (Somalia)	27916T>C	Coding	ALA390ALA(390GCT>GCC)	Silent	EA
SLC36A1	10	NC_009157	E. africanus (HOI)	27917G>A	Coding	VAL391ILE(391GTC>ATC)	Functional	HOI
SLC36A1	10	NC_009157	E. asinos (Miranda)	27917G>A	Coding	VAL391ILE(391GTC>ATC)	Functional	EA
SLC36A1	10	NC_009157	E. asinus (Oman)	27917G>A	Coding	VAL391ILE(391GTC>ATC)	Functional	EA
SLC36A1	10	NC_009157	E. asinus (Somalia)	27917G>A	Coding	VAL391ILE(391GTC>ATC)	Functional	EA
SLC45A2	1	NC_009164.2	E. caballus (2)	121G>G	Coding	ALA41THR(41GCC>ACC)	Functional	EC
SLC45A2	1	NC_009164.2	E. africanus (HOI)	159G>A	Coding	ALA53ALA(53GCG>GCA)	Silent	HOI
SLC45A2	1	NC_009164.2	E. africanus (HOI)	159G>A	Coding	ALA53ALA(53GCG>GCA)	Silent	HOI
SLC45A2	1	NC_009164.2	E. asinos (Miranda)	159G>A	Coding	ALA53ALA(53GCG>GCA)	Silent	EA
SLC45A2	1	NC_009164.2	E. asinos (Miranda)	159G>A	Coding	ALA53ALA(53GCG>GCA)	Silent	EA
SLC45A2	1	NC_009164.2	E. asinus (Oman)	159G>A	Coding	ALA53ALA(53GCG>GCA)	Silent	EA
SLC45A2	1	NC_009164.2	E. asinus (Oman)	159G>A	Coding	ALA53ALA(53GCG>GCA)	Silent	EA
SLC45A2	1	NC_009164.2	E. asinus (Somalia)	159G>A	Coding	ALA53ALA(53GCG>GCA)	Silent	EA
SLC45A2	1	NC_009164.2	E. caballus (Garrano)	165C>T	Coding	TYR55TYR(55TAC>TAT)	Silent	EC
SLC45A2	1	NC_009164.2	E. caballus (Garrano)	165C>T	Coding	TYR55TYR(55TAC>TAT)	Silent	EC
SLC45A2	1	NC_009164.2	E. asinos (Miranda)	26G>A	Coding	GLY9ASP(9GGC>GAC)	Functional	EA
SLC45A2	1	NC_009164.2	E. africanus (HOI)	318T>C	Coding	ILE106ILE(106ATT>ATC)	Silent	HOI
SLC45A2	1	NC_009164.2	E. asinos (Miranda)	318T>C	Coding	ILE106ILE(106ATT>ATC)	Silent	EA
SLC45A2	1	NC_009164.2	E. asinus (Oman)	318T>C	Coding	ILE106ILE(106ATT>ATC)	Silent	EA
SLC45A2	1	NC_009164.2	E. asinus (Somalia)	318T>C	Coding	ILE106ILE(106ATT>ATC)	Silent	EA
SLC45A2	1	NC_009164.2	E. africanus (HOI)	32C>A	Coding	PRO11HIS(11CCC>CAC)	Functional	HOI
SLC45A2	1	NC_009164.2	E. asinos (Miranda)	32C>A	Coding	PRO11HIS(11CCC>CAC)	Functional	EA
SLC45A2	1	NC_009164.2	E. asinus (Oman)	32C>A	Coding	PRO11HIS(11CCC>CAC)	Functional	EA
SLC45A2	1	NC_009164.2	E. asinus (Somalia)	32C>A	Coding	PRO11HIS(11CCC>CAC)	Functional	EA
SLC45A2	1	NC_009164.2	E. caballus (3)	32C>A	Coding	PRO11HIS(11CCC>CAC)	Functional	EC
SLC45A2	1	NC_009164.2	E. caballus (4)	32C>A	Coding	PRO11HIS(11CCC>CAC)	Functional	EC

GENE	Exon	REF.Sequence	Sequence name	Position	Region	Aminoacid change	Mutation	Code
SLC45A2	1	NC_009164.2	E. caballus (5)	32C>A	Coding	PRO11HIS(11CCC>CAC)	Functional	EC
SLC45A2	1	NC_009164.2	E. caballus (Garrano)	32C>A	Coding	PRO11HIS(11CCC>CAC)	Functional	EC
SLC45A2	1	NC_009164.2	E. caballus(Sorraia)	32C>A	Coding	PRO11HIS(11CCC>CAC)	Functional	EC
SLC45A2	1	NC_009164.2	E. africanus (HOI)	357G>C	Coding	LEU119LEU(119CTG>CTC)	Silent	HOI
SLC45A2	1	NC_009164.2	E. asinos (Miranda)	357G>C	Coding	LEU119LEU(119CTG>CTC)	Silent	EA
SLC45A2	1	NC_009164.2	E. asinus (Oman)	357G>C	Coding	LEU119LEU(119CTG>CTC)	Silent	EA
SLC45A2	1	NC_009164.2	E. asinus (Somalia)	357G>C	Coding	LEU119LEU(119CTG>CTC)	Silent	EA
SLC45A2	1	NC_009164.2	E. caballus (2)	60C>T	Coding	GLY20GLY(20GGC>GGT)	Silent	EC
SLC45A2	1	NC_009164.2	E. caballus (2)	72C>T	Coding	SER24SER(24TCC>TCT)	Silent	EC
SLC45A2	1	NC_009164.2	E. asinos (Miranda)	96G>A	Coding	THR32THR(96ACG>ACA)	Silent	EA
SLC45A2	1	NC_009164.2	E. asinus (Somalia)	96G>A	Coding	THR32THR(96ACG>ACA)	Silent	EA
SLC45A2	1	NC_009164.2	E. caballus (2)	99C>A	Coding	GLY33GLY(33GGC>GGA)	Silent	EC
SLC45A2	2	NC_009164.2	E. africanus (HOI)	2253G>A	Coding	ASP134GLN(134CGA>CAA)	Functional	HOI
SLC45A2	2	NC_009164.2	E. asinos (Miranda)	2253G>A	Coding	ASP134GLN(134CGA>CAA)	Functional	EA
SLC45A2	2	NC_009164.2	E. asinus (Oman)	2253G>A	Coding	ASP134GLN(134CGA>CAA)	Functional	EA
SLC45A2	2	NC_009164.2	E. asinus (Somalia)	2253G>A	Coding	ASP134GLN(134CGA>CAA)	Functional	EA
SLC45A2	2	NC_009164.2	E. caballus (5)	2309G>A	Coding	ASP153ASN(153GAT>AAT)	Functional	EC
SLC45A2	2	NC_009164.2	E. grevyi(GZ)	2396C>T	Coding	HIS182HIS(182CAC>TAC)	Functional	GZ
SLC45A2	3	NC_009164.2	E. africanus (HOI)	17010A>G	Non-coding	-	-	HOI
SLC45A2	3	NC_009164.2	E. asinos (Miranda)	17010A>G	Non-coding	-	-	EA
SLC45A2	3	NC_009164.2	E. asinus (Oman)	17010A>G	Non-coding	-	-	EA
SLC45A2	3	NC_009164.2	E. asinus (Somalia)	17010A>G	Non-coding	-	-	EA
SLC45A2	3	NC_009164.2	E. asinos (Miranda)	17050C>T	Coding	GLY198GLY(198GGC>GGT)	Silent	EA
SLC45A2	3	NC_009164.2	E. africanus (HOI)	17186G>C	Coding	ASP244HIS(244GAT>CAT)	Functional	HOI
SLC45A2	3	NC_009164.2	E. asinos (Miranda)	17186G>C	Coding	ASP244HIS(244GAT>CAT)	Functional	EA
SLC45A2	3	NC_009164.2	E. asinus (Oman)	17186G>C	Coding	ASP244HIS(244GAT>CAT)	Functional	EA
SLC45A2	3	NC_009164.2	E. asinus (Somalia)	17186G>C	Coding	ASP244HIS(244GAT>CAT)	Functional	EA
SLC45A2	3	NC_009164.2	E. hemionus(WAI)	17186G>C	Coding	ASP244HIS(244GAT>CAT)	Functional	EH
SLC45A2	3	NC_009164.2	E. africanus (HOI)	17203A>T	Coding	ILE249ILE(249ATA>ATT)	Silent	HOI
SLC45A2	3	NC_009164.2	E. asinos (Miranda)	17203A>T	Coding	ILE249ILE(249ATA>ATT)	Silent	EA
SLC45A2	3	NC_009164.2	E. asinus (Oman)	17203A>T	Coding	ILE249ILE(249ATA>ATT)	Silent	EA
SLC45A2	3	NC_009164.2	E. asinus (Somalia)	17203A>T	Coding	ILE249ILE(249ATA>ATT)	Silent	EA
SLC45A2	3	NC_009164.2	E. caballus (1)	17203A>T	Coding	ILE249ILE(249ATA>ATT)	Silent	EC
SLC45A2	3	NC_009164.2	E. caballus (2)	17203A>T	Coding	ILE249ILE(249ATA>ATT)	Silent	EC
SLC45A2	3	NC_009164.2	E. caballus (3)	17203A>T	Coding	ILE249ILE(249ATA>ATT)	Silent	EC
SLC45A2	3	NC_009164.2	E. caballus (4)	17203A>T	Coding	ILE249ILE(249ATA>ATT)	Silent	EC
SLC45A2	3	NC_009164.2	E. caballus (5)	17203A>T	Coding	ILE249ILE(249ATA>ATT)	Silent	EC
SLC45A2	3	NC_009164.2	E. caballus (Garrano)	17203A>T	Coding	ILE249ILE(249ATA>ATT)	Silent	EC
SLC45A2	3	NC_009164.2	E. caballus(Sorraia)	17203A>T	Coding	ILE249ILE(249ATA>ATT)	Silent	EC
SLC45A2	3	NC_009164.2	E. hemionus(WAI)	17203A>T	Coding	ILE249ILE(249ATA>ATT)	Silent	EH
SLC45A2	3	NC_009164.2	E. asinos (Miranda)	17311G>T	Coding	LEU285LEU(285CTG>CTT)	Silent	EA
SLC45A2	3	NC_009164.2	E. asinus (Oman)	17311G>T	Coding	LEU285LEU(285CTG>CTT)	Silent	EA
SLC45A2	3	NC_009164.2	E. africanus (HOI)	17328C>A	Coding	THR291LYS(291ACA>AAA)	Functional	HOI
SLC45A2	3	NC_009164.2	E. asinos (Miranda)	17328C>A	Coding	THR291LYS(291ACA>AAA)	Functional	EA
SLC45A2	3	NC_009164.2	E. asinus (Oman)	17328C>A	Coding	THR291LYS(291ACA>AAA)	Functional	EA
SLC45A2	3	NC_009164.2	E. asinus (Somalia)	17328C>A	Coding	THR291LYS(291ACA>AAA)	Functional	EA
SLC45A2	3	NC_009164.2	E. caballus (1)	17328C>A	Coding	THR291LYS(291ACA>AAA)	Functional	EC
SLC45A2	3	NC_009164.2	E. caballus (2)	17328C>A	Coding	THR291LYS(291ACA>AAA)	Functional	EC
SLC45A2	3	NC_009164.2	E. caballus (3)	17328C>A	Coding	THR291LYS(291ACA>AAA)	Functional	EC
SLC45A2	3	NC_009164.2	E. caballus (4)	17328C>A	Coding	THR291LYS(291ACA>AAA)	Functional	EC
SLC45A2	3	NC_009164.2	E. caballus (5)	17328C>A	Coding	THR291LYS(291ACA>AAA)	Functional	EC
SLC45A2	3	NC_009164.2	E. caballus (Garrano)	17328C>A	Coding	THR291LYS(291ACA>AAA)	Functional	EC
SLC45A2	3	NC_009164.2	E. caballus(Sorraia)	17328C>A	Coding	THR291LYS(291ACA>AAA)	Functional	EC
SLC45A2	3	NC_009164.2	E. grevyi(GZ)	17328C>A	Coding	THR291LYS(291ACA>AAA)	Functional	GZ
SLC45A2	3	NC_009164.2	E. hemionus(WAI)	17328C>A	Coding	THR291LYS(291ACA>AAA)	Functional	EH
SLC45A2	3	NC_009164.2	E.kiang(EK)	17328C>A	Coding	THR291LYS(291ACA>AAA)	Functional	EK
SLC45A2	3	NC_009164.2	E. caballus (1)	17338A>T	Coding	GLN294HIS(294CAA>CAT)	Functional	EC

GENE	Exon	REF.Sequence	Sequence name	Position	Region	Aminoacid change	Mutation	Code
SLC45A2	3	NC_009164.2	E. caballus (Garrano)	17338A>T	Coding	GLN294HIS(294CAA>CAT)	Functional	EC
SLC45A2	3	NC_009164.2	E. caballus(Sorraia)	17338A>T	Coding	GLN294HIS(294CAA>CAT)	Functional	EC
SLC45A2	3	NC_009164.2	E. africanus (HOI)	17346G<A	Non-coding	-	-	HOI
SLC45A2	3	NC_009164.2	E. asinos (Miranda)	17346G<A	Non-coding	-	-	EA
SLC45A2	3	NC_009164.2	E. asinus (Oman)	17346G<A	Non-coding	-	-	EA
SLC45A2	3	NC_009164.2	E. asinus (Somalia)	17346G<A	Non-coding	-	-	EA
SLC45A2	3	NC_009164.2	E. caballus (1)	17346G<A	Non-coding	-	-	EC
SLC45A2	3	NC_009164.2	E. caballus (2)	17346G<A	Non-coding	-	-	EC
SLC45A2	3	NC_009164.2	E. caballus (3)	17346G<A	Non-coding	-	-	EC
SLC45A2	3	NC_009164.2	E. caballus (4)	17346G<A	Non-coding	-	-	EC
SLC45A2	3	NC_009164.2	E. caballus (5)	17346G<A	Non-coding	-	-	EC
SLC45A2	3	NC_009164.2	E. caballus (Garrano)	17346G<A	Non-coding	-	-	EC
SLC45A2	3	NC_009164.2	E. caballus(Sorraia)	17346G<A	Non-coding	-	-	EC
SLC45A2	3	NC_009164.2	E. grevyi(GZ)	17346G<A	Non-coding	-	-	GZ
SLC45A2	3	NC_009164.2	E. hemionus(WAI)	17346G<A	Non-coding	-	-	EH
SLC45A2	3	NC_009164.2	E.kiang(EK)	17346G<A	Non-coding	-	-	EK
SLC45A2	4	NC_009164.2	E. hemionus(WAI)	21210G>T	Non-coding	-	-	EH
SLC45A2	4	NC_009164.2	E.caballus(EU272794.1)	21310G>A	Coding	ALA329THR(329GCC>ACC)	Functional	EC
SLC45A2	5	NC_009164.2	E. africanus (HOI)	22067C>G	Non-coding	-	-	HOI
SLC45A2	5	NC_009164.2	E. asinos (Miranda)	22067C>G	Non-coding	-	-	EA
SLC45A2	5	NC_009164.2	E. asinus (Oman)	22067C>G	Non-coding	-	-	EA
SLC45A2	5	NC_009164.2	E. asinus (Somalia)	22067C>G	Non-coding	-	-	EA
SLC45A2	5	NC_009164.2	E. hemionus(WAI)	22145C>T	Coding	ILE361ILE(361ATC>ATT)	Functional	EH
SLC45A2	6	NC_009164.2	E. africanus (HOI)	26083T>C	Coding	TYR393TYR(393TAT>TAC)	Silent	HOI
SLC45A2	6	NC_009164.2	E. asinus (Oman)	26083T>C	Coding	TYR393TYR(393TAT>TAC)	Silent	EA
SLC45A2	6	NC_009164.2	E. grevyi(GZ)	26083T>C	Coding	TYR393TYR(393TAT>TAC)	Silent	GZ
SLC45A2	6	NC_009164.2	E.kiang(EK)	26083T>C	Coding	TYR393TYR(393TAT>TAC)	Silent	EK
SLC45A2	6	NC_009164.2	E.kiang(EK)	26084G>A	Coding	VAL394ILE(394GTT>ATT)	Functional	EK
SLC45A2	6	NC_009164.2	E. africanus (HOI)	26113G>A	Coding	GLY403GLY(403GGG>GGA)	Silent	HOI
SLC45A2	6	NC_009164.2	E. africanus (HOI)	26194C>T	Coding	PHE430PHE(430TTC>TTT)	Silent	HOI
SLC45A2	6	NC_009164.2	E. asinus (Oman)	26194C>T	Coding	PHE430PHE(430TTC>TTT)	Silent	EA
SLC45A2	6	NC_009164.2	E. asinus (Somalia)	26194C>T	Coding	PHE430PHE(430TTC>TTT)	Silent	EA
SLC45A2	6	NC_009164.2	E. grevyi(GZ)	26194C>T	Coding	PHE430PHE(430TTC>TTT)	Silent	GZ
SLC45A2	6	NC_009164.2	E.kiang(EK)	26194C>T	Coding	PHE430PHE(430TTC>TTT)	Silent	EK
SLC45A2	7	NC_009164.2	E. africanus (HOI)	28629G>T	Coding	ARG458LEU(458CGG>CTG)	Functional	HOI
SLC45A2	7	NC_009164.2	E. asinos (Miranda)	28629G>T	Coding	ARG458LEU(458CGG>CTG)	Functional	EA
SLC45A2	7	NC_009164.2	E. asinus (Oman)	28629G>T	Coding	ARG458LEU(458CGG>CTG)	Functional	EA
SLC45A2	7	NC_009164.2	E. asinus (Somalia)	28629G>T	Coding	ARG458LEU(458CGG>CTG)	Functional	EA
SLC45A2	7	NC_009164.2	E. africanus (HOI)	28660T>C	Coding	SER468>SER(468AGT>AGC)	Functional	HOI
SLC45A2	7	NC_009164.2	E. asinos (Miranda)	28660T>C	Coding	SER468>SER(468AGT>AGC)	Functional	EA
SLC45A2	7	NC_009164.2	E. asinus (Oman)	28660T>C	Coding	SER468>SER(468AGT>AGC)	Functional	EA
SLC45A2	7	NC_009164.2	E. asinus (Somalia)	28660T>C	Coding	SER468>SER(468AGT>AGC)	Functional	EA
SLC45A2	7	NC_009164.2	E. grevyi(GZ)	28660T>C	Coding	SER468>SER(468AGT>AGC)	Functional	GZ
SLC45A2	7	NC_009164.2	E. hemionus(WAI)	28660T>C	Coding	SER468>SER(468AGT>AGC)	Functional	EH
SLC45A2	7	NC_009164.2	E. africanus (HOI)	28767T>C	Coding	VAL504ALA(504GTT>GCT)	Functional	HOI
SLC45A2	7	NC_009164.2	E. asinos (Miranda)	28767T>C	Coding	VAL504ALA(504GTT>GCT)	Functional	EA
SLC45A2	7	NC_009164.2	E. asinus (Oman)	28767T>C	Coding	VAL504ALA(504GTT>GCT)	Functional	EA
SLC45A2	7	NC_009164.2	E. asinus (Somalia)	28767T>C	Coding	VAL504ALA(504GTT>GCT)	Functional	EA
SLC45A2	7	NC_009164.2	E. hemionus(WAI)	28804G>C	Coding	LEU516LEU(516CTG>CTC)	Functional	EH

Table S4 -Polymorphic nucleotide positions observed among MC1R, SILV, ASIP, SLC36A1 and SLA45A2 genes. The different haplotypes are listed on the left column (H). NP – nucleotide position.

	ASIP	
	NP	
	163	168
H1	G	G
H2	A	A
H3	A	.

	MC1R															
	E1A		E1B								E1C			E1D	E1E	
	NP		NP								NP			NP	NP	
	93	111	243	247	248	250	303	326	355	409	497	557	577	629	822	922
H1	G	A	C	T	C	G	G	A	A	A	T	C	G	T	C	A
H2	A	G	.	.	T	A	.	.	.	.	C	.	.	C	A	.
H3	.	G	.	.	.	.	A	G	G	G	C	.	A		.	C
H4			T	.	.	.	A	G	G	G	C	T	A			
H5			.	T	.	.	.	.	.	.						
H6																

	SILV																																					
	E1	E2					E3						E5					E6A		E6B			E6C				E7			E8	E9		E10			E11		
	Np	NP					NP						NP					NP		NP			NP				NP			NP	NP		NP			NP		
	44	2041	2051	2061	2071	2112	3259	3268	3300	3349	3357	3377	3759	3761	3773	3842	3965	4319	4356	4418	4461	4622	4761	4812	4819	4929	6195	6256	6258	6544	6745	6763	7037	7080	7083	7701	7721	7752
H1	T	C	G	C	C	A	G	T	G	G	G	A	C	A	G	A	C	C	C	C	T	C	G	G	G	C	G	C	T	C	A	C	A	C	C	G	C	G
H2	C	.	A	.	T	.	.	A	C	.	.	.	.	.	.	G	.	T	.	T	.	.	C	A	A	T	.	T	.	T	G	.	G	.	.	G	.	.
H3		.	.	.	.	C	.	A	C	.	.	G	.	.	.	G	A	T	T	.	A	T					A	.	.		G	A	.	.	T	.	A	.
H4		.	.	.	.	.	A	A	C	.	.	.	A	C	C	.	.										.	.	A				G	A	.	.	.	A
H5		.	.	T	.	.	A	A	C	C	A	.																										
H6		A	.	.	.	.																																

	SLC36A1																																			
	E1					E2			E3				E4			E6					E7		E8				E9							E10			
	NP					NP			NP				NP			NP					NP		NP				NP							NP			
	60	66	115	120	1147	5614	5671	5676	6598	6712	6717	6719	7141	7170	7248	9750	9817	9840	9946	9978	13429	13477	16162	16253	16278	16284	19184	19226	19241	19284	19325	19361	19365	27901	27902	27916	27917
H1	C	C	C	G	A	A	T	C	A	A	A	C	A	C	G	C	G	T	A	G	G	C	G	T	G	G	C	C	A	A	C	G	C	T	G	T	G
H2	T	.	.	A	.	.	.	G	G	G	G	.	C	T	A	.	A	.	G	A	.	T	.	C	A	.	.	.	G	.	T	.	.	C	A	C	A
H3	T	T	.	A	.	.	C	.	G	G	G	A				.	A	C	G	A	A	T	.	C	.	.	.	A	.	.	.	.	.				
H4	.	.	A	.	C	G	C	.								T	.	.	.	A			A	.	.	.	T	.	G	G	T	.	.				
H5	.	.	A	.	C																		.	.	A	.	T	.	G	G	T	C	G				
H6																							.	.	.	A											

	SLC45A2																																				
	E1												E2			E3A			E3B				E4	E5		E6								E7			
	NP												NP			NP			NP				NP	NP		NP								NP			
	26	32	60	72	96	99	121	159	165	175	328	367	2253	2309	2396	17222	17358	17375	17311	17328	17338	17346	21310	22067	22145	26083	26084	26113	26117	26123	26124	26180	26194	28629	28660	28767	28804
H1	G	C	C	C	G	C	G	G	C	C	T	G	G	G	C	C	G	A	G	C	A	G	G	C	C	T	G	G	T	T	T	C	C	G	T	T	G
H2	.	A	.	.	.	.	.	A	.	.	C	C	A	.	.	.	C	T	T	A	.	A	A	G	.	C	.	.	.	.	.	.	T	T	C	C	.
H3	A	A	.	.	A	.	.	A	.	.	C	C	.	.	T	T	C	T	.	A	.	A		.	T	C	A	.	.	.	.	.	T	.	C	.	.
H4	.	.	.	.	.	.	.	.	.	T	.	.	.	A	.	.	.	T	.	A	T	A				C	.	A	.	.	.	.	T	.	C	.	C
H5	.	A	.	.	.	.	.	.	T	T	.	.														.	.	.	A	C	A	A	.				
H6	.	.	T	T	.	A	A	.	.	.	.	.																									
H7	.	A	.	.	.	.	.	.	.	.	.	.																									
H8	.	A	.	.	A	.	.	A	.	.	C	C																									

CHAPTER 4

General discussion and main findings

This thesis study was the first to estimate population genetic parameters from natural populations of Kiang.

Despite the previous assumptions that kiang populations were probably highly fragmented, our data on five populations did not found any evidence of population fragmentation. This conclusion is highly supported by the low levels of structuration obtained and by the remaining analysis that exhibited very low evidence of variation among populations, which may validate the occurrence of inter-population breeding. This merging occur, probably during Kiang seasonal dislocations between different regions (St-louis & Côte 2009). According to our results, and taking in to account the high distance between some of the populations (see Fig. 7), kiang might scroll very large distances, despite no migratory movements have been described for this species (Schaller 1998). The obtained results also shown that despite low structuration levels, some differentiation is observed between the studied populations, but curiously the most genetically divergent populations are not the most geographically apart, evidencing the possibility of the occurrence of topographic barriers that were not previously considered. It is very important to take into account that this work was based on noninvasive sampling and that despite all the associated handicaps and genotyping errors; this rested the only way to access this species at molecular level.

Nonetheless, it will be interesting that in future studies the number of samples, markers, populations (including those from India, Nepal and Pakistan) can be extended. Also a landscape genetics approach would allow the detection of possible cryptic boundaries that can eventually be breaking gene flow by constrain long distance migrations (Radespiel *et al.* 2008; Quéméré *et al.* 2010; Kaczensky *et al.* 2011).

In the second chapter of this thesis, we report some interesting, though preliminary, genetic insights that contribute to highlight the complexity among coat color genes function, and the role that the artificial and natural selection in the evolutionary history of this genes. Nonetheless, this was just the first step of a more extended study. We have characterized polymorphisms found in these five genes for all equid species. One of most interesting findings is the fact that despite few mutations have been described as responsible for the large number of coat color patterns, on this study we found that the same genotypes (mutations) are associated to completely different coat colors when compared between species.

Here we have also demonstrated that mutations on those genes have occurred at different times, and therefore can be used to assess the evolutionary history of the Equus genus. In

fact, coat color genes have been used to infer phylogenetic relationships in other species like in chamois (Pérez *et al.* 2013) and Lepus (Koutsogiannouli *et al.* 2012). This information could be explored on a future phylogenetic approach since on equids the only study that used coat color genes to infer phylogenetic relationships only resorted to three genes: KIT, MC1R and EDNRB (Steiner & Ryder 2011; Steiner *et al.* 2012). The coat color genes can, indeed, provide important information to resolve several doubts about phylogeny of genus *Equus*, which yet still remain to be answered (Steiner et al., 2012). The fact of the most part of equid living species presents some level of conservation concern, the resolution of phylogenetic relationships between species and sub-species is essential for define conservation units on the wild, and then contribute to design appropriate conservation plans (Krüger *et al.* 2005) highlighting the importance of incorporate both population and phylogenetic analysis.

Both works, despite their different purposes, contributed to increase the knowledge on the *E. kiang* in particular, and on the *Equus* genus in general. Firstly by assessing genetic diversity of a wild species and secondly by assessing the molecular diversity and phylogenetic relations between the different coat color polymorphisms in both wild and domestic species. In fact, the second work can also contribute for resolve the taxonomic and phylogenetic doubts surrounding the validity of the three *E.kiang* subspecies, since one of the distinctive traits between the purposed subspecies is the coat color variation (Groves and Mazák, 1967; Groves 1974).

REFERENCES

Albaugh GP, Iyengar V, Lohani a *et al.* (1992) Isolation of exfoliated colonic epithelial cells, a novel, non-invasive approach to the study of cellular markers. *Journal international du cancer*, **52**, 347–50.

Beaumont MA (1999) Detecting population expansion and decline using microsatellites. *Genetics*, **153**, 2013–29.

Beja-Pereira A, England PR, Ferrand N *et al.* (2004) African Origins of the Domestic Donkey. *Science*, **304**, 1781.

Beja-Pereira A, Oliveira R, Alves PC, Schwartz MK, Luikart G (2009) Advancing ecological understandings through technological transformations in noninvasive genetics. *Molecular ecology resources*, **9**, 1279–301.

Bertorelle G, Bruford MW, Hauffe HC, Rizzoli A, Vernesi C (2009) *Population Genetics for Animal Conservation*. Cambridge University Press,

Bhatnagar YV, Wangchuk R, Prins HHT, Van Wieren SE, Mishra C (2006) Perceived conflicts between pastoralism and conservation of the kiang Equus kiang in the Ladakh Trans-Himalaya, India. *Environmental management*, **38**, 934–41.

Broquet T, Berset-Braendli L, Emaresi G, Fumagalli L (2006) Buccal swabs allow efficient and reliable microsatellite genotyping in amphibians. *Conservation Genetics*, **8**, 509–511.

Broquet T, Ménard N, Petit E (2006) Noninvasive population genetics: a review of sample source, diet, fragment length and microsatellite motif effects on amplification success and genotyping error rates. *Conservation Genetics*, **8**, 249–260.

Broquet T, Petit E (2004) Quantifying genotyping errors in noninvasive population genetics. *Molecular ecology*, **13**, 3601–8.

Brunberg E, Andersson L, Cothran G *et al.* (2006) A missense mutation in PMEL17 is associated with the Silver coat color in the horse. *BMC genetics*, **7**, 46.

Calonje M, Martín-Bravo S, Dobeš C *et al.* (2008) Non-coding nuclear DNA markers in phylogenetic reconstruction. *Plant Systematics and Evolution*, **282**, 257–280.

Chikhi L, Bruford.M (2005) Mammalian Population Genetics and Genomics. In: *Mamalian Genetics* , pp. 539–584.

Ficetola GF, Miaud C, Pompanon F, Taberlet P (2008) Species detection using environmental DNA from water samples. *Biology letters*, **4**, 423–5.

Foggin M (2012) Pastoralists and wildlife conservation in western China: collaborative management within protected areas on the Tibetan Plateau. *Pastoralism: Research, Policy and Practice*, **2**, 17.

Fox JL (1991) THe Mountain Ungulates of Ladakh, India. *Biological Conservation*, **58**, 167–190.

Frankham R (2005) Genetics and extinction. *Biological Conservation*, **126**, 131–140.

Frankham R, Ballou.J.D, Briscoe.D.A (2002) *Introduction to Conservation Genetics*. Cambridge University Press.

Frantzen M a, Silk JB, Ferguson JW, Wayne RK, Kohn MH (1998) Empirical evaluation of preservation methods for faecal DNA. *Molecular ecology*, **7**, 1423–8.

Gagnon M, Chew AE (2000) Dietary Preferences in Extant African Bovidae. *Journal of Mammalogy*, **81**, 490–511.

Galan M, Cossons J., Aulagnier S *et al.* (2003) Cross-amplification tests of ungulate primers in roe deer (Capreolus capreolus) to develop a multiplex panel of 12 microsatellite loci. *Molecular Ecology Notes*.

Geigl EM, Grange T (2012) Eurasian wild asses in time and space: morphological versus genetic diversity. *Anatomischer Anzeiger*, **194**, 88–102.

Goodwin W, Linacre A, Hadi S (2007) *An Introduction to Forensic Genetics*. JohnWiley & Sons Ltd.

Goossens B, Chikhi L, Ancrenaz M *et al.* (2006) Genetic signature of anthropogenic population collapse in orang-utans. *PLoS biology*, **4**, e25.

Groves CP (1974) *Horses, asses, and zebras in the wild*. R. Curtis Books.

Groves, C.P. and Mazák, V. (1967) On some taxonomic problems of Asiatic wild asses with a description of a new subspecies (Perissodactyla, Equidae). *Zeitschrift für Säugetierkunde*, **32**, 321–355.

Hammond R, Saccheri I, Ciofi C*, et al.* (1998) Isolation of Microsatellite Markers in Animals. In: *Molecular Tools for Screening Biodiversity* (eds. Karp A, Isaac P, Ingram D), pp. 279-285. Springer Netherlands.

Hare MP (2001) Prospects for nuclear gene phylogeography. *Trends in Ecology & Evolution*, **16**, 700–706.

Harris RB, Miller D. (1995) Overlap in summer habitats and diets of Tibetan Plateau ungulates. *Mammalia*, **59**, 197–212.

Hasegawa M, Kishino H, Yano T (1989) Estimation of branching dates among primates by molecular clocks of nuclear DNA which slowed down in Hominoidea Individual level Species tree. *Journal of Human Evolution*, **18**, 461–476.

Hoss M, Kohn M, Paabo S (1992) Excrement analysisi by PCR. *Nature*, **359**.

Kaczensky P, Kuehn R, Lhagvasuren B *et al.* (2011) Connectivity of the Asiatic wild ass population in the Mongolian Gobi. *Biological Conservation*, **144**, 920–929.

Kattan G, Hernández OL, Goldstein I *et al.* (2004) Range fragmentation in the spectacled bear Tremarctos ornatus in the northern Andes. *Oryx*, **38**, 155–163.

Kawai K, Shimizu M, Hugues NR, Takenaka O (2004) A non-invasive technique for obtaining DNA from marine intertidal snails. *Journal of the Marine Biological Association of the United Kingdom*, **84**, 773–774.

Kohn M, Knauer F, Stoffella a, Schröder W, Pääbo S (1995) Conservation genetics of the European brown bear--a study using excremental PCR of nuclear and mitochondrial sequences. *Molecular ecology*, **4**, 95–103.

Kohn M, Wayne R (1997) Facts from feces revisited. *Trends in ecology & evolution*, **12**, 223 - 227.

Koutsogiannouli E a., Moutou K a., Stamatis C, Mamuris Z (2012) Analysis of MC1R genetic variation in Lepus species in Mediterranean refugia. *Mammalian Biology - Zeitschrift für Säugetierkunde*, **77**, 428–433.

Krüger K, Gaillard C, Stranzinger G, Rieder S (2005) Phylogenetic analysis and species allocation of individual equids using microsatellite data. *Zeitschrift für Tierzüchtung und Züchtungsbiologie*, **122**, 78–86.

Laird NM, Lange C (2011) *The Fundamentals of Modern Statistical Genetics*. Springer.

Larson G, Burger J (2013) A population genetics view of animal domestication. *Trends in genetics : TIG*, **29**, 197–205.

Lewontin RC (1985) Population genetics. *Annual review of genetics*, **19**, 81–102.

Luikart G, England PR, Tallmon D, Jordan S, Taberlet P (2003) The power and promise of population genomics: from genotyping to genome typing. *Nature reviews. Genetics*, **4**, 981–94.

Luikart G, Zundel S, Rioux D *et al.* (2008) Low Genotyping Error Rates and Noninvasive Sampling in Bighorn Sheep. *Journal of Wildlife Management*, **72**, 299–304.

Manel S, Schwartz MK, Luikart G, Taberlet P (2003) Landscape genetics: combining landscape ecology and population genetics. *Trends in Ecology & Evolution*, **18**, 189–197.

Mariat D, Taourit S, Guérin G (2003) A mutation in the MATP gene causes the cream coat colour in the horse. , **35**, *Genetics Selection Evolution*, **35**, 119–133.

Meeuwig MH, Guy CS, Kalinowski ST, Fredenberg W a (2010) Landscape influences on genetic differentiation among bull trout populations in a stream-lake network. *Molecular ecology*, **19**, 3620–3633.

Mills SL (2007) *Conservation of Wildlife Populations - Demography, Genetics, and Management*, Blackwall Publishing

Moehlman PD (2002) *Equids : Zebras , Asses and Horses*. IUCN, Gland, Switzerland and Cambridge, UK

Morin PA, Archer F, Pease V *et al.* (2012) Empirical comparison of single nucleotide polymorphisms and microsatellites for population and demographic analyses of bowhead whales. *Endangered Species Research*, **19**, 129–147.

Morin PA, Wallis J, Moore JJ, Woodruff DS (1994) Paternity exclusion in a community of wild chimpanzees using hypervariable simple sequence repeats. *Molecular ecology*, **3**, 469–77.

Morin PA, Woodruff DS (1992) Paternity Exclusion using Multiple Hypervariable Microsatellite Loci Amplified from Nuclear DNA of Hair Cells in: *Paternity in Primates: Genetic Tests and Theories.* Basel, Karger,Switzerland.

Morin PA, Woodruff. D.S (1996) Non-invasive genotyping for vertebrate conservation. In: *Molecular Genetic Approaches in Conservation.* (eds Smith TB, Wayne RK), pp. 298–313. Oxford University Press.

Mullis K, Faloona F, Scharf S *et al.* (1986) Specific enzymatic amplification of DNA in vitro: the polymerase chain reaction. *Cold Spring Harbour Symposia on Quantitative Biology*, **51**, 263-273.

Murphy MA, Waits LP, Kendall KC, Higbee JA, Express A (2002) An evaluation of long-term preservation methods for brown bear (*Ursus arctos*) faecal DNA samples.*Conservation Genetics*, **3**, 435-440

Oakenfull EA, Lim HN, Ryder OA (2000) A survey of equid mitochondrial DNA : Implications for the evolution , genetic diversity and conservation of Equus.*Conservation Genetics* ,**1**, 341–355.

Oakenfull EA, Ryder OA (2002) Genetics of Equid Species and Subspecies. In: *Equids: zebras, asses and horses. Status survey and conservation action plan* , pp. 108–112.

Pérez T, Essler S, Palacios B, Albornoz J, Domínguez A (2013) Evolution of the melanocortin-1 receptor gene (MC1R) in chamois (Rupicapra spp.). *Molecular phylogenetics and evolution*, **67**, 621–5.

Piggott MP, Taylor AC (2003a) Remote collection of animal DNA and its applications in conservation management and understanding the population biology of rare and cryptic species. *Wildlife Research*, **30**, 1.

Piggott MP, Taylor AC (2003b) Extensive evaluation of faecal preservation and DNA extraction methods in Australian native and introduced species. *Australian Journal of Zoology*, **51**, 341.

Pluzhnikov A, Di Rienzo A, Hudson RR (2002) Inferences about human demography based on multilocus analyses of noncoding sequences. *Genetics*, **161**, 1209–18.

Pompanon F, Bonin A, Bellemain E, Taberlet P (2005) Genotyping errors: causes, consequences and solutions. *Nature reviews. Genetics*, **6**, 847–59.

Pool JE, Hellmann I, Jensen JD, Nielsen R (2010) Population genetic inference from genomic sequence variation. *Genome research*, **20**, 291–300.

Quéméré E, Crouau-Roy B, Rabarivola C, Louis EE, Chikhi L (2010) Landscape genetics of an endangered lemur (*Propithecus tattersalli*) within its entire fragmented range. *Molecular ecology*, **19**, 1606–21.

Quéméré E, Louis EE, Ribéron A, Chikhi L, Crouau-Roy B (2009) Non-invasive conservation genetics of the critically endangered golden-crowned sifaka (*Propithecus tattersalli*): high diversity and significant genetic differentiation over a small range. *Conservation Genetics*, **11**, 675–687.

Radespiel U, Rakotondravony R, Chikhi L (2008) Natural and anthropogenic determinants of genetic structure in the largest remaining population of the endangered golden-brown mouse lemur, *Microcebus ravelobensis*. *American journal of primatology*, **70**, 860–70.

Reddy PA, Bhavanishankar M, Bhagavatula J *et al.* (2012) Improved methods of carnivore faecal sample preservation, DNA extraction and quantification for accurate genotyping of wild tigers. *PloS one*, **7**, e46732.

Rosenbom S, Costa V, Steck B et al., (2011) Cross-species genetic markers : a useful tool to study the world ' s most threatened wild equid — *Equus africanus*.*European Journal of Wildlife Research*

Rudnick JA, Katzner TE, Bragin EA, DeWoody JA (2007) A non-invasive genetic evaluation of population size, natal philopatry, and roosting behavior of non-breeding eastern imperial eagles (*Aquila heliaca*) in central Asia. *Conservation Genetics*, **9**, 667–676.

Ryder OA, Chemnick LG (1990) Chromosomal and molecular evolution in Asiatic wild asses. *Genetica*, **83**, 67–72.

Scandura M (2005) Individual sexing and genotyping from blood spots on the snow: A reliable source of DNA for non-invasive genetic surveys. *Conservation Genetics*, **6**, 871–874.

Schaller GB (1998) *Wildlife of the Tibetan steppe*. University of Chicago Press, Chicago.

Schlötterer C (2004) The evolution of molecular markers--just a matter of fashion? *Nature reviews. Genetics*, **5**, 63–9.

Selkoe KA, Toonen RJ (2006) Microsatellites for ecologists: a practical guide to using and evaluating microsatellite markers. *Ecology letters*, **9**, 615–29.

Shah N (2002) Status and action plan for the kiang (*Equus kiang*). In: *Equids: zebras, asses and horses. Status survey and conservation action plan* (ed P. D. Moehlman), pp. 72–81. International Union for Conservation of Nature and Natural Resources/Species Survival Commission, Equid Specialist Group, Gland, Switzerland.

Shah, N. V. and Huibin, Z. (2000) Status of kiang (*Equus kiang*) in East Arjin Mountain Nature Reserve, Xinjiang Province, People's Republic of China. Report (Dec 1999-Jan 2000*). China Exploration and Research Society*, Hong Kong, China.

Sharma BD, Clevers J, Graaf R De, Chapagain NR (2004) Mapping Equus kiang (Tibetan Wild Ass) Habitat in Surkhang, Upper Mustang, Nepal. *Mountain Research and Development*, **24**, 149–156.

Sheehy DP, Miller D, Johnson DA (2006) Transformation of traditional pastoral livestock systems on the Tibetan steppe. , **17**, 142–151.

Soto-Calderón ID, Ntie S, Mickala P *et al.* (2009) Effects of storage type and time on DNA amplification success in tropical ungulate faeces. *Molecular ecology resources*, **9**, 471–9.

Steiner CC, Mitelberg A, Tursi R, Ryder O a (2012) Molecular phylogeny of extant equids and effects of ancestral polymorphism in resolving species-level phylogenies. *Molecular phylogenetics and evolution*, **65**, 573–81.

Steiner CC, Ryder O a. (2011) Molecular phylogeny and evolution of the Perissodactyla. *Zoological Journal of the Linnean Society*, **163**, 1289–1303.

St-louis A, Côte SD (2009) *Equus kiang* (Perissodactyla: Equidae). *American Society of Mammalogists*, **835**, 1–11.

Sunnucks P (2000) Efficient genetic markers for population biology. , **15**, 199–203.

Taberlet P, Bouvet J (1992) Bear conservation genetics. *Nature*, **358**.

Taberlet P, Camarra JJ, Griffin S *et al.* (1997) Noninvasive genetic tracking of the endangered Pyrenean brown bear population. *Molecular ecology*, **6**, 869–76.

Taberlet P, Griffin S, Goossens B *et al.* (1996) Reliable genotyping of samples with very low DNA quantities using PCR. *Nucleic acids research*, **24**, 3189–94.

Taberlet P, Luikart G (1999) Non-invasive genetic sampling and individual identification. *Biological Journal of the Linnean Society*, **68**, 41–55.

Taberlet P, Waits L, Luikart G (1999) Noninvasive genetic sampling: look before you leap. *Trends in ecology & evolution*, **14**, 323–327.

Valiere N, Taberlet P (2000) Urine collected in the field as a source of DNA for species and individual identificatio. *Molecular Ecology*, 2149–2154.

Vial L, Maudet C, Luikart G (2003) Thirty-four polymorphic microsatellites for European roe deer. *Molecular Ecology Notes*, **3**, 523–527.

Waits LP, Paetkau D (2005) Noninvasive Genetic sampling tools for wildlife biologists: A review of applications and recommendations for accurate data collection. *Journal of Wildlife Management*, **69**, 1419–1433.

Wang Y (2003) *A Complete Checklist of Mammal Species and Subspecies in China: A Taxonomic and Geographic Reference*. China Forestry Publishing House, Beijing.

Wasser SK, Houston CS, Koehler GM, Cadd GG, Fain SR (1997) Techniques for application of faecal DNA methods to field studies of Ursids. *Molecular ecology*, **6**, 1091–7.

Wilcox BA, Murphy DD (1985) The University of Chicago Conservation Strategy : The Effects of Fragmentation on Extinction. , **125**, 879–887.

Xiaofeng L, Yi Q, Diqiang L *et al.* (2011) Habitat evaluation of wild Amur tiger (Panthera tigris altaica) and conservation priority setting in north-eastern China. *Journal of Environmental Management*, **92**, 31–42.

Zhang D-X, Hewitt GM (2003) Nuclear DNA analyses in genetic studies of populations: practice, problems and prospects. *Molecular ecology*, **12**, 563–58