Evolution and Biogeography of Canids (*Canis* and *Vulpes*) in North-West Africa

João Vasco Leite
Biodiversidade, Genética e Evolução
Biologia
2012

Orientador
Raquel Godinho, Post-Doc, FCUP/CIBIO

Coorientadores
José Carlos Brito, Senior Scientist, Assoc. Researcher, FCUP/CIBIO
Guillermo Velo-Antón, Post-doc, CIBIO

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

O Presidente do Júri,

Porto, _____/_____/_____

Agradecimentos

Reservo este espaço a todos aqueles que de uma maneira ou de outra, contribuíram para a elaboração desta dissertação. A todos, os meus sinceros agradecimentos.

Em primeiro lugar agradeço à Doutora Raquel Godinho e Doutores José Carlos Brito e Guillermo Velo-Antón, meus orientadores, por terem concedido a oportunidade de desenvolver um tema tão cativante para a minha tese, mais ainda por envolver animais tão fascinantes como os canídeos. Superaram largamente as minhas expectativas e estou muito grato pela competência científica, a dedicação e disponibilidade demonstradas, para além de todo um inestimável espírito de entreajuda e amizade que esteve sempre presente durante este trabalho. Estendo os agradecimentos também ao Doutor Francisco Álvares por todo o acompanhamento e pelo interesse e entusiasmo que sempre manifestou durante a realização desta tese. Considero um privilégio único ter sido orientado por estes Professores.

Todo o trabalho laboratorial durante este último ano não teria sido possível sem o auxílio incondicional e generoso da Susana Lopes, a quem devo muito e deixo estas sinceras palavras de apreço. Não poderia também deixar de agradecer a toda a equipa do CTM, que se revelou um excelente espaço de trabalho e convívio, com pessoas únicas e sempre dispostas a ajudar.

Agradeço igualmente a todas as pessoas que contribuíram com amostras para este estudo, mas em especial a todos os membros do grupo de investigação Biodeserts, que constituíram uma plataforma de enriquecimento científico e de debate, sempre numa toada descontraída.

Sem dúvida que a elaboração desta tese teria sido bem mais difícil sem o apoio e companheirismo de todos os meus colegas de mestrado, em especial da Vanessa, com os quais foi possível uma discussão fértil, profícua e construtiva de ideias, abordagens e metodologias. O constante encorajamento e momentos de descontracção proporcionados pelos meus amigos foram também fundamentais para manter a motivação e empenho. Obrigado a todos – sim, vocês sabem quem são…

Finalmente, os devidos e mais sentidos agradecimentos à minha família, em particular aos meus pais e irmão, pelo constante esforço em proporcionar as melhores condições para o meu sucesso.

Last but not the least… dado o âmbito desta dissertação, também queria “agradecer” ao bom e fiel canídeo que sempre me fez companhia, e que se encontra exactamente a meu lado enquanto escrevo estas palavras.

Sumário

O Norte de África, sobretudo as zonas áridas do Saara-Sahel, possuí características biogeográficas únicas que têm vindo a ser associadas a importantes padrões filogeográficos. No entanto, a biodiversidade é no geral pobremente avaliada, assim como as forças evolutivas responsáveis por essa mesma diversidade. Espécies não-voadoras e com elevada capacidade de dispersão, como os canídeos (Mammalia; Carnivora) encontram-se particularmente pouco estudadas. O Norte de África faz parte da área de distribuição de cinco espécies de canídeos para as quais existe uma escassez de dados genéticos detalhados: o chacal dourado *Canis aureus*, e as raposas *Vulpes pallida* (endémica do Sahel), *V. rueppellii*, *V. vulpes* e *V. zerda*.

O principal objectivo deste estudo foi levar a cabo uma avaliação genética de canídeos do Norte de África, recorrendo a dois tipos de marcadores moleculares – sequências de ADN mitocondrial e um conjunto de 46 microssatélites autossómicos. Dados moleculares dos dois tipos de marcadores foram obtidos para 14 *C. aureus*, 15 *V. pallida*, sete *V. rueppellii*, 25 *V. vulpes* e sete *V. zerda*. Árvores filogenéticas, redes de haplótipos e índices de diversidade foram calculados com base em sequências de ADN mitocondrial. Foi avaliada a estruturação populacional e hibridização entre espécies simpátricas, bem como outros índices de diversidade e distâncias genéticas entre indivíduos e grupos, a partir de genótipos *multilocus* individuais.

Os resultados sugerem: 1) que as populações europeias e africanas de *C. aureus* são divergentes e cada uma monofilética; 2) a possível ocorrência no passado de um evento de introgressão mitocondrial de lobo em populações de *C. aureus* norte-africanas; 3) duas linhagens distintas dentro de *V. vulpes*, uma das quais pode representar um primeiro e mais antigo evento de colonização, presente no Norte de África e no Japão, a partir do qual *V. rueppellii* terá possivelmente divergido para uma espécie melhor adaptada a ambientes áridos, e outra linhagem que inclui espécimes da Europa e da Ásia (incluindo novamente o Japão); 4) níveis significativos de variabilidade molecular em *V. pallida*, que constituem os primeiros registos de dados moleculares alguma vez produzidos para esta espécie; e 5) reduzida evidência de estruturação biogeográfica da diversidade genética no Norte de África (excepto para *V. vulpes* no Magrebe), relacionada provavelmente com a elevada vagilidade dos canídeos, que permite a ocorrência de fluxo génico mesmo entre áreas relativamente distantes.

Este estudo realça padrões biogeográficos para espécies de elevada mobilidade e dispersão no Norte de África, que devem ser tidos em consideração em futuras

medidas de conservação dirigidas tanto a essas mesmas espécies como aos seus habitats.

Palavras-chave: *Canis aureus*, evolução, introgressão, microssatélites, mtDNA, Norte de África, filogenia, Saara, Sahel, sistemática, *Vulpes pallida*, *Vulpes rueppellii*, *Vulpes vulpes*, *Vulpes zerda*

Abstract

North Africa, and especially the arid areas of the Sahara-Sahel, is characterised by unique biogeographical features that have been associated to major phylogeographic patterns. Nevertheless, biodiversity is generally poorly assessed as well as the evolutionary drivers regulating such diversity. High dispersal and non-volant species, such as canids (Mammalia; Carnivora) are particularly understudied. North Africa encompasses the range of five canid species that lack detailed genetic assessments: the golden jackal *Canis aureus*, and the foxes *Vulpes pallida* (Sahelian endemic), *V. rueppellii*, *V. vulpes* and *V. zerda*.

The main objective of this study was to conduct a genetic assessment of North African canids, using two types of molecular markers - mitochondrial DNA sequences and a set of 46 autosomal microsatellite *loci*. Molecular data for the two types of markers was obtained for 14 *C. aureus*, 15 *V. pallida*, seven *V. rueppellii*, 25 *V. vulpes* and seven *V. zerda*. Phylogenetic trees, networks and diversity indexes were estimated based on mitochondrial DNA sequences. Population structuring and hybridization between sympatric species was assessed, as well as further diversity indexes and genetic distances among individuals and groups, from individual *multilocus* genotypes.

Results suggest: 1) European and African populations of *C. aureus* are divergent and each monophyletic; 2) a possible past introgression of wolf mitochondria into North African *C. aureus* populations; 3) two distinct lineages within *V. vulpes*, one that may account for an early colonization stage that is present in North Africa and Japan, from which *V. rueppellii* possibly diverged into a more arid adapted species, and another lineage that includes specimens from Europe and Asia (including again Japan); 4) significant levels of molecular variability for *V. pallida*, which constitute the first ever molecular data on the species; and 5) little evidence for clear biogeographical structuring of genetic diversity within North Africa (except for *V. vulpes* in the Maghreb), probably related to the high vagility of canid species that allows considerable gene flow between relatively distant areas.

This study enlightens biogeographic patterns for widely ranged and mobile species in North Africa that should be taken into account in future conservation efforts directed to both species and habitats.

Keywords: *Canis aureus*, evolution, introgression, microsatellites, mtDNA, North Africa, phylogeny, Sahara, Sahel, systematics, *Vulpes pallida*, *Vulpes rueppellii*, *Vulpes vulpes*, *Vulpes zerda*

Index

Figure Index

Table Index

1. Introduction

1.1. Biodiversity loss

The perception of biodiversity - the variety of genes, species, and ecosystems that constitute life on Earth (Rands *et al.* 2010) - has suffered tremendous changes throughout the years, with the amount of available knowledge, and awareness given to it by policy makers and the general public, never being so plentiful. In the past ten years, we have seen world leaders committing to significantly reduce the rates of biodiversity loss by 2010 through a series of measures. These include the Convention on Biological Diversity (CBD), the incorporation of this same target into the United Nations Millennium Development Goals (Butchart *et al.* 2010), and reaching a peak in the aforementioned year, considered the International Year of Biodiversity by the United Nations (UN), and building up on that the UN declaring 2011-2020 the United Nations Decade on Biodiversity (www.cbd.int/). Additionally, there has been a constant growing number of national, regional and local scale conservation societies and communities (Rands *et al.* 2010). In spite of this increased notoriety, we are still observing an unprecedented rate of biodiversity loss. The steady overall decline of wild species populations' size, range, condition, and connectivity to other populations and/or patches of suitable habitat (Butchart *et al.* 2010; Rands *et al.* 2010), contrasts with the intensifying pressures, such as human induced climate change, resource overexploitation, land-use transformation, pollution or exotic species introduction (Dirzo & Raven 2003; Díaz *et al.* 2006; Butchart *et al.* 2010).

Good examples exist at the local scale and/or aiming specific taxa, demonstrating that the rate of biodiversity loss can be reduced and even reversed when there is political will and correct application of the available resources (Butchart *et al.* 2010). One of those resources is knowledge itself. For both researchers and policy makers, the available biodiversity data continues to be, in the grand majority of cases, very insufficient and inaccurate (Botkin *et al.* 2007). Even for the better described taxonomic groups, there is still a lack of basic information on patterns of distribution, abundance and diversity, such as genetic or morphological variation (Botkin *et al.* 2007). Additionally, the existing data is very concentrated and biased, in the sense that is correlated to political geographical units (states), each of which having put forth, throughout history, different scales of effort when assessing biodiversity (Whittaker *et*

al. 2005). Hence, our overall knowledge on the world's biodiversity is still sadly inadequate and incomplete (Whittaker *et al.* 2005).

So drawing from the increased notoriety, we must halt this biodiversity crisis and to do so it is essential to increase scientific research to fill the knowledge gaps, mainly in developing regions of the world, where biodiversity is facing the greatest threats and there is limited capacity to respond (Rands *et al.* 2010).

1.2. Lack of knowledge in undeveloped regions and arid environments

In order to slow down the current rate of biodiversity loss we need in first-hand to know the more accurately and precisely possible: (1) what does in fact exist out there; (2) how to qualify and quantify it; (3) how each aspect of biological diversity is distributed in space and time; (4) how each biological entities interact with one and another (5) and how its future survival can be assured with minimal human influence (Dirzo & Raven 2003; Díaz *et al.* 2006).

While in developed regions of the world this type of knowledge is available at distinct degrees and conservation programs are operative, the same cannot be said of less favoured regions. In the majority of times, such regions hold very important and unique biodiversity hotspots, and basic information on species distribution and abundance is either lacking or very incomplete (Botkin *et al.* 2007). The situation becomes even bleaker when we look into the arid regions of the globe, especially in desert ecosystems, which in comparison with other biomes have been somewhat neglected in terms of scientific research and conservation effort (Durant *et al.* 2012).

Deserts and arid regions are commonly associated to bare and monotonous landscapes, wastelands possessing relatively fewer species than the better-watered biomes, that attract little attention by both the scientific and non-scientific communities (Ward 2009; Durant *et al.* 2012). Yet, within these "wastelands" many kinds of endangered micro-hotspots of biodiversity exist (Trape 2009; Wilson & Pitts 2012; Migliore *et al.* 2012; Murphy *et al.* 2012), and for each biological entity lost from an arid region, the percentage of loss for the region's biodiversity is much higher than in more species-rich biomes, thus making arid regions higher priority targets of conservation (McNeely 2003). Deserts constitute unique opportunities to study some of the most extreme adaptive outcomes in order for fauna and flora to subsist in such harsh climatic conditions (Ward 2009). Deserts and arid regions are also predicted with the

world's fastest velocities of climate change (Loarie *et al.* 2009); the development of conservation strategies based on biodiversity knowledge is thus urgently needed.

Historically, deserts and arid regions have been especially subjected to harsh climatic cycles, such as prolonged droughts, that affect the productivity and availability of resources for human populations, and more than often are responsible for hunger, disease, poverty and conflict (McNeely 2003; Thornton & Owiyo 2008). One of the tragedies of this situation is that the biological resources which could help make these lands productive often are abused and have been seriously reduced, whilst efforts at development, such as road building, borehole drilling, and food importation for people and livestock, have typically caused more overexploitation rather than less (McNeely 2003). North Africa and the Sahara-Sahel regions arise as clear examples of all mentioned above.

1.3. North-West Africa – a natural biodiversity laboratory

North Africa is, for the most part, associated with the Sahara desert (Ward 2009). Due to its remoteness and frequent political instability, scientific exploration encounters great difficulties in the region - access to key biogeographical areas, for instance mountain systems that may hold endemic/relict biological entities, are often denied by further local conflicts (McNeely 2003; Ward 2009; Olivier & Denis 2010; Ewi 2010). As a consequence, North African biodiversity is still poorly sampled and understood, resulting in knowledge gaps in species ranges, ecological requirements, and genetic diversity (Sillero-Zubiri *et al.* 2004; Brito *et al.* 2009; Durant *et al.* 2012).

In particular, North-West Africa is an incredible region to conduct biological research for various reasons. First of all, it constitutes a biogeographic crossroad between the Palearctic and Afro-tropical regions (Olson *et al.* 2001), encompassing three main biomes that are constantly influenced by the Atlantic Ocean along the West coast of Africa. Starting from north, we have the relatively more humid Mediterranean climate; as we begin to move south we see a significant change to the extreme arid settings of the Sahara, and at the end we see a slight increase in the availability of water as we reach the arid Sahel (Le Houérou 1992, 1997; Olson *et al.* 2001).

A second reason comes from North-West Africa's relative proximity to Europe and the fact that it belongs to the Mediterranean region, one of the world's most important biodiversity hotspots (Myers *et al.* 2000), where significant fauna and flora exchange has been occurring between the two regions, greatly influencing each region's biota

and phylogeographic patterns (Busack 1986; Le Houérou 1992, 1997; Pleguezuelos *et al.* 2008).

A third reason is related to North-West Africa's own set of great scale geological events, such as the uplift of the Atlas mountain chain, sea transgressions and the Messinian salinity crisis (Le Houérou 1992, 1997), as well as strong past climatic shifts, mainly a period of hyper-aridity during the Last Glacial Maximum (LGM) and a humid period during the Holocene (Kröpelin *et al.* 2008; Drake *et al.* 2011). The combined effects of such events provided the region with astonishing geographical features, from high mountains, to wide river systems and seas of sand dunes, which have been responsible for shaping current species ranges and for driving the creation of biological diversity (Le Houérou 1997).

Additionally the fast growing human population requires increasingly more resources, and reconciling the fulfilment of basic human needs with the conservation of biodiversity is now one of the major challenges for Africa (Newby 2007). Furthermore, future climate predictions indicate a constant rise of temperature and a decrease of precipitation in the region, which will undoubtedly originate new biogeographical shifts and ultimately extinction of populations and species (Foley *et al.* 2003; Loarie *et al.* 2009). For all these reasons, North Africa has become an outstanding region to study biogeographic patterns in the distribution of biodiversity and essential to assure its preservation (Durant *et al.* 2012).

1.3.1. The Mediterranean basin

The Mediterranean Sea together with the European Southern Peninsulas, North Africa and the Levant region is located approximately between latitudes 30 ºN and 40ºN, and extends 2,500,000 km² covering three continents (Africa, Asia and Europe) (Hewitt 2004; Temple & Cuttelod 2011). Many forces have helped shape the current biogeographic patterns and make the Mediterranean one of the world's biodiversity hotspots (Myers *et al.* 2000; Migliore *et al.* 2012). It is a region where fauna and flora of North European, Saharan, Afro-Tropical and Irano-Turanian origins meet, besides holding an impressive set of endemic species (Weiss & Ferrand 2007; Migliore *et al.* 2012). Complex tectonic and geographic shifts in this area, predominantly during the Messinian salinity crisis and around the Strait of Gibraltar during the past 14 million years (De Jong 1998; Govers 2009) have, at different levels through time, allowed expansions of European-originated species into Africa and vice-versa (Le Houérou 1992, 1997), has we can still presently observe in today's Morocco and Iberian

Peninsula's fauna and flora, respectively (Bons *et al.* 1996; Cosson *et al.* 2005; Sindaco & Jeremčenko 2008; Gaubert *et al.* 2011).

During the Quaternary Ice Ages, the Southern European Peninsulas of Iberia, Italy and the Balkans have acted as refugia for more warmer/temperate climate species during times of ice sheet expansion (Hewitt 2004; Weiss & Ferrand 2007). All of this together with a historically intense human presence that has been responsible for creating unique combinations of natural and cultural habitat patches, but also at the same time endangering many species (Temple & Cuttelod 2011), has made the Mediterranean basin one of the most interesting region to study biodiversity patterns.

1.3.2. The Sahara-Sahel regions

As the main biogeographic feature of all of Africa, a special focus must be given to the Sahara desert, the largest warm desert in the world, occupying around 8,000,000 km^2 between latitudes 16 and 32°N, circumscribed within the 100 ± 50 mm isohyets of mean annual rainfall (Le Houérou 1997). Within such an enormous area we find a diverse set of geographic elements, from sand dunes, salt pans, rocky plains and grasslands, to high elevation mountain chains, scattered oasis and rock pools (locally known as *gueltas*) (Le Houérou 1997; Anthelme *et al.* 2008). An important biogeographic zone is also the neighbouring Sahel region - a narrow band of semi-arid savannah and shrubland that stretches from the West African Atlantic coast to the Red Sea shores, acting as a transition zone between the Sahara desert and Afro-Tropical savannah (Le Houérou 1992, 1997; Olson *et al.* 2001). Although the timing of Sahara onset is relatively uncertain, there is evidence to place the start of the desertification around 7 Mya during the late Miocene (Schuster *et al.* 2006), or even more recently, at around 6-2.5 Mya in western areas (Swezey 2009). But the Sahara and the Sahel have experienced many range constrictions and expansions (Figure 1), shifting respectively between more humid and more arid phases (Rognon 1993; Foley *et al.* 2003; Kröpelin *et al.* 2008; Drake *et al.* 2011). Such shifts occurred in response to climatic oscillations caused by fluctuations in the Earth's orbital processes, sea surface temperature, and feedback mechanisms between rainfall reduction and vegetation cover (Wang *et al.* 2007; Claussen 2009). The frequency at which these Saharan range cycles occurred are estimated to have ranged from approximately 100,000–20,000 years, around eight to ten major humid phases just in the last 3 My (Le Houérou 1992; Rognon 1993). One of the most interesting biogeographical questions partially unanswered is to which degree has the Sahara desert functioned has a barrier between areas of higher and

lower latitude, being responsible for the emergence of genetically and morphologically divergent taxa (Douady *et al.* 2003; Geraads 2010). The Sahara currently constitutes a strong barrier to dispersal for organisms that are not adapted to live in the most arid parts of the deserts (Geraads 2010). But during the "Green-Sahara", many regions that are now arid would harbour large lakes and river courses, surrounded by steppe, savannah or even forest environments, allowing the dispersal of more water-dependent species and genetic exchange between populations from the South to the North and vice-versa, that were previously separated by a wide gap of desert (Douady *et al.* 2003; Drake *et al.* 2008, 2011). Indeed, this climatic instability led to periodic modifications of North African biota. The basic premise is that shifts in climate alter the ecological composition of the landscapes, which in turn impose new selective pressures and/or geographical isolation leading to genetic diversification, adaptation and ultimately speciation (Mouline *et al.* 2008).

Although not many phylogeographic studies have been conducted within the region, it is starting to become apparent that the Sahara throughout its history has promoted

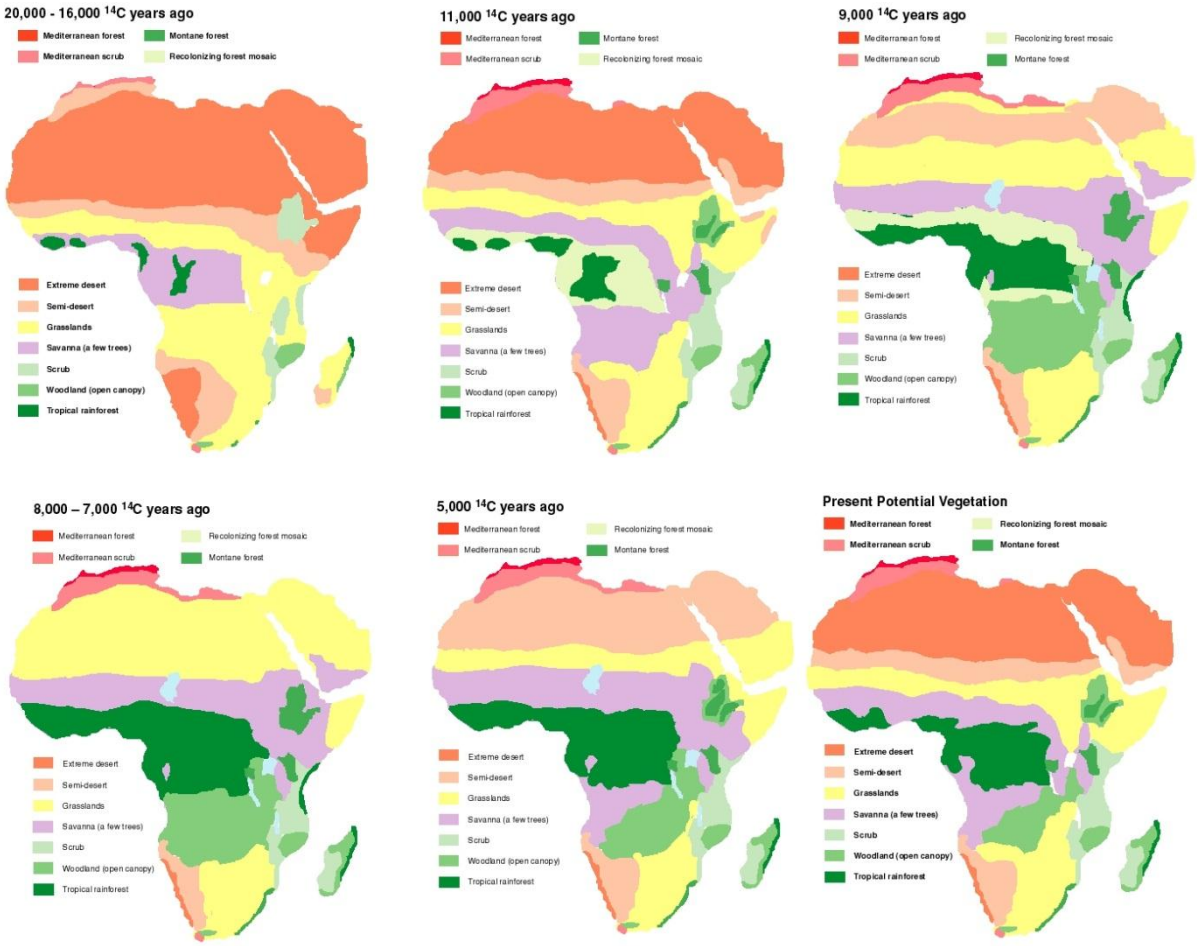


Figure 1 - Temporal changes in the distribution of African ecosystems since the Last Glacial Maximum until present-day [adapted from Adams & Faure (2004)].

taxa diversification through vicariance events since the onset of the arid conditions (Douady *et al.* 2003; Carranza *et al.* 2008) and during the successive range shifts that followed: (1) phases of greater aridity and expansion of the desert, when concerning more water-dependent species (Le Houérou 1997; Dobigny *et al.* 2005; Guillaumet *et al.* 2006, 2008; Mouline *et al.* 2008; Nicolas *et al.* 2009); (2) humid and desert constriction phases for more arid adapted species (Le Houérou 1997; Carranza *et al.* 2008; Boratynski *et al.* 2012). Furthermore, species that exhibit continuous distributions North and/or South of the Sahara are often found in isolated areas on the opposite edge of the desert, or even in oasis and Saharan mountain ranges that display suitable habitat patches and are able to host relict populations of more water dependent species from both Mediterranean or Afro-Tropical origins (Messerli & Winiger 1992; Le Houérou 1997; Anthelme *et al.* 2008; Shaibi & Moritz 2010; Brito *et al.* 2011; Migliore *et al.* 2012). Concerning specifically the Sahel region, the situation in terms of published research is even bleaker. The few research available as suggested Lake Chad, the Niger River and other Sahelian river courses as linked to population and even species-level diversification events (Flagstad *et al.* 2001; Alpers *et al.* 2004; Dobigny *et al.* 2005; Brouat *et al.* 2009).

The focus of the great majority of studies in North-West Africa has been on small, non-volant vertebrates with relative low dispersal capacity, for which barrier effects may be more pronounced (Dobigny *et al.* 2005; Mouline *et al.* 2008; Carranza *et al.* 2008; Boratynski *et al.* 2012). As a result, there are still considerable knowledge gaps regarding highly-mobile meso/macro fauna, but see Lerp *et al.* (2011); Godinho *et al.* (2012) as an example for ungulates and Gaubert *et al.* (2012) for canids.

Given this context, there is a clear need to improve and expand our knowledge on all levels of biodiversity and its distribution in North Africa.

1.4. Assessing biodiversity

Biological diversity and its distribution over wide regions of the globe are often measured solely as species diversity (numbers, proportions and distinctness). However, genetic studies, especially phylogenetic and phylogeographic research, have brought new insights, supporting and demonstrating the existence of geographically separated subspecies, divergent populations and cryptic lineages that are essential to preserve (Hewitt 2004).

1.4.1 Phylogenetic and phylogeographic research

Phylogenetics can be simply described as the study of evolutionary relationships between two or more biological entities (can be molecules, breeds, species, families and so on), that are normally depicted branching out within a tree-like diagram where closest related entities share a more recent common node, constituting a clade (Brinkman & Leipe 2001). But the reasons behind these branching patterns are often much more complicated to discern. So, the biogeography of a group and the bases of its biological diversity at any level are essential in the reconstruction of phylogenies and the understanding of the dynamic patterns of evolution (Avise 2000). With that in mind, the term phylogeography as risen in the last two decades and can be described as the "phylogenetic analysis of geographically contextualized genetic data for testing hypotheses regarding the causal relationship among geographic phenomena, species distributions, and the mechanisms driving speciation" (Hickerson *et al.* 2010). Nowadays, the term phylogeography implies a wide spectrum of methodologies and techniques, which combine phylogenies and summary statistics, providing high statistical support which in turn allows the formulation of models and *a priori* hypothesis testing (Hickerson *et al.* 2010). So this recent branch of biological knowledge has contributed considerably to the estimation of demographic histories, divergence times, migration rates, historical hybridization events, hybrid zones, introgression occurrences, among many other aspects (Hickerson *et al.* 2010), thus being of key importance to our understanding of biodiversity.

1.4.2. Molecular tools

The growing relevance of phylogeography, and overall genetic studies, is due largely to a continuous methodological development, whether thanks to the invention of the polymerase chain reaction (PCR), the appearance of a variety of molecular markers (Zhang & Hewitt 2003) and/or the increasing computational power accompanied by ever more sophisticated software (Hewitt 2004).

The most important data source has been without a doubt the genealogical research based on mitochondrial DNA (mtDNA) (Hickerson *et al.* 2010). Animal mitochondrial genome consists of a ~16000 base pair (bp) circular chromosome, containing about 37 genes and a non-coding control region (Wan *et al.* 2004). Each gene and the control region have different evolutionary rates, so depending on the hypothesis that is being tested, phylogenetic trees of different segments allow drawing conclusions at different

time scales and assess different taxonomic levels (Wan *et al.* 2004). The advantages of using mtDNA include its lack of recombination, putative neutrality, and smaller effective population size (Hickerson *et al.* 2010). Also, segments of mtDNA can be treated phylogenetically as taxonomic units, allowing researchers to directly link geographic patterns displayed by a certain gene's genealogy with a population's or species' own genealogy (Hickerson *et al.* 2010). Still, some precautions have to be taken into account and since evolutionary rates are not homogeneous throughout the mitochondrial genome, mtDNA sequences should be at least two coding genes in size to infer concordant species tree (Wan *et al.* 2004). These advantages come, in part, from the fact that in animals mtDNA is maternally inherited (Zhang & Hewitt 2003). However, this means that we are looking into a single *locus*, which gives us only one side of the phylogenetic history that may not be representative of the rest of the genome (Zhang & Hewitt 2003). Moreover, this molecular marker is not the most appropriate when assessing more recent, population and individual level, phylogenetic events (Wan *et al.* 2004). Conceptually, using mtDNA data is more valid in more distant speciation events or simply when evaluating an overall genetic diversity of a grouping (Hickerson *et al.* 2010).

For more recent phylogeographic events at the intraspecific level, such as population structuring and hybridization, microsatellites have a greater utility (Zhang & Hewitt 2003). These molecular markers consist of tandem repeated short DNA sequences of one to six nucleotides (dinucleotide motifs being the more commonly used) and are widely spread throughout the nuclear genome of eukaryotes (Wan *et al.* 2004). Microsatellites follow Mendelian co-dominant inheritance, are individually bi-allelic and multi-allelic in a population, and have high mutational rates (Wan *et al.* 2004). Such high levels of polymorphism allow researchers with even a small set of markers to assign unique individual *multilocus* genotypes (combination of alleles from different *loci* belonging to one organism) (Wan *et al.* 2004). Microsatellites are commonly applied in the assessment of genetic diversity and structure, gene flow between different populations, or species in cases of hybridization, and recent population history (Zhang & Hewitt 2003). However, these molecular markers are not exempted from shortcomings. They cannot be used to infer evolutionary relationships, since allele size and variation are not correlated with divergence (Zhang & Hewitt 2003). The statistical confidence is related to sample size, the number of *loci* available for research, and the respective levels of polymorphism of each *locus* (most studies so far do not surpass 20 *loci*, at times less than 10) (Wan *et al.* 2004). A challenging issue is related with the entire process of choosing and designing a microsatellite *locus* data set, which for most

species must be created almost from scratch since noncoding regions where microsatellites are normally found have high mutation rates that make it difficult for universal primers to be used (Wan *et al.* 2004). The degree of polymorphism of one *locus* may vary considerably even between closely related species, because of nucleotide substitutions within the repeated motifs (Wan *et al.* 2004). Additionally, problems may arise associated with allele scoring errors, unconformity with Hardy-Weinberg equilibrium (HWE), null alleles and linkage disequilibrium (LD), among others (Selkoe & Toonen 2006).

The development of molecular markers is always accompanied by a better understanding of the theoretical background, which allows the emergence of ever more powerful and statistically valid methods of analysis of the data extracted from those same markers (Brinkman & Leipe 2001; Hewitt 2004). Not to mention ways and means of preparing data, such as sequence alignments for example (Brinkman & Leipe 2001). Each analytical method suits differently depending on the type of interpretation a researcher wants to make. Sequence data can be analysed and displayed as a tree or as a network depending if you want to investigate evolutionary relationships and divergence times, or look at the overall haplotype diversity and how it is geographically distributed (Brinkman & Leipe 2001; Huson & Bryant 2006). In turn, many different approaches exist to build: 1) phylogenetic trees, like distance-based methods (for example neighbour-joining) or character-based methods such as Bayesian inference, Maximum-likelihood or Maximum-Parsimony; 2) and networks, such as median-joining or statistical parsimony, among many others (Clement *et al.* 2000; Brinkman & Leipe 2001; Ronquist & Huelsenbeck 2003; Hewitt 2004; Tamura *et al.* 2011). On the other hand, clustering methods have aided immensely more population based studies, allowing the detection and assignment of individuals to different groupings using *multilocus* genotypes, but also to evaluate the dimension of hybridization events and so on (Pritchard *et al.* 2000; Randi & Lucchini 2002; Falush *et al.* 2003; Godinho *et al.* 2011).

The methodological advancement of genetic studies has been continuously moving forward and every now and then a new organism has his complete genome sequenced (Hewitt 2004). Insights from model organisms, mainly domesticated species, can help develop the investigation and application of molecular markers in their closest wild relatives (Hewitt 2004). A clear example comes from the dog and remaining canids, that can now be considered “genome enabled” (Wayne & Ostrander 2007). Due to their close kinship, many molecular tools developed for the dog can be and have been used successfully in phylogenetic and conservation studies of their wild counterparts, such

as microsatellites and expression microarrays (Lindblad-Toh *et al.* 2005; Wayne & Ostrander 2007).

1.5. Canids

1.5.1. General characteristics

The canids belong to a diverse group of mostly predatory mammals, the Order Carnivora, which are characterized by having a pair of carnassial teeth, allowing maximized cutting efficiency (Sillero-Zubiri *et al.* 2004). The Carnivora are divided into the Suborders Feliformia (felids, hyenas, mongooses, among others) and the Caniformia, where canids are included (alongside bears, mustelids, pinnipeds, and others alike), constituting the Family Canidae, in which the distinctive feature is the presence of an inflated entotympanic bulla (a bony chamber that surrounds the middle ear region) (Macdonald *et al.* 2004). The Canidae are in turn further divided into three subfamilies - the Hesperocyoninae (most ancient lineage) and the Borophaginae, which are only represented by fossil forms; and the Caninae, the present day canids, containing about 16 genera and 36 species, with also a considerable fossil record (Van Gelder 1978; Sillero-Zubiri *et al.* 2004).

Contemporary canids are the most widely distributed family of the Order Carnivora, with members on every continent except Antarctica (Sillero-Zubiri *et al.* 2004). Africa is the continent that harbours the greatest diversity with 13 species, of which eight are endemics (Sillero-Zubiri *et al.* 2004). Canids occupy a very wide and continuous spectrum of habitats – from Arctic to Tropic biomes, from sea level to high altitudes, including all kinds of forest, jungle, prairie, savannah, mountain, tundra, desert, and coastline, even thriving in highly anthropomorphized areas, such as agricultural fields, cities or even in households (Wandeler *et al.* 2003; Macdonald *et al.* 2004).

Canids display an incredible dispersing capacity in comparison to many other non-volant species, roaming for very long distances and surpassing topographic and environmental barriers (Macdonald *et al.* 2004). Additionally, many canids are habitat generalist, so they can easily cross even the most highly anthropomorphized areas – a case example would be the presence of red fox (*Vulpes vulpes*) in city centres (Wandeler *et al.* 2003). So it is common to find weak patterns of intraspecific differentiation even between geographically distant regions (Roy *et al.* 1994; Vilà *et al.* 1999). On the other hand, structured genetic variability has been found related to density effects (Roemer *et al.* 2001), kinship (Girman *et al.* 2001), prey specialization

and movement (Carmichael *et al.* 2001; Musiani *et al.* 2007), and natal-habitat-biased dispersal (Sacks *et al.* 2004, 2008; Pilot *et al.* 2006; Carmichael *et al.* 2007; Musiani *et al.* 2007). Canids' high vagility and habitat plasticity has to be taken into account when studying evolutionary, ecological and behavioural aspects.

Another important facet of canid biology is intraguild animosity between sympatric species. Whether it is because of overlapping niches (diets and even den sites) or directly through spatial displacement and intraguild predation, the consequences of these hostile dynamics can affect patterns of habitat use, social structure, and ultimately population size (Tannerfeldt *et al.* 2002; Sillero-Zubiri *et al.* 2004; Roth *et al.* 2008). Many cases have been reported and the tendency is for species more similar in size and with overlapping ranges to have the strongest competition, as well as being the largest species in size to, in most cases, harass the smaller ones - red foxes excluding arctic foxes (*Vulpes lagopus*) (Tannerfeldt *et al.* 2002; Dalen *et al.* 2004; Rodnikova *et al.* 2011), grey wolves (*Canis lupus*) excluding coyotes (*Canis latrans*) (Arjo & Pletscher 1999), African wild dogs (*Lycaon pictus*) predating on bat-eared foxes (*Otocyon megalotis*) (Rasmussen 1996), among others – but there are also cases where, against expectations, no aggressive behaviour is observed (Sunquist 1989) or the smaller species dominates over the larger one (Loveridge & Macdonald 2003).

1.5.2. Origin and evolution

The family Canidae first originated in North America, during the Late Eocene (around 50 to 40 Mya). This estimated date is supported by both the fossil record, and molecular clock calculations (Wayne, Van Valkenburgh, *et al.* 1989; Macdonald *et al.* 2004).

All contemporary canids (subfamily Caninae) are the result of a final radiation that began approximately 12 to 10 Mya in North America, from an ancient genus – *Leptocyon*, that was already known from the early Oligocene (Macdonald *et al.* 2004). Wolf-like and fox-like canids started diverging almost immediately after, with both lineages colonizing independently the Old World. During the late Miocene they were able to leave North America through the Bering Strait and into Eurasia, and thereafter Africa. About 3 Mya, canids finally arrive to South America after the formation of the Isthmus of Panama. All of this means that the Caninae reached the great diversity displayed today in only a few millions of years (Macdonald *et al.* 2004; Sillero-Zubiri *et al.* 2004).

At the mid-Miocene (12 to 9 Mya) species of the wolf-like clade, or Tribe Canini, begin to appear in the form of the taxon *Eucyon*. The more advanced members of this Tribe, the genus *Canis* (and *Lycaon*) are first recorded around the boundary between the Miocene and the Pliocene (6 to 5 Mya) in the North American continent. While expanding their range into Eurasia and Africa, these canids went through an extensive radiation, resulting in a series of closely related species with a predominant circum-arctic distribution, enduring range expansions and contractions according to the warmer/colder climate cycles (Macdonald *et al.* 2004; Sillero-Zubiri *et al.* 2004).
The fox clade, or Tribe Vulpini, arises in the late Miocene (9 to 5 Mya). Shortly after a modest diversification, the two genera that comprise this group begin to occur in the fossil record - *Vulpes* and *Urocyon*. The *Vulpes* genus was able to widely spread and diversify throughout Eurasia during the Pliocene, eventually returning to North America in the forms of the red fox and Arctic fox during the Pleistocene (Macdonald *et al.* 2004; Sillero-Zubiri *et al.* 2004).

1.5.3. Phylogeny

There has always been a great interest in deciphering the phylogenetic relationships within the Canidae, and important progresses have been made in the past three decades. All kinds of data have been used to achieve this purpose: morphologic data (Tedford *et al.* 1995; Lyras & Van Der Geer 2003); molecular data, including comparative karyology (Wayne *et al.* 1987a; b), allozyme electrophoresis (Wayne & O'Brien 1987), DNA–DNA hybridization (Wayne, Benveniste, *et al.* 1989), mitochondrial DNA (Geffen *et al.* 1992; Wayne *et al.* 1997) and nuclear DNA sequences (Bardeleben *et al.* 2005a; b; Tsubouchi *et al.* 2012); and the combination of both morphologic and molecular data (Zrzavy & Ricankova 2004). The most recent comprehensive molecular phylogenetic analysis of the Canidae was achieved by Lindblad-Toh *et al.* (2005), in which the authors performed a genome comparison analysis of the complete dog genome, in order to obtain a set of rapidly evolving nuclear genes. The ensuing phylogenetic tree (Figure 2) is in agreement with previous results when regarding the existence of three major groupings within the Canidae: (1) the fox-like canids (in red); (2) the South American canids (in green); (3) and the wolf-like canids (in blue) (Lindblad-Toh *et al.* 2005; Wayne & Ostrander 2007). This study was also able to resolve with strong statistical support many issues that had been remaining unclear in previous studies, in terms of branching order and closest affinities, such as: (1) the grey fox (*Urocyon cinereoargenteus*) appearing as the most primitive

lineage, adding more evidence to a North American origin for all canids; (2) the closest affinity of the South American clade to the wolf-like clade; (3) the incorporation of the raccoon dog (*Nyctereutes procyonoides*) and bat-eared fox in the fox-like group; (4) and finally suggesting an African origin for the wolf-like canids, since the basal taxa for this group are two sister species endemic from this continent, the side-striped jackal (*Canis adustus*) and the black-backed jackal (*Canis mesomelas*) (Lindblad-Toh *et al.* 2005).

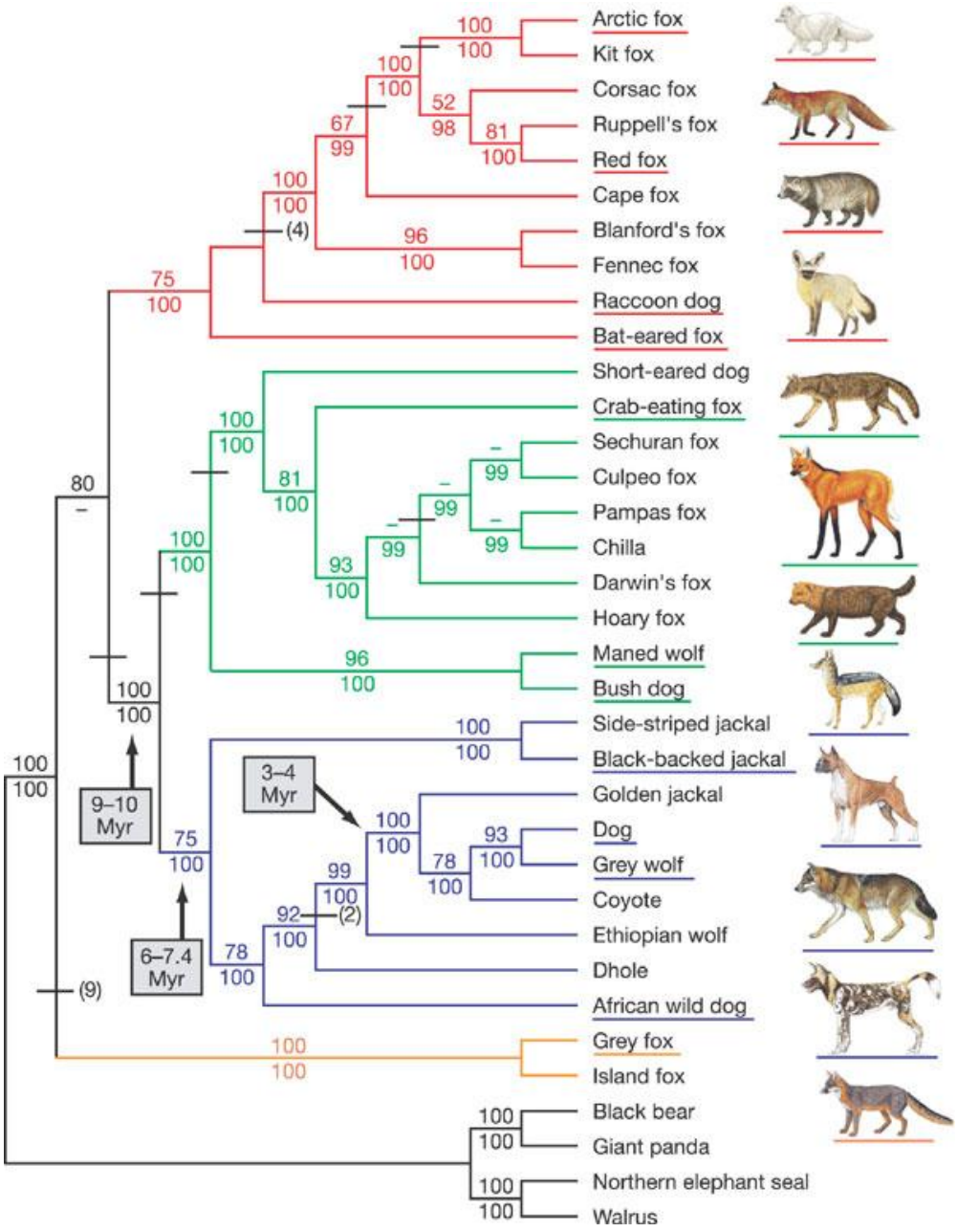


Figure 2 - Phylogenetic tree of living Canidae based on ~15 kb of exon and intron sequences, as in Lindblad-Toh *et al.* (2005). Branch colors identify the fox-like clade (red), the South American clade (green), the wolf-like clade (blue) and the grey and island fox clade (orange). Divergence times in millions of years (Myr). Copy right permissions for illustrations are listed in the Supplementary Information of Lindblad-Toh *et al.* (2005).

1.5.4. Hybridization and Introgression

Hybridization and introgression are strong evolutionary forces occurring naturally among many taxa that can be responsible for important speciation events and local adaptation of populations (Barton & Hewitt 1989). However, they can also become a serious threat to the genetic heritage of endangered species and populations, especially in "human induced scenarios" that result in population declines, habitat changes, translocation of novel species (or individuals in the case of populations) and contact with closely related domestic species (Rhymer & Simberloff 1996; Allendorf *et al.* 2001; Seehausen *et al.* 2008).

This phenomena is probably more dramatic in the Canidae then in any other group of carnivores (Sillero-Zubiri *et al.* 2004). The crown group of the wolf-like canids containing the closest wild relatives of the dog – Ethiopian wolf (*Canis simensis*), golden jackal (*Canis aureus*), coyote and grey wolf – is especially subjected to hybridization issues because: (1) it is a set of closely related and recently diverged species (around three to four million years ago) (Lindblad-Toh *et al.* 2005; Wayne & Ostrander 2007) and because of that reproductive isolation may not be entirely established (Vilà *et al.* 2003); (2) they are highly mobile species and, for the exception of *C. simensis*, habitat generalists, so hybrid zones are wide and can expand rapidly (Macdonald *et al.* 2004; Sillero-Zubiri *et al.* 2004); (3) these species are strongly affected by habitat modifications and reduction of connectivity, and the Ethiopian and grey wolves in particular have been directly persecuted by humans and seen their population sizes decline substantially during the last centuries, which decreases the chances of same species pairs encountering each other to mate and consequently increases the chances of crossing between different species (Wayne *et al.* 1992; Gottelli *et al.* 1994; Randi *et al.* 2000; Vilà *et al.* 2003; Godinho *et al.* 2011). As a consequence of the combination of all these factors we witness the formation of admixture zones that can be responsible for rapid adaptive evolution (coyotes through hybridization with wolves - Kays *et al.* 2010), but for most cases it can compromise the conservation of an endangered species or population (Gottelli *et al.* 1994; Randi *et al.* 2000; Godinho *et al.* 2011). Clearly, one of the major threats to canid conservation is hybridization with free-ranging dogs. Due to the very recent domestication process (Savolainen *et al.* 2002), reproductive isolation is not completely developed and since dogs coexist with many of their wild relatives in most of their ranges (Vilà *et al.* 2003), examples abound of hybridization events between dogs and wolves (Randi *et al.* 2000; Randi & Lucchini 2002; Vilà *et al.* 2003; Godinho *et al.* 2011), coyotes (Adams *et al.*

2003; Kays *et al.* 2010; Bohling & Waits 2011) and Ethiopian wolves (Gottelli *et al.* 1994).

1.5.5. Taxonomic Uncertainties and Knowledge Gaps

While major clades and deeper lineage affiliations within the Canidae seem to be well resolved, many doubts still remain concerning the more recent splits. Due to the inherent phylogenetic, ecological and behavioural characteristics of this group, describing and identifying clear taxonomic units can become a problematic question to tackle.

For instance in North America, controversy still exists regarding the current taxonomic status of the Great Lakes wolf (*Canis lycaon* or *Canis lupus lycaon*) and the red wolf (*Canis rufus*), which seem to have a mixed ancestry as a result of past admixture between coyotes and grey wolves, and currently hybridization still occurs between these four wolf-like canids (vonHoldt *et al.* 2011). Another example of taxonomic uncertainty within the Canidae comes from North Africa, where what was thought to be until now North African golden jackals seem in fact to be the first ever African grey wolves (Rueness *et al.* 2011; Gaubert *et al.* 2012). The case of the African wolf illustrates well the lack of even the most basic biological knowledge in North Africa – a large carnivore such as a wolf has been able to remain unnoticed throughout its entire range and for such a long time (Gaubert *et al.* 2012).

Additionally, North Africa is designated as one of the priority regions in terms of canid research, primarily concerning the smaller desert foxes, for which the available data is extremely scarce (Macdonald *et al.* 2004).

1.5.6. North African Canids

A total of five canid species can be found in North-West Africa and all of them were the focus of this study:

1) The golden jackal, *Canis aureus* (Linnaeus, 1758), is a medium-sized canid and a well representative of the genus *Canis*. The species is widespread across North Africa, from Mauritania in the West coast to Egypt in the East, including Morocco, Algeria and Libya to the North; Senegal, Nigeria, Chad and Tanzania to the South (Jhala & Moehlman 2008). It is also present in South-Eastern Europe, the Middle East, and all the way to Central/Southeast Asia (Sillero-Zubiri *et al.* 2004; Jhala & Moehlman 2008). As a result of their tolerance to dry environments and their omnivorous diet, golden

jackals can occupy a wide variety of habitats, ranging from semi-desert and savannahs, to forests, to agricultural and urban surroundings (Macdonald & Sillero-Zubiri, 2004; Jhala & Moehlman, 2008). Still, more accurate patterns of habitat selection are unknown, so it is most likely that their distribution ranges are over-estimated, and to the exception of protected areas, golden jackal populations are gradually declining, mainly due to the replacement of wilderness areas and traditional land-uses for more agricultural intensive practices, industrialized and urbanized settlements (Sillero-Zubiri *et al.* 2004; Jhala & Moehlman 2008). Moreover, there is a clear need to better clarify phylogenetic relationships, population structure and assess levels of genetic diversity for this species using modern molecular techniques, as well as other aspects such as dispersal capacity, survival and mortality rates (Sillero-Zubiri *et al.* 2004; Jhala & Moehlman 2008).

The golden jackal has always been considered a monophyletic species (Wayne *et al.* 1997; Lindblad-Toh *et al.* 2005) and up to 12 subspecies have been described throughout its continuous range, based mainly on morphology and coat colour (Sillero-Zubiri *et al.* 2004; Jhala & Moehlman 2008), five of which in North Africa: *C.a. algirensis* found in Algeria, Morocco and Tunisia; *C.a. anthus* in Senegal; *C.a. lupaster* in Egypt; *C.a. soudanicus* in Sudan and Somaliland; and *C. a. riparius* also in Somaliland and coastal Ethiopia and Eritrea (Wilson & Reeder 2005). However, past research had already noticed similarities in terms of skull morphometrics along with overall size between the Egyptian jackal subspecies, *C. a. lupaster*, and grey wolves (Ferguson 1981), and a recent phylogenetic paper by Rueness *et al.* (2011) brought back this issue into the spotlight. Based on mtDNA, Rueness *et al.* (2011) placed the Egyptian jackal not within the golden jackal clade (with samples from Kenya, Serbia and India) but within the wolf clade (with samples covering the wide distribution of this species), suggesting that together with two new wolf species described in India (Aggarwal *et al.* 2007), it would represent ancient lineages that would have given rise to the Holarctic grey wolf clade. So the conclusions were that the Egyptian jackal is in fact the first African wolf, *C. lupus lupaster*, a cryptic subspecies in need of urgent conservation status assessment and additionally, expanding its known distribution range southwards to Ethiopia (Rueness *et al.* 2011). In 2012, the African wolf lineage's range was further extended to North-West Africa, confirmed also by analyses of mtDNA sequences (Gaubert *et al.* 2012). The two novel aspects of this last study are the morphological and behavioural distinction between the African wolf and the sympatric golden jackal, but on the other hand, all the golden jackals identified by morphology belonged to the African wolf mtDNA lineage (Gaubert *et al.* 2012). The

authors suggested two explanations for their findings: (1) occurrence of hybridization between these canid forms; (2) the African golden jackals represent a smaller eco-morphological variant within the African wolf lineage (Gaubert *et al.* 2012).
Besides these two studies in North Africa, the only assessment of genetic variability of the golden jackal was conducted in Serbia (Zachos *et al.* 2009). Analysis of mtDNA sequences and a set of eight microsatellites revealed low levels of genetic differentiation, along with a strong founder effect, which was an expected result taking into account the recent time of expansion of the golden jackal into the Balkans (Zachos *et al.* 2009).

2) The pale fox, *Vulpes pallida* (Cretzschmar, 1827), is one of the least known canid species, being classified as Data Deficient (DD) by the IUCN (Sillero-Zubiri 2008). It is endemic to the semi-arid Sahelian region, covering Mauritania and Senegal, across Nigeria, Cameroon and Chad to the Red Sea, inhabiting dry sandy and stony sub-Saharan desert and semi-desert areas, but also savannah in the more southern limits of its range, and humanized areas where food is more available (Sillero-Zubiri *et al.* 2004; Sillero-Zubiri 2008). Even though it has an extensive distribution in most parts of their range, it is locally rare and more knowledge is needed on abundance, population status and trends, ecological requirements and specific threats to its survival, especially for a species that dwells in such an unstable and fluctuating habitat, lying between true desert and savannah (Macdonald *et al.* 2004; Sillero-Zubiri *et al.* 2004). Presently, the evolutionary position of the pale fox within the Canidae is known solely based on morphological characters (Clutton-Brock *et al.* 1976). The only recent phylogenetic studies to include this species are ones who use the same reference of 1976 to produce Carnivora supertrees (Bininda-Emonds *et al.* 1999; Nyakatura & Bininda-Emonds 2012) or combine morphology with molecular data (Zrzavy & Ricankova 2004). So no molecular data (DNA sequence or even allozymes) has ever been produced for the pale fox and it has never been considered in the various phylogenetic studies specific of the Canidae (Wayne *et al.* 1997; Lindblad-Toh *et al.* 2005) or fox-like canids (Geffen *et al.* 1992), so its evolutionary relationships remain poorly clarified.

3) The Rüppell's fox, *Vulpes rueppellii* (Schinz, 1825), can be found widespread in desert and semi-desert regions of North Africa, above 17°N latitude up to the northern fringes of the Sahara Desert, from Morocco and Mauritania to Egypt and Somalia, being also present across the Arabian Peninsula and Middle East up to Pakistan

(Cuzin *et al.* 2008). Its habitats include sand and stone deserts, with sparse vegetation, avoiding extreme desert and large sand dunes, displaying higher abundance on the fringes of these areas, mountain massifs and oasis (Sillero-Zubiri *et al.* 2004). Again, there is still a great lack of data on population sizes and trends, and overall ecology, especially in North Africa, but recent threats to the species' survival have been described, such as habitat loss, fragmentation and degradation, and human persecution, either by hunting or poisoning (Macdonald *et al.* 2004). Another major threat is competitive exclusion by the red fox. In Israel, Rüppell's fox are in risk of extinction as red fox increases its range to the desert fringes, following the establishment of new human settlements. Rüppell's foxes are only able to out compete the red foxes in harsher desert or in protected areas where red fox populations are better controlled (Yom-Tov & Mendelssohn 1988; Sillero-Zubiri *et al.* 2004). The same kind of competitive behaviour has also been noticed in Morocco (Cuzin 2003). Besides overall phylogenetic studies of the Canidae or foxes, there has not been molecular research focused specifically on the Rüppell's fox.

4) The red fox, *Vulpes vulpes* (Linnaeus, 1758) has a long historical association with Man, and its biology is well understood. It is a canid of medium size, but the largest of the *Vulpes* genus. The species displays the widest geographical range of any member of the Carnivora order, across the entire Northern Hemisphere, to the exception of areas of extreme cold or aridity – in Africa it is restricted to the Mediterranean region, north of the Sahara (Cuzin 2003; Macdonald & Reynolds 2008). Such a wide distribution comes accompanied by a substantial geographical variation in adult body size - red foxes from the Middle East and Algeria are smaller than the European ones (Sillero-Zubiri *et al.* 2004, and references therein). Being a very resilient and adaptable species, it is able to populate a broad variety of habitats, through a wide spectrum of latitudes and altitudes, including tundra, desert, open plains and forests, but also agricultural areas and city centres (Wandeler *et al.* 2003; Macdonald & Reynolds 2008). It has been observed many times that red foxes thrive in areas of intense human presence, being able to colonize new areas that otherwise would not be capable of, to the expense of the range of other species better suited to that particular environment before the anthropogenic disturbance (see text above about the Rüppell's fox) (Cuzin 2003; Macdonald *et al.* 2004). Habitat degradation, loss and fragmentation, as well as direct and indirect persecution (in many countries red fox hunting is legal and in some cases it is considered a pest) count as the most relevant pressures for their survival. Still, their versatility and opportunistic nature allows the red fox to

successfully persist, in fairly large numbers, all through its distribution area (Sillero-Zubiri *et al.* 2004; Macdonald & Reynolds 2008).

In terms of genetic research, the red fox is a fairly well studied species, mainly in North America where two distinct clades have been described due to different Pleistocene refugia (Aubry *et al.* 2009) and there are some issues related to native and non-native (European and Asian) ancestry of some populations (Sacks *et al.* 2011; Statham *et al.* 2012). In the Hokkaido Island, Japan, interesting results were achieved. Native red foxes possess unique mtDNA haplotypes in comparison to adjacent islands and the main continent (Inoue *et al.* 2007), and further research within the island using microsatellite data indicated the existence of at least one phylogeographically distinct subpopulation (Oishi *et al.* 2011). In North Africa there has not been a genetic assessment of the red fox, and even in Europe only two studies have focused on the Mediterranean basin with somewhat contradictory results. While Frati *et al.* (1998), based on allozyme and mtDNA data, suggests the existence of two main groups originated as a consequence of different refugia and colonization routes during Quaternary glaciation cycles, Teacher *et al.* (2011), using mtDNA of ancient samples and sequence data of Frati *et al.* (1998), found no signs of genetic structure or variation of variability through space and time.

5) The fennec fox, *Vulpes zerda* (Zimmermann, 1780), is the smallest of all canids, although the most striking characteristic of this species are the extremely large ears, giving it the greatest ear-to-body ratio in the family (Sillero-Zubiri *et al.* 2004). It has a wide geographic range throughout North-African sandy deserts and semi-deserts, up to the northern Sinai Peninsula, and in the South to 14°N, in the more northern Sahelian areas (Sillero-Zubiri *et al.* 2004; Asa *et al.* 2008). The fennec fox is the only Saharan carnivore able to live in areas where there is no clear water source, and so it is able to subsist in extreme arid environments, preferably in stable sand dunes, but can also be found in more vegetated dunes along the Atlantic coast. On the northern fringes of its distribution, annual rainfall is less than 100mm per year (Sillero-Zubiri *et al.* 2004; Asa *et al.* 2008). The major threats seem to be trapping for commercial use and habitat loss. Although data on current population trends is deficient, overall it is assumed that the fennec fox is still relatively common throughout its range, based on observations on the number of individuals trapped and sold commercially (Sillero-Zubiri *et al.* 2004; Asa *et al.* 2008). Population declines have been reported in northern Moroccan Sahara, due to the construction of new permanent human settlements (Cuzin 2003). Again, as in the

case of *V. rueppellii* and *V. pallida*, there is no published genetic research dedicated exclusively to the fennec fox.

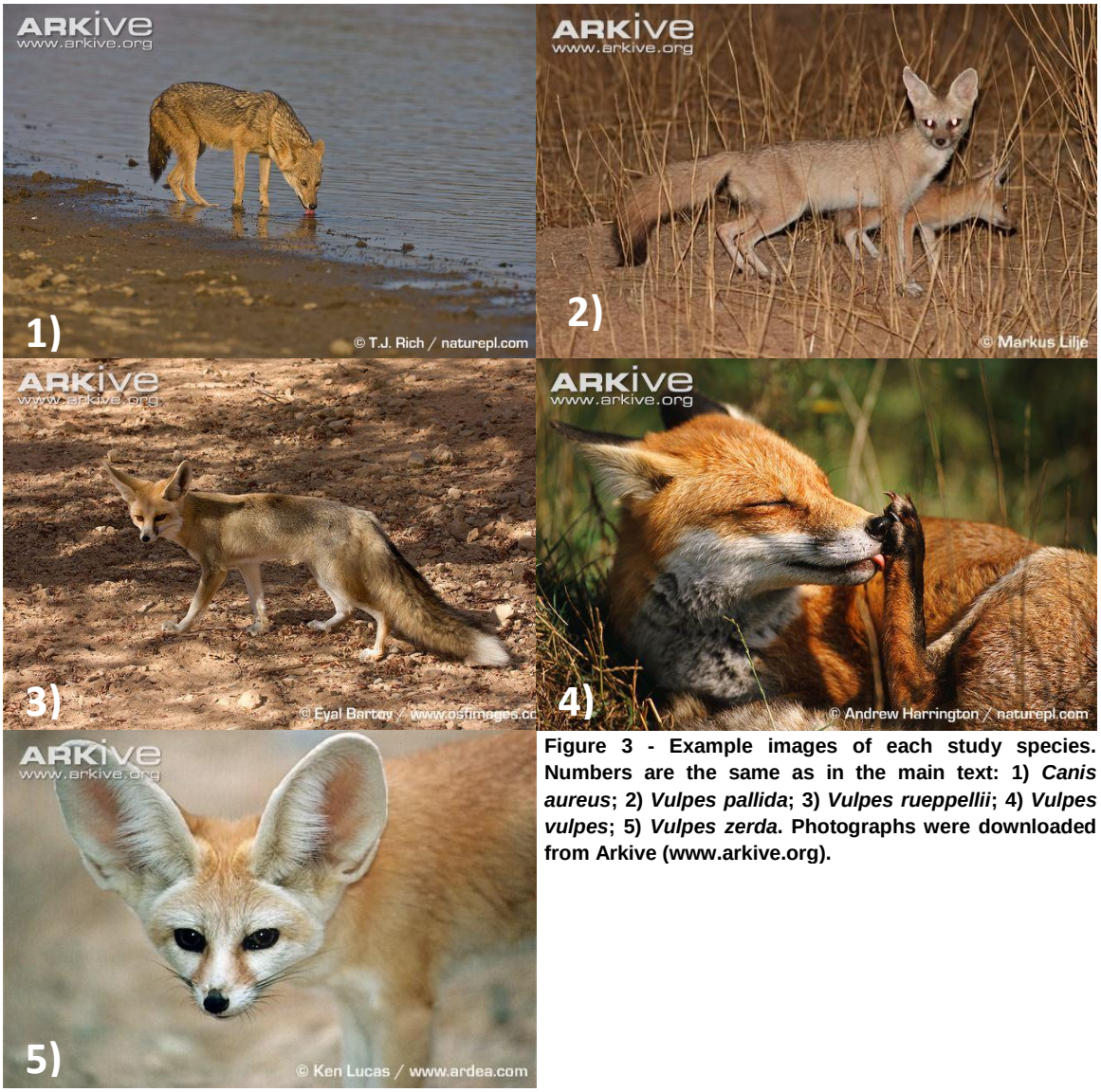


Figure 3 - Example images of each study species. Numbers are the same as in the main text: 1) *Canis aureus*; 2) *Vulpes pallida*; 3) *Vulpes rueppellii*; 4) *Vulpes vulpes*; 5) *Vulpes zerda*. Photographs were downloaded from Arkive (www.arkive.org).

To the exception of the golden jackal, and the red fox in other regions other than North Africa, none of these species has been the main focus of phylogenetic and population studies, and existing molecular data is scarce, as a result of contributions just by general evolutionary studies on the entire Canidae. The few studies of *V. vulpes* in the Mediterranean basin have never included samples from North Africa, and the most extreme case comes from the pale fox for which there is no reported molecular data, and its systematic position within the fox-like canids remains a mystery regarding comparison of DNA sequences.

Furthermore, still very little research has been conducted aiming at the evaluation of the putative barrier effect that the Sahara could constitute for highly vagile species

occurring north and south of the desert, as are the cases of *C. aureus* and *V. rueppellii*. In addition, the Niger River could also be pointed out as a possible force in structuring genetic diversity for the pale fox throughout the Sahel region.

North-West African canid species constitute a guild, due to their overall similar sizes, weights and trophic niche exploitation, being essentially omnivores/carnivores, and because of that some degree of competition between species is bound to occur. Within this group of canids, special attention must be given to three morphologically-similar species, *V. pallida*, *V. rueppellii* and *V. vulpes*, that are found in distinct, yet successive, biogeographic regions, namely in the Sahelian, the Saharo-Sindian and Mediterranean regions, respectively (Sillero-Zubiri 2009). Ecological niche-based models predicted broadly parapatric ranges and potential sympatry zones between *V. vulpes* and *V. rueppellii*, and between *V. rueppellii* and *V. pallida*, where gene flow between sympatric populations, and/or species competition and spatial exclusion may be happening, due to environmental and human-related factors (Brito *et al.* 2009). While competition in the first case has been described (Cuzin 2003), in the later pair of species more observations and ecological studies are needed.

1.6. Objectives

The central aim of this study is to infer evolutionary and phylogeographic patterns at inter and intraspecific levels, in North-West African canids, namely *Canis aureus*, *Vulpes pallida*, *V. rueppellii*, *V. vulpes* and *V. zerda*. Two types of molecular markers were applied and objectives were set according to the different types of questions that each marker allows addressing:

1) Analyses of mitochondrial DNA sequences:
    - How are the study species phylogenetically related with each other? And with the remaining species of the Canidae? Specifically: 1) Are *C. aureus* in this region in fact African wolves?; 2) Does *V. rueppellii* constitute a true sister taxa or does it happens to be a desert ecotype of *V. vulpes*?; 3) Which are the closest relatives of *V. pallida*?
    - What is the degree of genetic diversity of each species and how it is geographically distributed?
    - Is there cryptic diversity?

2) Analyses of microsatellite *loci*:
    - Is intraspecific genetic variability geographically structured?
    - If populations are identified, is there gene flow between them? And if not, which graphical features are acting as a barrier? It will be tested the role of the Sahara desert for North and South populations of *C. aureus* and *V. rueppellii*, and the Mauritanian mountain ranges and Niger river for *V. pallida*;
    - Is gene flow occurring between sympatric species? Are there signs of hybridization or introgression, namely between 1) *V. rueppellii* and *V. vulpes* and 2) *V. rueppellii* and *V. pallida*?

Hopefully the results of this study will aid in increasing our understanding on the evolutionary relationships, the distribution of genetic diversity and identification of threats to the sustainable continuity of North African canids, in order to improve our capabilities for management and conservation of these and other endangered species, as well to contribute to the knowledge on biodiversity as a whole, in such a remote and poorly comprehended region of the world.

2. Methods

2.1. Samples

A total of 120 tissue samples from *Canis aureus* (n=37), *Vulpes pallida* (n=29), *Vulpes rueppellii* (n=11), *Vulpes vulpes* (n=30) and *Vulpes zerda* (n=11) were available for this study. Two further samples of side-striped jackal (*Canis adustus*) from Senegal and Ethiopia were provided by collaborators and also used in this research. All samples were collected from road-kill animals during various overland field missions in North-West Africa or offered by other researchers and collaborators. Samples were collected between November 2003 and December 2011. Tissue samples were preserved in ethanol 96% for genetic analyses at the moment of collection. Geographical location of each sample was recorded with a Global Positioning System (GPS) on the WGS84 datum and represented using the Geographical Information System ArcMap 10.0 (Figure 4).

2.2. DNA extraction

Total genomic DNA extractions were performed using QIAGEN's QIAmp® DNA Micro Kit, following the respective protocol for tissue samples. Then, the success of DNA extractions, both in terms of quantity and quality, was assessed by electrophoresis in 0.8% agarose gels stained with gel red (Biotium), in TBE 0.5x (Tris-Borate-EDTA buffer: Tris 89mM, boric acid 89mM, EDTA 2mM, pH 8.0) at 300V for approximately 10 to 15 minutes. Agarose gel visualisation through UV radiation was done in a BioRad Universal Hood II Quantity One 4.4.0.

When quantity and quality of DNA extraction was high, they were properly diluted in ultra-pure water. Otherwise, they were used directly for the following PCR and sequencing reactions. Extractions were stored at -20ºC.

Two distinctive molecular analyses were carried out in this work based on different molecular markers – mitochondrial DNA sequences and microsatellite genotyping.

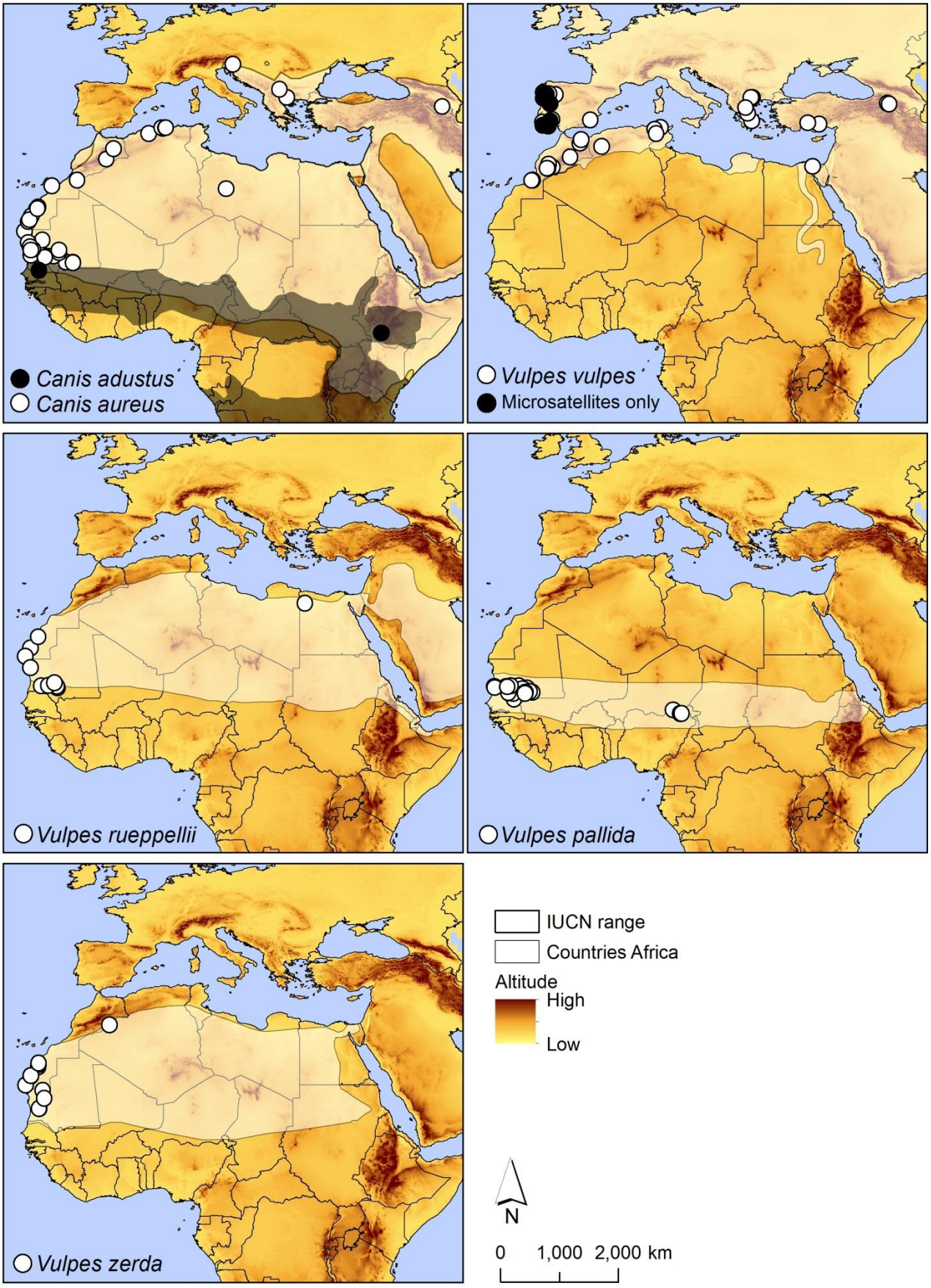


Figure 4 - Geographic distribution of taxa under study according to the IUCN (http://www.iucnredlist.org/) and location of all samples available for this study. For *Vulpes vulpes*, samples for which we only had microsatellite data are represented by black circles.

2.3. Mitochondrial DNA analysis

2.3.1. Laboratory procedure

Genetic diversity of North-West African canids was first assessed at the mitochondrial genome using Cytochrome b (Cyt-b) and control region (D-loop) fragments, for the entire set of samples. Foxes were amplified for a Cyt-b fragment situated in the middle of the gene [14.639 – 15.002 base pair (bp) of the dog mtDNA genome (accession number U96639)], and jackals (*C. aureus* and the two *C. adustus*) were amplified for the left domain [14.202 – 14.594 bp of the dog mtDNA genome (U96639)]. This was because: 1) the success of amplification of the first mentioned fragment was considerably low for the golden jackal; 2) available sequences in GenBank (National Resource for Molecular Biology Information, http://www.ncbi.nlm.nih.gov/Genbank) to use for comparison in the jackals are only from the left domain.

For a smaller set of samples comprising eight golden jackals and the two side-striped jackals, partial sequences of the 12S rDNA (12S) and 16S rDNA (16S) genes were also amplified to compare with the available sequence data of *Canis lupus lupaster* (Rueness *et al.* 2011) and other wolf-like canids. The golden jackal samples were selected based on the quality of the data obtained for the other mtDNA genes and microsatellite *loci*, as well as best representing the geographic range of our sample collection.

The primers used were the following: (1) Cyt-b (foxes) - cb2F (forward) and CB3-H (reverse) (Palumbi *et al.* 2002); (2) Cyt-b (jackals) - Gludg (forward) and cb2R (reverse) (Palumbi *et al.* 2002); (3) D-loop – DLH (forward) and THR (reverse) (Kocher *et al.* 1989); (4) 12S – 12Sa (forward) and 12Sb (reverse) (Palumbi *et al.* 2002); 16S – 16Sa (forward) and 16Sb (reverse) (Palumbi *et al.* 2002).

Polymerase chain reactions (PCRs) were performed in approximately 10 µL reaction volumes that comprised: 5 µL of MyTaq (MyTaq™ Mix, Bioline) for Cyt-b, 12S and 16S, and Master Mix (Taq PCR Master Mix, QIAGEN) for D-loop; 0.4 µL of both forward and reverse primers; 3.2 µL of pure water; and 1 µL of DNA (for lower quality DNA samples higher quantities were used, up to 4 µL in some cases). A negative control was used in all PCRs to check for contaminations. All PCRs were carried out in a BioRad T100 Thermal Cycler (for specific amplification conditions of each fragment see section A1 in Annexes). Amplifications were evaluated by electrophoresis in 2% agarose gels, applying a mass DNA ladder (Marker 5, Eurogenetec), and otherwise following the

same conditions pointed out in the extraction electrophoresis. Gels were visualized in a BioRad Universal Hood II Quantity One 4.4.0.

Before sequencing, PCR products were purified using ExoSap (USB® ExoSAP-IT® PCR Product Cleanup, Affymetrix), as specified by the manufacturer, to remove any excess of primers and other amplification reagents. It was then followed by standard protocol of Big-Dye cycle sequencing kit (BigDye® Terminator v3.1 Cycle Sequencing Kits, AB Applied Biosystems), using the forward primers for each respective gene, in a BioRad T100 Thermal Cycler. Sequencing reaction products were cleaned with Sephadex following the manufacturer's protocol, which removes dNTPs (deoxynucleotide triphosphates) and other reagents. Finally, the sequencing reaction was performed on an ABI 3130xl Genetic Analyser (AB Applied Biosystems).

Resulting electropherograms were verified and aligned using SeqScape v2.5 (Applied Biosystems, 2004). The final alignments consisted of 363 bp (foxes) and 393 bp (jackals) for Cyt-b; 394-406 bp for D-loop; 386 bp for 12S and 504 bp for 16S.

2.3.2. Data analysis

The GENEIOUS PRO software v4.8.2 (Drummond *et al.* 2010) was used to download further available sequence data of our study species and other related canids at GenBank (see sections A5-7 with accession numbers in Annexes), generate further alignments with the built in MUSCLE alignment tool, and create concatenated sequences.

Phylogenetic relationships of our samples and other downloaded sequences were assessed by Bayesian inference (BI) and Maximum Parsimony (MP). For *C. aureus*, three concatenated datasets were analysed – Cyt-b and D-loop; the latter two genes plus 16S; and all the mentioned genes with added 12S. The more genes were incorporated into the analysis the less sequences were available. For the foxes only a concatenated dataset comprising Cyt-b and D-loop was used. D-loop sequence sizes varied for each dataset depending on: 1) the extent of trimming at the beginning and end of sequences in order to avoid areas that are very difficult to align in cases of multiple taxa (as a non-coding region of the mitochondrial genome, the chances of mutations fixating is higher); 2) trying to overlap as much as possible our sequences with the available data online. The best-fit model of evolution for each gene was selected using JMODELTEST v0.1.1 (Posada 2008), with the Akaike information criterion (AIC). Tajima´s nonparametric test and likelihood ratio test of molecular clock were implemented in MEGA v.5 (Tamura *et al.* 2011) in order to test the clocklike evolution of

the sequences. Both tests did not reject the strict molecular clock (P > 0.05) and thus it was assumed in the phylogenetic analyses. BI was conducted using MRBAYES v3.1.2 (Ronquist & Huelsenbeck 2003). When analysing the concatenated sequences, the dataset was partitioned by gene to be run under the corresponding evolutionary model. Markov Chain Monte Carlo (MCMC) analyses were run for 5^6 generations with four chains, with a sampling frequency of 500 generations, and discarding in the end the first 2500 trees as burn-in. TRACER v1.5 (http://tree.bio.ed.ac.uk/software/tracer/) was then used to check if that burn-in should be applied [if all parameter values of ESS (effective sample size) were above 200]. Resulting trees were drawn with FIGTREE 1.3.1 (Rambaut 2009). MP analysis was run using a close-neighbour-interchange (CNI) search and 1,000 bootstrap replications in MEGA v.5 (Tamura *et al.* 2011).

Networks of each concatenated dataset, as well as for Cyt-b alone for *C. aureus* and wolf-like canids (for comparison with other jackal sequences for which only Cyt-b data is available), were generated using parsimony calculations in TCS v1.21 (Clement *et al.* 2000), and considering gaps as a fifth state. A preliminary analysis was first conducted to see how the different clades observed in the phylogenetic tree linked to each other, with the connection limit set to 95% of probability (below that threshold the program does not draw any connection between haplotypes). Since only closely related individuals of each clade appeared associated we fixed the connection limit at 80 steps in order to link the more unrelated groups and see the number of mutational steps between them.

For the concatenated sequences, additional neighbour-net networks based on uncorrected patristic distances and bootstrap analysis with 1,000 replicates, as implemented in SPLITSTREE v4.6 (http://www.splitstree.org/), were also produced.

Finally, haplotype and nucleotide diversity were calculated using DNASP v5 (Librado & Rozas 2009), and sequence divergence based on Kimura 2-parameter model were calculated in MEGA v5.0 (Tamura *et al.* 2011).

2.4. Microsatellite analysis

2.4.1. Laboratory procedure

We selected a total set of 46 autosomal microsatellite *loci*, previously developed for the dog and wolf (Godinho et al., 2011), to determine individual *multilocus* genotypes. Microsatellite *loci* were distributed in different sets (from now on named kits), a procedure known as multiplexing, which allows simultaneously to amplify several *loci* in one single polymerase chain reaction, and to subsequently analyse and sequence multiple molecular markers at the same time. Eighteen *loci* were amplified using a commercial kit (Kit 3 - FINNZYMES) and the remaining 28 in four kits that were previously developed in the laboratory [Kit 2, Kit 4, Kit AP1 and AP2 (the last two are mixed in the same sequencing reaction)] using the Multiplex PCR Kit (QIAGEN) (see section A2 in Annexes for distribution of *loci* to each multiplex kit).

PCRs were prepared with a negative control, in approximately 10 µL reaction volumes: 5 µL of Master Mix (Taq PCR Master Mix, QIAGEN), 1 µL of primers (kits), 3 µL of pure water and 1 µL of DNA (again for lower quality DNA samples higher quantities were used). Success of amplification was also evaluated by electrophoresis in 2% agarose gels, following the same procedure explained above in the mtDNA analysis.

Amplification conditions and genotype reading had already been optimized in the laboratory in a preliminary study for the aimed canid species. PCRs were performed in a BioRad T100 Thermal Cycler, with programs varying according to each kit (see section A4 in Annexes).

Afterwards, amplification products were diluted according to their quality observed in the agarose gels, and finally analysed on an ABI 3130xl Genetic Analyser (AB Applied Biosystems), using the 350ROX size standard. Allele data was obtained using GENEMAPPER v4.1 (Applied Biosystems, 2009) and then manually checked.

2.4.2. Data analysis

For this analysis a data set comprising *multilocus* genotypes of 15 red foxes from Portugal was added to the remaining genotyped red foxes (black dots in Figure 4).

First of all, in order to avoid any possible bias in the final results, some pre-evaluation and editing of the microsatellite data sets was done before further analyses (in the following order): (1) *loci* that did not amplified for any of the individuals of a species, or were monomorphic, were rejected; (2) samples that presented a considerable amount

of non-amplifying *loci* (missing data >20%) were not taken into account; (3) assessment of Hardy-Weinberg equilibrium (HW) - *loci* that were deviated from a HW equilibrium were excluded; (4) test for linkage disequilibrium – if non-random combination of alleles at two or more *loci* occurred, those *loci* were omitted.

GENEALEX v6.4.1 (Peakall & Smouse 2006) add-in for Microsoft® Office Excel® 2010 was used to calculate measures of genetic diversity and variation, such as allele frequencies and patterns, mean number of alleles per *locus* (Na), observed (Ho) and expected (He) heterozygosities. The same software was used to evaluate deviations from Hardy-Weinberg (HW) equilibrium, as well as to calculate pairwise F-statistics and individual-by-individual genetic distances that were then used to conduct Principal Coordinate Analyses (PCA). ARLEQUIN v3.5 (Excoffier & Lischer 2010) was used to test pairwise linkage disequilibrium for all *loci* (16 000 permutations) based on the exact test of Guo & Thompson (1992).

The Bayesian clustering software STRUCTURE v2.3.3 (Pritchard et al. 2000; Falush et al. 2003) was used to infer population structure and assign individuals to putative populations. This program tries to detect signs of population genetic structuring and allocate, probabilistically, an individual *multilocus* genotype to a putative source population, based on the assumption that there will exist a K number of populations (where K can be already known or not), each of which characterized by a different set of allelic frequencies for each given *locus* (Randi & Lucchini 2002).

The golden jackal was analysed for two *multilocus* genotype datasets: 1) using our entire golden jackal set of samples (North African and two European samples); 2) same as in the previously mentioned analysis but adding a group of 14 wolves, from various origins [Portugal, Slovenia, Romania, Russia (Chukotka) and Siberia] and eight domestic dogs. The aim of the last analyses was to see if the segregation of each clade would follow a similar pattern to the one found at the mitochondrial level (as a proxy for sequencing at the nuclear genome level).

Foxes were analysed as follows: 1) individual analysis of each species dataset to check for population structuring; 2) assessment of any putative signal of hybridization between *Vulpes rueppellii* and *V. vulpes*, and again *V. rueppellii* with *V. pallida* (for both cases only including in the analysis *loci* common for both species).

All analyses were run using the admixture model with correlated allele frequencies. No priors of individual geographic collection site were used. STRUCTURE was run for a number of clusters (K) between one and ten, with ten repetitions of 10^6 MCMC iterations for each K value, following a burn-in period of 10^5 steps in order to guarantee the achievement of similar posterior probabilities of the data in each run and to

ascertain confidence in the model fit. STRUCTURE software analyses were performed on Bioportal [Kumar *et al.* 2009 (https://www.bioportal.uio.no/)]. Results were then submitted to STRUCTURE HARVESTER v0.6.92 online software (Earl & vonHoldt 2011) to assess the probability of each K-value. The most likely number of clusters was estimated using the mean values of likelihood [L(K)] and the Evanno method (Delta K - Evanno *et al.* 2005).

3. Results

3.1. Laboratory overview

Out of 120 samples, 68 were successfully amplified for all molecular markers: 14 of *Canis aureus* (including selected samples for 12S and 16S); 15 of *Vulpes pallida*; seven of *Vulpes rueppellii*; 25 of *Vulpes vulpes*; and seven of *Vulpes zerda*. Data from all markers were also obtained for one *Canis adustus* sample (including 12S and 16S). Sequencing and genotyping was not possible for three out of 120 specimens, while the success of the remaining 48 specimens varied depending on each marker. For more detailed information see sections A8-10 in Annexes.

3.2. Mitochondrial DNA

3.2.1. Golden jackal and wolf-like canids

The phylogenetic analysis of wolf-like canids based on a combined 508 bp mtDNA sequence dataset (307 bp of Cyt-b and 201 bp of D-loop) yielded 155 polymorphic sites and a total number of mutations of 189 (excluding sites with gaps and missing data).

Bayesian and Maximum Parsimony phylogenetic analyses showed the same tree topologies and relationships between clades. However statistical support was very low in the MP analysis (to the exception of some clades; bootstrap values not shown).

We obtained seven well-supported clades (Figure 5): 1) *Canis adustus* (Side-striped jackal - in grey); 2) *C. latrans* (Coyote - in dark blue); 3) Eurasian *C. aureus* (Eurasian golden jackal - in purple); 4) *C. himalayensis* and/or *C. lupus chanco* (Himalayan wolf - in green); 5) *C. lupus lupaster*, including our North African *C. aureus* samples (African wolf - in yellow); 6) *C. indica* and/or *C. lupus pallipes* (Indian wolf - in cyan); 7) *C. lupus* sp. (Holarctic wolf – in red). *Cuon alpinus* (Dhole – in brown) appeared more closely related to the side-striped jackals, and *C. simensis* (Ethiopian wolf – in orange) grouped with Eurasian golden jackals, coyotes and grey wolves. However, we were not able to resolve phylogenetic relationships within this group. Eurasian golden jackals, coyotes and Ethiopian wolves were all more or less equally distant from grey wolves (Figure 6). Two additional subgroups were found in the grey wolf clade: African wolves and golden jackals were more closely related with the Himalayan grey wolves (but with

low BPP = 0,6), and Indian grey wolves clustered with the Holarctic wolves (well supported, BPP = 1). Further subdivisions were also observed in the: 1) *C. adustus* clade, with one group comprising individuals from Senegal and Guinea, and another with specimens from Benin and Ethiopia; and 2) Eurasian golden jackals, with distinct European and Indian lineages.

The smaller set of samples for which we also amplified the 12S and 16S genes allowed constructing two concatenated datasets: 1) 836 bp sequence set (239 bp 16S, 317 bp

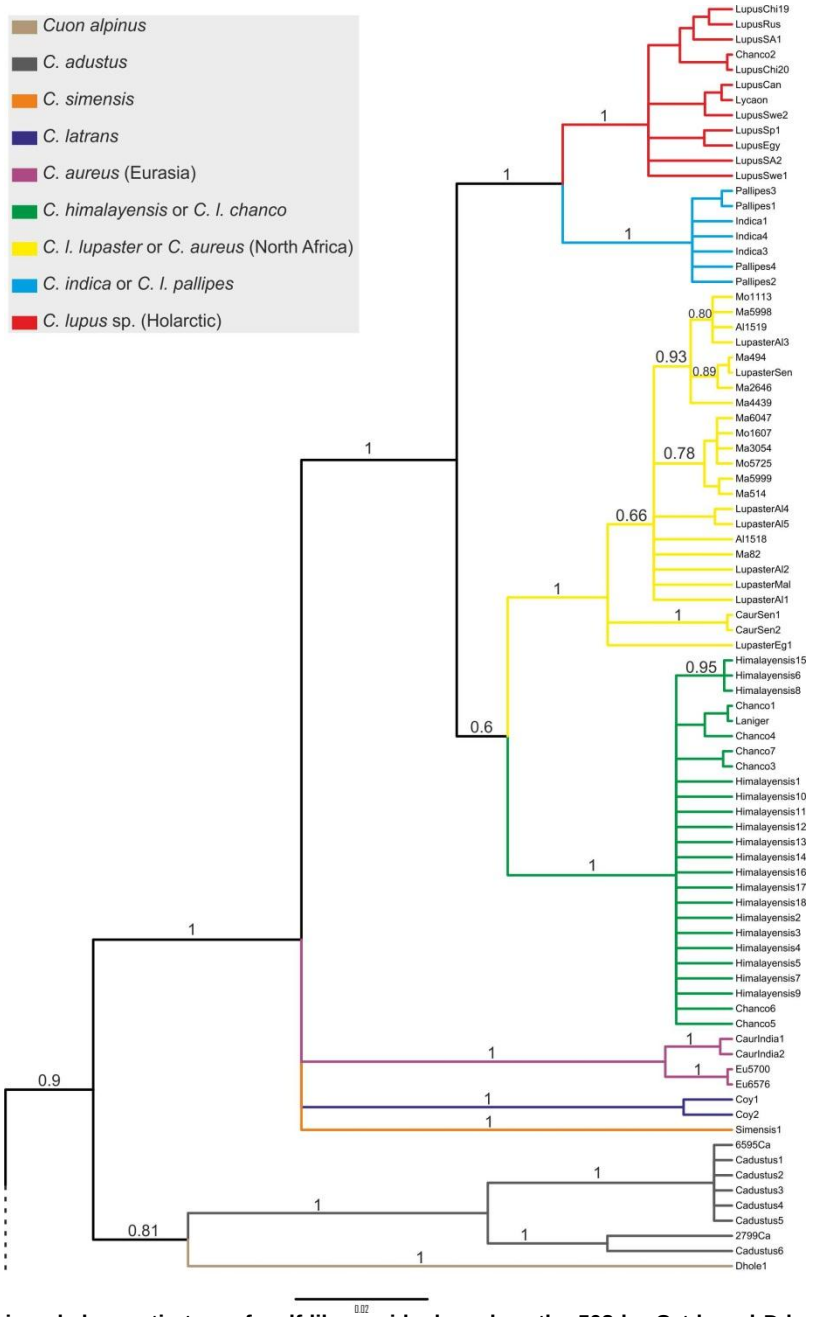


Figure 5 - Bayesian phylogenetic tree of wolf-like canids, based on the 508 bp Cyt-b and D-loop concatenated dataset. The HKY+G and HKY+I+G evolutionary models were applied to Cyt-b and D-loop partitions, respectively. *V. vulpes* (AM181037) was used as outgroup (not shown). Numbers above branches indicate Bayesian posterior probabilities (BPP). Scale bar represents 2% sequence divergence. Colors highlight clades/lineages referred in the main text: *Cuon alpinus* (dhole - brown); *Canis adustus* (side-striped jackal - grey); *C. simensis* (Ethiopian wolf – orange); *C. latrans* (coyote – dark blue); Eurasian *C. aureus* (golden jackal – purple); *C. himalayensis* or *C. lupus chanco* (Himalayan wolf – green); *C. l. lupaster* or North African *C. aureus* (African wolf or golden jackal – yellow); *C. indica* or *C. l. pallipes* (Indian wolf – cyan); *C. lupus sp.* (Holarctic wolf – red).

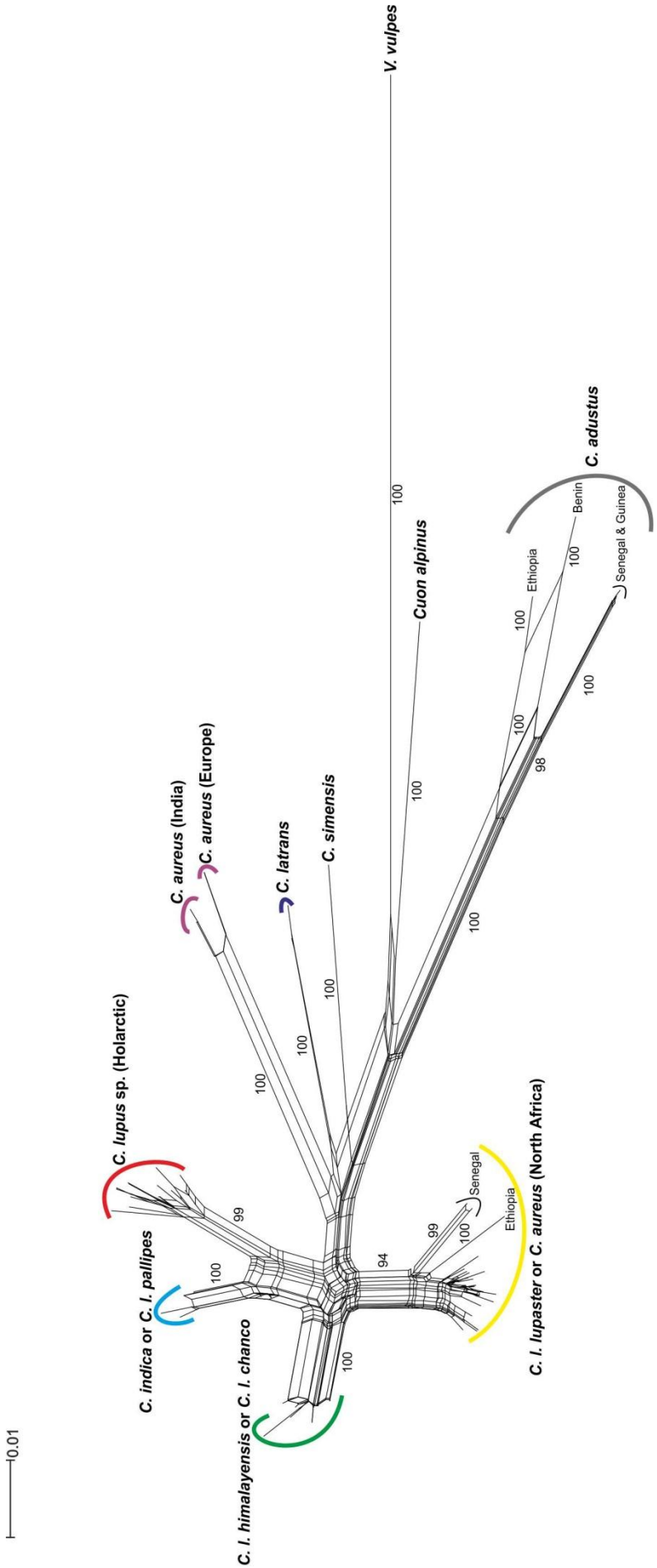


Figure 6 - Neighbor-net network using the 508 bp concatenated dataset, and based on uncorrected patristic distances as implemented in SplitsTree. Numbers indicate bootstrap values. Scale bar represents 1% sequence divergence. Highlighted are the clades referred in the main text as well as some relevant geographic subdivisions: *Cuon alpinus* (dhole); *Canis adustus* (side-striped jackal - grey); *C. simensis* (Ethiopian wolf); *C. latrans* (coyote – dark blue); Eurasian *C. aureus* (golden jackal – purple); *C. himalayensis* or *C. lupus chanco* (Himalayan wolf – green); *C. l. lupaster* or North African *C. aureus* (African wolf or golden jackal – yellow); *C. indica* or *C. l. pallipes* (Indian wolf – cyan); *C. lupus* sp. (Holarctic wolf – red).

Cyt-b and 230 Dloop); 2) 1168 bp sequence set (adding to the beginning of the previous set 332 bp of 12S). Both Bayesian phylogenetic trees retrieved the same topology and clades (although in the 1168 bp we did not had any sequence representing the Indian wolf lineage). Noteworthy is the fact that these trees (sections A11-12 in Annexes) show complete statistical support for the first subdivision within the grey wolf clade, splitting African wolves/ golden jackals and Himalayan wolves to one side, and Indian and Holarctic wolves to the other.

For the 508 bp concatenated dataset, we can discriminate a total of 56 haplotypes, of which 11 are entirely new - ten of *C. l. lupaster* / North African *C. aureus*, and one from Ethiopia of *C. adustus*. Relationships between haplotypes are depicted in Figure 7. Within the *C. l. lupaster*/ North African *C. aureus* clade haplotypes were represented by only one or two individuals and do not display clear geographical coherence, with some

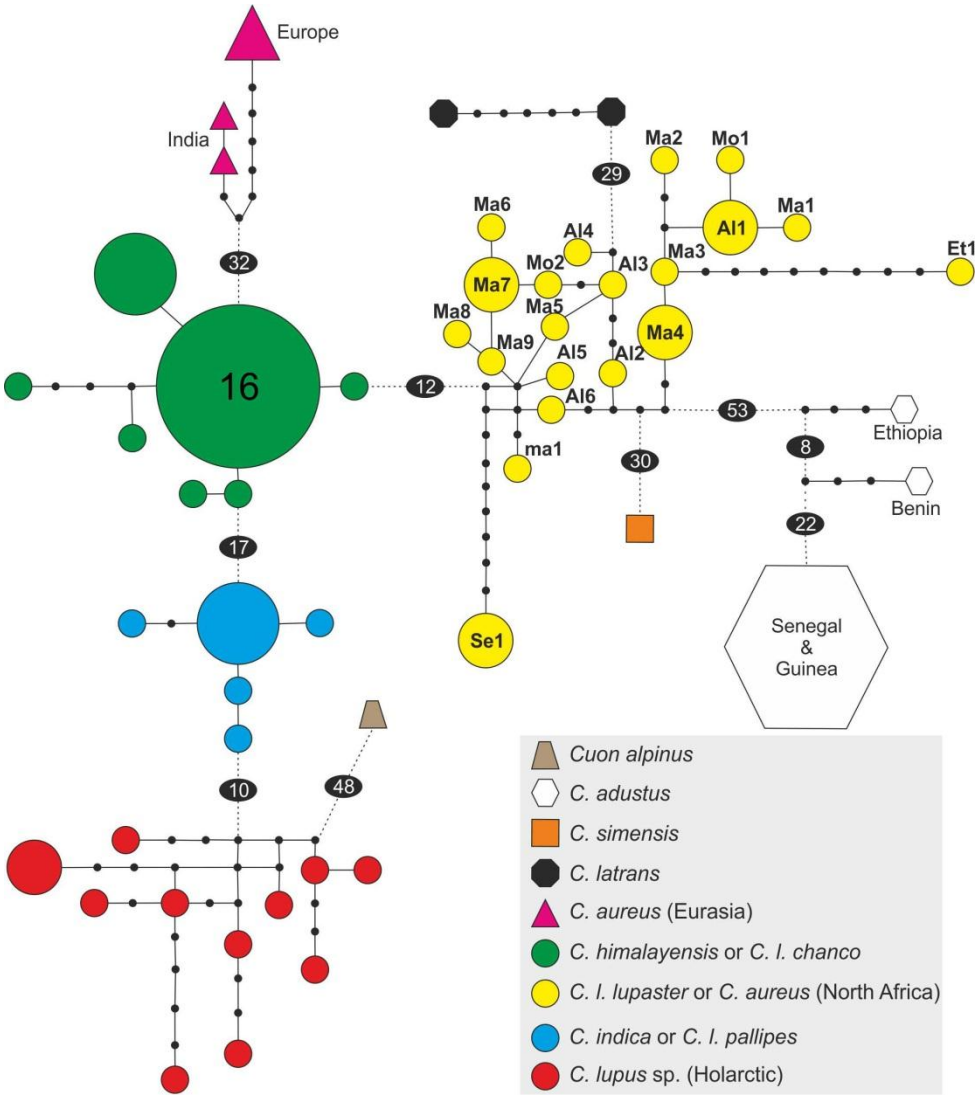


Figure 7 - Statistical parsimony haplotype network based on the 508 bp concatenated dataset and constructed by TCS. Figure sizes and branches are proportional to the number of shared individuals per haplotype (except for the most common haplotype of *C. himalayensis* where the number of sequences is indicated – 16) and number of mutational steps among haplotypes, respectively. Numbers in black background refer to mutation steps separating main clades. *C. l. lupaster*/ North African *C. aureus* haplotype names are displayed above or within corresponding circle. See section A13 in Annexes for a list of corresponding sequences.

haplotypes being shared by individuals from Mauritania and either Morocco (Ma7) or Senegal (Ma4). Sequences from Ethiopia (Et1) and Senegal (Se1) represent more distinct lineages, separated of the main haplotype cluster by eight and seven mutational steps, respectively. Genetic diversity of each clade is summarised in Table 1, where we highlight the significant levels of genetic diversity displayed by the *C. l. lupaster* / *C. aureus* (North Africa) clade.

A 301 bp Cyt-b dataset, for which we were able to incorporate more GenBank sequences (especially those of other jackal species), was used to construct an additional haplotype network. In Figure 8 we are able to identify all the same main groups of previous analyses. Additionally, we can point out a golden jackal sequence from Kenya completely isolated from both North-African and Eurasian specimens, and the fact that the two sequences of black-backed jackal (*C. mesomelas*) appeared in opposite sides of the network. The 301 bp Cyt-b dataset was also used to calculate genetic distances between the main clades (Table 2), corroborating the previous results and clearly showing the genetic distinctiveness of African jackal sequences and close relationship between wolf clades.

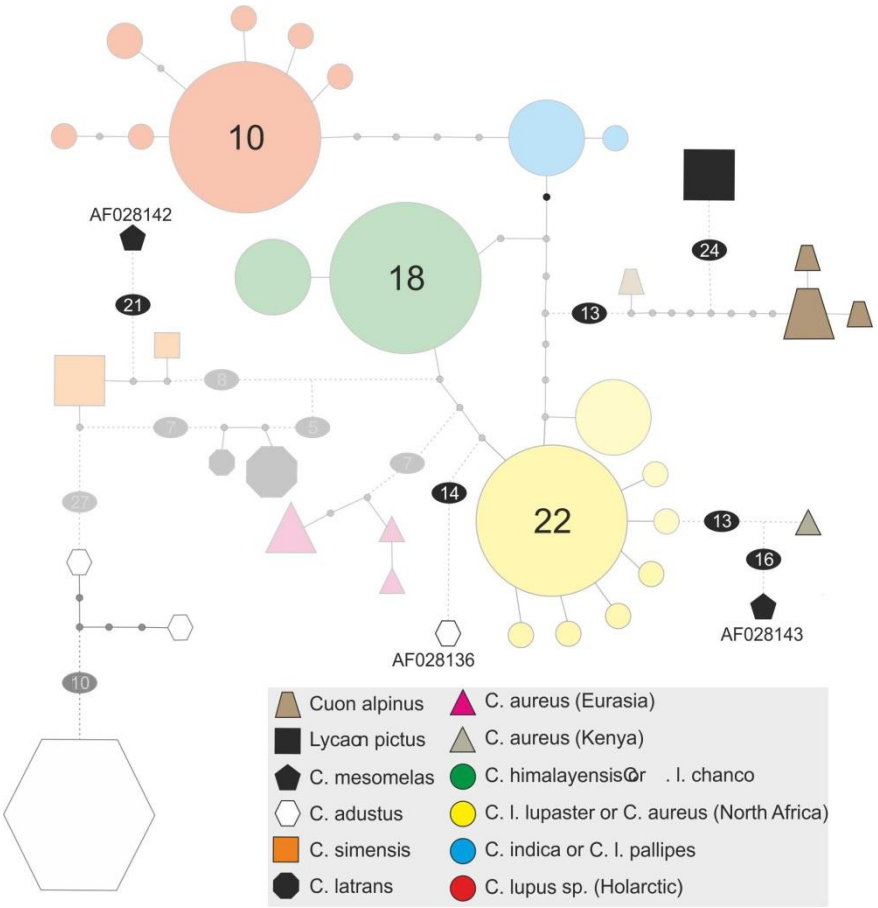


Figure 8 - Statistical parsimony haplotype network based on the 301 bp Cyt-b sequence dataset and constructed by TCS. Figures sizes and branches are proportional to the number of shared individuals per haplotype (except for most common haplotypes where number of sequences are indicated) and number of mutational steps among haplotypes, respectively. Numbers in black background also refer to mutational steps. Highlighted are the added GenBank sequences and number of mutational steps linking them.

Table 1 - Sequence diversity of the 508 bp concatenated sequence dataset (excluding sites with gaps and missing data). N – Number of sequences; S – Polymorphic sites; η - Number of mutations; π – Nucleotide diversity; K – Average number of nucleotide differences; H – Number of haplotypes; Hd – Haplotype diversity.

Taxon	N	S	η	π	K	H	Hd
C. adustus	8	36	37	0,025	12,79	3	0,464
***C. aureus* (Eurasia)**	4	9	9	0,012	5,833	3	0,833
C. himalayensis* / *C. l. chanco	25	9	9	0,002	1,147	7	0,587
***C. l. lupaster* / *C. aureus* (North Africa)**	24	33	33	0,012	5,938	20	0,986
C. indica* / *C. l. pallipes	7	5	5	0,003	1,169	5	0,857
***C. lupus* sp. (Holarctic)**	12	21	23	0,013	6,364	11	0,985

Table 2 – Mean genetic distances between wolf-like canids (below diagonal), based on the 301 bp Cyt-b dataset. Analysis was conducted in MEGA using the Kimura 2-parameter model. Standard error estimates are shown above the diagonal.

	Taxon	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1	*C. lupus* (Holarctic)		0,007	0,011	0,010	0,015	0,017	0,014	0,014	0,020	0,027	0,025	0,018	0,020	0,021
2	*C. indica* (India)	0,019		0,008	0,007	0,012	0,014	0,012	0,012	0,018	0,024	0,024	0,016	0,020	0,019
3	*C. l. lupaster* (North Africa)	0,041	0,022		0,007	0,011	0,013	0,012	0,012	0,018	0,023	0,022	0,013	0,019	0,019
4	*C. himalayensis* (Himalayas)	0,033	0,015	0,015		0,011	0,013	0,011	0,011	0,018	0,024	0,024	0,014	0,019	0,020
5	*C. aureus* (Eurasia)	0,067	0,048	0,042	0,041		0,015	0,013	0,012	0,019	0,024	0,023	0,015	0,019	0,021
6	*C. aureus* (Kenya)	0,079	0,060	0,053	0,053	0,065		0,016	0,017	0,020	0,023	0,023	0,015	0,020	0,018
7	*C. latrans*	0,063	0,045	0,046	0,040	0,050	0,072		0,010	0,019	0,024	0,024	0,016	0,019	0,020
8	*C. simensis*	0,061	0,042	0,043	0,040	0,047	0,082	0,039		0,020	0,025	0,021	0,016	0,018	0,019
9	*Cuon alpinus*	0,103	0,090	0,090	0,090	0,100	0,110	0,100	0,105		0,020	0,022	0,020	0,021	0,021
10	*Lycaon pictus*	0,158	0,137	0,129	0,137	0,140	0,130	0,139	0,139	0,107		0,024	0,023	0,022	0,022
11	*C. adustus*	0,162	0,151	0,135	0,143	0,137	0,142	0,146	0,121	0,132	0,145		0,022	0,023	0,022
12	*C. adustus* (AF028136)	0,099	0,079	0,057	0,064	0,064	0,067	0,083	0,078	0,106	0,130	0,136		0,020	0,022
13	*C. mesomelas* (AF028142)	0,111	0,099	0,092	0,091	0,097	0,106	0,101	0,082	0,117	0,123	0,140	0,109		0,019
14	*C. mesomelas* (AF028143)	0,118	0,099	0,090	0,098	0,112	0,086	0,109	0,090	0,112	0,126	0,129	0,121	0,094	

3.2.2. Fox-like canids

For developing phylogenetic analyses, we used a 586 bp concatenated sequence dataset that consisted of 364 bp of Cyt-b and 222 bp of D-loop, holding 167 polymorphic sites and a total number of mutations of 188 (excluding sites with gaps or missing data).

Again, MP revealed the same tree topology as in the Bayesian analysis but with lower support values (bootstrap values not shown). In Figures 9 and 10 we can delineate three highly supported clades corresponding to three of our study species: *V. zerda*, *V. pallida* and *V. vulpes*. Branches connecting these three species were not very well supported, most likely due to the rapid genetic differentiation between them (Figures 9 and 10). The unexpected outcome, however, is seeing that *V. rueppellii* samples constitute a clear separated lineage but within the red fox group. Moreover, the *V. rueppellii* appeared more closely related to one of the three Japanese red foxes groups, with high BPP value. Looking closer to the *V. vulpes* clade, we notice a first split separating: 1) North Africa, Hokkaido 2 and *V. rueppellii* on one side (although branch support for this clade was low); 2) Eurasian red foxes and also one sample from Egypt (with full statistical support). Within the Maghreb there were signs of at least two, possibly three, further subdivisions, with samples from the Western Atlas [Maghreb 2 (Western Atlas)] clustering together, and then two more widespread lineages (Maghreb 1). A European lineage, which includes sequences from the Iberian Peninsula, Sweden, Greece and Turkey, is also supported but inside this clade there were no clear biogeographical patterns. European and similar haplotypes found in Japan (Hokkaido 1a) belong all to the same clade and appear within the European lineage. The other two clades from Japan (Hokkaido 1b and Honshu/Kyushu) constitute, together with samples from far East Russia (Primorsky Krai) and one from China, an Asian group that however includes one sample from Egypt and does not include a Korean and Armenian sequence.

The 586 bp concatenated data set was also used to construct a TCS haplotype network (Figure 11), where seven haplotypes belonged to *V. zerda*, 14 to *V. pallida* and 49 to *V. vulpes*, including eight of *V. rueppellii*, giving a total number of 70 haplotypes. All clades previously mentioned were retrieved and high numbers of differences between *V. zerda*, *V. pallida* and *V. vulpes*/*V. rueppellii* were observed. The amount of mutational steps separating groups within the red fox was also significant.

Genetic diversity of each species and main groupings of the red fox was also calculated (Table 3) and all species and clades, to the exception of Hokkaido 2 and Hokkaido 1a, show significant levels of genetic diversity. Kimura-2 parameter genetic distances were calculated based on the Cyt-b 364 bp partition of the concatenated dataset (Table 4), which further reinforces our results.

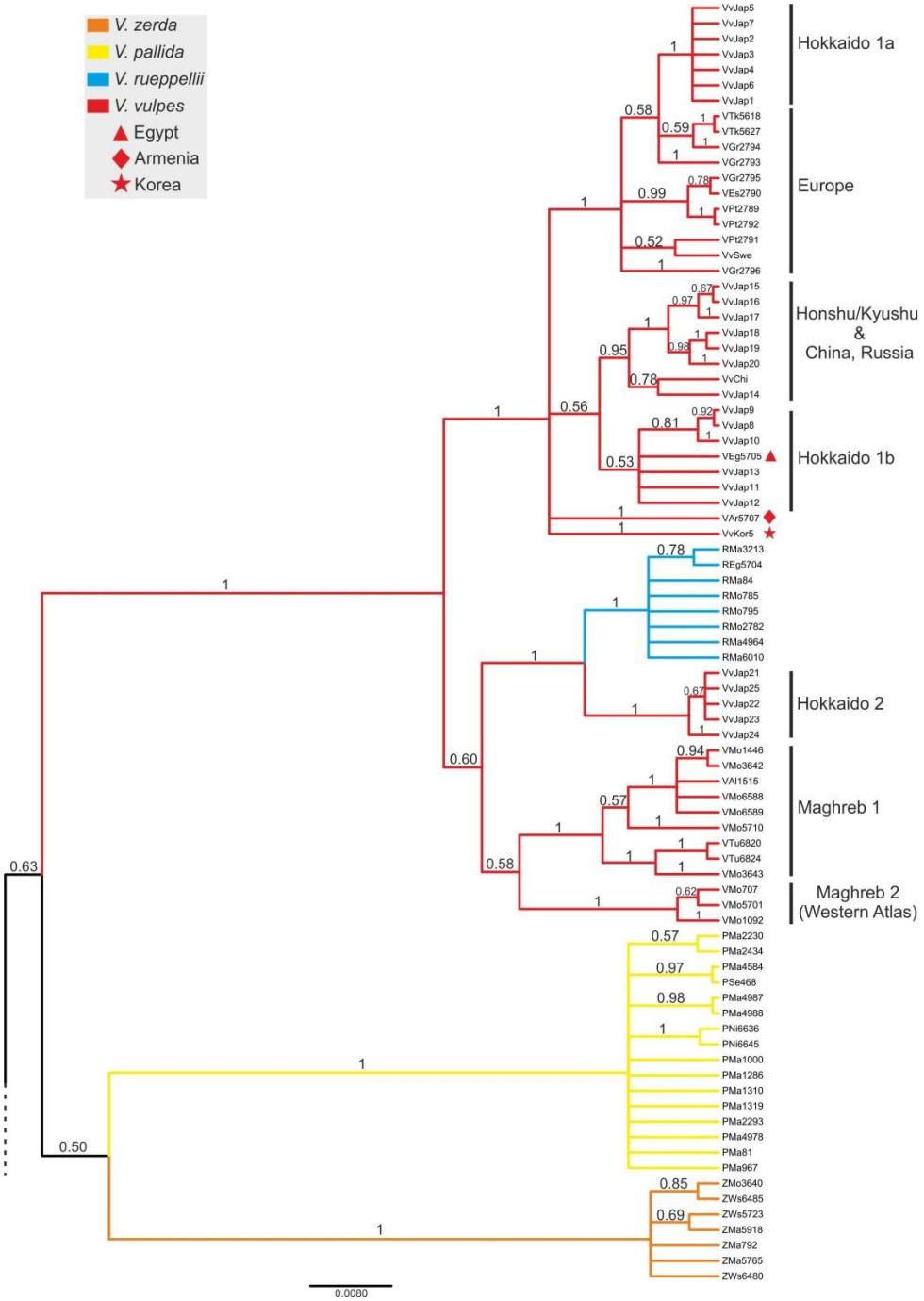


Figure 9 - Bayesian phylogenetic tree of fox-like canids, based on the 586 bp concatenated dataset. The GTR+I and GTR+I+G evolutionary models were applied to Cyt-b and D-loop partitions, respectively. *C. latrans* (DQ480509) was used as outgroup (not shown). Numbers above branches indicate Bayesian posterior probabilities (BPP). Scale bar represents 0.8% sequence divergence. Colors highlight the study species: *V. zerda* (Fennec fox - orange); *V. pallida* (Pale fox – yellow); *V. rueppellii* (Rüppell's fox – cyan); *V. vulpes* (Red fox – red). Black vertical bars (right side) with adjacent text limit *V. vulpes* lineages/clades discussed in the main text.

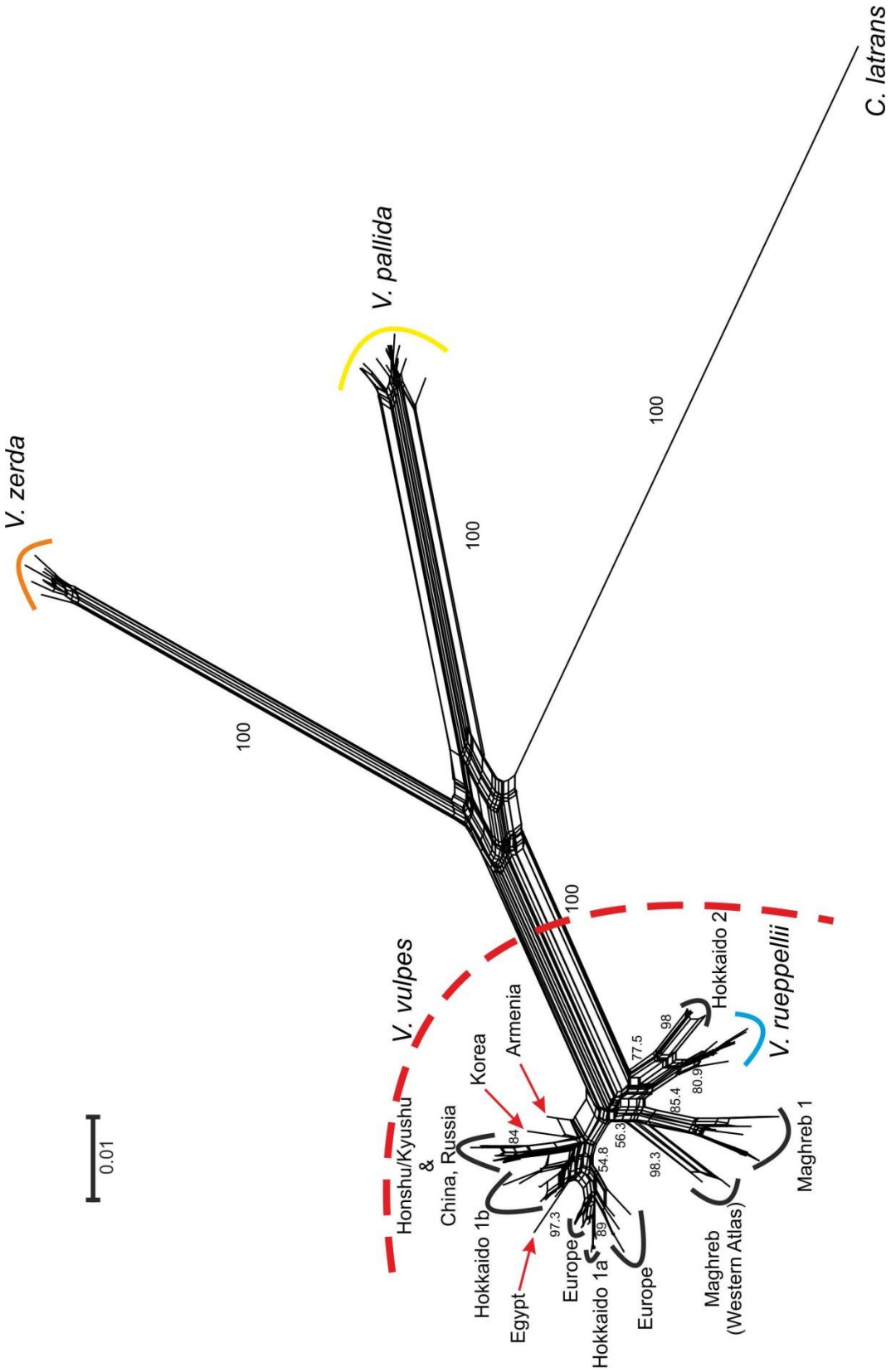


Figure 10 - Neighbor-net network using the 586 bp concatenated dataset, and based on uncorrected patristic distances as implemented in SplitsTree. Numbers indicate bootstrap values. Scale bar represents 1% sequence divergence. Highlighted are the clades referred in the main text.

Table 3 - Sequence diversity of the 586 bp concatenated sequence dataset (excluding sites with gaps and missing data). N – Number of sequences; S – Polymorphic sites; η - Number of mutations; π – Nucleotide diversity; K – Average number of nucleotide differences; H – Number of haplotypes; Hd – Haplotype diversity.

Species/Clade	N	S	η	π	K	H	Hd
V. zerda	7	13	13	0,007	4,095	7	1,000
V. pallida	16	21	22	0,006	3,592	14	0,983
V. vulpes/ V. rueppellii	60	60	61	0,025	13,10	46	0,985
V. rueppellii	8	11	12	0,006	3,321	8	1,000
Hokkaido 2	5	1	1	0,001	0,400	2	0,400
Maghreb 1	9	17	17	0,011	6,611	9	1,000
Maghreb 2 (Western Atlas)	3	2	2	0,003	1,333	3	1,000
Hokkaido 1b	6	4	4	0,003	1,800	5	0,933
Honshu/Kyoshu & China, Russia	8	15	15	0,010	5,893	7	0,964
Europe	11	14	14	0,008	4,400	8	0,945
Hokkaido 1a	7	2	2	0,001	0,571	3	0,524

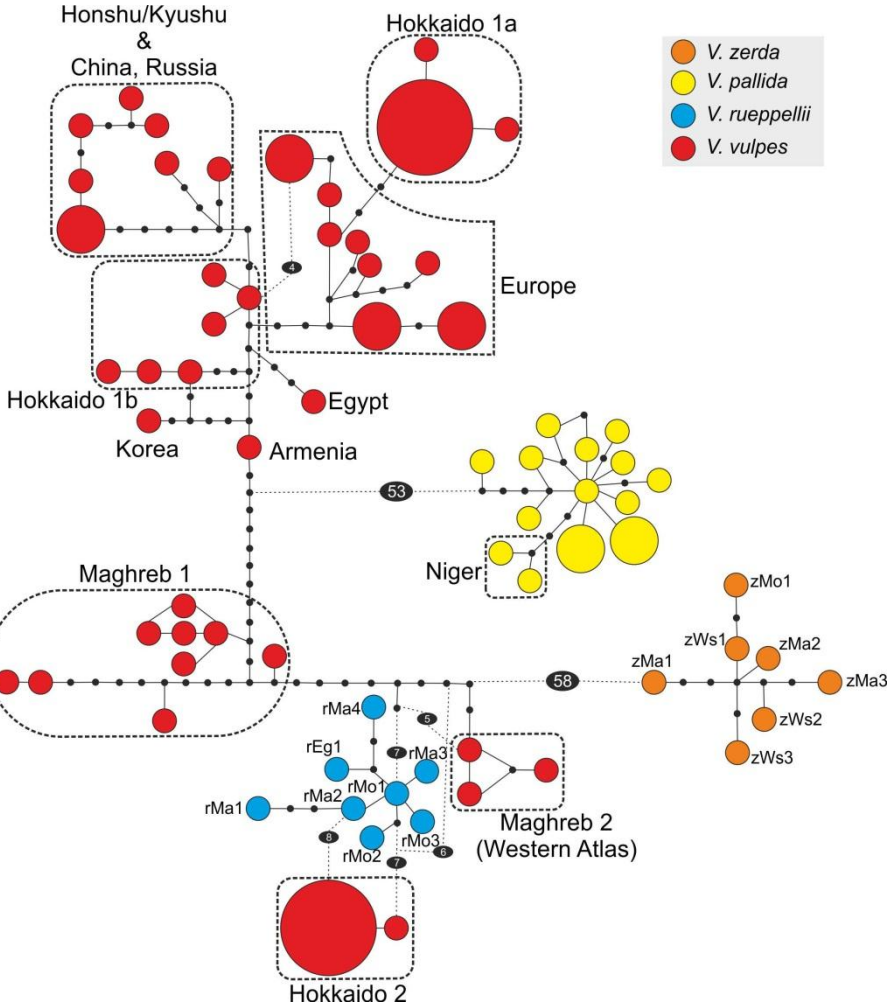


Figure 11 - Statistical parsimony haplotype network based on the 586 bp concatenated dataset and constructed by TCS. Figure sizes and branches are proportional to the number of shared individuals per haplotype and number of mutational steps among haplotypes, respectively. Numbers in black background also refer to number of mutation steps. Within *V. pallida* all haplotypes are found in Mauritania, except the two Niger ones that appear signalized. *V. rueppellii and V. zerda* haplotype names are displayed. See section A13 in Annexes for a list of corresponding sequences. Also highlighted are the main *V. vulpes* clades mentioned in the main text.

Table 4 - Mean genetic distances between fox-like canids and V. vulpes clades (below diagonal), based on the 364 bp Cyt-b partition of the concatenated dataset. Analysis was conducted in MEGA using the Kimura 2-parameter model. Standard error estimates are shown above the diagonal.

	Taxon	1	2	3	4	5	6	7	8	9	10	11	12	13
1	Hokkaido 1a		0,003	0,005	0,004	0,006	0,006	0,006	0,007	0,009	0,007	0,009	0,022	0,026
2	Europe	0,004		0,005	0,003	0,006	0,006	0,005	0,007	0,009	0,007	0,008	0,023	0,025
3	Honshu/Kyushu	0,014	0,012		0,004	0,006	0,006	0,006	0,007	0,009	0,008	0,009	0,022	0,026
4	Hokkaido 1b	0,007	0,006	0,010		0,005	0,005	0,003	0,006	0,008	0,007	0,008	0,022	0,025
5	Egypt	0,014	0,012	0,017	0,010		0,004	0,006	0,008	0,009	0,007	0,008	0,022	0,025
6	Armenia	0,014	0,012	0,017	0,010	0,006		0,006	0,007	0,009	0,006	0,008	0,022	0,025
7	Korea	0,011	0,009	0,014	0,005	0,014	0,014		0,007	0,009	0,008	0,009	0,023	0,025
8	*V. rueppellii*	0,021	0,019	0,024	0,016	0,024	0,024	0,020		0,005	0,006	0,006	0,021	0,025
9	Hokkaido 2	0,031	0,029	0,034	0,026	0,034	0,034	0,031	0,013		0,008	0,008	0,022	0,027
10	Maghreb 1	0,024	0,022	0,026	0,022	0,023	0,017	0,026	0,018	0,029		0,007	0,022	0,026
11	Western Atlas	0,028	0,026	0,031	0,024	0,025	0,025	0,028	0,017	0,027	0,019		0,022	0,025
12	*V. pallida*	0,144	0,149	0,143	0,145	0,140	0,140	0,151	0,135	0,140	0,144	0,142		0,024
13	*V. zerda*	0,174	0,170	0,174	0,167	0,162	0,162	0,166	0,164	0,178	0,170	0,159	0,165	

3.3. Microsatellites

Genotyping success was not identical for all studied species (see section A3 in Annexes). While for *C. aureus* all 46 *loci* were successfully amplified, the number of *loci* for the foxes varied between species. Information on the datasets of each species and pre-evaluation procedures are represented in Table 5.

3.3.1. Golden jackal

For our *C. aureus* dataset, comprising North African and European individuals, the number of alleles per *locus* varied between 6 (FHC2010, INU005) and 18 (FH2161). Genetic indexes are summarised in Table 6, showing a relatively high level of Ho.

Structure Harvester results point out K=2 as the most probable number of clusters to explain our dataset (Figure 12A). Structure bar plots showed (Figure 12B) a clear separation between North African *C. aureus* (in red, average proportion of membership Q_1=0,997) and the two European samples (in green, Q_2=0,996). Within North Africa no further subdivisions of genetic diversity were discernible (K>2, section A14 in Annexes). Principle coordinates analysis (PCA) based on genetic distances suggested a considerable genetic distance between the two clusters, and lack of coherence between genetic and geographic distance within North Africa (Figure 12C).

So taking now into account these two clusters, and despite the obvious statistical limitations of only having two individuals representing the European cluster, we observed that the number of alleles per locus varied between 6 (FHC2010, INU005,

Table 5 – Summary of microsatellite genotyping and pre-evaluation. N – Number of samples; N (MD) – Number of samples excluded with missing data above 20%; Final N – Final number of samples; L – Number of loci genotyped; L (HW) – Number of loci excluded in the Hardy – Weinberg equilibrium test; L (LD) – Number of loci excluded in the linkage disequilibrium assessment; Final L – Final number of loci.

Species	N	N (MD)	Final N	L	L (HW)	L (LD)	Final L
C. aureus	35	6	29	46	12	0	34
V. pallida	28	7	21	36	8	0	28
V. rueppellii	10	3	7	34	5	0	29
V. vulpes	43	13	30	34	8	0	26
V. zerda	11	2	9	31	3	0	28

Table 6 – Mean Heterozygosity, F-statistics and Polymorphism data for *C. aureus*. SE – Standard error; N - Sample Size; Na – No. of Alleles; Ne - No. of Effective Alleles; Ho - Observed Heterozygosity; He - Expected Heterozygosity; UHe - Unbiased Expected Heterozygosity; F - Fixation Index.

		N	Na	Ne	Ho	He	UHe	F
C. aureus	Mean	27,735	9,588	6,180	0,727	0,819	0,834	0,112
	SE	0,287	0,435	0,349	0,019	0,011	0,011	0,020

INU055) and 14 (C22.763) within North African samples, while many *loci* were monomorphic for European individuals and the maximum possible number of four alleles per *locus* was found for two markers (AHT111, CPH9). Genetic diversity was again calculated as well as F-statistics (Table 7), which indicated a great differentiation

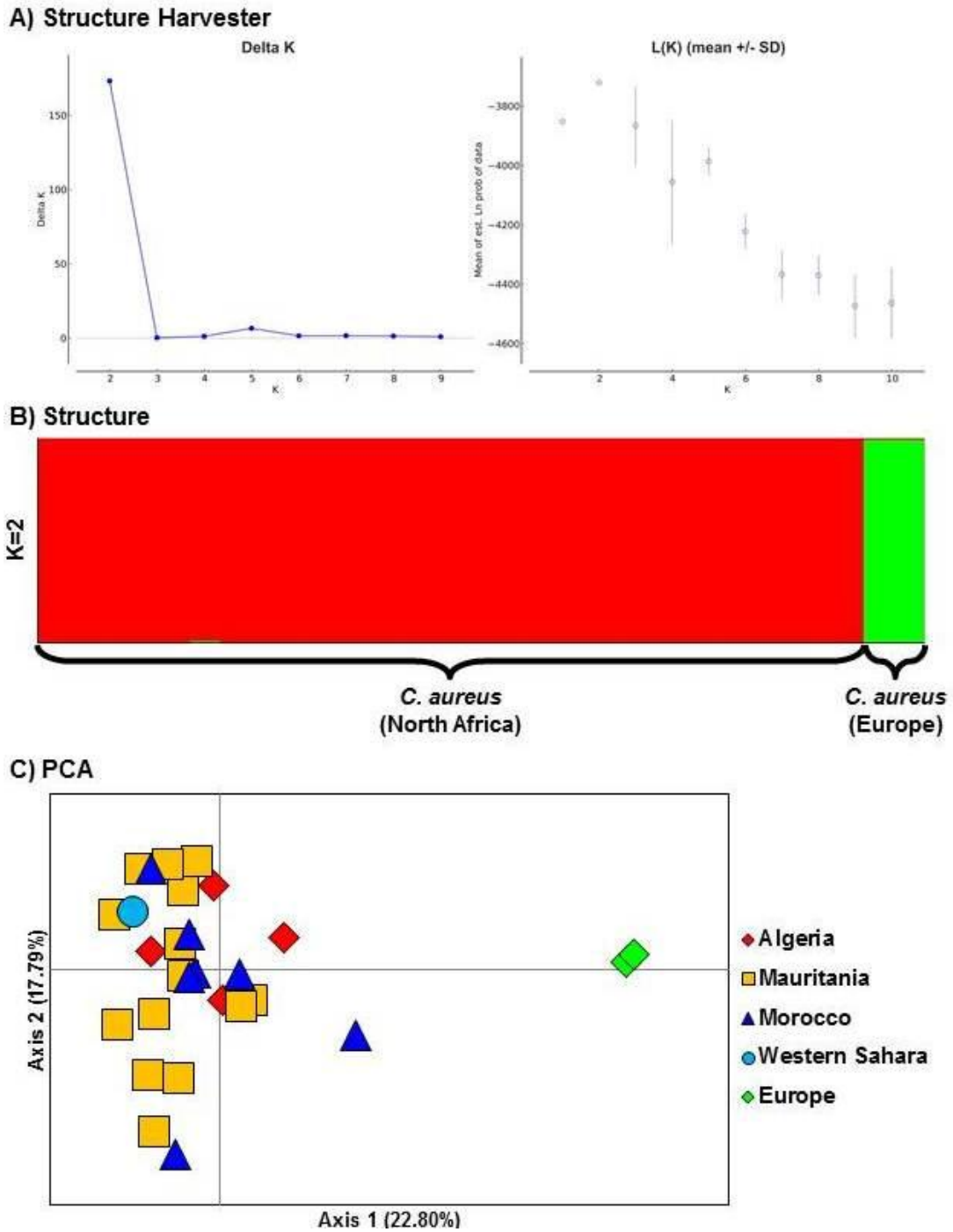


Figure 12 - Population structure analyses of 35 *C. aureus* samples based on 34 microsatellite *loci*. A) STRUCTURE HARVESTER graphic output of Delta K and Mean L(K); B) STRUCTURE bar plot of Bayesian assignment of individuals to two clusters (K=2). Vertical bars represent individuals, while colors within a bar represent probability of assignment of each individual to a cluster – North Africa (red) and Europe (green). C) Principal coordinate analysis (PCA) based on individual-by-individual genetic distances.

between North Africa and Europe, with 21.7% of the genetic variability being explained due to differences between the two clusters (again we have to be very cautious with these results because of the low sample size representing Europe).

Given the results obtained with the mitochondrial DNA sequences, a structure analysis (27 *loci*) was conducted adding to our golden jackal dataset a group of 14 grey wolves and eight dogs, to see how the different groups would segregate. STRUCTURE HARVESTER results suggested K=2 as the most likely scenario to explain this dataset. Looking at STRUCTURE bar plots of the indicated number of clusters it corresponds to a first separation between *C. aureus* (North Africa and Europe) to one side and *C. lupus/familiaris* to the other (Figure 13). Also for K=2 it became evident that one of our samples from Africa (initially attributed to a jackal) was clearly identified as a dog, which further strengthens our analysis. As we increased the number of clusters, grey wolves tended to separate from dogs at K=3, and then finally European golden jackals segregated from North African ones.

The PCA (Figure 14) reinforced the STRUCTURE results, where we see North African and European C. aureus closer together than any of these two groups with grey wolves or dogs. It also visibly clustered our supposed jackal specimen with the dogs. All around these results seem to refute the previous outcomes from the mtDNA analysis.

Table 7 – Mean Heterozygosity, F-statistics and Polymorphism data for *C. aureus* – North Africa, Europe and Total dataset. SE – Standard error; N - Sample Size; Na – No. of Alleles; Ho - Observed Heterozygosity; He - Expected Heterozygosity; F_{IS} - Inbreeding Coefficient; F_{IT} – Total Fixation Index; F_{ST} – Fixation Index.

Cluster		N	Na	Ho	He	F_{IS}	F_{IT}	F_{ST}
North Africa	Mean	25,79	9,206	0,751	0,812	-	-	-
	SE	0,273	0,393	0,020	0,012	-	-	-
Europe	Mean	1,941	1,941	0,397	0,338	-	-	-
	SE	0,041	0,146	0,069	0,044	-	-	-
Total	Mean	13,868	5,574	0,574	0,575	0,002	0,219	0,217
	SE	1,463	0,490	0,042	0,037	0,040	0,038	0,020

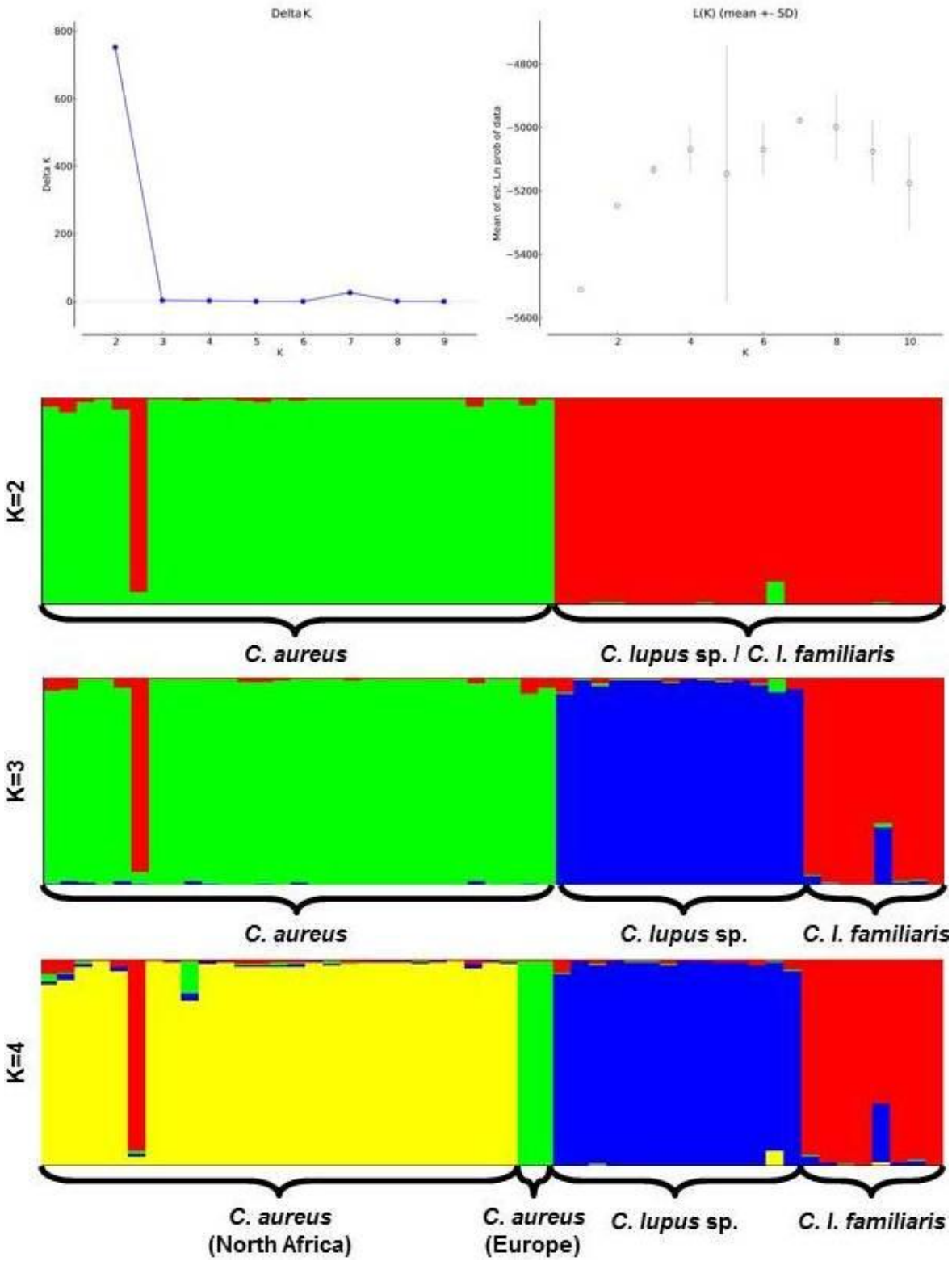


Figure 13 – STRUCTURE HARVESTER graphic output of Delta K and Mean L(K), and STRUCTURE bar plot of Bayesian assignments of *C. aureus* (North Africa and Europe), *C. lupus* sp. and *C. l. familiaris* individuals to two (K=2), three (K=3) and four (K=4) clusters. Analysis was based on 27 microsatellite *loci*. Vertical bars represent individuals, while colors within a bar represent probability of assignment of each individual to a cluster.

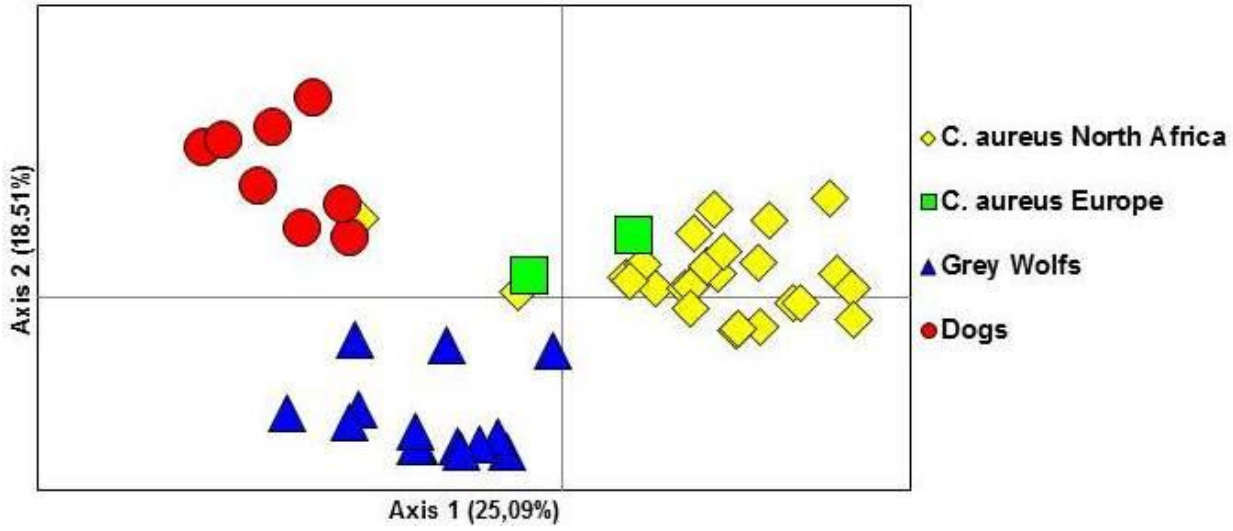


Figure 14 - PCA based on individual-by-individual genetic distances of 27 microsatellite loci of the combined dataset of *C. aureus* (North Africa and Europe), *C. lupus* sp. (grey wolves) and *C. l. familiaris* (dogs).

3.3.2. Desert foxes (V. pallida, V. rueppellii and V. zerda)

The number of alleles per *locus* varied between two (INU005, C14.866) and 12 (PEZ3, AHT121, FH2054FZ, REN54P11) in *V. pallida*, between two (AHTh171, INU030, REN169D01) and nine (PEZ3) in *V. rueppellii*, and between two (FHC2010, C08.140, CPH9) and 12 (AHTk253) in *V. zerda*. Genetic diversity indexes were calculated and are depicted in Table 8. Some precautions must be taken concerning the results for *V. rueppellii* and *V. zerda* due to the small sample size.

Population structure analysis conducted with STRUCTURE did not yield any signs of genetic subdivision for any of the desert foxes, *V. pallida*, *V. rueppellii* and *V. zerda*. For all cases, although Delta K does indicate a K number of clusters, when we check the bar plot for the respective K value, no population structuring is discernible.

Table 8 – Mean Heterozygosity, F-statistics and Polymorphism data for *V. pallida*, *V. rueppellii* and *V. zerda*. SE – Standard error; N - Sample Size; Na – No. of Alleles; Ne - No. of Effective Alleles; Ho - Observed Heterozygosity; He - Expected Heterozygosity; Uhe - Unbiased Expected Heterozygosity; F - Fixation Index.

Species		N	Na	Ne	Ho	He	UHe	F
V. pallida	Mean	19.786	8.036	5.089	0.686	0.737	0.757	0.068
	SE	0.483	0.588	0.438	0.039	0.032	0.033	0.036
V. rueppellii	Mean	6.517	4.655	3.285	0.660	0.635	0.689	-0.046
	SE	0.154	0.319	0.272	0.041	0.030	0.033	0.039
V. zerda	Mean	8.536	6.036	4.371	0.710	0.720	0.765	0.009
	SE	0.141	0.441	0.339	0.037	0.029	0.031	0.038

We show the example of *V. pallida* to illustrate this situation (Figure 15).

This lack of clear genetic subdivision was further enlightened with the PCA of each species (Figures 16, 17 and 18). Samples of all desert foxes are somewhat randomly placed and genetic distance does not seem coherent with geographic distance: samples of *V. pallida* from Niger are shown amongst Mauritanian ones; *V. rueppellii* from Mauritania are more genetically distant from each other than to individuals of Morocco or Egypt, and the same applies for *V. zerda*.

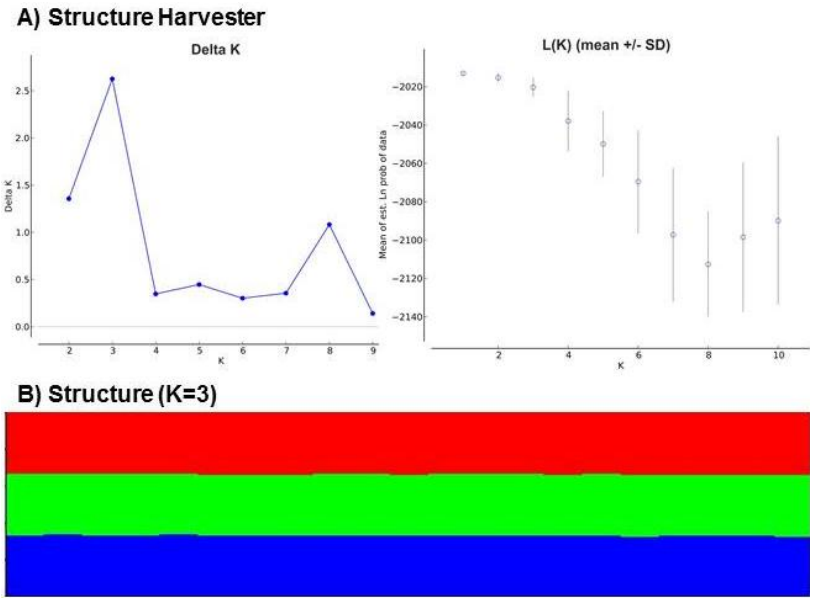


Figure 15 – Example for *V. pallida* of discrepancy between Delta K and Structure. Population structure analysis was conducted for 21 *V. pallida* samples based on 28 microsatellite loci. A) STRUCTURE HARVESTER graphic output of Delta K and Mean L(K); B) STRUCTURE bar plot of Bayesian assignment of individuals to three clusters (K=3). Vertical bars represent individuals, while colors within a bar represent probability of assignment of each individual to a cluster.

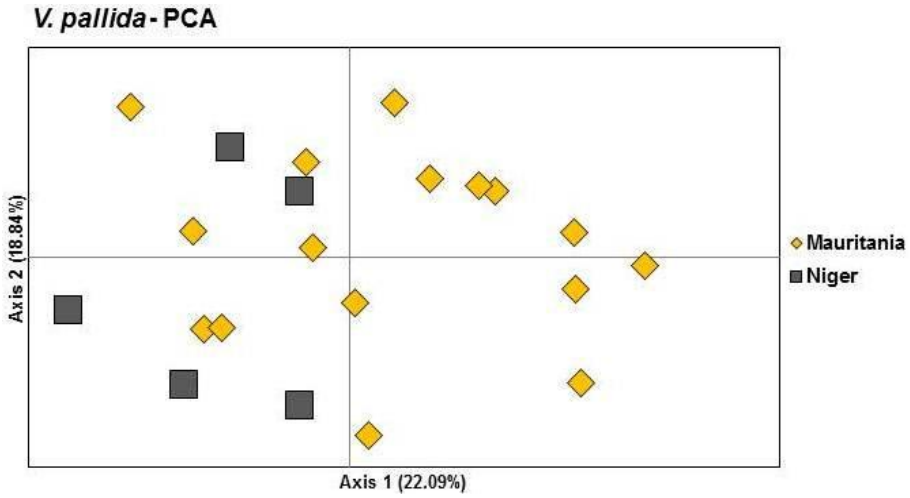


Figure 16 – PCA for *V. pallida* (28 *loci*) based on individual-by-individual genetic distances.

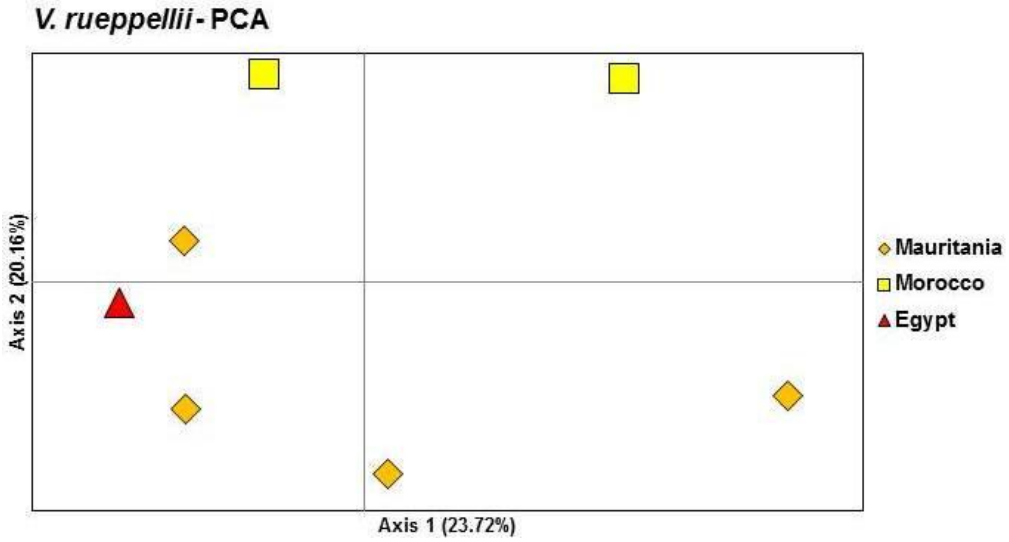


Figure 17 – PCA for *V. rueppellii* (29 *loci*) based on individual-by-individual genetic distances.

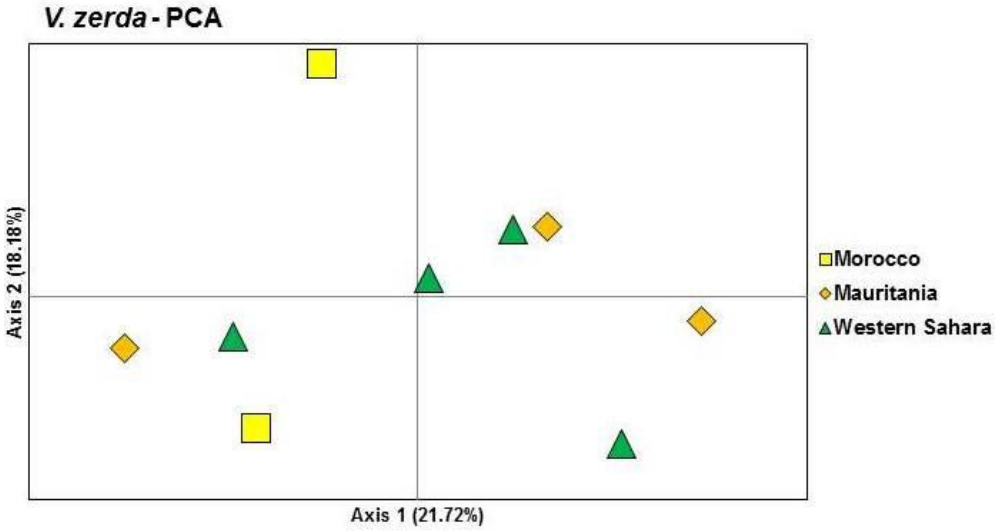


Figure 18 – PCA for *V. zerda* (28 *loci*) based on individual-by-individual genetic distances.

3.3.3. Red fox

For the *V. vulpes* dataset that included samples from both North Africa and Europe, number of alleles per *locus* varied between two (VWF, C08.140) and 16 (FH2054FZ, REN54P11), with genetic indexes displayed in Table 9.

Table 9 – Mean Heterozygosity, F-statistics and Polymorphism data for *V. vulpes*. SE – Standard error; N - Sample Size; Na – No. of Alleles; Ne - No. of Effective Alleles; Ho - Observed Heterozygosity; He - Expected Heterozygosity; UHe - Unbiased Expected Heterozygosity; F - Fixation Index.

		N	Na	Ne	Ho	He	UHe	F
V. vulpes	**Mean**	28,731	9,000	5,510	0,638	0,710	0,723	0,097
	SE	0,432	0,845	0,562	0,049	0,051	0,052	0,021

Population structure analysis of *V. vulpes* indicated K=2 as the most probable number of clusters to explain our dataset, either using Delta K or the mean L(K) (Figure 19A). The bar plot (Figure 19B) shows a cluster number one corresponding to a North African population that includes samples from Morocco, Algeria, Tunisia and Egypt (in red, Q_1=0.960) and cluster number two corresponding to an European group that encompasses individuals from the Iberian Peninsula, Greece, Turkey and Armenia (in green, Q_2=0.979). Higher K values did not showed clear further subdivisions. PCA analysis on genetic distance further reinforces a distinction between North African and European *V. vulpes*, however, a group of samples from Portugal appeared more distant from the remaining samples of Europe (including one more individual from Portugal), then this group to North African red foxes (Figure 19C).

Within North Africa, two loci were monomorphic (VWF, C08.140) and the maximum number of alleles per *locus* was 13 (FH2054FZ), while in Europe number of alleles per *locus* varied between two (AHTh171, VWF, C08.140, C20.446) and also 13 (FH2054FZ). Genetic diversity of each cluster (Table 10) showed similar allelic patterns, with a slightly higher proportion in European samples (Figure 19D). Mean F-statistics for the entire dataset (Table 10) showed low F_{IS} and F_{IT} values, indicating small differentiation between individuals of the same group, and low differentiation between North Africa and Europe, respectively. Mean value F_{ST} points out in the same direction suggesting little too moderate degree of genetic differentiation among the two groups of *V. vulpes*.

Table 10 - Mean Heterozygosity, F-statistics and Polymorphism data for *V. vulpes* – North Africa, Europe and Total dataset. SE – Standard Error; N - Sample Size; Na – No. of Alleles; Ne - No. of Effective Alleles; Ho - Observed Heterozygosity; He - Expected Heterozygosity; UHe – Unbiased Expected Heterozygosity; F_{IS} - Inbreeding Coefficient; F_{IT} – Total Fixation Index; F_{ST} – Fixation Index.

Cluster		N	Na	Ne	Ho	He	UHe	F_{IS}	F_{IT}	F_{ST}
North Africa	Mean	12,731	6,615	4,303	0,680	0,661	0.757	-	-	-
	SE	0,118	0,659	0,425	0,057	0,052	0.033	-	-	-
Europe	Mean	16,000	7,269	4,738	0,610	0,686	0.689	-	-	-
	SE	0,408	0,639	0,459	0,047	0,051	0.033	-	-	-
Total	Mean	14,365	6,942	4,521	0,645	0,673	0.765	0,036	0,089	0,055
	SE	0,311	0,457	0,311	0,037	0,036	0.031	0,016	0,020	0,012

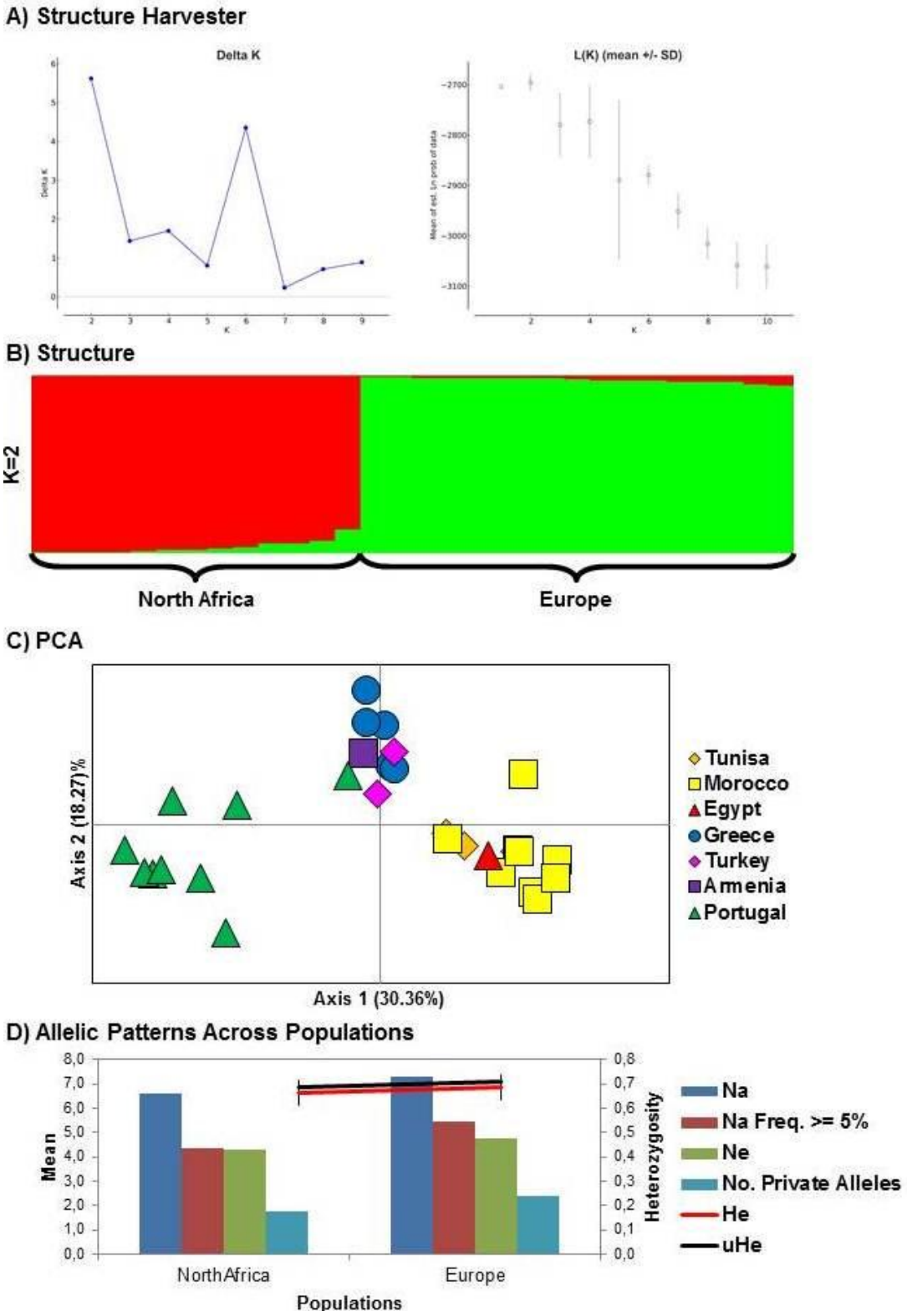


Figure 19 - Population structure analyses of 30 *V. vulpes* samples based on 26 microsatellite *loci*. A) STRUCTURE HARVESTER graphic output of Delta K and Mean L(K); B) STRUCTURE bar plot of Bayesian assignment of individuals to two clusters (K=2). Vertical bars represent individuals, while colors within a bar represent probability of assignment of each individual to a cluster – North Africa (red) and Europe (green). C) PCA based on individual-by-individual genetic distances; D) Allelic patterns across populations: Na – No. of different alleles; Na Freq. ≥ 5% - No. of different alleles with a frequency above 5%; Ne - No. of effective alleles; No. Private Alleles - No. of alleles unique to a single population; He – Expected heterozygosity; UHe – Unbiased expected heterozygosity.

3.3.4. Hybridization assessment

STRUCTURE analysis and PCA of genetic distances of a combined dataset of *V. rueppellii* and *V. pallida*, with 22 *loci*, showed no signs of mixture between these two species, each one representing clear distinct entities (Figure 20). Again we see some disagreement between STRUCTURE HARVESTER outputs and remaining analyses.

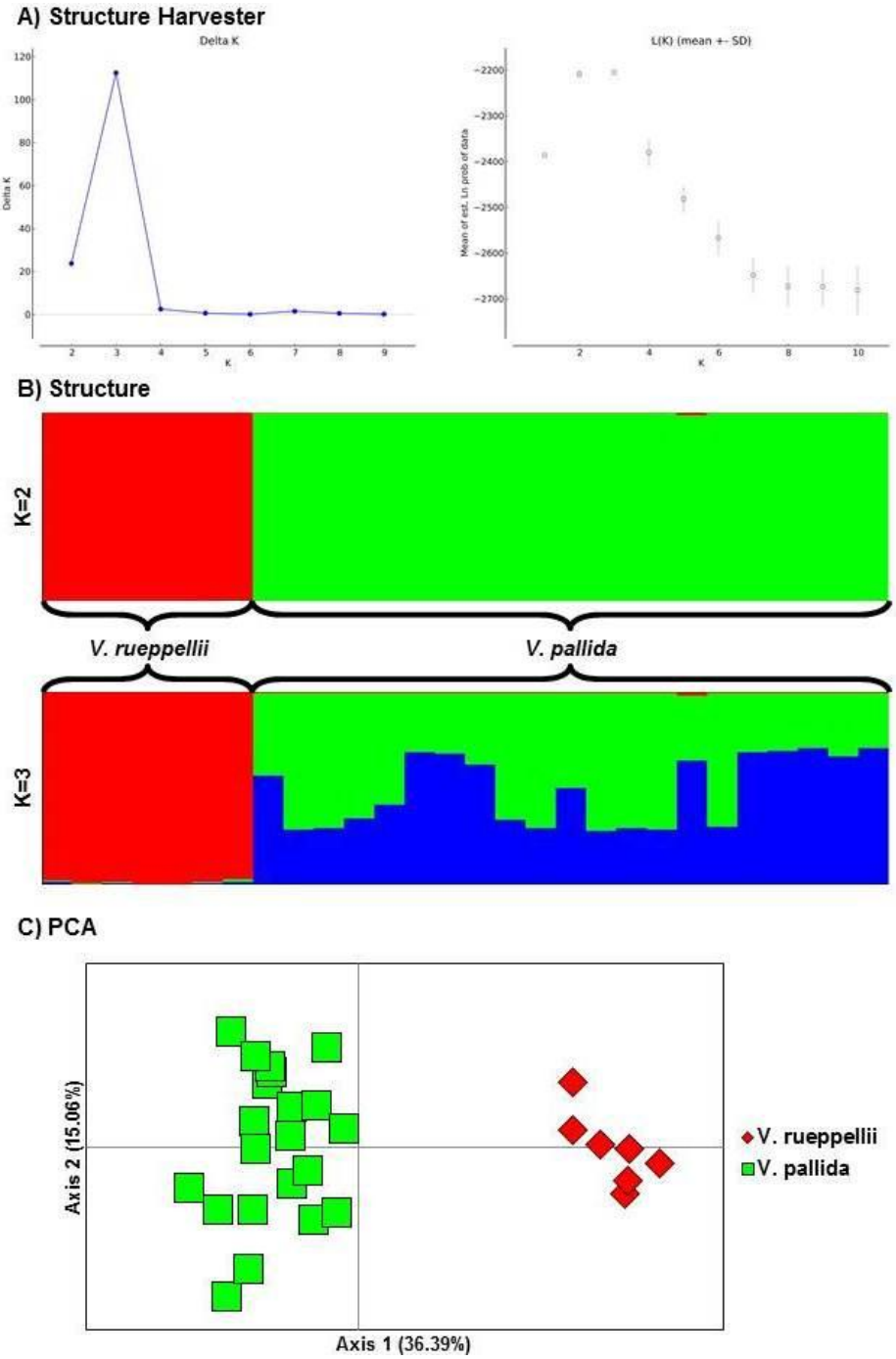


Figure 20 – Hybridization assessment between *V. rueppellii* and *V. pallida*, based on 22 *loci*. A) Structure Harvester graphic output of Delta K and Mean L(K); B) Structure bar plot of Bayesian assignment of individuals to two (K=2) and three (K=3) clusters. Vertical bars represent individuals, while colors within a bar represent probability of assignment of each individual to a cluster – *V. rueppellii* (red) and *V. pallida* (green). C) PCA based on individual-by-individual genetic distances.

For the case of *V. rueppellii* and *V. vulpes*, the dataset comprised 26 *loci*, and given the previous mtDNA results, we decided to use the entire set of samples of the red fox and not only from North Africa. STRUCTURE HARVESTER indicated K=3 as the most likely number of clusters, which matched with the bar plots (Figure 21) to *V. rueppellii* (in green), North African *V. vulpes* (in red) and European *V. vulpes* (in blue). In K=2 the first group to segregate corresponded to *V. rueppellii*. PCA of genetic distances yielded same results (Figure 21).

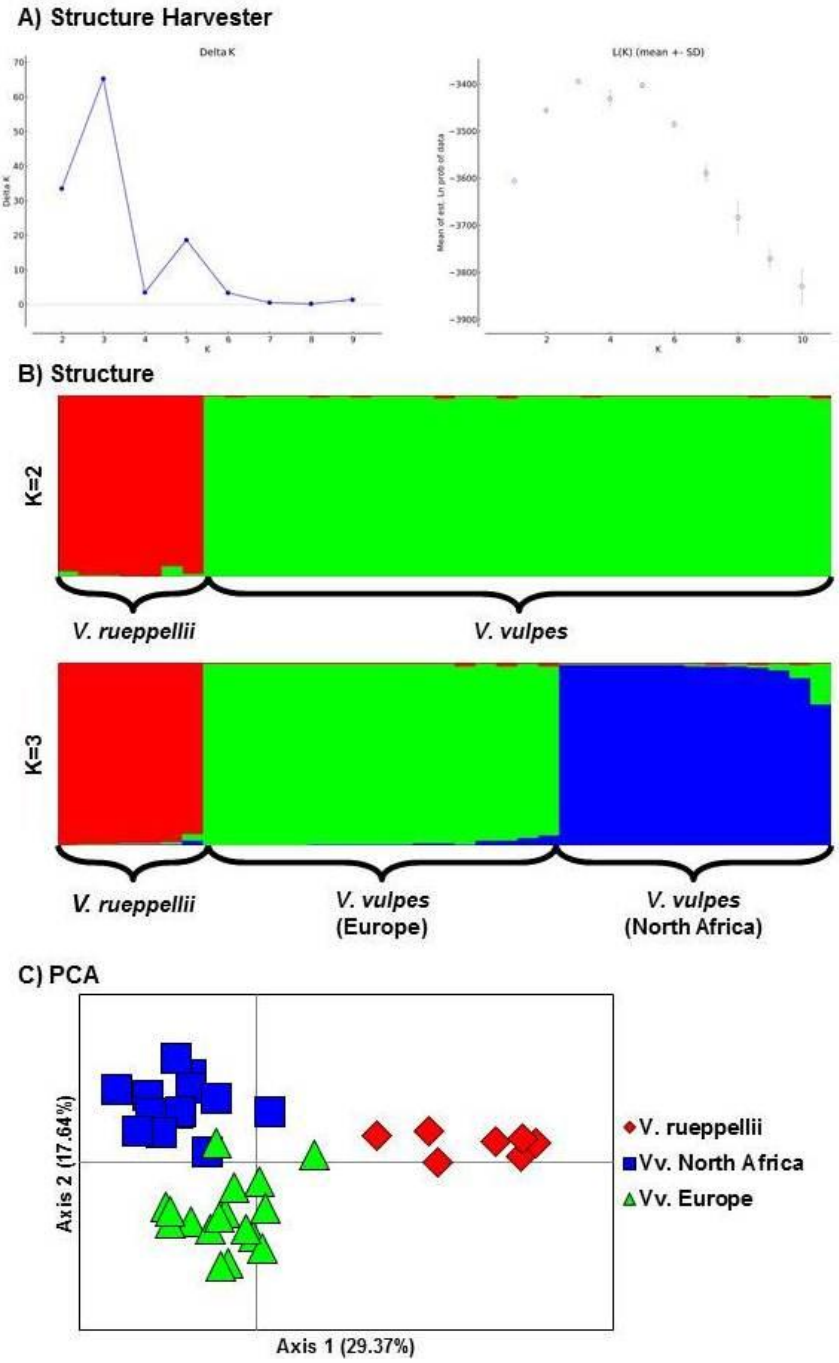


Figure 21 – Hybridization assessment between *V. rueppellii* and *V. vulpes*, based on 26 *loci*. A) STRUCTURE HARVESTER graphic output of Delta K and Mean L(K); B) STRUCTURE bar plot of Bayesian assignment of individuals to two (K=2) and three (K=3) clusters. Vertical bars represent individuals, while colors within a bar represent probability of assignment of each individual to a cluster – *V. rueppellii* (red), European *V. vulpes* (green) and North African *V. vulpes* (blue). C) PCA based on individual-by-individual genetic distances.

Two additional datasets, combining *V. rueppellii* and *V. vulpes* with either *V. pallida* (16 *loci*) or *V. zerda* (18 *loci*) were analysed to see how each group would segregate and compare pairwise F_{ST}. We also tried using all of the fox species together but the number of *loci* that could be used was very small and their variability was low (resulting, for example, in putative hybrids between pale fox and red fox, two species with clearly separated distribution ranges). Structure analysis (sections A15-16 in Annexes for the example with *V. zerda*) showed either *V. pallida* or *V. zerda* being the first group to segregate at K=2, then followed by *V. rueppellii* at K=3, and finally by *V. vulpes* groups at K=4 (we also see again putative hybrids due to the reason already mentioned above). Regarding F_{ST} in Table 11 it is noticeable that *V. rueppellii* appears even more genetically differentiated to the two groups of *V. vulpes*, than comparatively *V. pallida* and *V. zerda*.

Table 11 - Pairwise Population Fst Values for different datasets.

Dataset	Pop1	Pop2	F_{ST}
***V. rueppellii* and *V. vulpes* (26 *loci*)**	*V. rueppellii*	Vv. North Africa	0,118
	V. rueppellii	Vv. Europe	0,097
	Vv. North Africa	Vv. Europe	0,043
***V. rueppellii*, *V. vulpes* and *V. pallida* (16 *loci*)**	*V. rueppellii*	*V. pallida*	0,096
	V. rueppellii	Vv. North Africa	0,077
	V. pallida	Vv. North Africa	0,063
	V. rueppellii	Vv. Europe	0,067
	V. pallida	Vv. Europe	0,054
	Vv. North Africa	Vv. Europe	0,038
***V. rueppellii*, *V. vulpes* and *V. zerda* (18 *loci*)**	*V. rueppellii*	*V. zerda*	0,100
	V. rueppellii	Vv. North Africa	0,087
	V. zerda	Vv. North Africa	0,075
	V. rueppellii	Vv. Europe	0,075
	V. zerda	Vv. Europe	0,062
	Vv. North Africa	Vv. Europe	0,040

4. Discussion

4.1. Golden jackal

4.1.1. Phylogenetic relationships

Phylogenetic analysis of mitochondrial DNA sequences of this study points in the same direction as recently published papers (Rueness *et al.* 2011; Gaubert *et al.* 2012). All works suggest the presence of a grey wolf lineage in North Africa, *Canis lupus lupaster*, that together with two other wolf clades found in the Indian sub-continent, *C. himalayensis* (or *C. l. chanco*) and *C. indica* (or *C. l. pallipes*) (Sharma *et al.* 2004; Aggarwal *et al.* 2007), constitute a group of ancient lineages from which the more recent Holarctic grey wolf could have differentiated (Rueness *et al.* 2011; Gaubert *et al.* 2012). While the phylogenetic tree presented by Gaubert et al. (2012) depicts the African wolf lineage as the oldest within the grey wolf clade, our results suggest a separation into two further sub clades: one encompassing African and Himalayan wolves, and another covering the Indian and Holarctic wolves (Figures 5 and 6, as well as sections A11-12 in Annexes). Genetic distances corroborate these results since *C. l. lupaster* / *C. aureus* (North Africa) clade and Himalayan wolves diverged by 1.5%, both in turn diverging between 4.1 and 3.3%, respectively with the Holarctic wolves, respectively (Table 2).

All of *C. aureus* samples from North Africa were identified as *C. l. lupaster* and clustered together with available sequence data, while European samples appear grouped with golden jackals from India (Figure 5). As such, considering only mtDNA data, *C. aureus* is portrayed as a polyphyletic species when including specimens from North Africa. Looking deeply into the various subspecies of *C. aureus* described in the region, there is no clear evidence for further sub-specific divisions. Our results did not showed any clear biogeographical partition of genetic diversity at the mitochondrial level, probably due to the high dispersal capability and low habitat specificity of this canid, together with the existence of a putative Atlantic corridor that would allow avoidance of the most extreme arid regions of the Sahara desert and North-South genetic interchange.

Our results also extend the known range of mitochondrial *C. l. lupaster* in North-West Africa, mainly in Mauritania, Morocco and Libya, after it being previously confirmed in Algeria, Senegal, Mali (Gaubert *et al.* 2012) and Ethiopia (Rueness *et al.* 2011). Although no clear signs of biogeographical barriers were observed, overall there are considerable levels of genetic variability at the mtDNA level, higher than it is found for

the two Indian wolf clades. A total of 20 haplotypes were described in only 24 sequences (Hd=0.986). Despite the higher number of samples in Mauritania, this country holds a great genetic diversity for these canids, hosting nine unique haplotypes.

Patterns found for other wolf-like species indicated that the two *C. adustus* samples analysed were found in two distinct groups within a side-striped jackal clade (Figures 5 and 6). The sample from Senegal shares the same haplotype with sequences from Guinea, while our sample from Ethiopia clusters with a sequence from Benin in the phylogenetic tree. The network relating these *C. adustus* haplotypes suggests that they constitute relatively separated haplotypes that together are even more distant from the Senegal/Guinea haplotype. Thus, this pattern hints for the first time at an East-West differentiation of *C. adustus* that has already been described for some sub-Saharan ungulates (e.g. Lorenzen *et al.* 2012). Furthermore, Rueness *et al.* (2011) and Gaubert *et al.* (2012), pointed out that the available cytochrome-b sequence data on GenBank for African "jackal species" – *C. adustus*, *C. aureus* and *C. mesomelas* - do not cluster with newly produced data of the supposed same species. The most striking cases are the two available Cyt-b black-backed jackal sequences that do not even group together. These sequences may represent either cases of cryptic diversity in Africa that should be urgently assessed or just examples of taxonomic misidentification on GenBank.

4.1.2. Population structure

First of all, a quick mention to the assessment of linkage disequilibrium between pairs of *loci* (that also applies for the foxes) – although non-random combinations of *loci* were statistically assigned, there was no exclusions from the final datasets mainly because of the following issue: the probability of certain combinations of *loci* occurring in higher proportions than others increases considerably in scenarios where there is a small number of samples and the variability of those loci is low (or it is not entirely represented), and it may suggest that they are segregated together, which in fact is not the case (some of the combinations involved *loci* that are located in different chromosomes so they are segregated independently).

Population structure analysis of the golden jackal, based on 34 autosomal microsatellite *loci*, indicated a number of two clusters as the most likely scenario to explain our sample set, which matches with a mtDNA separation between individuals from North Africa and Europe. Given the previously mentioned results at the mitochondrial genome, this partition should reflect two fairly distinctive populations.

When a group of grey wolves and domestic dogs were added to the golden jackal dataset, we observe a separation between golden jackals (North African and European) and grey wolves / dogs. Only at K=4, after the split between wolves and dogs, we see golden jackals from Europe segregating from remaining North African specimens. These results contradict the ones obtained with the analysis of mitochondrial genome. Using the microsatellites as a proxy for the nuclear genome it suggests that North African golden jackals do belong to the *C. aureus* species and not to *C. lupus*.

No signs of population structuring within North African golden jackals was observed. However, given our low sample number these population level results have to be interpreted with some cautions. Given that STRUCTURE needs a minimum number of two individuals to constitute a cluster, genetic diversity may be structured but there may not be available two representative specimens from each putative population. As an example, the sample identified as a dog would not have been detected if we had not added a group of dogs. For our golden jackal sample set however this seems unlikely, since we have groups of samples from different geographic areas and PCA analysis did not show any particular samples considerably isolated from the main cluster.

4.1.3. Golden jackal in North Africa: C. aureus or C. lupus?

Until now, genetic assessments of the golden jackal in North Africa and comparisons with other wolf-like canids have only focused on the mitochondrial genome and researchers have concluded that a grey wolf lineage is widespread in the region (Rueness *et al.* 2011; Gaubert *et al.* 2012). Our present research analysed for the first time the nuclear genome through analysis of microsatellites genotypes. While our results from phylogenetic analyses of mtDNA sequences corroborate the previously mentioned studies, the output of the microsatellite analyses disputes these same results demonstrating that at the nuclear genome level, specimens of golden jackal from North Africa still cluster with the *C. aureus* clade. In terms of morphology, Ferguson (1981) had already stated the presence of wolves in Egypt, but also pointed to the existence of two forms, *C. aureus* and *C. l. lupaster* throughout the rest of North Africa. Gaubert *et al.* (2012) also mentions the existence of one larger, more bulker morphological type which would correspond to *C. l. lupaster*, and a smaller, more slender form that would be characteristic of *C. aureus*. The morphological analysis in Gaubert *et al.* (2012) is highly debatable given that it is based on a subjective colour pattern assessment from a limited set of photographs of these supposed taxa in one

locality of Senegal. Quantitative assessments, including of cranial features, are needed to understand geographic variation and possible structuring of morphological traits.

So what scenario could be hypothesized to explain these two contradictory genomes? Gaubert *et al.* (2012) proposes two possibilities: one is that hybridization occurs between the two *taxa* and the other one is that *C. aureus* from North Africa are a smaller eco-morphological variant of *C. l. lupaster*. The latter hypothesis was only based on slight phenotypic and behavioural differences, thus lacking clear support, while the former hypothesis is rejected by our analyses with nuclear markers as no evidences of hybridization between golden jackals and wolves were found. All of our samples, and the ones from Rueness *et al.* (2011) and Gaubert *et al.* (2012), were identified has *C. l. lupaster*, so it would be surprising if none of the studies had sampled a "genuine" North African *C. aureus*.

A possible explanation for the lack of agreement between mitochondrial and nuclear markers could be a huge introgression of wolf mitochondrial genome into North African golden jackals. We hypothesize that the beginning of this event could have occurred in North-East Africa, probably in Egypt [although the fossil record in Maghreb confirms the existence of jackals and a larger *Canis* sp., probably a wolf – Geraads (2011)], where a male golden jackal would have mated with a female wolf and resulting female offspring would have been successfully incorporated in golden jackal populations. The captured mtDNA would have reached the West Atlantic coast owing to the golden jackal's high vagility combined with a putative fast population growth (Sillero-Zubiri 2009). Although this would not be the first report of introgression among canids (Lehman *et al.* 1991; Adams *et al.* 2003), it would be one of the largest expansions of an introgressed lineage ever registered [but see Currat *et al.* (2008) for a review on the subject]. Although sequencing of nuclear genes from a wider location of samples is needed to support or reject this hypothesis, still the patterns found when analysing autosomal microsatellite genotypes strongly support the introgression hypothesis, especially since they regard a group of species where reproductive isolation is not complete.

4.2. Foxes

4.2.1. Phylogenetic relationships

This study provides the first ever molecular data on *V. pallida*, and for the remaining fox species it is also presents novel sequences for North Africa. More than anything, phylogenetic analysis helped to prove that *V. pallida*, *V. zerda* and *V. vulpes* constitute clearly distinct species, largely supported by the number of mutational differences observed (Figure 11) and genetic distances found (Table 4). The lack of mitochondrial sequence data of remaining fox-like species does not allow inferring on the closest relatives of the pale fox. If we look at the Carnivora supertree (Nyakatura & Bininda-Emonds 2012), we see that *V. pallida* appears together with *V. chama* (Cape fox, found in Southern Africa) as one the most basal taxa of the fox-like canids. This is also coherent with previous findings using mtDNA, only with *V. chama* (Geffen *et al.* 1992). Yet, the Cape fox is placed as a more recently diverged species in Lindblad-Toh *et al.* (2005), comparing with the fennec fox for example. Hence, there is a clear need to conduct a thorough research, with as much fox-like species as possible, to truly understand the evolutionary relationships of the pale fox.

Significant levels of genetic diversity were found for *V. pallida* and 14 haplotypes were described (Hd=0.983), although as in the case of *C. aureus*, no clear biogeographical patterns were discernible - haplotypes from Niger only differed by three nucleotide substitutions from the Mauritanian haplotypes in the concatenated dataset (Figure 11). Although there are sample size limitations, *V. zerda* also presented clear signs of genetic diversity and the distribution of haplotypes in the networks does not suggest any clear biogeographical arrangement. This pattern maybe is due to the high dispersal capability of *V. pallida* and *V. zerda*, associated with a still relatively low anthropogenic pressure that allows the existence of suitable habitat corridors between distant regions. In fact this is the case of the Sahelian desertic and semi-desertic areas for *V. pallida* and extreme Saharan sand desert for *V. zerda* (Brito *et al.* 2009).

An unexpected outcome of the present study was to see *V. rueppellii* as a distinct lineage of its own but clustered within the *V. vulpes* clade, more closely related to a divergent group of haplotypes found in the Hokkaido island of Japan (1.2% of genetic divergence). These two later groups, together with North African specimens, constitute one of two putative main clades within *V. vulpes*. The other clade encompasses individuals from Europe, Asia and Japan, but also a sample from Egypt. These results question the taxonomic status of *V. rueppellii* and suggest that this taxon may in fact be a desert ecotype of the red fox, although microsatellite data refute this idea. An

alternative hypothesis, very speculative, would be that *V. rueppellii* has diverged into an independent more arid adapted species from a first ancestral lineage of *V. vulpes*, possibly the same that was also responsible for the origin of the Hokkaido 2 clade. The present red fox in North Africa may be a result of another putative colonization wave.

Our analysis suggest that North Africa, especially in Morocco, may hold at least two, possibly three, distinct genetic entities of *V. vulpes*. Three haplotypes from the Western Atlas are shown very closely related to each other but significantly isolated from remaining North African haplotypes. A second lineage (Maghreb 1) seems to be more widely spread but when looking at the networks we can observe a group of five haplotypes that are found in Morocco and Algeria, while the haplotypes from Tunisia are also considerably far apart, even taking into account that only one sample from Algeria was available and that intermediate haplotypes may be found in this country.

Within Europe we did not find clear genetic differences in *V. vulpes* between Western (Iberian Peninsula) and Eastern (Greece and Turkey) Mediterranean Basin, since one haplotype was shared by one sample from Spain and Greece and it is not noticeable a significant number of mutational steps between haplotypes unique to each region. Even though it was not possible to compare directly our cytochrome b sequences with the haplotypes obtained by Frati *et al.* (1998), we can assume that we did not found haplotypes C and N (found in Italy and Spain), since none of the European samples clustered together with the Hokkaido 1a clade that comprises such haplotypes (or similar). Thus, our results point out more or less in the same direction of Teacher *et al.* (2011) when stating that the red fox, at least in Southern Europe does not show strong signs of genetic structuring. Finally, due to the fact that North African and Iberian Peninsula individuals belong to different putative main clades of the red fox, and that we did not find any shared haplotypes between the two regions (although we did not had samples from Southern Spain), the Strait of Gibraltar must have been acting as a dispersal barrier, possibly even during the Messinian salt crisis. Moreover, the fact that an individual from Egypt is shown genetically closer to Europe suggests that *V. vulpes* reached North Africa through the Middle East and Sinai Peninsula, and not directly crossing the Mediterranean from Europe, following one of the major colonizing routes of the Maghreb region (Dobson & Wright 2000).

4.2.2. Population structure

Population structure analysis of *V. pallida*, *V. rueppellii* and *V. zerda* did not found any evidence for segregation of genetic diversity. There was a discrepancy between Delta K and the bar plots, probably a consequence of the algorithm used in the Evanno

method – to calculate Delta K of a given number of clusters, it needs the probability of both the previous and the next K, hence in a range of K groups the first (in our case K=1) and the last (K=10) are not estimated. If we take a closer look to the mean L(K) we can observe that K=1 has the highest probability and/or more stable values (Figure 15).

In the case of *V. pallida*, again, not even five specimens from Niger were distinct from the remaining specimens of Mauritania which hints at the existence of one continuous population through in the Western part of the Sahel (Brito *et al.* 2009), and thus, neither the Niger River or the Mauritanian mountain chains seem to be acting as barriers for gene flow. For *V. rueppellii* and *V. zerda* one must analyse these results with some caution due to sample size limitations. While for *V. pallida* we include at least two representatives of many areas, the same cannot be said for *V. rueppellii* and *V. zerda*. As such, if population structure exists, it may remain undetected because of lack of enough genotypes to assign them to different clusters. PCA analysis further enlightens this issue and even though there is no coherence between geographic and genetic distances, each specimen is shown more or less isolated.

For *V. vulpes*, two clusters were statistically supported, resulting in a separation between individuals from North Africa (including one sample from Egypt) and Southern Europe. However genetic differentiation is low between the two groups, as indicated by Fst values. Within these two regions no further subdivisions were found. The results regarding Southern Europe are somewhat coherent with the ones obtained at the mitochondrial level, while in North Africa no evidence was found for different lineages within the Maghreb, suggesting that diversity of the red fox in that region is likely due to past biogeographical processes and that genetic interchange is currently occurring all over the region.

4.2.3. Hybridization assessment

We found no evidence of admixture between *V. rueppellii* and either *V. pallida* or *V. vulpes*. Still, interesting results were obtained when analysing *V. rueppellii* together with European and North African *V. vulpes*. Given the previous mtDNA results, one could suppose that the first segregation of the data at K=2 would consist in *V. rueppellii* and North African *V. vulpes* to one side, and European *V. vulpes* to the other. Yet, *V. rueppellii* is always the first to separate from remaining *V. vulpes*, and when adding another species to the dataset, such as *V. zerda*, the genetic distance of both populations of *V. vulpes* to either *V. zerda* or *V. rueppellii* are very similar. Once again we see microsatellites analysis contrasting mtDNA results and data at the nuclear level

supporting the taxonomic status of *V. rueppellii* as a species, or at least a taxonomic entity independent of *V. vulpes*.

4.2.4. Vulpes rueppellii: species or subspecies of V. vulpes?

Introgression of *V. vulpes* mitochondrial genome into a *V. rueppellii* population may be a hypothetical scenario to explain discrepancies between mitochondrial and nuclear markers. However, within the Canidae, hybridization and introgression events have only been reported in species belonging to the group of most recent wild relatives of the dog (*C. lupus*, *C. latrans* and *C. simensis*), and never in foxes. Comparing to the example of the golden jackal and the African wolf, *V. rueppellii* and *V. vulpes* are morphologically and phenotypically more distinct species (Clutton-Brock *et al.* 1976; Osborn Dale J. & Helmy 1980), and occupy different types of habitats (Sillero-Zubiri *et al.* 2004). Additionally, there are no evidences for hybridization and whenever there is range overlap between the two species, it has been reported that the red fox competitively excludes the Rüppell's fox (Yom-Tov & Mendelssohn 1988; Cuzin 2003). So given the circumstances, it is hard to find a plausible scenario to explain our results. One could propose that *V. rueppellii* could have diverged from an ancient North African *V. vulpes* lineage by geographic isolation during one of various paleoclimatic shifts in North Africa, possibly during a more humid period. Under humid condition, *V. vulpes* would have been able to expand its range into the Sahara, but after the transition into an arid period, some populations could have been retained in small suitable areas within the desert (like mountains or oasis) and diverged into a more arid adapted species, *V. rueppellii*. This new taxon would have afterwards been able to colonize further arid regions through the Arabian Peninsula and reaching Pakistan. Nevertheless, *V. rueppellii* appears related with a Japanese lineage of red foxes, clearly distinct even within the rest of the Japanese archipelago and closest continent. Inoue *et al.* (2007) estimated a time of divergence of this distinct lineage of 1.4 Mya, suggesting long-term geographic structure of genetic diversity within the Hokkaido island, and does not report any case of introduced red foxes of different origin in Japan, much less *V. rueppellii*. This lineage then could have also originated from the same ancient lineage that was present in North Africa and from which *V. rueppellii* diverged, while remaining lineages found by Inoue *et al.* (2007) in Japan are related to more recent colonization events during the Quaternary Ice Ages, where the Japanese island could have been connected with the main continent.

Although a North American origin of the Tribe Vulpini is mostly certain, the more recent set of present-day species could have been originated in Africa, since the oldest fossil

known of a fox-like canid (*V. riffautae*) was found in Chad dated to 7 Mya. Such evidence suggests first an African colonization, followed by a South-North movement reaching Europe (De Bonis *et al.* 2007). Additionally, the only fossil of a *V. rueppellii*-like canid was retrieved in Morocco, dating 0.8 Mya (Geraads 2011), while the earliest fossils assigned to *V. vulpes* in North Africa are commonly found since the upper/middle Pleistocene (Aouraghe 2000; Geraads 2011).

4.3. Contributions of this study and future research

The innovative aspects of the present study are related with the combination of two different types of molecular markers to conduct for the first time a phylogenetic and population genetic analysis in North African canids: *C. aureus*, *V. pallida*, *V. rueppellii*, *V. vulpes* and *V. zerda*. This study has produced 186 new sequences, including the first ever for *V. pallida*, this way contributing for an increase of available genetic data of fox-like canids. It also further reinforces the interest and relevance of applying molecular markers initially developed for domestic or model organisms, in their wild counterparts. Furthermore, this study has contributed with a total of 115 autosomal microsatellite genotypes using an initial set of 46 loci that already had been developed and applied to the grey wolf.

This study was based on a relative small and untargeted sampling effort, taking advantage of occasional gathering of biological material from deceased animals along roads. Future sampling efforts should, ideally, try to be more focused on these species and on areas of particular interest such as central and eastern Sahara and Sahel, as well as the Middle East (specially for *C. aureus*, *V. rueppellii* and *V. vulpes*). Non-invasive samples such as scats and looking at museum catalogues to search for specimens of areas that are presently now less accessible are also very viable options for forthcoming research.

More microsatellite loci are needed, especially for the foxes for which other sets of markers have been successfully used in red fox population studies [for example Oishi *et al.* (2011)]. Yet most importantly, future research should include nuclear markers, with a set of slow and fast evolving genes, as well as taking advantages of the recently developed genomic tools, such as the canine SNP (Single Nucleotide Polymorphisms) microarray. Another interesting aspect would be to apply molecular results to a Geographical Information System (GIS) and follow a landscape genetics approach. Finally and regarding specifically *V. vulpes*, future research on the Mediterranean basin should include analysis of the left domain of the cytochrome b gene in order to compare with the results obtained by other authors (e.g. Frati *et al.* 1998).

5. Conclusions

1) North African biodiversity is still poorly assessed, especially for the case of highly vagile and non-volant species such as the canids. This study exemplifies that applying different kinds of data and analyses, even with a relatively small sampling effort, provides interesting and novel outcomes, such as the first molecular data for *V. pallida;*

2) Data for *C. aureus* produced during this research has brought new insights and raised new questions relative to the recent African wolf issue (Rueness *et al.* 2011; Gaubert *et al.* 2012). According to nuclear markers, *C. aureus* seem to be a valid taxonomic unit across their entire range, independent from wolves. However, mtDNA data shows a puzzling pattern of a widespread ancient grey wolf mitochondrial lineage in North African jackals, suggesting past extensive introgression between jackals and wolves in this region;

3) The taxonomic status of *V. rueppellii* does not seem to be clear-cut and there are evidences for complex genetic patterns in *V. vulpes* from North Africa. Probably, *V. rueppellii* may result from an earlier colonization of North Africa by an ancient lineage of *V. vulpes* (possibly similar to the one that colonized Hokkaido in Japan) that became well adapted to arid conditions, as well as ecologically and reproductively isolated from *V. vulpes*. The present occurrence of *V. vulpes* in North Africa may result of this first colonization (as the different lineages from Morocco suggest) and/or from a second and more recent colonization event, apparently through the Sinai Peninsula (or even on-going colonization as the sample from Egypt suggests);

4) There is lack of evidence for strong biogeographical structuring of genetic diversity and population differentiation within North Africa, with the exceptions of *V. vulpes* in Morocco;

5) Main conclusions of this study are key issues that should be considered in future conservation efforts for North African canids.

6. References

Adams JM, Faure H (2004) Review and Atlas of Palaeovegetation - Preliminary land ecosystem maps of the world since the Last Glacial Maximum. *Quaternary Environment Network (http://www.esd.ornl.gov/projects/qen/adams4.html)*.

Adams JR, Leonard JA, Waits LP (2003) Widespread occurrence of a domestic dog mitochondrial DNA haplotype in southeastern US coyotes. *Molecular Ecology*, 12, 541–546.

Aggarwal RK, Kivisild T, Ramadevi J, Singh L (2007) Mitochondrial DNA coding region sequences support the phylogenetic distinction of two Indian wolf species. *Journal of Zoological Systematics and Evolutionary Research*, 45, 163–172.

Allendorf FW, Leary RF, Spruell P, Wenburg JK (2001) The problems with hybrids: setting conservation guidelines. *Trends in Ecology & Evolution*, 16, 613–622.

Alpers DL, Van Vuuren BJ, Arctander P, Robinson TJ (2004) Population genetics of the roan antelope (*Hippotragus equinus*) with suggestions for conservation. *Molecular Ecology*, 13, 1771–84.

Anthelme F, Waziri Mato M, Maley J (2008) Elevation and local refuges ensure persistence of mountain specific vegetation in the Nigerien Sahara. *Journal of Arid Environments*, 72, 2232–2242.

Aouraghe H (2000) Les carnivores fossiles d'El Harhoura 1, Temara, Maroc. *L'Anthropologie*, 104, 147–171.

Arjo WM, Pletscher DH (1999) Behavioral responses of coyotes to wolf recolonization in northwestern Montana. *Canadian Journal of Zoology*, 77, 1919–1927.

Asa CS, Valdespino C, Cuzin F, De Smet K, Jdeidi T (2008) *Vulpes zerda*. *In: IUCN 2012. IUCN Red List of Threatened Species. Version 2012.1. <www.iucnredlist.org>.*

Aubry KB, Statham MJ, Sacks BN, Perrine JD, Wisely SM (2009) Phylogeography of the North American red fox: vicariance in Pleistocene forest refugia. *Molecular Ecology*, 18, 2668–86.

Avise JC (2000) *Phylogeography: The History and Formation of Species*. Harvard University Press.

Bardeleben C, Moore RL, Wayne RK (2005a) Isolation and molecular evolution of the selenocysteine tRNA (Cf TRSP) and RNase P RNA (Cf RPPH1) genes in the dog family, Canidae. *Molecular Biology and Evolution*, 22, 347–59.

Bardeleben C, Moore RL, Wayne RK (2005b) A molecular phylogeny of the Canidae based on six nuclear loci. *Molecular Phylogenetics and Evolution*, 37, 815–31.

Barton NH, Hewitt GM (1989) Adaptation, speciation and hybrid zones. *Nature*, 341, 497–503.

Bininda-Emonds OR, Gittleman JL, Purvis A (1999) Building large trees by combining phylogenetic information: a complete phylogeny of the extant Carnivora (Mammalia). *Biological Reviews of the Cambridge Philosophical Society*, 74, 143–175.

Bohling JH, Waits LP (2011) Assessing the prevalence of hybridization between sympatric Canis species surrounding the red wolf (*Canis rufus*) recovery area in North Carolina. *Molecular Ecology*, 20, 2142–56.

De Bonis L, Peigné S, Likius A, *et al.* (2007) The oldest African fox (*Vulpes riffautae* n. sp., Canidae, Carnivora) recovered in late Miocene deposits of the Djurab desert, Chad. *Die Naturwissenschaften*, 94, 575–80.

Bons J, Geniez P, Montori A, Roca V, Española AH (1996) *Amphibians & reptiles of Morocco (including Western Sahara):* Asociación Herpetológica Española.

Boratynski Z, Brito JC, Mappes T (2012) The origin of two cryptic species of African desert jerboas (Dipodidae: *Jaculus*). *Biological Journal of the Linnean Society*, 105, 435–445.

Botkin DB, Saxe H, Araújo MB, *et al.* (2007) Forecasting the Effects of Global Warming on Biodiversity. *BioScience*, 57, 227–236.

Brinkman FS, Leipe DD (2001) Phylogenetic Analysis. In: *Bioinformatics: A Practical Guide to the Analysis of Genes and Proteins* , pp. 323–58. John Wiley & Sons, Inc.

Brito JC, Acosta AL, Álvares F, Cuzin F (2009) Biogeography and conservation of taxa from remote regions: An application of ecological-niche based models and GIS to North-African canids. *Biological Conservation*, 142, 3020–3029.

Brito JC, Fahd S, Geniez P, *et al.* (2011) Biogeography and conservation of viperids from North-West Africa: An application of ecological niche-based models and GIS. *Journal of Arid Environments*, 75, 1029–1037.

Brouat C, Tatard C, Bâ K, *et al.* (2009) Phylogeography of the Guinea multimammate mouse (*Mastomys erythroleucus*): a case study for Sahelian species in West Africa. *Journal of Biogeography*, 36, 2237–2250.

Busack SD (1986) Biogeographic analysis of the herpetofauna separated by the formation of the Strait of Gibraltar. *National Geographic Research*, 2, 17–36.

Butchart SHM, Walpole M, Collen B, *et al.* (2010) Global biodiversity: indicators of recent declines. *Science*, 328, 1164–8.

Carmichael LE, Krizan J, Nagy JA, *et al.* (2007) Historical and ecological determinants of genetic structure in arctic canids. *Molecular Ecology*, 16, 3466–83.

Carmichael LE, Nagy JA, Larter NC, Strobeck C (2001) Prey specialization may influence patterns of gene flow in wolves of the Canadian Northwest. *Molecular Ecology*, 10, 2787–98.

Carranza S, Arnold EN, Geniez P, Roca J, Mateo JA (2008) Radiation, multiple dispersal and parallelism in the skinks, Chalcides and Sphenops (Squamata: Scincidae), with comments on Scincus and Scincopus and the age of the Sahara Desert. *Molecular phylogenetics and evolution*, 46, 1071–94.

Claussen M (2009) Late Quaternary vegetation-climate feedbacks. *Climate Of The Past*, 5, 203–216.

Clement M, Posada D, Crandall KA (2000) TCS: a computer program to estimate gene genealogies. *Molecular Ecology*, 9, 1657–1659.

Clutton-Brock J, Corbet GG, Hills M (1976) A review of the family Canidae, with a classification by numerical methods. *Bull Brit Museum Nat Hist*, 29, 117–199.

Cosson J-F, Hutterer R, Libois R, *et al.* (2005) Phylogeographical footprints of the Strait of Gibraltar and Quaternary climatic fluctuations in the western Mediterranean: a case study with the greater white-toothed shrew, Crocidura russula (Mammalia: Soricidae). *Molecular Ecology*, 14, 1151–62.

Currat M, Ruedi M, Petit RJ, Excoffier L (2008) The hidden side of invasions: massive introgression by local genes. *Evolution; international journal of organic evolution*, 62, 1908–20.

Cuzin F (2003) Les Grands Mammiferes du Maroc Meridional (Haut Atlas , Anti Atlas et Sahara): Distribution , écologie et conservation. *PhD thesis. Laboratoire de Biogéographie et Écologie des Vertébrés. EPHE, Université Montpellier II*.

Cuzin F, Lenain DM, Jdeidi T, *et al.* (2008) *Vulpes rueppellii*. *In: IUCN 2012. IUCN Red List of Threatened Species. Version 2012.1. <www.iucnredlist.org>.*

Dalen L, Elmhagen B, Angerbjorn A (2004) DNA analysis on fox faeces and competition induced niche shifts. *Molecular Ecology*, 13, 2389–92.

Dirzo R, Raven PH (2003) Global State of Biodiversity and Loss. *Annual Review of Environment and Resources*, 28, 137–167.

Dobigny G, Aniskin V, Granjon L, Cornette R, Volobouev V (2005) Recent radiation in West African Taterillus (Rodentia, Gerbillinae): the concerted role of chromosome and climatic changes. *Heredity*, 95, 358–68.

Dobson M, Wright A (2000) Faunal relationships and zoogeographical affinities of mammals in north-west Africa. *Journal of Biogeography*, 27, 417–424.

Douady CJ, Catzeflis F, Raman J, Springer MS, Stanhope MJ (2003) The Sahara as a vicariant agent, and the role of Miocene climatic events, in the diversification of the mammalian order Macroscelidea (elephant shrews). *Proceedings of the National Academy of Sciences*, 100, 8325–30.

Drake N a, Blench RM, Armitage SJ, Bristow CS, White KH (2011) Ancient watercourses and biogeography of the Sahara explain the peopling of the desert. *Proceedings of the National Academy of Sciences*, 108, 458–62.

Drake N a. A, El-Hawat a. SS, Turner P, *et al.* (2008) Palaeohydrology of the Fazzan Basin and surrounding regions: The last 7 million years. *Palaeogeography, Palaeoclimatology, Palaeoecology*, 263, 131–145.

Drummond A, Ashton B, Buxton S, *et al.* (2010) Geneious V4.8. *Geneious V4.8*.

Durant SM, Pettorelli N, Bashir S, *et al.* (2012) Forgotten Biodiversity in Desert Ecosystems. *Science*, 336, 1379–80.

Díaz S, Fargione J, Chapin FS, Tilman D, Chapin III FS (2006) Biodiversity Loss Threatens Human Well-Being. *PLoS Biol*, 4, e277.

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

Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Molecular Ecology*, 14, 2611–2620.

Ewi M (2010) A decade of kidnappings and terrorism in West Africa and the Trans-Sahel region. *African Security Review*, 19, 64–71.

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, 564–567.

Falush D, Stephens M, Pritchard JK (2003) Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. *Genetics*, 164, 1567–1587.

Ferguson WW (1981) The systematic position of *Canis aureus lupaster* (Carnivora: Canidae) and the occurrence of *Canis lupus* in North Africa, Egypt and Sinai. *Mammalia*, 45, 459.

Flagstad O, Syvertsen PO, Stenseth NC, Jakobsen KS (2001) Environmental change and rates of evolution: the phylogeographic pattern within the hartebeest complex as related to climatic variation. *Proceedings of the Royal Society B Biological Sciences*, 268, 667–677.

Foley J a., Coe MT, Scheffer M, Wang GL (2003) Regime shifts in the Sahara and Sahel: interactions between ecological and climatic systems in northern Africa. *Ecosystems*, 6, 524–539.

Frati F, Hartl GBB, Lovari S, Delibes M, Markov G (1998) Quaternary radiation and genetic structure of the red fox Vulpes vulpes in the Mediterranean Basin, as revealed by allozymes and mitochondrial DNA. *Journal of Zoology*, 245, 43–51.

Gaubert P, Bloch C, Benyacoub S, *et al.* (2012) Reviving the African Wolf *Canis lupus lupaster* in North and West Africa: A Mitochondrial Lineage Ranging More than 6,000 km Wide (S-O Kolokotronis, Ed,). *PLoS ONE*, 7, e42740.

Gaubert P, Machordom A, Morales A, *et al.* (2011) Comparative phylogeography of two African carnivorans presumably introduced into Europe: disentangling natural versus human-mediated dispersal across the Strait of Gibraltar. *Journal of Biogeography*, 38, 341–358.

Geffen E, Mercure A, Girman DJ, Macdonald DW, Wayne RK (1992) Phylogenetic relationships of the fox-like canids: mitochondrial DNA restriction fragment, site and cytochrome b sequence analyses. *Journal of Zoology*, 228, 27–39.

Van Gelder RG (1978) *A Review of Canid Classification*. American Museum of Natural History.

Geraads D (2010) Biogeographic relationships of Pliocene and Pleistocene North-western African mammals. *Quaternary International*, 212, 159–168.

Geraads D (2011) A revision of the fossil Canidae (Mammalia) of north-western Africa. *Palaeontology*, 54, 429–446.

Girman DJ, Vilà C, Geffen E, *et al.* (2001) Patterns of population subdivision, gene flow and genetic variability in the African wild dog (*Lycaon pictus*). *Molecular Ecology*, 10, 1703–23.

Godinho R, Abáigar T, Lopes S, *et al.* (2012) Conservation genetics of the endangered Dorcas gazelle (*Gazella dorcas* spp.) in Northwestern Africa. *Conservation Genetics*, 13, 1003–1015.

Godinho R, Llaneza L, Blanco JC, *et al.* (2011) Genetic evidence for multiple events of hybridization between wolves and domestic dogs in the Iberian Peninsula. *Molecular Ecology*, 20, 5154–66.

Gottelli D, Sillero-Zubiri C, Applebaum GD, *et al.* (1994) Molecular genetics of the most endangered canid: the Ethiopian wolf *Canis simensis*. *Molecular Ecology*, 3, 301–12.

Govers R (2009) Choking the Mediterranean to dehydration: The Messinian salinity crisis. *Geology*, 37, 167–170.

Guillaumet A, Crochet P-A, Pons J-M (2008) Climate-driven diversification in two widespread Galerida larks. *BMC Evolutionary Biology*, 8, 32.

Guillaumet A, Pons J-M, Godelle B, Crochet P-A (2006) History of the Crested Lark in the Mediterranean region as revealed by mtDNA sequences and morphology. *Molecular Phylogenetics and Evolution*, 39, 645–56.

Guo SW, Thompson EA (1992) Performing the exact test of Hardy-Weinberg proportion for multiple alleles. *Biometrics*, 48, 361–372.

Hewitt GM (2004) The structure of biodiversity - insights from molecular phylogeography. *Frontiers in Zoology*, 1, 4.

Hickerson MJ, Carstens BC, Cavender-Bares J, *et al.* (2010) Phylogeography's past, present, and future: 10 years after Avise, 2000. *Molecular Phylogenetics and Evolution*, 54, 291–301.

Le Houérou HN (1992) Outline of the biological history of the Sahara. *Journal of Arid Environments*, 22, 3–30.

Le Houérou HN (1997) Climate, flora and fauna changes in the Sahara over the past 500 million years. *Journal of Arid Environments*, 37, 619–647.

Huson DH, Bryant D (2006) Application of phylogenetic networks in evolutionary studies. *Molecular biology and evolution*, 23, 254–67.

Inoue T, Nonaka N, Mizuno A, *et al.* (2007) Mitochondrial DNA phylogeography of the red fox (*Vulpes vulpes*) in northern Japan. *Zoological Science*, 24, 1178–1186.

Jhala YV, Moehlman PD (2008) *Canis aureus*. *In: IUCN 2011. IUCN Red List of Threatened Species. Version 2011.2. <www.iucnredlist.org>.*

De Jong H (1998) In search of historical biogeographic patterns in the western Mediterranean terrestrial fauna. , 99–164.

Kays R, Curtis A, Kirchman JJ (2010) Rapid adaptive evolution of northeastern coyotes via hybridization with wolves. *Biology Letters*, 6, 89–93.

Kocher TD, Thomas WK, Meyer a, *et al.* (1989) Dynamics of mitochondrial DNA evolution in animals: amplification and sequencing with conserved primers. *Proceedings of the National Academy of Sciences of the United States of America*, 86, 6196–200.

Kröpelin S, Verschuren D, Lézine A-M, *et al.* (2008) Climate-driven ecosystem succession in the Sahara: the past 6000 years. *Science*, 322, 1326; author reply 1326.

Kumar S, Skjæveland Å, Orr RJ, *et al.* (2009) AIR: A batch-oriented web program package for construction of supermatrices ready for phylogenomic analyses. *BMC Bioinformatics*, 10, 357.

Lehman N, Eisenhawer A, Hansen K, *et al.* (1991) Introgression of coyote mitochondrial DNA into sympatric North American gray wolf populations. *Evolution*, 45, 104–119.

Lerp H, Wronski T, Pfenninger M, Plath M (2011) A phylogeographic framework for the conservation of Saharan and Arabian Dorcas gazelles (Artiodactyla: Bovidae). *Organisms Diversity & Evolution*, 11, 317–329.

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

Lindblad-Toh K, Wade CM, Mikkelsen TS, *et al.* (2005) Genome sequence, comparative analysis and haplotype structure of the domestic dog. *Nature*, 438, 803–19.

Loarie SR, Duffy PB, Hamilton H, *et al.* (2009) The velocity of climate change. *Nature*, 462, 1052–5.

Lorenzen ED, Heller R, Siegismund HR (2012) Comparative phylogeography of African savannah ungulates[1]. *Molecular Ecology*, no–no.

Loveridge a. J, Macdonald DW (2003) Niche separation in sympatric jackals (*Canis mesomelas* and *Canis adustus*). *Journal of Zoology*, 259, 143–153.

Lyras GA, Van Der Geer AAE (2003) External brain anatomy in relation to the phylogeny of Caninae (Carnivora: Canidae). *Zoological Journal of the Linnean Society*, 138, 505–522.

Macdonald DW, Reynolds JC (2008) *Vulpes vulpes*. *In: IUCN 2012. IUCN Red List of Threatened Species. Version 2012.1. <www.iucnredlist.org>.*

Macdonald DW, Sillero-Zubiri C, Nicholson K (2004) *The Biology and Conservation of Wild Canids*. Oxford University Press.

McNeely J a. (2003) Biodiversity in arid regions: values and perceptions. *Journal of Arid Environments*, 54, 61–70.

Messerli B, Winiger M (1992) Climate, environmental change, and resources of the African mountains from the Mediterranean to the equator. *Mountain research and development*, 12, 315–336.

Migliore J, Baumel A, Juin M, Médail F (2012) From Mediterranean shores to central Saharan mountains: key phylogeographical insights from the genus *Myrtus*. *Journal of Biogeography*, 39, 942–956.

Mouline K, Granjon L, Galan M, *et al.* (2008) Phylogeography of a Sahelian rodent species *Mastomys huberti*: a Plio-Pleistocene story of emergence and colonization of humid habitats. *Molecular Ecology*, 17, 1036–53.

Murphy NP, Breed MF, Guzik MT, Cooper SJB, Austin AD (2012) Trapped in desert springs: phylogeography of Australian desert spring snails. *Journal of Biogeography*, no–no.

Musiani M, Leonard J a, Cluff HD, *et al.* (2007) Differentiation of tundra/taiga and boreal coniferous forest wolves: genetics, coat colour and association with migratory caribou. *Molecular Ecology*, 16, 4149–70.

Myers N, Mittermeier RA, Mittermeier CG, Da Fonseca GA, Kent J (2000) Biodiversity hotspots for conservation priorities. *Nature*, 403, 853–858.

Newby J (2007) Four perspectives on conservation in Africa. *Oryx*, 41, 135.

Nicolas V, Granjon L, Duplantier J-M, Cruaud C, Dobigny G (2009) Phylogeography of spiny mice (genus *Acomys*, Rodentia: Muridae) from the south-western margin of the Sahara with taxonomic implications. *Biological Journal of the Linnean Society*, 29–46.

Nyakatura K, Bininda-Emonds OR (2012) Updating the evolutionary history of Carnivora (Mammalia): a new species-level supertree complete with divergence time estimates. *BMC Biology*, 10, 12.

Oishi T, Uraguchi K, Takahashi K, Masuda R (2011) Population structures of the red fox (*Vulpes vulpes*) on the Hokkaido Island, Japan, revealed by microsatellite analysis. *The Journal of Heredity*, 102, 38–46.

Olivier W, Denis R (2010) Sahara or Sahel? The fuzzy geography of terrorism in West Africa. *CEPS/INSTEAD Working Paper Series*.

Olson DM, Dinerstein E, Wikramanayake ED, *et al.* (2001) Terrestrial Ecoregions of the World: A New Map of Life on Earth. *BioScience*, 51, 933.

Osborn Dale J., Helmy I (1980) *The Contemporary Land Mammals of Egypt (including Sinai)*. Field Museum of Natural History, Chicago.

Palumbi SS, Martin A, Romano S, *et al.* (2002) The simple fool's guide to PCR (H University Of Hawaii, Ed,). , 96822, 1–45.

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

Pilot M, Jedrzejewski W, Branicki W, *et al.* (2006) Ecological factors influence population genetic structure of European grey wolves. *Molecular Ecology*, 15, 4533–53.

Pleguezuelos J, Fahd S, Carranza S (2008) El papel del Estrecho de Gibraltar en la conformación de la actual fauna de anfibios y reptiles en el Mediterráneo Occidental. *Boletín de la Asociación Herpetológica Española*, 2–17.

Posada D (2008) jModelTest: phylogenetic model averaging. *Molecular Biology and Evolution*, 25, 1253–1256.

Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. *Genetics*, 155, 945–959.

Rambaut A (2009) FigTree, a graphical viewer of phylogenetic trees. *Institute of Evolutionary Biology University of Edinburgh*.

Randi E, Lucchini V (2002) Detecting rare introgression of domestic dog genes into wild wolf (*Canis lupus*) populations by Bayesian admixture analyses of microsatellite variation. *Conservation Genetics*, 3, 31–45.

Randi E, Lucchini V, Christensen M (2000) Mitochondrial DNA variability in Italian and East European wolves: detecting the consequences of small population size and hybridization. *Conservation*, 14, 464–473.

Rands MRW, Adams WM, Bennun L, *et al.* (2010) Biodiversity conservation: challenges beyond 2010. *Science*, 329, 1298–303.

Rasmussen GSA (1996) Predation on bat-eared foxes *Otocyon megalotis* by Cape hunting dogs *Lycaon pictus*. *Koedoe*, 39, 127–129.

Rhymer JM, Simberloff D (1996) Extinction By Hybridization and Introgression. *Annual Review of Ecology and Systematics*, 27, 83–109.

Rodnikova A, Ims R a., Sokolov A, *et al.* (2011) Red fox takeover of arctic fox breeding den: an observation from Yamal Peninsula, Russia. *Polar Biology*, 34, 1609–1614.

Roemer G, Coonan T, Garcelon D (2001) Feral pigs facilitate hyperpredation by golden eagles and indirectly cause the decline of the island fox. *Animal Conservation*, 307–318.

Rognon P (1993) *Biogéographie d'un désert*. L'Harmattan, Paris.

Ronquist F, Huelsenbeck JP (2003) MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics*, 19, 1572–1574.

Roth JD, Murray DL, Steury TD (2008) Spatial dynamics of sympatric canids: Modeling the impact of coyotes on red wolf recovery. *Ecological Modelling*, 214, 391–403.

Roy MS, Geffen E, Smith D, Ostrander E a, Wayne RK (1994) Patterns of differentiation and hybridization in North American wolflike canids, revealed by analysis of microsatellite loci. *Molecular Biology and Evolution*, 11, 553–70.

Rueness EK, Asmyhr MG, Sillero-Zubiri C, *et al.* (2011) The cryptic African wolf: *Canis aureus lupaster* is not a golden jackal and is not endemic to Egypt. *PloS ONE*, 6, e16385.

Sacks BN, Bannasch DL, Chomel BB, Ernest HB (2008) Coyotes demonstrate how habitat specialization by individuals of a generalist species can diversify populations in a heterogeneous ecoregion. *Molecular Biology and Evolution*, 25, 1384–94.

Sacks BN, Brown SK, Ernest HB (2004) Population structure of California coyotes corresponds to habitat-specific breaks and illuminates species history. *Molecular Ecology*, 13, 1265–75.

Sacks BN, Moore M, Statham MJ, Wittmer HU (2011) A restricted hybrid zone between native and introduced red fox (*Vulpes vulpes*) populations suggests reproductive barriers and competitive exclusion. *Molecular Ecology*, 20, 326–41.

Savolainen P, Zhang Y, Luo J, Lundeberg J, Leitner T (2002) Genetic evidence for an East Asian origin of domestic dogs. *Science*, 298, 1610–3.

Schuster M, Duringer P, Vignaud P (2006) The Age of the Sahara Desert. , 2006.

Seehausen O, Takimoto G, Roy D, Jokela J (2008) Speciation reversal and biodiversity dynamics with hybridization in changing environments. *Molecular Ecology*, 17, 30–44.

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

Shaibi T, Moritz RF a. (2010) 10,000 years in isolation? Honeybees (*Apis mellifera*) in Saharan oases. *Conservation Genetics*, 11, 2085–2089.

Sharma DK, Maldonado JE, Jhala Y V, Fleischer RC (2004) Ancient wolf lineages in India. *Proceedings. Biological sciences / The Royal Society*, 271 Suppl , S1–4.

Sillero-Zubiri C (2008) *Vulpes pallida*. *In: IUCN 2012. IUCN Red List of Threatened Species. Version 2012.1. <www.iucnredlist.org>.*

Sillero-Zubiri C (2009) Family Canidae (Dogs). In: *Wilson, D.E., Mittermeier, R.A. (Eds.), Handbook of the Mammals of the World, Carnivores, vol. 1.* (ed Editions L), pp. 352–446. Barcelona.

Sillero-Zubiri C, Hoffmann M, Macdonald DW (2004) *Canids: Foxes, Wolves, Jackals and Dogs. Status Survey and Conservation Action Plan* (C Sillero-Zubiri, M Hoffmann, D Macdonald, Eds,). IUCN/SSC Canid Specialist Group.

Sindaco R, Jeremčenko VK (2008) *The reptiles of the Western Palearctic: Annotated checklist and distributional atlas of the turtles, crocodiles, amphisbaenians and lizards of Europe, North Africa, Middle East and Central Asia*. Edizioni Belvedere, Latina.

Statham MJ, Sacks BN, Aubry KB, Perrine JD, Wisely SM (2012) The origin of recently established red fox populations in the United States: translocations or natural range expansions? *Journal of Mammalogy*, 93, 52–65.

Sunquist M (1989) Comparison of spatial and temporal activity of red foxes and gray foxes in north-central Florida. *Florida Field Naturalist*, 17, 11–18.

Swezey CS (2009) Cenozoic stratigraphy of the Sahara, Northern Africa. *Journal of African Earth Sciences*, 53, 89–121.

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, 2731–2739.

Tannerfeldt M, Elmhagen B, Angerbjörn A (2002) Exclusion by interference competition? The relationship between red and arctic foxes. *Oecologia*, 132, 213–220.

Teacher AGF, Thomas J a, Barnes I (2011) Modern and ancient red fox (*Vulpes vulpes*) in Europe show an unusual lack of geographical and temporal structuring, and differing responses within the carnivores to historical climatic change. *BMC evolutionary biology*, 11, 214.

Tedford RiH, Taylor BE, Wang X (1995) Phylogeny of the Caninae (Carnivora, Canidae): The Living Taxa. *American Museum Novitates*, 3146, 1–37.

Temple HJ, Cuttelod A (2011) *The Status and Distribution of Mediterranean Mammals* (HJ Temple, Ed,). IUCN.

Thornton PK, Owiyo T (2008) Climate change and poverty in Africa: Mapping hotspots of vulnerability Philip K THORNTON ∗. , 2, 24–44.

Trape S (2009) Impact of climate change on the relict tropical fish fauna of central sahara: threat for the survival of adrar mountains fishes, mauritania. *PloS ONE*, 4, e4400.

Tsubouchi A, Fukui D, Ueda M, *et al.* (2012) Comparative molecular phylogeny and evolution of sex chromosome DNA sequences in the family Canidae (mammalia: carnivora). *Zoological science*, 29, 151–61.

Vilà C, Amorim I, Leonard J, *et al.* (1999) Mitochondrial DNA phylogeography and population history of the grey wolf canis lupus. *Molecular Ecology*, 8, 2089–103.

Vilà C, Walker C, Sundqvist a-K, *et al.* (2003) Combined use of maternal, paternal and bi-parental genetic markers for the identification of wolf-dog hybrids. *Heredity*, 90, 17–24.

vonHoldt BM, Pollinger JP, Earl D a, *et al.* (2011) A genome-wide perspective on the evolutionary history of enigmatic wolf-like canids. *Genome Research*, 21, 1294–305.

Wan Q-H, Wu H, Fujihara T, Fang S-G (2004) Which genetic marker for which conservation genetics issue? *Electrophoresis*, 25, 2165–76.

Wandeler P, Funk SM, Largiadèr CR, Gloor S, Breitenmoser U (2003) The city-fox phenomenon: genetic consequences of a recent colonization of urban habitat. *Molecular Ecology*, 12, 647–56.

Wang Y, Notaro M, Liu Z, *et al.* (2007) Detecting vegetation-precipitation feedbacks in mid-Holocene North Africa from two climate models. *Climate of the Past Discussions*, 3, 961–975.

Ward D (2009) *The Biology of Deserts*. Oxford University Press.

Wayne RK, Benveniste RE, O'Brien SJ (1989) Molecular and biochemical evolution of the Carnivora. In: *Carnivore Behavior, Ecology and Evolution* (ed J.L. Gittleman E), pp. 465–494.

Wayne RK, Geffen E, Girman DJ, *et al.* (1997) Molecular systematics of the Canidae. *Systematic Biology*, 46, 622–653.

Wayne RK, Lehman N, Allard MW, Honeycutt RL (1992) Mitochondrial DNA variability of the gray wolf: genetic consequences of population decline and habitat fragmentation. *Conservation Biology*, 6, 559–569.

Wayne RK, Nash WG, O'Brien SJ (1987a) Chromosomal evolution of the Canidae. I. Species with high diploid numbers. *Cytogenetics and Cell Genetics*, 44, 123–133.

Wayne RK, Nash WG, O'Brien SJ (1987b) Chromosomal evolution of the Canidae. II. Divergence from the primitive carnivore karyotype. *Cytogenetics and Cell Genetics*, 44, 134–141.

Wayne RK, Ostrander E a (2007) Lessons learned from the dog genome. *Trends in genetics : TIG*, 23, 557–67.

Wayne RK, O'Brien SJ (1987) Allozyme Divergence within the Canidae . *Systematic Biology* , 36 , 339–355.

Wayne RK, Van Valkenburgh B, Kat PW, *et al.* (1989) Genetic and morphological divergence among sympatric canids. *The Journal of heredity*, 80, 447–54.

Weiss S, Ferrand N (2007) Current perspectives in phylogeography and the significance of South European refugia in the creation and maintenance of European biodiversity. In: *Phylogeography of Southern European Refugia* (eds Weiss S, Ferrand N), pp. 341–357. Springer.

Whittaker RJ, Araujo MB, Paul J, *et al.* (2005) Conservation Biogeography: assessment and prospect. *Diversity and Distributions*, 11, 3–23.

Wilson JS, Pitts JP (2012) Identifying Pleistocene refugia in North American cold deserts using phylogeographic analyses and ecological niche modelling. *Diversity and Distributions*, no–no.

Wilson DE, Reeder DM (2005) *Mammal species of the world: a taxonomic and geographic reference* (DE Wilson, DM Reeder, Eds,). Johns Hopkins University Press.

Yom-Tov Y, Mendelssohn H (1988) Changes in the distribution and abundance of vertebrates in Israel during the 20th century. In: *In: Yom-Tov, Y., Tchernov, E. (Eds.), The Zoogeography of Israel* (ed Publishers DWJ), pp. 515–547. The Netherlands.

Zachos FE, Cirovic D, Kirschning J, *et al.* (2009) Genetic variability, differentiation, and founder effect in golden jackals (*Canis aureus*) from Serbia as revealed by mitochondrial DNA and nuclear microsatellite loci. *Biochemical Genetics*, 47, 241–50.

Zhang D-X, Hewitt GM (2003) Nuclear DNA analyses in genetic studies of populations: practice, problems and prospects. *Molecular Ecology*, 12, 563–84.

Zrzavy J, Ricankova V (2004) Phylogeny of Recent Canidae (Mammalia , Carnivora): relative reliability and utility of morphological and molecular datasets. *Zoologica Scripta*, 33, 311–333.

Annexes

A1) PCR programs for amplification of mtDNA sequences;

Cyt-b (both fragments)

Amplification Step	Nr. Of Cycles	Temperature (ºC)	Duration
Initial Denaturation	1	95	10 minutes
Denaturation	5 (touchdown -0.5ºC)	95	30 seconds
Annealing		50	45 seconds
Extension		72	30 seconds
Denaturation	35	95	30 seconds
Annealing		48	45 seconds
Extension		72	30 seconds
Final Extension	1	72	10 minutes

D-loop

Amplification Step	Nr. Of Cycles	Temperature (ºC)	Duration
Initial Denaturation	1	95	15 minutes
Denaturation	40	95	30 seconds
Annealing		52	45 seconds
Extension		72	45 seconds
Final Extension	1	60	10 minutes

12S and 16S

Amplification Step	Nr. Of Cycles	Temperature (ºC)	Duration
Initial Denaturation	1	95	10 minutes
Denaturation	40	95	30 seconds
Annealing		52	30 seconds
Extension		72	40 seconds
Final Extension	1	72	5 minutes

A2) List of analysed microsatellite *loci* along with repeat motif, multiplex and reference;

Name	Repeat	Multiplex	Reference
AHT103	Di	Kit 2	Holmes et al, 1995
AHT111	Di	Kit 2	Holmes et al, 1993
AHT121	Di	Kit 3	Holmes et al, 1995
AHT132	Di	Kit AP	by N. Holmes
AHT137	Di	Kit 3	Holmes et al, 1995
AHTh171	Di	Kit 3	Breen et al. 2001
AHTh260	Compond	Kit 3	Breen et al. 2001
AHtk211	Di	Kit 3	Thomas et al. 1997
AHTk253	Di	Kit 3	Thomas et al. 1997
C04.140	Di	Kit 2	Ostrander et al. 1993
C08.140	Di	Kit 4	Ostrander et al. 1993
C08.618	Di	Kit 4	Ostrander et al. 1995
C09.173	Di	Kit 2	Ostrander et al. 1993
C09.474	Di	Kit 4	Ostrander et al. 1995
C13.758	Di	Kit 2	Mellersh et al. 1997
C14.866	Di	Kit 2	Mellersh et al. 1997
C20.253	Di	Kit 2	Ostrander et al. 1993
C20.446	Di	Kit 4	Ostrander et al. 1995
C22.763	Di	Kit 4	Mellersh et al. 1997
C27.442	Di	Kit AP	Ostrander et al. 1995
CPH05	Di	Kit 4	Fredholm & Wintero 1995
CPH14	Di	Kit 2	Fredholm & Wintero 1995
CPH2	Di	Kit 4	Fredholm & Wintero 1995
CPH9	Di	Kit 4	Fredholm & Wintero 1995
Cxx.459	Di	Kit 4	Ostrander et al. 1995
Cxx279	Di	Kit 3	Ostrander et al. 1993
FH2001	Tetra	Kit 2	Francisco et al. 1996
FH2054FZ	Tetra	Kit 3	Francisco et al. 1996
FH2161	Tetra	Kit 4	Francisco et al. 1996
FH2848	Di	Kit 3	Breen et al. 2001
FH2010	Tetra	Kit AP	Francisco et al. 1996
FH2079	Tetra	Kit AP	Francisco et al. 1996
INRA21	Di	Kit 3	Mariat et al. 1996
INU005	Di	Kit 3	Finnzymes, Inc
INU030	Di	Kit 3	Finnzymes, Inc
INU055	Di	Kit 3	Finnzymes, Inc
PEZ1	Tetra	Kit AP	Neff et al. 1999
PEZ3	Tri	Kit AP	Neff et al. 2000
PEZ5	Tetra	Kit AP	Neff et al. 2001
REN162C04	Di	Kit 3	Guyon et al. 2003
REN169D01	Di	Kit 3	Guyon et al. 2004
REN169O18	Di	Kit 3	Guyon et al. 2005
REN247M23	Di	Kit 3	Guyon et al. 2006
REN54P11	Di	Kit 3	Guyon et al. 2007
REN64E19	Di	Kit 4	Breen et al. 2001
VWF	Hexa	Kit 2	Shibuya et al. 1994

A3) Allele ranges of microsatellite *loci* for all study species;

Name	Allele Range					
	C. a.	C. l.	V. p.	V. r.	V. v.	V. z.
AHT103	88-102	98-112	86-108	94-110	94-106	96-108
AHT111	94-116	96-104	96-112	102-116	96-122	n/a
AHT121	80-114	74-108	88-112	86-108	80-112	106-140
AHT132	188-200	188-202	n/a	188-202	188-202	n/a
AHT137	132-152	124-154	130-146	130-134	132-148	126-148
AHTh171	216-238	216-232	206-210	220-224	202-212	216-228
AHTh260	233-265	229-257	n/a	n/a	n/a	n/a
AHtk211	84-96	84-92	84-98	88-94	88-104	88-100
AHTk253	272-302	280-298	275-282	276-282	288-302	281-304
C04.140	152-180	152-168	158-186	164-178	152-180	164-176
C08.140	113-135	115-145	115-123	129-145	112-115	115-117
C08.618	210-224	208-224	204-228	208-222	208-224	206-218
C09.173	122-132	238-254	140-156	140-144	128-156	156-164
C09.474	220-244	131-151	n/a	n/a	n/a	n/a
C13.758	240-262	259-271	236-248	246-256	244-260	242-262
C14.866	257-275	115-133	241-243	253-269	253-277	263-271
C20.253	113-137	120-138	n/a	n/a	n/a	n/a
C20.446	196-216	202-212	194	194	192-194	194
C22.763	210-236	218-231	n/a	n/a	n/a	n/a
C27.442	180-204	180-192	184-194	190-198	184-202	184-202
CPH05	127-149	127-143	122-142	n/a	n/a	n/a
CPH14	203-221	203-219	209-225	n/a	n/a	n/a
CPH2	112-132	112-124	n/a	n/a	n/a	n/a
CPH9	156-176	156-182	163-183	165	166-175	169-170
Cxx.459	170-183	161-183	181-201	179-196	175-200	189-199
Cxx279	110-128	112-130	104-114	110-122	112-136	118-126
FH2001	146-176	145-169	154-174	157-158	134-165	159-167
FH2054FZ	121-181	145-177	157-184	156-184	144-191	155-173
FH2161	234-281	247-275	n/a	n/a	n/a	n/a
FH2848	220-244	224-242	212-242	210-238	218-242	232-244
FH2010	233-253	233-253	225-249	233-245	233-249	225-229
FH2079	272-308	280-296	n/a	n/a	n/a	n/a
INRA21	85-103	87-99	91-113	91-101	87-109	91-105
INU005	112-132	110-132	120-124	n/a	n/a	n/a
INU030	136-154	136-152	123-146	133-135	127-152	n/a
INU055	198-212	196-210	n/a	198-210	196-220	202-214
PEZ1	113-149	121-141	n/a	n/a	n/a	n/a
PEZ3	124-169	127-163	133-163	127-160	127-166	139-157
PEZ5	115-135	115-135	127-139	131	115-131	119-131
REN162C04	191-211	189-213	189-207	195-205	189-209	193-209
REN169D01	192-218	198-222	192-210	198-200	194-198	196
REN169O18	145-171	153-169	157-176	n/a	n/a	n/a
REN247M23	263-279	267-279	259-283	265-279	259-283	265-283
REN54P11	227-241	223-241	22-256	236-244	218-256	240-256
REN64E19	155-173	159-177	189-205	185-189	171-189	186-198
VWF	157-205	157-193	139	139	139-157	139

C. a. – *Canis aureus*; C. l. – *C. lupus* and *C. l. familiaris*; V. p. – *Vulpes pallida*; V. r. – *V. rueppellii*; V. v. - *V. vulpes*; V. z. - *V. zerda*; n/a – did not amplify.

A4) PCR programs for each microsatellite multiplex (continues next page);

Kit 2

Amplification Step	Nr. Of Cycles	Temperature (ºC)	Duration
Initial Denaturation	1	95	15 minutes
Denaturation	35	95	30 seconds
Annealing		56	45 seconds
Extension		72	45 seconds
Denaturation	8	95	30 seconds
Annealing		53	45 seconds
Extension		72	45 seconds
Final Extension	1	60	30 minutes

Kit 3 (Finnzymes)

Amplification Step	Nr. Of Cycles	Temperature (ºC)	Duration
Initial Denaturation	1	98	3 minutes
Denaturation	40	98	15 seconds
Annealing		60	1.15 minutes
Extension		72	45 seconds
Final Extension	1	72	5 minutes

Kit 4

Amplification Step	Nr. Of Cycles	Temperature (ºC)	Duration
Initial Denaturation	1	95	15 minutes
Denaturation	7 (touchdown -0.5ºC)	95	30 seconds
Annealing		60	45 seconds
Extension		72	45 seconds
Denaturation	22	95	30 seconds
Annealing		57	45 seconds
Extension		72	45 seconds
Denaturation	8	95	30 seconds
Annealing		53	45 seconds
Extension		72	45 seconds
Final Extension	1	60	30 minutes

Kit AP (1 and 2)

Amplification Step	Nr. Of Cycles	Temperature (ºC)	Duration
Initial Denaturation	1	95	15 minutes
Denaturation	20 (touchdown -0.1ºC)	95	30 seconds
Annealing		58	45 seconds
Extension		72	45 seconds
Denaturation	15	95	30 seconds
Annealing		56	45 seconds
Extension		72	45 seconds
Denaturation	10	95	30 seconds
Annealing		53	45 seconds
Extension		72	45 seconds
Final Extension	1	60	30 minutes

A5) List of Cyt-b and D-loop sequences of wolf-like canids available on GenBank and used in this study (continues on next page);

Taxon (as in GenBank)	Analysis Code Name	Accession Number		Country (Region)	Reference
		Cyt-b	D-loop		
Canis aureus	CaurIndia1	AY291427	AY289996	India	Aggarwal et al. (2007)
Canis aureus	CaurIndia2	AY291428	AY289997	India	Aggarwal et al. (2007)
Canis aureus	CaurKenya	AF028138	-	Kenya	Wayne et al. (1997)
Canis aureus	CaurSen1	JQ088656	JQ088675	Senegal	Gaubert et al. (2012)
Canis aureus	CaurSen2	JQ088657	JQ088676	Senegal	Gaubert et al. (2012)
Canis adustus	Cadustus	AF028136	-	East Africa	Wayne et al. (1997)
Canis adustus	Cadustus1	JQ088650	JQ088669	Guinea	Gaubert et al. (2012)
Canis adustus	Cadustus2	JQ088651	JQ088670	Guinea	Gaubert et al. (2012)
Canis adustus	Cadustus3	JQ088652	JQ088671	Guinea	Gaubert et al. (2012)
Canis adustus	Cadustus4	JQ088653	JQ088672	Guinea	Gaubert et al. (2012)
Canis adustus	Cadustus5	JQ088654	JQ088673	Guinea	Gaubert et al. (2012)
Canis adustus	Cadustus6	JQ088655	JQ088674	Benin	Gaubert et al. (2012)
Canis himalayensis	Himalayensis1	AY291406	AY289985	India	Aggarwal et al. (2007)
Canis himalayensis	Himalayensis2	AY291407	AY289977	India	Aggarwal et al. (2007)
Canis himalayensis	Himalayensis3	AY291408	AY289983	India	Aggarwal et al. (2007)
Canis himalayensis	Himalayensis4	AY291409	AY289978	India	Aggarwal et al. (2007)
Canis himalayensis	Himalayensis5	AY291410	AY289979	India	Aggarwal et al. (2007)
Canis himalayensis	Himalayensis6	AY291421	AY289980	India	Aggarwal et al. (2007)
Canis himalayensis	Himalayensis7	AY291411	AY289981	India	Aggarwal et al. (2007)
Canis himalayensis	Himalayensis8	AY291422	AY289982	India	Aggarwal et al. (2007)
Canis himalayensis	Himalayensis9	AY291412	AY289992	India	Aggarwal et al. (2007)
Canis himalayensis	Himalayensis10	AY291413	AY289986	India	Aggarwal et al. (2007)
Canis himalayensis	Himalayensis11	AY291414	AY289993	India	Aggarwal et al. (2007)
Canis himalayensis	Himalayensis12	AY291415	AY289994	India	Aggarwal et al. (2007)
Canis himalayensis	Himalayensis13	AY291416	AY289995	India	Aggarwal et al. (2007)
Canis himalayensis	Himalayensis14	AY291417	AY289987	India	Aggarwal et al. (2007)
Canis himalayensis	Himalayensis15	AY291423	AY289990	India	Aggarwal et al. (2007)
Canis himalayensis	Himalayensis16	AY291418	AY289988	India	Aggarwal et al. (2007)
Canis himalayensis	Himalayensis17	AY291419	AY289989	India	Aggarwal et al. (2007)
Canis himalayensis	Himalayensis18	AY291420	AY289991	India	Aggarwal et al. (2007)
Canis indica	Indica1	AY291424	AY289974	India	Aggarwal et al. (2007)
Canis indica	Indica3	AY291426	AY289976	India	Aggarwal et al. (2007)
Canis indica	Indica4	AY291425	AY289984	India	Aggarwal et al. (2007)
Canis latrans	Coy1	DQ480509		EUA	Bjornerfeldt et al. (2006)
Canis latrans	Coy2	DQ480510		EUA	Bjornerfeldt et al. (2006)
Canis latrans	Coy3	AF028140	-	EUA	Wayne et al. (1997)
Canis lupus	LupusCan	DQ480508		Canada	Bjornerfeldt et al. (2006)
Canis lupus	LupusCan2	AY928668	-	Canada	Koepfli et al. (2006)
Canis lupus	LupusChi19	EU789787		China	Pang et al. (2009)
Canis lupus	LupusChi20	EU789788		China	Pang et al. (2009)
Canis lupus	LupusEgy	JQ088658	JQ088677	Egypt (Sinai)	Gaubert et al. (2012)
Canis lupus	LupusPt	EF689057	-	Portugal	Fernandes et al. (2008)
Canis lupus	LupusSp1	DQ480505		Spain	Bjornerfeldt et al. (2006)
Canis lupus	LupusSp2	AJ441334	-	Spain	Palomares et al. (2002)
Canis lupus	LupusMong	JX013645		Mongolia	unpublished
Canis lupus	LupusRus	DQ480503		Russia	Bjornerfeldt et al. (2006)
Canis lupus	LupusSA1	DQ480506		Saudi Arabia	Bjornerfeldt et al. (2006)
Canis lupus	LupusSA2	DQ480507		Saudi Arabia	Bjornerfeldt et al. (2006)
Canis lupus	LupusSwe1	DQ480504		Sweden	Bjornerfeldt et al. (2006)

Canis lupus	LupusSwe2	AM711902		Sweden	Arnason et al. (2006)
Canis lupus bayleyi	Bayleyi	HM222711	-	EUA	Naidu et al. (2012)
Canis lupus chanco	Chanco1	EU442884		-	unpublished
Canis lupus chanco	Chanco2	GQ374438		China (Inner Mongolia)	unpublished
Canis lupus chanco	Chanco3	AY333748	AY333738	Indian Sub-Continent	Sharma et al. (2003)
Canis lupus chanco	Chanco4	*	AY333739	Indian Sub-Continent	Sharma et al. (2003)
Canis lupus chanco	Chanco5	*	AY333740	Indian Sub-Continent	Sharma et al. (2003)
Canis lupus chanco	Chanco6	*	AY333741	Indian Sub-Continent	Sharma et al. (2003)
Canis lupus chanco	Chanco7	*	AY333742	Indian Sub-Continent	Sharma et al. (2003)
Canis lupus laniger	Laniger	FJ032363		Tibete	Meng et al. (2009)
Canis lupus lupaster	LupasterAl1	JQ088659	JQ088678	Algeria	Gaubert et al. (2012)
Canis lupus lupaster	LupasterAl2	JQ088660	JQ088679	Algeria	Gaubert et al. (2012)
Canis lupus lupaster	LupasterAl3	JQ088661	JQ088680	Algeria	Gaubert et al. (2012)
Canis lupus lupaster	LupasterAl4	JQ088662	JQ088681	Algeria	Gaubert et al. (2012)
Canis lupus lupaster	LupasterAl5	JQ088663	JQ088682	Algeria	Gaubert et al. (2012)
Canis lupus lupaster	LupasterSen	JQ088664	JQ088683	Senegal	Gaubert et al. (2012)
Canis lupus lupaster	LupasterMal	JQ088665	JQ088684	Mali	Gaubert et al. (2012)
Canis lupus lupaster	LupasterEg1	HQ845258	HQ845259	Ethiopia	Rueness et al. (2011)
Canis lupus pallipes	Pallipes1	AY333749	AY333743	Indian Sub-Continent	Sharma et al. (2003)
Canis lupus pallipes	Pallipes2	*	AY333744	Indian Sub-Continent	Sharma et al. (2003)
Canis lupus pallipes	Pallipes3	*	AY333745	Indian Sub-Continent	Sharma et al. (2003)
Canis lupus pallipes	Pallipes4	*	AY333746	Indian Sub-Continent	Sharma et al. (2003)
Canis lycaon	Lycaon	JF342907		-	unpublished
Canis mesomelas	Mesomelas1	AF028142	-	East Africa	Wayne et al. (1997)
Canis mesomelas	Mesomelas2	AF028143	-	East Africa	Wayne et al. (1997)
Canis simensis	Simensis1	HQ845262	HQ845261	Ethiopia	Rueness et al. (2011)
Canis simensis	Simensis2	L29415	-	Ethiopia	Gottelli et al. (1994)
Canis simensis	Simensis4	AF028144	-	Ethiopia	Wayne et al. (1997)
Cuon alpinus	Dhole1	GU063864		China	Zhang & Chen (2011)
Cuon alpinus	Dhole2	AY291429	-	India	Aggarwal et al. (2007)
Cuon alpinus	Dhole3	AY291430	-	India	Aggarwal et al. (2007)
Cuon alpinus	Dhole4	JN709955	-	Vietnam	unpublished
Cuon alpinus	Dhole5	AF028137	-	-	Wayne et al. (1997)
Lycaon pictus	Lpictus1	AF028147	-	-	Wayne et al. (1997)
Lycaon pictus	Lpictus2	S69130	-	-	Girman et al. (1993)

*** - In Sharma et al. (2003) only one Cyt-b haplotype was found for *C. l. chanco* and *C. l. pallipes*.**

A6) List of 12S and 16S sequences of wolf-like canids available on GenBank and used in this study;

Taxon (as in GenBank)	Analysis Code Name	Accession Number		Country (Region)	Reference
		12S	16S		
Canis aureus	CaurIndia1	-	AY289969	India	Aggarwal et al. (2007)
Canis aureus	CaurIndia2	-	AY289970	India	Aggarwal et al. (2007)
Canis himalayensis	Himalayensis1	-	AY289946	India	Aggarwal et al. (2007)
Canis himalayensis	Himalayensis2	-	AY289947	India	Aggarwal et al. (2007)
Canis himalayensis	Himalayensis3	-	AY289948	India	Aggarwal et al. (2007)
Canis himalayensis	Himalayensis4	-	AY289949	India	Aggarwal et al. (2007)
Canis himalayensis	Himalayensis5	-	AY289950	India	Aggarwal et al. (2007)
Canis himalayensis	Himalayensis6	-	AY289951	India	Aggarwal et al. (2007)
Canis himalayensis	Himalayensis7	-	AY289952	India	Aggarwal et al. (2007)
Canis himalayensis	Himalayensis8	-	AY289953	India	Aggarwal et al. (2007)
Canis himalayensis	Himalayensis9	-	AY289954	India	Aggarwal et al. (2007)
Canis himalayensis	Himalayensis10	-	AY289955	India	Aggarwal et al. (2007)
Canis himalayensis	Himalayensis11	-	AY289956	India	Aggarwal et al. (2007)
Canis himalayensis	Himalayensis12	-	AY289957	India	Aggarwal et al. (2007)
Canis himalayensis	Himalayensis13	-	AY289958	India	Aggarwal et al. (2007)
Canis himalayensis	Himalayensis14	-	AY289959	India	Aggarwal et al. (2007)
Canis himalayensis	Himalayensis15	-	AY289960	India	Aggarwal et al. (2007)
Canis himalayensis	Himalayensis16	-	AY289961	India	Aggarwal et al. (2007)
Canis himalayensis	Himalayensis17	-	AY289962	India	Aggarwal et al. (2007)
Canis himalayensis	Himalayensis18	-	AY289963	India	Aggarwal et al. (2007)
Canis indica	Indica1	-	AY289964	India	Aggarwal et al. (2007)
Canis indica	Indica3	-	AY289966	India	Aggarwal et al. (2007)
Canis indica	Indica4	-	AY289968	India	Aggarwal et al. (2007)
Canis latrans	Coy1	DQ480509		EUA	Bjornerfeldt et al. (2006)
Canis latrans	Coy2	DQ480510		EUA	Bjornerfeldt et al. (2006)
Canis lupus	LupusCan	DQ480508		Canada	Bjornerfeldt et al. (2006)
Canis lupus	LupusChi19	EU789787		China	Pang et al. (2009)
Canis lupus	LupusChi20	EU789788		China	Pang et al. (2009)
Canis lupus	LupusSp1	DQ480505		Spain	Bjornerfeldt et al. (2006)
Canis lupus	LupusRus	DQ480503		Russia	Bjornerfeldt et al. (2006)
Canis lupus	LupusSA1	DQ480506		Saudi Arabia	Bjornerfeldt et al. (2006)
Canis lupus	LupusSA2	DQ480507		Saudi Arabia	Bjornerfeldt et al. (2006)
Canis lupus	LupusSwe1	DQ480504		Sweden	Bjornerfeldt et al. (2006)
Canis lupus	LupusSwe2	AM711902		Sweden	Arnason et al. (2006)
Canis lupus chanco	Chanco1	EU442884		-	unpublished
Canis lupus chanco	Chanco2	GQ374438		China (Inner Mongolia)	unpublished
Canis lupus laniger	Laniger	FJ032363		Tibete	Meng et al. (2009)
Canis lupus lupaster	LupasterEg1	HQ845256	HQ845257	Ethiopia	Rueness et al. (2011)
Canis lycaon	Lycaon	JF342907		-	unpublished
Cuon alpinus	Dhole1	GU063864		China	Zhang & Chen (2011)

A7) List of Cyt-b and D-loop sequences of fox-like canids available on GenBank and used in this study;

Taxon (as in GenBank)	Analysis Code Name	Accession Number		Country (Region)	Reference
		Cyt-b	D-loop		
Urocyon cinereoargenteus	Urocyon	JF489121	-	-	Naidu et al. (2012)
Alopex lagopus	Vlagopus	AY598511	-	-	Delisle & Strobeck (2005)
Vulpes macrotis	Vmacrotis	JF489127	-	-	Naidu et al. (2012)
Vulpes vulpes	VvJap1	AB292741		Japan (Hokkaido)	Inoue et al. (2007)
Vulpes vulpes	VvJap2	AB292742		Japan (Hokkaido)	Inoue et al. (2007)
Vulpes vulpes	VvJap3	AB292743		Japan (Hokkaido)	Inoue et al. (2007)
Vulpes vulpes	VvJap4	AB292744		Japan (Hokkaido)	Inoue et al. (2007)
Vulpes vulpes	VvJap5	AB292745		Japan (Hokkaido)	Inoue et al. (2007)
Vulpes vulpes	VvJap6	AB292746		Japan (Hokkaido)	Inoue et al. (2007)
Vulpes vulpes	VvJap7	AB292747		Japan (Hokkaido)	Inoue et al. (2007)
Vulpes vulpes	VvJap8	AB292748		Japan (Hokkaido)	Inoue et al. (2007)
Vulpes vulpes	VvJap9	AB292749		Japan (Hokkaido)	Inoue et al. (2007)
Vulpes vulpes	VvJap10	AB292750		Japan (Hokkaido)	Inoue et al. (2007)
Vulpes vulpes	VvJap11	AB292751		Japan (Hokkaido)	Inoue et al. (2007)
Vulpes vulpes	VvJap12	AB292752		Russia (Primorsky Krai)	Inoue et al. (2007)
Vulpes vulpes	VvJap13	AB292753		Japan (Hokkaido)	Inoue et al. (2007)
Vulpes vulpes	VvJap14	AB292754		Russia (Primorsky Krai)	Inoue et al. (2007)
Vulpes vulpes	VvJap15	AB292755		Japan (Aomori)	Inoue et al. (2007)
Vulpes vulpes	VvJap16	AB292756		Japan (Aomori)	Inoue et al. (2007)
Vulpes vulpes	VvJap17	AB292757		Japan (Aomori)	Inoue et al. (2007)
Vulpes vulpes	VvJap18	AB292758		Japan (Kanagawa)	Inoue et al. (2007)
Vulpes vulpes	VvJap19	AB292759		Japan (Miyazaki)	Inoue et al. (2007)
Vulpes vulpes	VvJap20	AB292760		Japan (Kanagawa)	Inoue et al. (2007)
Vulpes vulpes	VvJap21	AB292761		Japan (Hokkaido)	Inoue et al. (2007)
Vulpes vulpes	VvJap22	AB292762		Japan (Hokkaido)	Inoue et al. (2007)
Vulpes vulpes	VvJap23	AB292763		Japan (Hokkaido)	Inoue et al. (2007)
Vulpes vulpes	VvJap24	AB292764		Japan (Hokkaido)	Inoue et al. (2007)
Vulpes vulpes	VvJap25	AB292765		Japan (Hokkaido)	Inoue et al. (2007)
Vulpes vulpes	VvChi	GQ374180		China	unpublished
Vulpes vulpes	VvKor1	DQ498124	-	Korea	unpublished
Vulpes vulpes	VvKor2	DQ498125	-	Korea	unpublished
Vulpes vulpes	VvKor3	DQ498126	-	Korea	unpublished
Vulpes vulpes	VvKor4	DQ498127	-	Korea	unpublished
Vulpes vulpes	VvKor5	JN711443		Korea	Yu et al. (2012)
Vulpes vulpes	VvMong1	JX013646	-	Mongolia	unpublished
Vulpes vulpes	VvMong2	JX013647	-	Mongolia	unpublished
Vulpes vulpes	VvMong3	JX013648	-	Mongolia	unpublished
Vulpes vulpes	VvPt1	EF689058	-	Portugal	Fernandes et al. (2008)
Vulpes vulpes	VvPt2	EF689059	-	Portugal	Fernandes et al. (2008)
Vulpes vulpes	VvSp2	EF689060	-	Spain	Fernandes et al. (2008)
Vulpes vulpes	VvSp3	EF689061	-	Spain	Fernandes et al. (2008)
Vulpes vulpes	VvUkraine	EF689062	-	Ukraine (Kharkiv)	Fernandes et al. (2008)
Vulpes vulpes	VvSwe	AM181037		Sweden	Arnason et al. (2006)
Vulpes vulpes	VvUk	AY928669	-	United Kingdom	Koepfli et al. (2006)

A8) List of *C. aureus* and *C. adustus* samples that amplified at least for one type of molecular marker;

C. aureus

Code	Country	Province	Latitude	Longitude	Markers
82	Mauritania	Hodh El Gharbi	16,489217	-10,643133	Cyt-b; D-loop; M
116	Mauritania	Brakna	17,392517	-13,45285	M
142	Morocco	Dakhla	24,422417	-15,141083	D-loop; M
332	Libya	Ash Shati	27,499233	13,776617	D-loop
494	Mauritania	Trarza	17,551683	-16,028833	Cyt-b; D-loop; M
514	Mauritania	Inchiri	19,8448	-14,246483	Cyt-b; D-loop; M
682*	Morocco	Midelt-Er Rachidia	31,951183	-4,471917	Cyt-b; M
793	Mauritania	Dakhlet-Nouâdhibou	21,154933	-16,983	D-loop; M
931	Algeria	Kabylia	36,583183	4,147183	Cyt-b; M
1005	Morocco	Dakhla - Fort Guerguerat	22,53985	-16,389883	D-loop
1113	Morocco	Akka - Guelmim	28,8129	-8,907967	Cyt-b; D-loop; M
1282	Morocco	Dakhla - Fort Guerguerat	23,422183	-15,975567	Cyt-b; M
1516	Algeria	Tissemsilt	35,841617	1,9866	Cyt-b; M
1517	Algeria	Tizi-Ouzou	36,6973	3,985167	Cyt-b; M
1518	Algeria	Tizi-Ouzou	36,721183	4,5347	Cyt-b; D-loop; 12S; 16S; M
1519	Algeria	Tizi-Ouzou	36,711667	4,59	Cyt-b; D-loop; 12S; 16S; M
1592	Morocco	Tan-Tan - Laayoune	27,95904	-12,83306	Cyt-b; M
1606	Morocco	Boujdour - Dakhla	24,79455	-14,8639	D-loop; M
1607	Morocco	Dakhla - Fort Guerguerat	22,86779	-16,21969	Cyt-b; D-loop
2202	Mauritania	Hodh El Gharbi	16,42812	-9,561713	Cyt-b; M
2204	Mauritania	Hodh El Gharbi	16,425255	-9,562955	D-loop
2646	Mauritania	Brakna	17,415093	-12,528163	Cyt-b; D-loop; 12S; 16S; M
2800	Azerbaijan	Nagorno Karabakh	39,91777	46,74149	D-loop
3054	Mauritania	Tagant	17,929848	-11,745102	Cyt-b; D-loop; 12S; 16S; M
3549	Mauritania	Dakhlet-Nouâdhibou	19,400027	-16,397685	D-loop
3644	Morocco	Boulemane	33,55234	-3,4247	D-loop; M
4439	Mauritania	Trarza	18,35084	-16,00829	D-loop; M
4982	Mauritania	Brakna	17,1883	-13,8541	D-loop; M
4992	Mauritania	Trarza	18,374573	-15,737555	Cyt-b; M
5700	Serbia	Southern Serbia	42,547378	21,933753	Cyt-b; D-loop; 12S; 16S; M
5725	Western Sahara	Rio de Oro	24,576168	-14,988853	Cyt-b; D-loop; M
5998	Mauritania	Trarza	18,795397	-16,13914	Cyt-b; D-loop; 12S; 16S; M
5999	Mauritania	Trarza	18,259342	-16,01784	Cyt-b; D-loop; M
6047	Mauritania	Tagant	18,292045	-11,728962	Cyt-b; D-loop; M
6576	Slovenia	Savinjska	46,28	14,81	Cyt-b; D-loop; 12S; 16S; M
6577	India	Madhya Pradesh	22,283	80,618	D-loop

*** - dog sample; M – microsatellites**

C. adustus

Code	Country	Province	Latitude	Longitude	Markers
2799	Ethiopia	Southern Nations	5,789065	37,446685	Cyt-b; D-loop; 12S; 16S
6595	Senegal	Louga	15,28402	-14,76846	Cyt-b; D-loop; 12S; 16S

A9) List of *V. pallida* and *V. rueppellii* samples that amplified at least for one type of molecular marker;

V. pallida

Code	Country	Province	Latitude	Longitude	Markers
81	Mauritania	Hodh El Gharbi	16,44145	-10,466417	Cyt-b; D-loop; M
126	Mauritania	Brakna	17,278717	-13,7699	M
127	Mauritania	Brakna	17,278717	-13,7699	Cyt-b
130	Mauritania	Trarza	17,252583	-14,27515	Cyt-b; M
468	Senegal	Saint-Louis	15,356267	-13,07555	Cyt-b; D-loop
966	Mauritania	Assaba	16,997	-11,142983	Cyt-b
967	Mauritania	Assaba	16,517833	-10,812583	Cyt-b; D-loop; M
968	Mauritania	Assaba	16,514817	-10,779683	Cyt-b; M
1000	Mauritania	Tagant	17,291133	-12,355817	Cyt-b; D-loop; M
1001	Mauritania	Tagant	17,347933	-12,432883	Cyt-b; M
1285	Mauritania	Trarza	17,242067	-16,1067	Cyt-b
1286	Mauritania	Trarza	17,056467	-16,0854	Cyt-b; D-loop
1310	Mauritania	Brakna	16,30315	-13,88225	Cyt-b; D-loop; M
1319	Mauritania	Gorgol	16,224017	-13,231067	Cyt-b; D-loop; M
1395	Mauritania	Assaba	16,899167	-11,870033	Cyt-b
2230	Mauritania	Hodh El Gharbi	16,389973	-10,305747	Cyt-b; D-loop; M
2293	Mauritania	Hodh El Gharbi	16,487108	-10,680413	Cyt-b; D-loop
2432	Mauritania	Assaba	16,42379	-11,362582	Cyt-b; M
2434	Mauritania	Assaba	16,211545	-11,419975	Cyt-b; D-loop; M
3483	Mauritania	Brakna	17,126187	-14,06545	M
4584	Mauritania	Brakna	16,3266	-13,9302	Cyt-b; D-loop; M
4978	Mauritania	Brakna	17,3434	-13,625315	Cyt-b; D-loop; M
4987	Mauritania	Brakna	17,1776	-14,1466	Cyt-b; D-loop; M
4988	Mauritania	Brakna	17,1776	-14,1466	Cyt-b; D-loop; M
6635	Niger	Diffa	13,709617	11,154947	Cyt-b; M
6636	Niger	Diffa	13,709617	11,154947	Cyt-b; D-loop; M
6644	Niger	Diffa	13,184333	12,247653	Cyt-b; M
6645	Niger	Diffa	13,194893	12,406035	Cyt-b; D-loop; M
6646	Niger	Diffa	13,228687	12,439328	M

M – microsatellites

V. rueppellii

Code	Country	Province	Latitude	Longitude	Markers
84	Mauritania	Assaba	17,03975	-11,967483	Cyt-b; D-loop; M
131	Mauritania	Trarza	17,339167	-14,388583	Cyt-b
542	Morocco	Dakhla - Fort Guerguerat	23,091183	-16,077217	Cyt-b
785	Morocco	Dakhla - Fort Guerguerat	21,80795	-16,88215	Cyt-b; D-loop; M
795	Morocco	Dakhla - Fort Guerguerat	21,806583	-16,8824	Cyt-b; D-loop
2782	Morocco	Boujdour - Dakhla	24,730292	-14,87428	Cyt-b; D-loop; M
3213	Mauritania	Tagant	17,333602	-12,077818	Cyt-b; D-loop; M
4964	Mauritania	Brakna	17,4445	-13,3037	Cyt-b; D-loop; M
5704	Egypt	Western Desert	29,78349	25,80814	Cyt-b; D-loop; M
6010	Mauritania	Tagant	17,840183	-12,410738	Cyt-b; D-loop; M

M – microsatellites

A10) List of *V. vulpes* and *V. zerda* samples that amplified at least for one type of molecular marker;

V. vulpes

Code	Country	Province	Latitude	Longitude	Markers
305	Tunisia	Beja	36,754617	9,195733	Cyt-b; D-loop; M
707	Morocco	Oukaimeden	31,200967	-7,855367	Cyt-b; D-loop; M
1092	Morocco	Ouarzazate - Foum Zguid	30,765767	-7,28775	Cyt-b; D-loop; M
1446	Morocco	Guelmim - Tan-Tan	28,8444	-10,23587	Cyt-b; D-loop; M
1515	Algeria	El Bayadh	33,8918	0,268567	Cyt-b; D-loop; M
2789	Portugal	Montalegre	41,835154	-7,778467	Cyt-b; D-loop; M
2790	Spain	Murcia	37,907917	-1,45875	Cyt-b; D-loop
2791	Portugal	Serpa	37,932017	-7,36225	Cyt-b; D-loop
2792	Portugal	Odemira	37,56805	-8,736783	Cyt-b; D-loop; M
2793	Greece	Peloponnese	37,71977	23,12499	Cyt-b; D-loop; M
2794	Greece	?	38,84853	22,51256	Cyt-b; D-loop; M
2795	Greece	?	41,25476	23,24988	Cyt-b; D-loop; M
2796	Greece	Macedonia	41,28124	22,98333	Cyt-b; D-loop; M
2797	Greece	?	39,70193	22,20824	D-loop; M
3642	Morocco	Ksar es Souk	32,316932	-4,5471	Cyt-b; D-loop; M
3643	Morocco	Nador	34,9352	-2,87944	Cyt-b; D-loop; M
5618	Turkey	Anatolia	37,28284	33,4711698	Cyt-b; D-loop; M
5627	Turkey	Anatolia	37,423177	31,6583369	Cyt-b; D-loop; M
5701	Morocco	Marrakech	31,21788	-7,84258	Cyt-b; D-loop
5705	Egypt	Port Said	30,9021	32,41613	Cyt-b; D-loop; M
5706	Armenia	Aragatsotn	40,384038	43,83411	D-loop
5707	Armenia	Yerevan	40,298833	44,1	Cyt-b; D-loop; M
5710	Morocco	Tan-Tan	28,799507	-10,418638	Cyt-b; D-loop; M
6588	Morocco	Agadir	30,55939	-7,98936	Cyt-b; D-loop; M
6589	Morocco	Agadir	30,59963	-8,08246	Cyt-b; D-loop; M
6820	Tunisia	Jendouba	36,49005	8,36994	Cyt-b; D-loop; M
6824	Tunisia	Le Kef	35,82826	8,56041	Cyt-b; D-loop; M
VV-06	Portugal	PNVG	?	?	M
VV-09	Portugal	Parque Biológico de Gaia	?	?	M
VV-10	Portugal	PNPG	?	?	M
VV-12	Portugal	PNSE	?	?	M
VV-13	Portugal	PNSE	?	?	M
VV-14	Portugal	PNSE	?	?	M
VV-15	Portugal	?	?	?	M

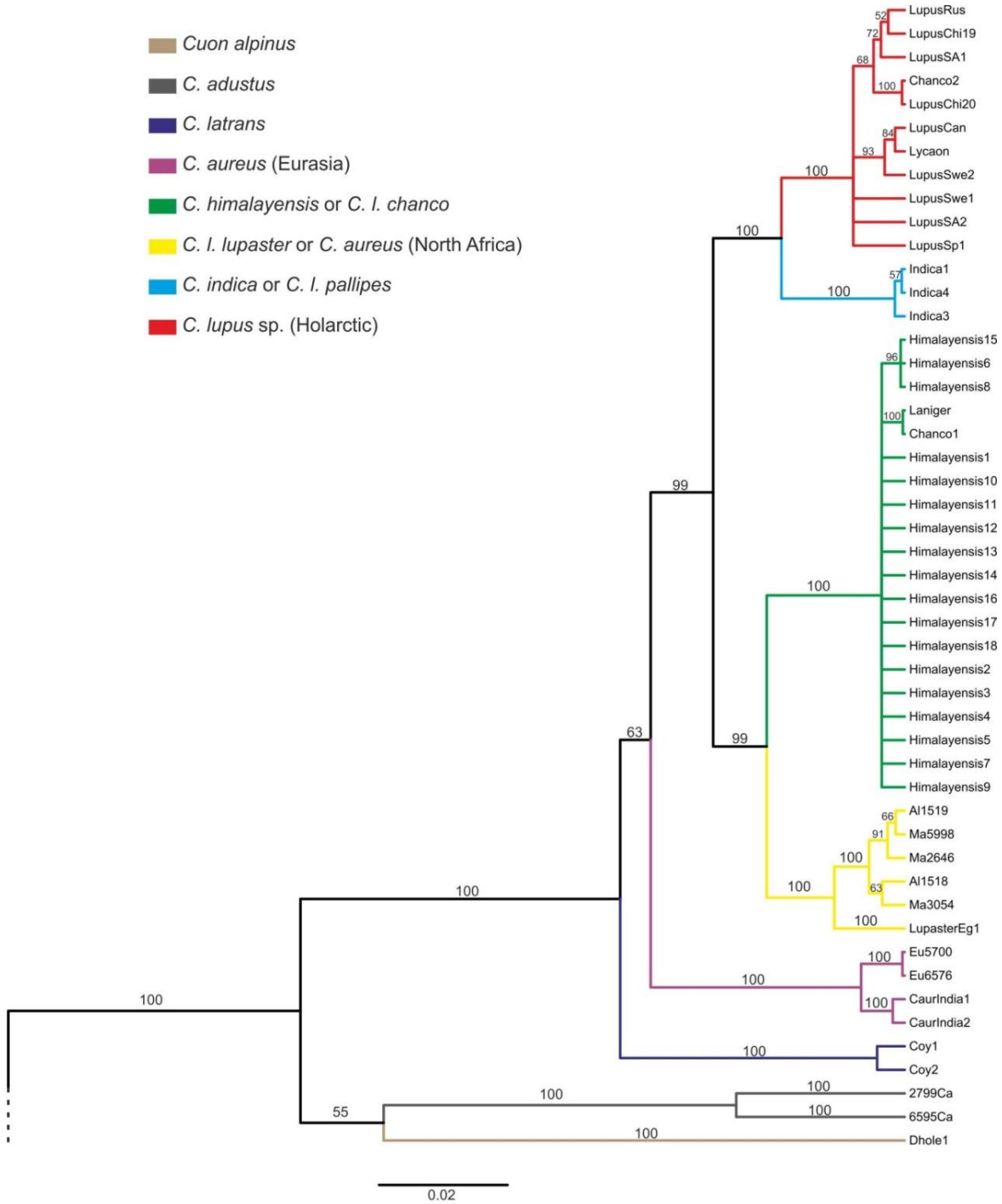
M – microsatellites; PNVG – Parque Nacional do Vale do Guadiana; PNPG – Parque Nacional Peneda Gerês; PNSE – Parque Nacional da Serra da Estrela

V. zerda

Code	Country	Province	Latitude	Longitude	Markers
544	Morocco	Dakhla - Fort Guerguerat	23,380483	-15,99735	Cyt-b
548	Morocco	Boujdour - Dakhla	24,659867	-14,877617	Cyt-b; M
791	Mauritania	Trarza	18,471833	-14,6911	Cyt-b
792	Mauritania	Trarza	18,45295	-14,733633	Cyt-b; D-loop; M
3640	Morocco	Ksar es Souk	31,05349	-4,00795	Cyt-b; D-loop; M
5723	Western Sahara	Rio de Oro	25,03705	-14,807872	Cyt-b; D-loop; M
5765	Mauritania	Dakhlet-Nouâdhibou	21,205183	-14,182682	Cyt-b; D-loop; M
5918	Mauritania	Adrar	19,984442	-13,94555	Cyt-b; D-loop; M
6480	Western Sahara	Rio de Oro	22,012438	-16,82393	Cyt-b; D-loop; M
6483	Western Sahara	Rio de Oro	23,415322	-15,97861	Cyt-b; M
6485	Western Sahara	Rio de Oro	25,280352	-14,8179	Cyt-b; D-loop; M

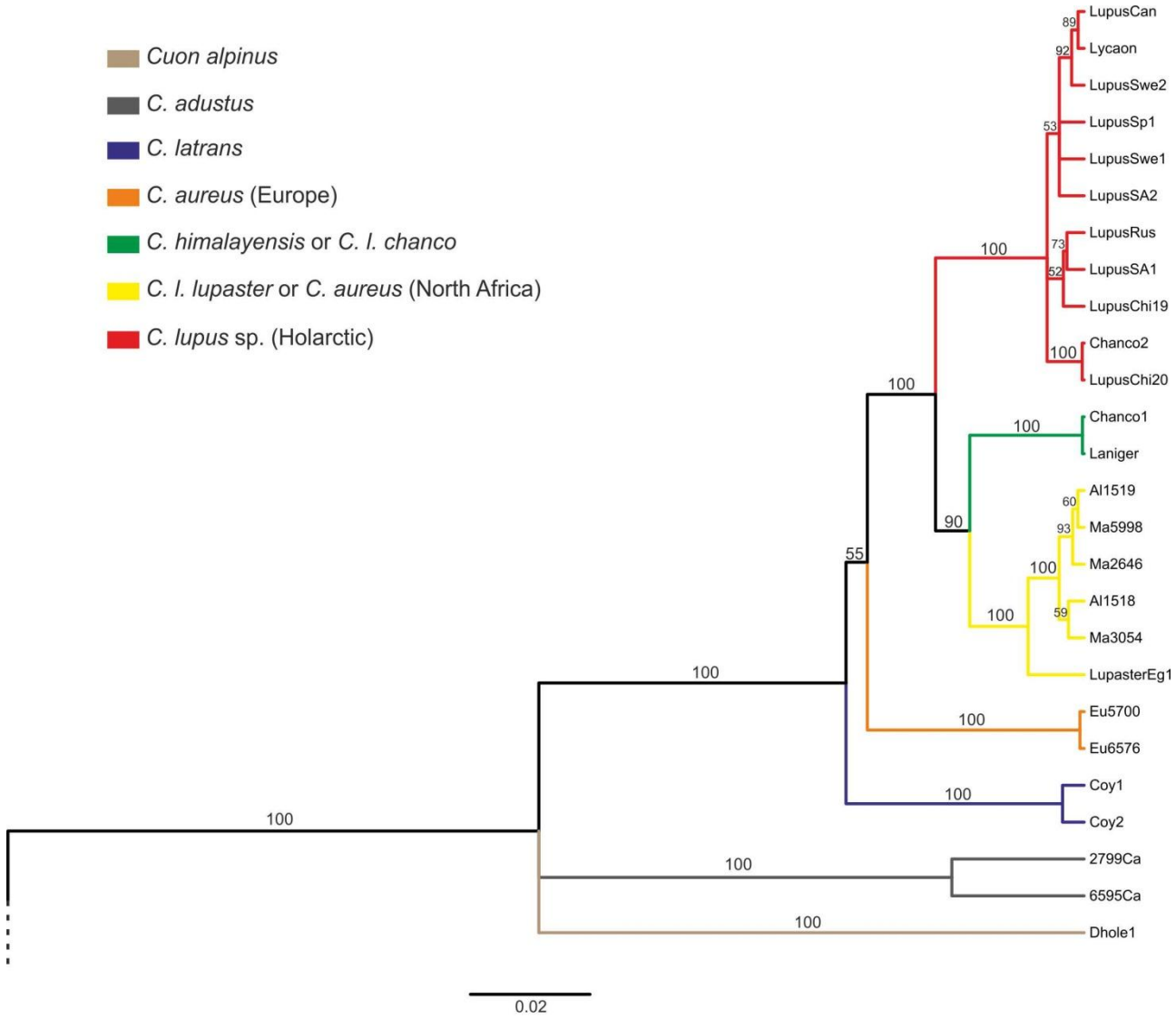
M – microsatellites

A11) Bayesian phylogenetic tree of wolf-like canids (16S, Cyt-b and D-loop);



Bayesian phylogenetics tree of wolf-like canids, based on the 836 bp 16S (239 bp), Cyt-b (317 bp) and D-loop (230 bp) concatenated dataset. The HKY+G evolutionary model was applied to the Cyt-b partition while the HKY+I+G to 16S and D-loop. *V. vulpes* (AM181037) was used as outgroup (not shown). Numbers above branches indicate Bayesian posterior probabilities (BPP, in percentage). Scale bar represents 2% sequence divergence. Colors highlight clades/lineages referred in the main text: *Cuon alpinus* (brown); *Canis adustus* (grey); C. latrans (dark blue); Eurasian *C. aureus* (purple); *C. himalayensis* or *C. lupus chanco* (green); *C. l. lupaster* or North African *C. aureus* (yellow); *C. indica* or *C. l. pallipes* (cyan); *C. lupus* sp. (red).

A12) Bayesian phylogenetic tree of wolf-like canids (12S, 16S, Cyt-b and D-loop);



Bayesian phylogenetics tree of wolf-like canids, based on the 1168 bp 12S (332 bp), 16S (239 bp), Cyt-b (317 bp) and D-loop (230 bp) concatenated dataset. The HKY+I, GTR+G, HKY+G, HKY+I+G evolutionary models were applied to 12S, 16S, Cyt-b and D-loop, respectively. *V. vulpes* (AM181037) was used as outgroup (not shown). Numbers above branches indicate Bayesian posterior probabilities (BPP, in percentage). Scale bar represents 2% sequence divergence. Colors highlight clades/lineages referred in the main text: *Cuon alpinus* (brown); *Canis adustus* (grey); *C. latrans* (dark blue); European *C. aureus* (orange); *C. himalayensis* or *C. lupus chanco* (green); *C. l. lupaster* or North African *C. aureus* (yellow); *C. lupus* sp. (red).

A13) List of haplotype names and corresponding sequences (new and published data) for *C. aureus* (Figure 7), *V. rueppellii* and *V. zerda* (Figure 11);

C. aureus

Haplotype	New Data	Published Data (Cyt-b / D-loop)
Al1	1519	JQ088661 / JQ088680
Al2	-	JQ088660 / JQ088679
Al3	-	JQ088663 / JQ088682
Al4	-	JQ088662 / JQ088681
Al5	-	JQ088659 / JQ088678
Al6	1518	-
Et1	-	HQ845258 / HQ845259
Ma1	5998	-
Ma2	4439	-
Ma3	2646	-
Ma4	494	JQ088664 / JQ088683
Ma5	82	-
Ma6	6047	-
Ma7	3054 / 5725	-
Ma8	5999	-
Ma9	514	-
ma1	-	JQ088665 / JQ088684
Mo1	1113	-
Mo2	1607	-
Se1	-	JQ088656 / JQ088675 & JQ088657 / JQ088676

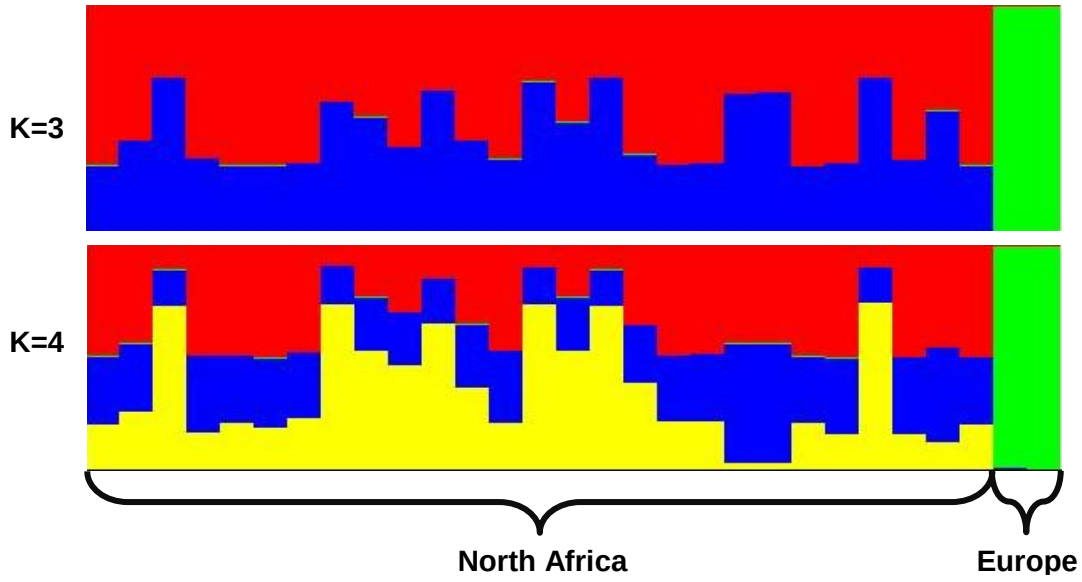
V. rueppellii

Haplotype	Sample Code
rEg1	5704
rMa1	4964
rMa2	84
rMa3	6010
rMa4	3213
rMo1	785
rMo2	2782
rMo3	795

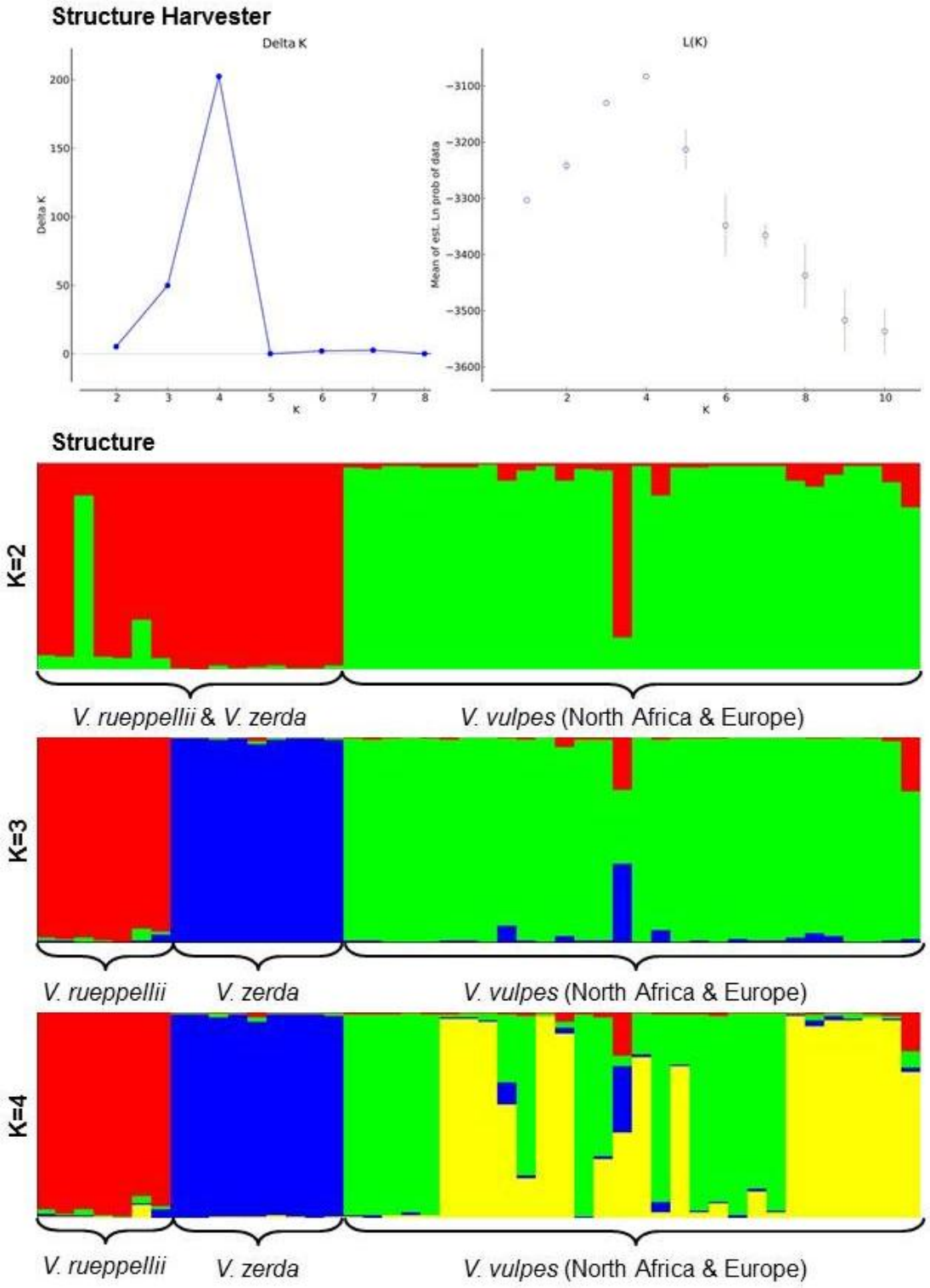
V. zerda

Haplotype	Sample Code
zMa1	792
zMa2	5765
zMa3	5918
zMo1	3640
zWs1	6485
zWs2	5723
zWs3	6480

A14) Population structure analyses of 35 *C. aureus* samples based on 34 microsatellite *loci* – Structure bar plot of Bayesian assignment of individuals to three (K=3) and four clusters (K=4). Vertical bars represent individuals, while colours within a bar represent probability of assignment of each individual to a cluster;



A15) Structure Harvester graphic output of Delta K and Mean L(K), and Structure bar plot of Bayesian assignments of *V. rueppellii*, *V. zerda* and *V. vulpes* (North Africa and Europe) individuals to two (K=2), three (K=3) and four (K=4) clusters. Analysis was based on 18 microsatellite *loci*. Vertical bars represent individuals, while colors within a bar represent probability of assignment of each individual to a cluster;



A16) PCA based on individual-by-individual genetic distances of *V. rueppellii*, *V. zerda* and *V. vulpes* (North Africa and Europe). Analysis was based on 18 microsatellite *loci*;