KTH Biotechnology

Analysis of the origin and spread of the domestic dog using Y-chromosome DNA and mtDNA sequence data

Mattias Oskarsson

KTH Royal Institute of Technology
School of Biotechnology
Gene Technology

Stockholm 2012

Printed by [Universitetsservice US-AB]

ISBN 978-91-7501-364-0

Mattias Oskarsson (2012): **Analysis of the origin and spread of the domestic dog using Y-chromosome DNA and mtDNA sequence data**
Division of Gene Technology, School of Biotechnology, Royal Institute of Technology (KTH), Stockholm, Sweden
ISBN 978-91-7501-364-0, TRITA-BIO Report 2012:6, ISSN 1654-2312

Abstract

The domestic dog was probably the first domesticated animal, and the only one to spread to all continents in ancient times. The dog is one of the most phenotypically diverse animals, a result of human selection throughout dog history. Studies of the genetic origins and early spread of domestic dogs is important to gather information about biological and cultural mechanisms behind domestication but also to investigate early human history. The step from a hunter and gatherer lifestyle to farming is one of the most important steps in human history. In this thesis I will present work aimed at understanding both domestic dog origins and dispersal. In order to be able to investigate dog origins based on a second haploid chromosome we identified 14,437 bp of Y-chromosomal DNA sequence. Based on this we show that dogs in Asia south of Yangtze River (ASY) has the highest genetic diversity and was founded from a large number of wolf founders confirming earlier mtDNA results. Early dog dispersal is tightly coupled to human history with the dog brought along as a cultural item. We have for the first time investigated the dog dispersal into Polynesia and Australia and our data can be used as evidence for a more complex settlement of Polynesia than earlier indicated from archaeological and linguistic studies. Analysis of Y-chromosome SNPs in Australian dingoes confirms earlier mtDNA genetic studies that the dingo is part of the domestic dog phylogeny and was founded from a small population of domestic dogs. We have also for the first time investigated the dog population on Madagascar and our data strongly indicates a mainland African origin for the Madagascan dogs. Finally, we have investigated the American dog population sampled from throughout the continent and also for the first time included putative indigenous breed dogs such as Chihuahua and Pero Sín Pelo del Peru, and the free-ranging Carolina dogs from southern USA. Our data clearly indicates a primarily Old World origin for the indigenous breed dogs and also for the free-ranging Carolina dogs in USA. We can also for the first time present evidence for continuity between the ancient and extant dog population with e.g. exclusive sharing of a haplotype between a modern sample of Chihuahua and an ancient Mexican sample.

Keywords: mtDNA, domestication, Y-chromosome, SNP, ASY, dog, dingo

LIST OF PUBLICATIONS

I. Natanaelsson, C., **Oskarsson, MCR**., Angleby, H., Lundeberg, J., Kirkness, E & Savolainen, P. Dog Y chromosomal DNA sequence: identification, sequencing and SNP discovery. *BMC Genetics* **7,** 45 (2006).

II. Ding, Z-L*., **Oskarsson, M*.**, Ardalan, A., Angleby, H., Dahlgren, L-G., Tepeli, C., Kirkness, E., Savolainen, P. & Zhang, Y-P. Origins of domestic dog in Southern East Asia is supported by analysis of Y-chromosome DNA. *Heredity* **108,** 507-514 (2012). [*These authors contributed equally to this work].

III. **Oskarsson, MCR.**, Klütsch, CFC., Boonyaprakob, U., Wilton, A., Tanabe, Y. & Savolainen, P. Mitochondrial DNA data indicate an introduction through Mainland Southeast Asia for Australian dingoes and Polynesian domestic dogs. *Proceedings of The Royal Society B* Epub ahead of print (2011).

IV. Ardalan, A*., **Oskarsson, M*.**, Natanaelsson, C., Wilton, A.N., Ahmadian, A. & Savolainen, P. Narrow genetic basis for the Australian dingo confirmed through analysis of paternal ancestry. *Genetica* In Press (2012). [*These authors contributed equally to this work].

V. **Oskarsson, M.**, Ardalan, A., Rabakonandriania, E. & Savolainen, P. African origin for Madagascan dogs revealed by mtDNA analysis. *Manuscript*.

VI. van Asch, B., Zhang, A-B., **Oskarsson, MC.**, Klütsch, C., Amorim, A. & Savolainen, P. mtDNA analysis confirms early Pre-Colombian origins of Native American dogs. *Manuscript.*

CONTENTS

LIST OF ABBREVIATIONS

ASY	Asia south of Yangtze River
BP	before present
bp	base pairs
DHPLC	denaturing high performance liquid chromatography
DLA	dog leukocyte antigen
D-loop	displacement loop
DNA	deoxyribonucleic acid
HG	haplogroup
HT	haplotype
ISEA	island southeast Asia
MHC	major histocompatibility complex
MYA	million years ago
mtDNA	mitochondrial DNA
NGSD	New Guinea singing dog
NRY	non-recombining region of the Y-chromosome
NUMT	nuclear mitochondrial DNA
PrASE	protease-mediated allele-specific extension
RFLP	restriction fragment length polymorphism
RNA	ribonucleic acid
rRNA	ribosomal RNA
SE	southeast
SNP	single nucleotide polymorphism
SW	southwest
tRNA	transfer RNA
UT	Universal haplotype

1 BACKGROUND

1.1 Introduction

The domestic dog (*Canis familiaris)* is one of our most beloved animals. The Fédération Cynologique Internationale currently recognizes 343 breeds sorted into 10 different breed groups, based on the breeds morphological appearance or special utility (Fédération Cynologique Internationale).

The dog is unique among our domestic animals in their exceptional variation with dogs as morphologically different as Great Dane and Chihuahua which differ nearly 40-fold in size (Parker, Shearin and Ostrander 2010). The domestic dog is also unique in its many different uses, compared to other domestic animals having only a few specific uses. For example, dogs are used as assistants to disabled people, by police to find missing persons or illegal substances, by hunters to track down prey, by the military to find land mines, or they are used as sled dogs, as guards or to herd flocks of sheep or even as a source of food. Beyond all these qualities they are also our loyal companions and are treated as family members.

There is evidence, both archaeological and genetic, that the dog was one of the first or possibly the first animal to be domesticated thousands of years before the transition from a hunter and gatherer to an agricultural lifestyle (Savolainen et al. 2002, Pang et al. 2009, Dayan 1994, Clutton-Brock 1995, Chaix 2000). Other domesticates emerged later, from approximately 10,000 years ago to the present day. At 10,000 years ago the dog was already domesticated and had started to spread around the world. The dog is also unique among domesticates in being the only animal which accompanied humans to all continents in ancient time.

However, where and when the dog was domesticated and whether it was domesticated from a single wolf population or from multiple populations is debated.

Archaeological and genetic studies tend to give different answers to these questions. Also different genetic studies do not agree on these questions; see for example (Pang et al. 2009) and (Vonholdt et al. 2010). Within the past 35 years major leaps in the field of molecular biology has made it possible to read the genetic code (Sanger, Nicklen and

Coulson 1977, Sanger 1988) and to copy specific segments of DNA (Saiki et al. 1985, Mullis and Faloona 1987). The development of these techniques and their use have made it possible to investigate DNA to trace domestication events and subsequent spread of domesticates.

In this thesis I will investigate the origin and early spread of the domestic dog using data obtained from dog mtDNA and Y-chromosome DNA.

The analysis of Y-chromosome DNA sequence in domestic dog agree with data obtained from earlier studies of mtDNA (Savolainen et al. 2002, Pang et al. 2009) indicating a single origin in Asia South of Yangtze River (ASY) from numerous wolves (paper I, paper II). Investigation of mtDNA in Australian dingoes, ancient Polynesian dogs and the extant dog population in Island Southeast Asia (ISEA) indicates an introduction route for dingoes and Polynesian dogs through ISEA (paper III). Y-chromosome SNP analysis in Australian dingoes indicates an ancestry from few domestic dog founders (paper IV) in agreement with mtDNA data (Savolainen et al. 2004). Investigation of mtDNA in the dog population of Madagascar indicates a largely mainland African origin for Madagascan dogs (paper V) and the mtDNA in the extant American dog population indicates an Old World origin and also continuity between the ancient and extant dog populations (paper VI).

1.2 From wolf/coyote to dog

Morphological, behavioral and genetic evidence clearly indicates an origin of the domestic dog from wolves (Clutton-Brock 1995, Wayne 1993, Vilà et al. 1997).

The lineages leading to coyote and wolf split from their common ancestor 1.5-4.5 Million years ago (MYA) according to fossil evidence (Nowak 2003, Tedford, Wang and Taylor 2009). After the split between coyotes and wolves on the North American continent one branch led to modern day coyotes (*Canis latrans*) and the other one led to modern day wolves (*Canis lupus*) (Nowak 2003).

A closer look at the branch leading to modern wolf suggests that the branch split again with one branch evolving in America and leading to the dire wolf, a large variant of wolf which later became extinct (Nowak 2003), and the other branch evolving in the

Old World leading to the modern day gray wolves (Nowak 2003) . The modern gray wolf seems to have repopulated the American continent less than 130,000 years ago (Nowak 2003).

The physical appearance of wolves in different regions and echological niches varies extensively (Nowak 1996). For example the Eurasian wolf (*Canis lupus lupus*) weighs between 32-80 kg while the Chinese wolf (*Canis lupus chanco*) is much smaller, rarely exceeding 45kg in weight (Heptner et al. 1967) .

1.3 The fossil record

Before the advent of molecular techniques the only means of studying dog domestication was through the fossil record.

Archaeological researchers use both metric and non-metric observations in fossilized bones from excavated material to distinguish between dogs and other canid species. Also, the material excavated is usually radiocarbon dated. Dating of archaeological material can be made either directly on the fossilized bones or indirectly by association to a cultural layer or radiocarbon dating of the layer. Obviously, direct dating on the specimen is preferred since objects can be transferred between different layers in an archaeological site (Clutton-Brock 1995).

For a long time the oldest specimen identified as a domestic dog was the Bonn-Oberkassel dog (Nobis 1979, Street 2002), based on a single jaw fragment which was radiocarbon dated at 14,000 years ago. Other old dog specimens have been excavated in Israel from the Natufian sites of Ein Mallah and Hayonim terrace dated at approximately 11,500 years ago (Davis and Valla 1978, Tchernov and Valla 1997, Dayan 1994) . Recently a cranial fragment and a tooth identified as domestic dog were detected in material from Kesslerloch cave, Switzerland and dated at 14,100-14,600 years ago (Napierala and Uerpmann 2010).

Recently, canid remains excavated in Eliseevichi I, Russia, radiocarbon dated at 13,000-17,000 years ago (Sablin and Khlopachev 2002) , Goyet, Belgium, 31,700 years BP (Germonpré et al. 2009), and Razboinichya cave, Siberia, 33,000 years ago (Ovodov et al. 2011) have been reported. This means that the earliest findings of

domestic dogs can be pushed back approximately 19,000 years. However, the status of these canids is not entirely clear and archaeologists have contested some of the findings (Napierala and Uerpmann 2010, Pionnier-Capitan et al. 2011, Wang and Tedford 2008). Therefore, the earliest reasonably firm evidence for domestic dogs is about 14,000 years old and is found on the European continent (Napierala and Uerpmann 2010). However, Europe and SW Asia are well investigated while other areas such as the southern part of China remains relatively poorly investigated, and therefore early domestic dogs might remain undetected in underrepresented areas.

An origin of dogs in China earlier than 10,000 years BP has been suggested (Underhill 1997). However, the morphological features of the putative dog are not presented and the dating is not entirely reliable. An interesting trait shared between dogs and Chinese wolves is the "turned back" apex of the coronoid process of the ascending ramus. This might be an indication that the domestic dog has an ancestry from Chinese wolves (Olsen and Olsen 1977). However, data for the occurrence of the trait in a world-wide collection of domestic dogs was not presented.

Except for the earlier mentioned bias in the fossil record, another problem with the fossil record is the preservation of ancient material which is largely dependent on soil conditions and humidity in the environment, with humid environment and acidic soils degrading ancient specimens.

Another obvious problem is how to distinguish between wolf and dog remains in archaeological excavations. This problem has led to debates among archaeologists (see below). Except for a diminution in size, some other traits usually used to distinguish between wolves and dog are shortened snout and mandibles, crowded teeth, increased snout width to length ratio and smaller carnassial teeth (Clutton-Brock 1995, Musil 2000, Raisor 2005). However, crowding of teeth has also been detected in wolves (Musil 2000). Also, differences in size might be an effect of climatic oscillations, a phenomenon called the Bergmann rule (Dayan 1994), or sexual dimorphism (Musil 2000).

1.3.1 Europe

For a long time the earliest dated archaeologically excavated dog remains from the European continent was a single mandible from Bonn-Oberkassel in Germany, radiocarbon dated at 14,000 years ago (Nobis 1979, Street 2002). Measurements of the mandible indicates that it is different from a wolf mandible.

Recently during a reexamination of faunal remains from the Kesslerloch cave, Switzerland, dog remains were found and a single maxilla fragment was radiocarbon dated at 14,100-14,600 years ago, the earliest reasonably firm evidence for dogs on the European continent (Napierala and Uerpmann 2010). The body size and the teeth from the specimen are significantly smaller than among wolves (Napierala and Uerpmann 2010).

Other early canid remains on the European continent have been excavated in France, St. Thibaud, dated at 10,050 ±100 years BP (Chaix 2000) and numerous small canid remains in northern and southern France, dating between approximately 11,500-15,000 years ago based on dating of layers associated with the dog remains (Pionnier-Capitan et al. 2011). One of the small canids was directly dated at 12,451-12,952 years ago. The canid remains were claimed to be early dogs based on measurements of the bones clearly falling below the size variation among extant and ancient European wolves (Pionnier-Capitan et al. 2011).

Recently, some very old canid specimens have been reported (Germonpré et al. 2009, Ovodov et al. 2011, Sablin and Khlopachev 2002). These canids have been excavated in Eliseevichi (Russia), dated at 13,000-17,000 years BP (Sablin and Khlopachev 2002); Goyet (Belgium), the oldest skull dated at approximately 31,700 years BP (Germonpré et al. 2009) and Razboinichya cave in Siberia, directly dated at skull and mandible to an average age of 33000 years ago (Ovodov et al. 2011).

However, the findings from Eliseevichi and Goyet consisted of large wolf-like specimens, and have been contested by other archaeologists (Napierala and Uerpmann 2010, Pionnier-Capitan et al. 2011, Wang and Tedford 2008). The canids from Goyet are identified as dogs based on the resemblance with specimens from Eliseevichi I, having short and broad snouts but being large and wolf like and also having wolf sized teeth. According to (Napierala and Uerpmann 2010) there are no indications that the canids from Eliseevichi I are dogs and analysis shows that these canids fit well into the

size variability of wolf teeth and these can therefore not be used as reference for other findings of early dogs such as the Goyet findings (Napierala and Uerpmann 2010). According to Napierala and Uerpmann (2010) the shorter and broader snout of the canids can be an effect of changes in predator-prey relations, where a shorter and therefore stronger snout is favored if the prey becomes larger (Napierala and Uerpmann 2010). The dog from Razboinichya cave has a short snout and the coronoid process of the mandible has a hooked profile but the tooth crowding index is wolflike and the carnassial tooth size is within the modern and ancient wolf size interval (Ovodov et al. 2011). The authors suggest that this is an early dog that became extinct and did not contribute to the contemporary dog population (Ovodov et al. 2011).

1.3.2 South west Asia

The earliest dated canid remains supposedly from dog in SW Asia have been excavated in northern Israel (Davis and Valla 1978, Dayan 1994, Tchernov and Valla 1997) and dated at approximately 11,500 years ago. The remains have been classified as dogs based on their short snout, small carnassials and overall small size.

1.3.3 East Asia

The earliest remains excavated in East Asia supposedly from dog were found in northern China and dated older than 10,000 years before present (BP). However, the dating is not calibrated and the morphological features of the dog not presented (Underhill 1997).

The earliest well described dog remains in China has been dated at approximately 7,000 years BP and were excavated in the northern part of China (Olsen and Olsen 1977) . The oldest dog remains from Asia South of Yangtze River (ASY) have been excavated at the site of Hemudu. This site was occupied 6,500-7,000 years BP by rice cultivators (Underhill 1997). Archaeological remains show that the domestic dog was established along the gulf of Thailand 4,000 years ago (Higham 1996) . The introduction of the domestic dog to SE Asia could possibly be connected with the spread of rice agriculture from China about 4,000-5,000 years ago (Bellwood 2011). Domestic dog appears in the archaeological record of Island Southeast Asia approximately 3,000 years BP, from remains excavated in the Uattamdi cave in northern Molluccas (Bellwood 1997, Bellwood 2005).

1.3.4 Australia and the Pacific Islands

The earliest substantiated evidence for dingoes on the Australian continent comes from remains excavated in the south part of the continent dated at approximately 3,500 years ago (Milham and Thompson 1976, Gollan 1984). Geological investigations have shown that the island of Tasmania was separated from the Australian continent 12,000 years ago (Gollan 1984). The absence of dingo remains from Tasmania makes an introduction of dingoes between 3,500 and 12,000 years ago plausible. The dog was introduced to the Pacific islands relatively recently in connection with the first human populations reaching the islands. Indisputable evidence for dogs in Polynesia appears in the archaeological record from 2,000 years ago (Matisoo-Smith 2007).

1.3.5 America

The formerly claimed earliest dated dog remains on the North American continent was excavated in Jaguar cave, Idaho, and radiocarbon dated at 10,370 years BP (Raisor 2005). The maxillary and mandibular fragments from the remains were metrically between wolves and coyotes (Raisor 2005). However, a reexamination of the dog remains dating directly on the bones, revealed a much more recent dating than previously estimated; the maxilla was dated at 3,200 ± 80 BP and the mandible to 940 ± 80 BP (Clutton-Brock 1995, Raisor 2005).

Excavations at the Koster site in Illinois have resulted in three complete adult canids categorized as dogs, based on their wide palates and cranial vaults and crowding of teeth (Morey and Wiant 1992). The site has been radiocarbon dated at 8,130 ± 90 BP – 8,480 ± 110 BP (Morey and Wiant 1992). The Koster site has received special attention since the dogs seem to have been intentionally buried by humans and thus indicating a close relationship between dogs and humans (Morey and Wiant 1992, Raisor 2005). The earliest putative dog remains from South America were excavated in Ecuador and dated at approximately 5,100 BP (Raisor 2005). No dog remains older than 3500 BP have been excavated elsewhere in South America (Raisor 2005).

Importantly, no dog remains older than remains found in the Old World have been excavated and thus an Old World origin for New World dogs seems most probable (Raisor 2005).

1.4 Domestication

Farming and domestication of plants and animals is perhaps one of the most important "inventions" in human history. The consequence of farming led to a surplus of food and spare time which ultimately generated armies, experts and kings and the evolution of complex societies. Numerous studies have tried to unravel the history of plant and animal domestication.

A domesticated species is defined by Jared Diamond as: "a species bred in captivity and thereby modified from its wild ancestors in ways making it more useful to humans who control its reproduction and (in the case of animals) its food supply" (Diamond 2002).

Domestication of plants and animals started almost simultaneously in the Fertile Crescent and north and south China approximately 10,000 – 13,000 years ago (Diamond 2002, Zeder 2008).

Why domestication arose at this time is not totally clear but an increase in competition between humans, new technologies for food collection and extinction of big game prey animals has been suggested (Diamond 2002). The first steps towards domestication must have been unintentional since knowledge about the outcome of domestication was unknown (Diamond 2002).

Farmers and their associated cultures then spread rapidly from their homelands into surrounding areas either through expansion of farmer populations into hunter and gatherer regions, resulting in admixture or replacement of human genes and culture, or through transfer of ideas and technology (Diamond 2002, Zeder 2008, Diamond and Bellwood 2003).

The first domestication of wolves must have involved selection for tameness and manageability. Possibly wolves gathered around human camps in search for food, and the animals least fearful towards humans and being manageable in size were the ones being "domesticated" (Crockford 2000).

A long going experiment on domestication of foxes has shown that extreme selection for tameness can create foxes showing a domesticated behavior compared to wild foxes within 10 generations (Trut, Plyusnina and Oskina 2004, Trut, Oskina and Kharlamova 2009). Also other traits typical to domestic animals, especially dog like traits, are detected within a few generations such as spotted coat colour, floppy ears and curly tail, and in later generations also craniological changes (Trut et al. 2004, Trut et al. 2009). This suggests that the same genes selected for by tameness are linked to genes controlling development leading to the phenotypic variation detected (Trut et al. 2004, Trut et al. 2009).

A study of the entire mitochondrial genome in dog, wolf and coyote has indicated that the accumulation of nonsynonomous mutations in dog lineages has been faster than among wolf lineages (Björnerfeldt, Webster and Vilà 2006). The accumulation has been interpreted as a relaxation of selection following the domestication of dog (Björnerfeldt et al. 2006).

In the following sections I will mention a few genetic studies of some of our most important domestic animals; goat, sheep, cattle, chicken, pig and horse. Dogs will be discussed in a later section.

1.4.1 Goat

Early mtDNA (see section 7 for more information on mtDNA) studies of domestic goats suggested that goats were domesticated in at least two geographically separate regions (Luikart et al. 2001). Luikart et al. (2001) detected three divergent clusters A, B and C among goats of which A was widespread and found in domestic goats from all over the Old World, while the two other clusters were geographically restricted B to Asia and C to Mongolia, Switzerland and Slovenia (Luikart et al. 2001). This was interpreted to suggest at least two domestication centers for domestic goats, one early in the Fertile Crescent and one later in Asia (Luikart et al. 2001). However, the possibility that the smaller clusters could be due to introgression of wild species was mentioned (Luikart et al. 2001). Later mtDNA studies have revealed that the world goat mtDNA gene pool actually consists of 6 divergent clusters with the majority of samples belonging to haplogroup A (Naderi et al. 2007). In a study using the Bezoar as a representative for wild goat populations of the Holocene, all previously detected domestic goat haplogroups were found in the Bezoar population (Naderi et al. 2008).

Haplotype sharing between domestic goat and Bezoar and the genetic distribution of Bezoar haplogroups suggested an origin in Eastern Anatolia for haplogroup A and C. The rest of the haplogroups are probably an effect of introgression of wild lineages.

1.4.2 Sheep

Evidence from mtDNA studies of domestic and wild sheep initially suggested multiple domestication events for domestic sheep (Hiendleder et al. 2002, Pedrosa et al. 2005).

A later study suggests that two haplogroups expanded at approximately the same time from a Near Eastern homeland (Tapio et al. 2006) with two other haplogroups being later introgressions of wild sheep into the domestic stock (Tapio et al. 2006). Analysis of a longer region of the mtDNA genome has revealed an additional haplogroup giving totally five in domestic sheep (Meadows et al. 2007).

With the two largest haplogroups emerging approximately simultaneously in the Near East together with archaeological evidence and the restricted geographic distribution of other sheep haplogroups a single origin for the domestic sheep in the Near east from a large wild population, with later introgression of wild lineages, cannot be excluded.

1.4.3 Cattle

mtDNA studies indicates multiple domestication of cattle with two centers of origin, one in the Fertile Crescent and one in the Indus Valley (Ajmone-Marsan et al. 2010). The highest genetic diversity of Taurine cattle in the Fertile Crescent suggests an origin there with subsequent spread to Europe (Anderung et al. 2005, Troy et al. 2001). The highest mtDNA genetic diversity for the Zebu cattle is detected in the northern part of India indicating an origin for Zebu cattle in this area (Chen et al. 2010). Obviously, there were two different domestication events for cattle from two different subspecies of the wild ancestor.

1.4.4 Chicken

Multiple maternal origins and separate domestication events in Southeast Asia and the Indus valley have been proposed by mtDNA studies (Tixier-Boichard, Bed'hom and Rognon 2011). In a study of 834 domestic chickens sampled throughout Eurasia it was revealed that the chickens clustered in eight divergent clades (Liu et al. 2006). Three of

the clades were distributed throughout the Eurasian continent while the rest of the clades were restricted to South or Southeast Asia (Liu et al. 2006).

1.4.5 Pig

Studies of mtDNA in domestic pigs and wild boars from Europe and Asia, revaled two distinct mtDNA pig clades, one clade containing wild boars from Europe and Israel and the majority of European domestic pigs and another, Asian, clade consisting of Asian wild boars and Asian domesticates and some European domestic pigs. This was interpreted as at least two separate domestications, one in Asia and one in Europe (Giuffra et al. 2000). An extended study of mtDNA in wild and domestic pig sampled throughout Eurasia revealed several wildboar mtDNA clades throughout the continent (Larson et al. 2005). Most of the clades consisted of both domestic pigs and local wild pigs and was interpreted as multiple local domestications of pigs throughout Eurasia (Larson et al. 2005). However, the occurrence of geographically restricted clades among the domestic pigs can also be an indication of introgression of wild boar into an already domesticated population, similar to clade D and F in domestic dog (Pang et al. 2009, Klütsch et al. 2011, Ardalan et al. 2011, Ishiguro, Inoshima and Shigehara 2009).

1.4.6 Horse

Studies of mtDNA control region sequences in wild and domestic horses indicates multiple maternal origins (Vilà et al. 2001) from at least 77 breeding mares (Jansen et al. 2002). Complete sequencing of the domestic and wild horse mtDNA in samples from throughout the world reveals 18 major haplogroups in horses of which one haplogroup was restricted to wild horses (Achilli et al. 2012). The highest haplogroup variation was detected in Asia with representation of all haplogroups (Achilli et al. 2012), thus indicating Asia as a center for domestication of horses. The high diversity of horse matrilines is not reflected by Y-chromosome studies. A study of 14.3 kb of Y-chromosome sequence in 52 horses did not reveal a single SNP. Thus all horses shared a single haplotype (Lindgren et al. 2004). This pattern could be explained by a sex-bias in the breeding practices of domestic horses and a generally low Y-chromosome diversity in pre-domesticated horses or a single origin of horses with extensive introgression of matrilineages after domestication (Lindgren et al. 2004). Recently, a study of ancient wild and domestic horses detected a high Y-chromosome diversity

among ancient wild horses and also a second Y-chromosome haplotype in an ancient domestic horse sample dated at 2,800 years ago (Lippold et al. 2011).

The data together indicates a single origin for the domestic horse with extensive introgression of wild maternal lineages after domestication, possibly the horse was domesticated in Asia, as indicated by the high diversity and presence of all haplogroups.

To summarize, even though most of the studies interpret data as indicating multiple domestication events, a single origin for some of the domestic animals cannot be ruled out, for example the goat, sheep and horse. A single origin for some of the domestic animals means that the ancestral wild populations must have been genetically quite diverse and that when humans started to domesticate it was not a single small incident but rather a common and widespread activity. Dogs were domesticated thousands of years before the other domesticates discussed here and mtDNA (Savolainen et al. 2002, Pang et al. 2009) together with Y-chromosome (paper II) indicate a single origin for the domestic dog in ASY from at least 51 female and 13 male wolf founders.

Studies of population history mainly focus on DNA variants in the genomes of different individuals. Here I will mention only two kinds of variations important in studies of evolutionary relationships between populations, SNPs and microsatellites.

1.5 Genetic investigations

I here go through some types of polymorphisms frequently used in population genetic studies.

1.5.1 Analysis of DNA sequence variation

The hallmark of DNA sequencing has long been the Sanger sequencing method (Sanger et al. 1977). However, new and improved sequencing technologies have recently revolutionized the world of genome sequencing by sequencing millions of bases in a single run from small amounts of DNA and without the necessity of cloning prior to sequencing (Stoneking and Krause 2011). While these new methods provide powerful future perspectives most of the early population genetic studies were based on

the Sanger sequencing technique, for example the early studies on human origin and migration (Vigilant et al. 1991, Ingman et al. 2000, Torroni et al. 1993).

The most non-biased way of detecting variation between different genomes is to resequence the same locus in several individuals from different parts of the world (Morin et al. 2004). Single Nucleotide Polymorphisms (SNPs), microsatellites, deletions, insertions or other variations are then easily observed in an alignment of the sequences.

However, resequencing of a large set of samples has been both time-consuming and expensive and therefore faster and cheaper methods of detecting variations between individuals or populations, such as DHPLC (Xiao and Oefner 2001) has been used in population studies (Thomson et al. 2000, Underhill et al. 2000).

As an alternative to the often expensive and time-consuming resequencing, analysis of just SNP positions is common. There are multitudes of SNP genotyping methods available (Kim and Misra 2007). A potential problem using SNP genotyping methods is the bias that can be introduced when the panel used for discovery of SNPs is different from the panel genotyped. Examples of bias have recently been discussed (Schuster et al. 2010, Klütsch and Savolainen 2011). Therefore, in the SNP discovery phase it is very important to use a large panel of samples from all populations intended to be studied. Schuster et al. (2010) could even show that bias in the Illumina SNP array indicated that Europeans have higher heterozygosity than Africans (Schuster et al. 2010).

With the new sequencing technologies, millions of DNA sequences can be read in parallel in a single sequencing run. The necessity of a targeted amplification step is avoided; instead the DNA is turned into a sequence library, with ligated adaptor sequences to both ends of each fragment in the library. The entire library of DNA sequences can then be amplified using primers directed at the adaptor sequences (Stoneking and Krause 2011).

The technology has been used to sequence almost the entire genomes of extinct species such as Neanderthals (Stoneking and Krause 2011, Green et al. 2010) and mammoths (Miller et al. 2008). To sequence the entire genome of an extinct species was

unthinkable only a few years ago. The possibility to obtain almost the entire nuclear genome from ancient and degraded material opens up for studies of genes in ancient populations. Recently, the nuclear genome from 5,300 year old human remains was investigated giving clues about the phenotype of the specimen (Keller et al. 2012).

Although large amounts of sequence data can be obtained relatively fast, these new methods are still relatively expensive, at least for analysis of many samples. Parallell targeted sequencing in many samples has been a problem because of difficulties in linking reads to specific samples. This has been solved by introducing identifying tags to the sequences either through PCR or ligation (Binladen et al. 2007, Meyer, Stenzel and Hofreiter 2008). Recently, a method reducing the complexity of analysis of a targeted genomic region in a multitude of samples run in parallel in a single run was developed employing a two-tag system positioning the sequence to a specific position on a specific 96 well plate (Neiman et al. 2011).

1.5.2 SNP

A single nucleotide polymorphism (SNP) is a single nucleotide difference between two individual genomes. SNPs are commonly used in studies of evolutionary history of populations (Morin et al. 2004).

There are various causes for mutations in the genome, endogenous (replication errors or oxidative free radicals) or exogenous (different kinds of radiation and chemical mutagens). The damage to the DNA molecule varies from misincorporations during DNA replication to chemical modification of bases or even breakage or cross linking of the DNA strands (Jobling, Hurles and Tyler-Smith 2004).

A mutation is a very rare event in the nuclear genome, so rare that it can mostly be assumed to have happened only once in the history of a specific lineage. However, considering the large amount of humans in the world today recurrent mutations exist also in the nuclear genome (Jobling et al. 2004). However, the effect is thought to be minor and easy to detect (Jobling et al. 2004). The mtDNA genome is different; because of the extremely high mutation rate (Ingman et al. 2000) recurrent mutations are not uncommon (Jobling et al. 2004). However, these recurrences can be detected in phylogenetic studies; see for example (Savolainen et al. 2004).

1.5.3 Microsatellites

Microsatellites are hypervariable segments of DNA scattered throughout the genome, usually in short mono, di-, tri-, tetra-, penta-, or hexanucleotide motifs tandemly repeated (Ellegren 2004). Polymorphism of microsatellites is mainly due to allelic length differences caused by different number of tandemly repeated units between alleles (Ellegren 2000). Microsatellites can arise from a base substitution (Messier, Li and Stewart 1996) or through insertions of a duplicated sequence (Zhu, Strassmann and Queller 2000). Once a microsatellite is born it tends to mutate in a stepwise fashion, with insertion or deletion of a single repeat unit at a time (Ohta and Kimura 1973), through a process called replication slippage (Schlötterer and Tautz 1992). However, the stepwise mutation model cannot explain why there is an upper limit for microsatellite expansion and other models have been suggested. One model suggests that there are two opposing mutational forces acting on the microsatellite, length mutations leading to longer microsatellites and point mutations breaking down the microsatellites (Kruglyak et al. 1998). Microsatellite mutation rates are not uniform and can differ between loci and perhaps also between species (Ellegren 2000). A problem with microsatellites in studies of population history are that they are very prone to mutation and can mutate back and forth within a few generations which can lead to homoplasy, i.e. identity by state rather than identity by descent. Therefore, microsatellites are mostly used when investigating recent evolutionary events like breed history (Bannasch et al. 2005) and hybridization in dog or wolf populations (Iacolina et al. 2010), or else microsatellites can be used in combination with SNPs to increase phylogenetic resolution (Brown et al. 2011, Sundqvist et al. 2006).

1.6 The Y-chromosome

The Y-chromosome does not recombine and is therefore a valuable tool in population studies since the chromosome is inherited intact from father to son without shuffling between each generation (Jobling and Tyler-Smith 2003, Underhill and Kivisild 2007). The Y-chromosome has been used to study human migrations (Underhill and Kivisild 2007) but has also been used in animal studies. The animal studies have mostly focused on Y-chromosome microsatellites (Bannasch et al. 2005) but also SNPs have been analyzed (Lindgren et al. 2004), mostly in combination with microsatellites (Brown et al. 2011, Vilà et al. 2003).

The Y-chromosome is a male specific chromosome containing the MSY gene which induces maleness and several genes important in spermatogenesis (Jobling and Tyler-Smith 2003, Hughes and Rozen 2012). The human Y-chromosome is highly repetitive; the ampliconic sequence can be up to 1.5 Mb long and consist of repeat units having almost 100% identity (Skaletsky et al. 2003). Due to the repetitive nature of the Y-chromosome and difficulties associated with sequencing and assembling of repetitive sequences only three Y-chromosomes have been sequenced to date: The human (Skaletsky et al. 2003), the chimpanzee (Hughes et al. 2010, Hughes et al. 2005) and the rhesus monkey (Hughes et al. 2012).

The human Y-chromosome consists of 78 different protein coding genes (Skaletsky et al. 2003) and the euchromatic sequence of the human Y-chromosome can be divided into three classes (Skaletsky et al. 2003), ampliconic sequence consisting of massive palindromes (most of the genes on the chromosome are found in these sequences), X-degenerate sequence and X-transposed sequence (Skaletsky et al. 2003). The X-transposed sequence stems from transposition from the X-chromosome to the Y-chromosome (Page et al. 1984), while the X-degenerate sequence is deteriorated X-chromosome sequences and the ampliconic sequence is gene dense and highly repetitive consisting of duplications and palindromes (Skaletsky et al. 2003).
The Y-chromosome is mostly non-recombining except for at the two ends of the chromosome, in the pseudoautosomal regions, where recombination with the X-chromosome occurs (Helena Mangs and Morris 2007, Jobling and Tyler-Smith 2003, Hughes and Rozen 2012). Even though the Y-chromosome largely escapes the recombination process, studies have shown that non-reciprocal recombination occurs within the chromosome (Skaletsky et al. 2003, Rozen et al. 2003).

The mutation rate is higher on the Y-chromosome than in the rest of the nuclear genome (Xue et al. 2009, Jobling et al. 2004). The balancing effect between high mutation rate and low effective population size (only ¼ of the autosomes) leads to Y-chromosomes having low diversity but being much differentiated between populations due to genetic drift (Jobling and Tyler-Smith 2003, Underhill and Kivisild 2007).

1.7 Mitochondrial DNA (mtDNA)

The high mutation rate and the maternal inheritance without recombination make mtDNA an excellent marker for population studies. The molecule is inherited intact from mother to offspring without shuffling between each generation making it easy to trace lineages. The highly variable and non-coding control region has been a popular target to analyze in population studies (Pakendorf and Stoneking 2005). One of the first studies on human population history in the pre-PCR era was aimed at the mtDNA molecule (Cann, Stoneking and Wilson 1987).

The abundance of mtDNA compared to nuclear DNA in cells also makes it an obvious choice for studies of ancient DNA (Pakendorf and Stoneking 2005, Hofreiter et al. 2001). A human diploid cell has normally two copies of the nuclear genome in the cell nucleus but can have several thousands of copies of the mitochondrial genome (White et al. 2008). The number of mitochondria in a cell varies depending on cell type. Cells requiring energy as muscle cells have thousands of mitochondria while other cells have only a few hundred. In addition, each mitochondrion has 2-10 copies of its genome (Jobling et al. 2004) explaining the discrepancy in copy number between the mtDNA and nuclear DNA in a cell.

Sperms contain about 100 mitochondria (White et al. 2008) but still paternal mitochondria are not inherited to the offspring. A mechanism for removal of paternal mtDNA during fertilization has been suggested in mice (Shitara et al. 1998) and recently removal of paternal mitochondria through autophagy was detected in Caenorhabditis elegans embryos (Sato and Sato 2011). However, defects in this mechanism can lead to paternal leakage; in humans a mutated mtDNA causing mitochondrial myopathy was determined to have a paternal origin (Schwartz and Vissing 2002). However, paternal inheritance of mtDNA is extremely rare and therefore mtDNA can still in most cases be treated as a maternally inherited molecule (Pakendorf and Stoneking 2005).

The main function of the mitochondrion is to provide energy to the cell through the oxidative phosporylation pathway (White et al. 2008). Mitochondria have their own DNA, a remnant from its early days as a bacterium (Gray, Burger and Lang 1999) later taken up by a proto-eukaryotic cell (Margulis 1981). They replicate semi-autonomously

from the nuclear genome, the protein machinery needed for replication is encoded by the nuclear genome (Ryan and Hoogenraad 2007).

Even though the mitochondrion has lost its independence through transfer of most of its genes to the nuclear genome (see below for a discussion how this might affect studies of population history) it still retains some of its genes and translation machinery. In total, both the human and the dog mitochondrion have 37 genes, most of which are involved in the protein translation machinery or in the oxidative phosporylation pathway. Thirteen of the genes encode subunits involved in the oxidative phosphorylation pathway, 22 encode tRNAs and 2 encode rRNAs (Anderson et al. 1981, Kim et al. 1998).
The mtDNA molecule is double stranded, and in contrast to the nuclear genome, a circular genome in almost all organisms (Gray et al. 1999). The mtDNA genome size in humans has been estimated at 16,569 bp (Anderson et al. 1981) and in dogs to approximately 16,728 bp, depending on length heteroplasmy in a repeat sequence of the control region (Kim et al. 1998).

The mutation rate of the mitochondrial genome is very high compared to the nuclear genome, $1.70*10^{-8}$ substitutions site^{-1} year^{-1} excluding the control region in humans (Ingman et al. 2000) and $6.4*10^{-9}$ – $1.92*10^{-8}$ substitutions site^{-1} year^{-1} in dogs (Pang et al. 2009). The difference in mutation rate between nuclear and mitochondrial DNA is an effect of the presence of by-products from oxidative phosphorylation, error-prone polymerase and inefficient DNA repair in the mitochondrion (White et al. 2008, Bogenhagen 1999).

The use of mtDNA data to estimate domestication dates has been criticized (Ho and Larson 2006). It was observed that the dates obtained from mtDNA data were much older than the dates obtained from the archaeological record (Ho and Larson 2006). However, these observations were obtained from mtDNA control region data. The low resolution of the mtDNA control region underestimates the number of founders in a population and therefore the dating will be older than the actual domestication event (Pang et al. 2009, Bandelt 2008). However, sequencing of the entire mitochondrial genome increases the resolution, giving a better estimation of the date of domestication. Sequencing of the entire mitochondrial genome in domestic dogs indicated a domestication date in good agreement with the archaeological record (Pang et al. 2009).

A potential problem investigating evolutionary history using mtDNA is NUMTs (Nuclear Mitochondrial DNA Sequences). The mechanism for incorporation of mtDNA segments into the nuclear genome is not known. It has been suggested that insertion of NUMTs close to repetitive regions might be caused by an open chromatin structure (Mishmar et al. 2004). The NUMTs have a slower mutation rate compared to mitochondrial sequences and a failure to detect them will lead to erroneous conclusions (Olson and Yoder 2002). NUMTs have a different history from the rest of the mitochondrial genome and can therefore be detected by their phylogenetic placement (Bensasson et al. 2001). NUMTs can also be detected by PCR ghost bands on agarose gels or through sequence ambiguities (Bensasson et al. 2001). Therefore, a NUMT can usually be detected and avoided during analysis of mtDNA.

1.8 Dog genetics

1.8.1 Genome sequencing

At the time of writing two dog genomes have been sequenced but with the rapid development of new sequencing technologies more dog genomes are expected. The first of the dog genomes to be sequenced was from a male Poodle. It was sequenced with 1.5x coverage, implying that there was on average 1.5 sequence reads for each nucleotide in the presented sequence. However, because of the random distribution of sequence reads large parts of the genome was not sequenced, resulting in a fragmented genome and the Y-chromosome sequences could therefore not be definitely identified (Kirkness et al. 2003). The second genome sequence was obtained from a female Boxer and was sequenced with 7.5x coverage (Lindblad-Toh et al. 2005). With this coverage almost the entire genome was obtained. Both dog genome sequences were obtained by Sanger sequencing (Kirkness et al. 2003, Lindblad-Toh et al. 2005).

The size of the euchromatic (genetically active) dog genome was estimated at 2.31 to 2.47 Gb (Kirkness et al. 2003) and 2.4 Gb (Lindblad-Toh et al. 2005). The dog genome consists of 39 pairs of chromosomes as opposed to the 23 pairs in humans but still the euchromatic genome size of humans, approximately 2.88 Gb (Consortium 2004), is larger than in dogs. This size discrepancy is explained by a lower proportion of repeat insertions in the dog genome compared to the human genome (Lindblad-Toh et al. 2005).

The SNP rate in the dog genome was investigated by searching for sites with different alleles within the Boxer genome and by comparison to the male Poodle genome and a set of 9 different breed dogs (sequenced with 0.02x coverage), 4 wolves and one coyote (sequenced with 0.004x coverage). The SNP rate between different breeds was estimated to approximately 1 SNP per 900 bp. The rate between the Boxer and gray wolves was estimated at 1 SNP per 580 bp and 1/420 bp for coyotes (Lindblad-Toh et al. 2005). The heterozygozity observed within the Boxer genome was lower than the observed variation between breeds, 1 SNP per 1600 bp (Lindblad-Toh et al. 2005).

The genome sequence of the domestic dog has had important implications in medical genetics (Parker et al. 2010) and population genetics (Vonholdt et al. 2010). The availability of large amounts of SNPs from the entire nuclear genome on hybridizing DNA chips makes mapping of traits in dogs or investigation of evolutionary relationships less time-consuming and cheaper.

1.8.2 Medical genetics

The start of intense breeding and the creation of breed clubs in the early 19th century with rules and standards on the appearance of dogs have led to strong selection of particular phenotypes such as size, shape and color. The strong selection and very limited genetic influx makes dog breeds a perfect model for studies of genes because of extensive linkage disequilibrium in the dog genome, and also large families available for analysis (Sutter and Ostrander 2004, Parker et al. 2010).

There are a couple of success stories where disease genes have been mapped in the dog genome and that have led to understanding of disease in humans. One such example is the location of the gene responsible for narcolepsy in Doberman pinschers (Lin et al. 1999). The study showed that disruption of the Hcrtr2 gene leads to narcolepsy in Doberman pinschers, indicating hypocretins to be involved in sleep disorders (Lin et al. 1999). This study highlighted a signalling pathway involved in causing narcolepsy not earlier detected in humans. Even though human narcolepsy rarely is caused by disruption of the Hcrtr2, decreased levels of hypocretin cells are usually observed (Parker et al. 2010).

The hairless dog phenotype has been mapped, to a region on chromosome 17 containing the FOXI3 gene and a possible duplication causing a frameshift and premature stop codon in the gene that might be involved in the phenotype (Drögemüller et al. 2008). The duplication was perfectly correlated with the hairless phenotype in 140 hairless dogs but was absent in 55 coated dogs from the same breeds and in 32 coated dogs from other breeds (Drögemüller et al. 2008).

Melanin production and coat colour variation in many vertebrates has been traced to the Melanocortin 1 receptor, MC1R gene (Parker et al. 2010). In dogs an additional locus has been shown to be important in pigment type switching (Candille et al. 2007). Melanism in dogs and wolves could be traced to a particular variant of the K locus (Anderson et al. 2009).

Another example of phenotypic variation in canines that has been investigated is the ridge in Rhodesian ridgeback (Salmon Hillbertz et al. 2007) which has been traced to a duplicated 133kb segment on chromosome 18 involving three fibroblast growth factor genes. Also genes for coat length and curl of hair have been identified (Cadieu et al. 2009). The combination of distinct mutations in three different genes was shown to explain most of the coat length and texture phenotypic diversity among 80 dog breeds (Cadieu et al. 2009).

1.9 Human population history (A short summary)

The history of dogs is tightly coupled to human history. Therefore, knowledge about human population history is important. Here I will present a short summary on human population history.

The first studies on human genetic diversity were based on the “classical markers” consisting of blood proteins (Cavalli-Sforza and Feldman 2003). The first seminal paper based on DNA sequence variation in humans was published in 1987, focusing on restriction length polymorphisms (RFLPs) in the mtDNA (Cann et al. 1987). After the advent of the PCR amplification method (Saiki et al. 1985, Mullis and Faloona 1987) most of the studies on human genetic diversity focused on DNA sequence variation in the non-coding mtDNA control region. There were two reasons for the popularity of the mtDNA as a marker to study human origins and spread. Firstly, mtDNA was

maternally inherited without recombination and secondly, mtDNA has a high mutation rate especially the hypervariable fragments of the control region (Pakendorf and Stoneking 2005). Because the mtDNA only reflects the maternal history analysis of a second independently inherited marker was desirable. The obvious choice of marker was the Y-chromosome, being paternally inherited and largely non-recombining except for the two pseudoautosomal regions in each end of the chromosome (Underhill and Kivisild 2007).

mtDNA, Y-chromosome as well as genome-wide autosomal SNP data indicate that modern humans evolved in Africa, and that only a few small founder groups subsequently dispersed out of Africa (Vigilant et al. 1991, Underhill et al. 2000, Li et al. 2008, Oppenheimer 2012). The branches closest to the root of the mtDNA phylogenetic tree were exclusively found on the African continent while branches found in the rest of the world contain only a subset of this variation (Vigilant et al. 1991, Reed and Tishkoff 2006).

Fossil evidence indicates that anatomically modern humans were first present on the African continent around 130,000 years ago (Jobling et al. 2004). An early expansion of human lineages within Africa where mtDNA lineages L2 and L3 largely replaced the original L1 lineages (Watson et al. 1997) successively led to the first successful human exodus out of Africa approximately 70,000 years ago (Oppenheimer 2012).

Modern humans appear in the archaeological record in Australia about 40,000 years BP (Bowler et al. 2003) and genetic evidence shows that aboriginal Australians and New Guineans carry the same set of haplogroups as all other non-Africans (Macaulay et al. 2005) and thus indicates a spread from Africa to Australia following the southern coast of Asia. An offspring of this first migration group later colonized northern Asia and Europe (Forster 2004).

Modern homo sapiens and Neanderthals coexisted during a couple of thousand years after the first arrival of modern humans in Europe approximately 30,000 – 35,000 years ago (Finlayson et al. 2006). The earliest studies of ancient Neanderthal DNA showed a highly divergent mtDNA haplotype not detected in any of the modern humans sequenced and thus proposing that the extent of admixture was low (Krings et al. 1997, Krings et al. 1999). Today, with the advent of the new sequencing technologies,

investigations can be made also on the nuclear DNA of the Neanderthals. These studies have shown that there was some admixture between modern humans and Neanderthals; between 1 and 4% of the genome is derived from Neanderthals for all non-Africans (Green et al. 2010). The uniform admixture between extant humans worldwide (except for sub Saharan Africans) and Neanderthals indicates that humans dispersed from Africa in a single exodus and mixed with Neanderthals (Stoneking and Krause 2011). A finding of a specimen of archaic human in Siberia, dubbed Denisova after the cave in Siberia where it was found, was sequenced and shown to be a sister group to Neanderthals (Reich et al. 2010). Admixture between this new group of hominids and modern humans has also been indicated to have occurred in southern Asia (Reich et al. 2010, Reich et al. 2011, Skoglund and Jakobsson 2011).

In America, five mtDNA haplogroups and two Y-chromosome haplogroups have been detected all of, which are present also in indigenous Siberian populations (Goebel, Waters and O'Rourke 2008). Genetic data suggest a single migration into America approximately 15,000-30,000 years ago (Goebel et al. 2008, Forster 2004).

The relatively recent advent of agriculture in several different areas of the world (Diamond 2002) led to migrations of farmers from their homelands into new territories (Diamond 2002, Diamond and Bellwood 2003). Here I will mention only two of these migrations.

The spread of tropical West African Bantu speaking farmers over much of eastern and southern Africa largely replaced the original hunter and gatherer inhabitants. Evidence from language (Diamond and Bellwood 2003) and genetics (Coelho et al. 2009) show that a west/central African population spread east and south approximately 4,000 years ago (Coelho et al. 2009).

The peopling of the Pacific Islands is debated and several different models describing the migration based on different evidence have been presented (Hurles et al. 2003). The most prominent are the Express train model (which is primarily based on archaeology) (Diamond 1988, Diamond 2001), the entangled bank model (based on archaeology) (Terrell 1988, Terrell, Kelly and Rainbird 2001) and the slow boat model (based on genetics) (Oppenheimer and Richards 2001). The details about these models are presented in section three of Present Investigation.

1.10 Origin and initial dispersal of the domestic dog

1.10.1 Origin

The origin of the domestic dog is debated but with the introduction of genetic studies giving a complement to the archaeological studies a pattern is starting to be revealed.

The archaeological evidence has favored a European and/or SW Asian origin for the domestic dog (Clutton-Brock 1995, Dayan 1994), primarily based on the earliest dog bones being excavated in Bonn-Oberkassel (Germany) (Nobis 1979, Street 2002) and in Ein Mallah and Hayonim Terrace (Israel) (Dayan 1994, Davis and Valla 1978, Tchernov and Valla 1997). However, difficulties in interpreting the archaeological evidence makes it hard to distinguish between bones from wolf and dog in archaeological assemblages (Musil 2000). This is not the only problem with the archaeological record. Some parts of the world such as Europe and SW Asia have been extensively investigated while other parts of the world have been largely neglected (Underhill 1997). East Asia is a region that has been poorly investigated. The oldest well described archaeological material from domestic dog has been dated at 7,000 BP and has been excavated in the northern part of China (Olsen and Olsen 1977, Underhill 1997), while the earliest remains from domestic dog in Southern China has been dated at 6,500 – 7,000 BP at the Hemudu site (Underhill 1997). However, poorly investigated remains in northern China claimed to be from dog has been dated at >10,000 years ago (Underhill 1997). Recently, some very old canid remains have been excavated in Belgium, and radiocarbon dated at 31,700 BP (Germonpré et al. 2009). These findings bear similarities to other early dog remains from Russia, Eliseviichi I, dated at 13,000 to 17,000 BP (Sablin and Khlopachev 2002), suggesting that dogs were domesticated long before the usually accepted date of approximately 14,000 BP (Clutton-Brock 1995).

However, these early remains of domestic dog have not been fully accepted by the archaeological community; see for example (Napierala and Uerpmann 2010). Napierala et al. writes that there is nothing indicating that the remains from Russia and Belgium are dogs and that the measurements obtained fits within the natural variation of wolves, and also cautions against using these specimens as reference for early domestic dog (Napierala and Uerpmann 2010).

Altogether, archaeologists seem to favour the theory that dogs have multiple origins in different parts of the world with the oldest being in Germany 14,000 years old and later in the Near East about 11,500 years old or ancient domestications of dogs later extinct and not contributing to the extant dog population (Ovodov et al. 2011).

The earliest genetic studies on the origin of the domestic dog focused on the mitochondrial DNA control region (Okumura et al. 1996, Vilà et al. 1997, Savolainen et al. 2002, Tsuda et al. 1997). mtDNA has a high mutation rate (Ingman et al. 2000, Pang et al. 2009) and a maternal inheritance without recombination between generations, which makes it easy to trace mtDNA lineages and also to obtain a lot of diversity from relatively short sequences compared to the nuclear genome. The first large scale study of mtDNA sequence in domestic dogs and wolves was published in 1997 (Vilà et al. 1997). Analysis of the mtDNA control region in 140 domestic dogs revealed that dog sequences grouped into 4 different clades numbered I-IV with wolf samples clustering in between (Vilà et al. 1997). Dog and wolf haplotypes differed with at most 12 mutations while dog and coyote/dhole haplotypes differed by at least 20 mutations. Thus, indicating that wolf is the founder for the domestic dog (Vilà et al. 1997).

Based on the divergence within the oldest clade (I) a domestication date was estimated to 135,000 BP, predating the appearance of modern humans outside Africa (Vilà et al. 1997). However, the calculation of this domestication date assumes an origin for all dog lineages in clade I from a single wolf lineage.

In a later study of mtDNA control region sequences in dogs, samples covering all parts of the world were included. All of the former domestic dog clades were detected I-IV (A-D) but also two new dog clades, E and F (Savolainen et al. 2002). Clades D, E and F were all small, consisting of a few haplotypes and restricted to particular regions of the world. Clade D to Scandinavia and the Near East, clade E to Korea, SE Asia and Japan and clade F to Japan (Savolainen et al. 2002) and thus indicating introgression of wolves rather than separate domestication events.

However, clade E being distributed over a relatively large area could be interpreted as having an origin together with clades A, B and C but with a restricted spread (Pang et

al. 2009). Separate domestication of clade D and F and subsequent isolation seems unlikely and these clades are probably an effect of wolf introgression into the domestic dog gene pool (see below).

With the observation that the domestic dog originated from several wolf lineages a new date of origin was calculated to 15,000-40,000 BP. The older age was calculated assuming a single origin for clade A while the younger age was calculated assuming an origin for clade A from several wolf lineages. Three different subgroups could be superficially detected (Savolainen et al. 2002). The younger age fitted very well with the archaeological record (Savolainen et al. 2002). The study also showed that the East Asian dog population had the highest genetic diversity and dog populations in other parts of the world carried only a subset of this diversity (Savolainen et al. 2002). The general conclusions were that the domestic dog had a single origin in East Asia approximately 15,000 years ago from numerous wolves (Savolainen et al. 2002).

The most comprehensive study on dog origins this far investigated mtDNA control region sequences in 1,543 domestic dogs and 40 wolves sampled throughout the world. In addition the entire mitochondrial genome was sequenced in 169 domestic dogs and 8 wolves selected to represent all parts of the mtDNA phylogeny (Pang et al. 2009).
The increased resolution obtained from sequencing of the entire mtDNA genome revaled 10 subclades among the three main clades A, B and C. The only region in the world with all 10 subclades represented was ASY (Asia South of Yangtze River) while the rest of the world carried only a subset of these subclades (Pang et al. 2009, Brown et al. 2011, Ardalan et al. 2011), Asia north of Yangtze had only 7 subclades, SW Asia and Africa 5, and Europe 4 (Pang 2009).

Almost all dogs sampled had a haplotype belonging to either clade A, B or C, indicating a single origin for the domestic dog (Pang et al. 2009). Most of the dogs in Europe and SW Asia had a Universal Haplotype (UT) which is a haplotype present at high frequency in all parts of the world. Simulation showed that for clade A, B and C to have the same proportion in all parts of the world, a single origin in ASY with a subsequent spread of dogs throughout the world is very likely while all other scenarios would need extremely high rates of migrations to explain the frequencies of clades and subclades in the extant dog population (Pang et al. 2009).

This study confirms earlier mtDNA studies showing that the highest genetic diversity is found among dogs in East Asia. With the improved sampling and sequencing of the entire mtDNA genome, the region with highest genetic diversity and the only region where all 10 subclades were detected could be pinpointed to ASY (Asia South of Yangtze River). Thus, indicating an origin for the domestic dog in ASY.

With the improved resolution from sequencing of the entire mitochondrial genome a new mutation rate was calculated, and a domestication date was estimated to between 5,000 and 16,000 years ago from at least 51 female wolf lineages (Pang et al. 2009). The domestic dogs origin from a large number of wolf founders is in agreement with studies of the MHC locus (at least 21 wolf founders) (Vilà, Seddon and Ellegren 2005) and autosomal SNPs in dogs and wolves (Gray et al. 2009a).

Dog-wolf crossbreeding does not seem to have contributed extensively to the domestic dog mtDNA gene pool (Pang et al. 2009). However, perhaps mating between male domestic dogs and female wolves is less likely because of the difference in size and also the offspring of such a mating would most likely be raised as a wolf rather than a dog and therefore other markers are more suitable for detecting crossbreeding, such as the Y-chromosome (Vilà et al. 2003). mtDNA data indicate at least three dog-wolf hybridizations, clade F in Japan, clade d2 in SW Asia and clade d1 in Scandinavia (Pang et al. 2009, Klütsch et al. 2011). All of these clades are found in geographically restricted areas and have a low frequency in the total dog population. A local dog domestication event in North America has been indicated by a group of North American dogs, dated at a few hundred years ago, clustering in a group together with local wolf (Koop et al. 2000). However this might be an example of local dog-wolf hybridization (Klütsch and Savolainen 2011). Recent wolf-dog hybridization has also been detected using Y-chromosome microsatellites in the Italian (Iacolina et al. 2010) and Scandinavian (Vilà et al. 2003) wolf populations.

Recently, more than 48,000 autosomal SNPs were investigated in domestic dog and wolf samples from most parts of the world (Vonholdt et al. 2010) but importantly samples from South Chinese wolves were not included, the possible source population for domestic dogs (Pang et al. 2009) and therefore an origin of dogs from Southern Chinese wolves could not be detected in this study. Also the collection of SNPs used in this study suffers from ascertainment bias (Morin et al. 2004) since most of the SNPs

were collected using European dogs and therefore a lot of genetic diversity in other populations will remain undetected (Morin et al. 2004). The study compared SNP haplotypes between 64 dog breeds and wolves from Europe, the Middle East and East Asia and based on the largest fraction of shared haplotypes, the Middle East was indicated as a primary source for domestic dog genome diversity but also Europe and East Asia for a few breeds (Vonholdt et al. 2010). However, the results should be treated with caution because of the possibility of missing the correct source population due to ascertainment bias and lack of representative samples.

A study of mtDNA in a comprehensive sample from SW Asia indicated as a possible source region for the domestic dog (Dayan 1994, Vonholdt et al. 2010) detected no signs of a domestication event (Ardalan et al. 2011). On the contrary, the data confirmed the data from Pang et al. (2009), indicating that the SW Asian dog population was created from a subset of the East Asian genetic diversity and that subsequent bottlenecks and drift have created the genetic structure of the SW Asian dog population (Ardalan et al. 2011).

In a study investigating the mtDNA genetic diversity in African village dogs (Boyko et al. 2009), the mtDNA genetic diversity was claimed to be similar to that observed in East Asian village dogs, contradictory to the (Savolainen et al. 2002, Pang et al. 2009) results. However, Pang et al. (2009) disprove this showing that the genetic diversity is actually higher in the East Asian village dogs which is the only dog population in the world with the full set of mtDNA diversity. Recently, mtDNA diversity in village dogs from the Middle East and Southeast Asia (mainly consisting of samples from Bali and surrounding islands) confirmed previous results of a higher genetic diversity in Southeast Asian village dogs (Brown et al. 2011). Thus, even though the samples were mainly from a region outside ASY (the area where the domestic dog has been proposed to have its origin), the genetic diversity was higher than in Middle Eastern village dogs (Brown et al. 2011).

The first study of the domestic dog Y-chromosome sequence phylogeny (paper I) identified 14 SNPs in 10 dog Y-chromosomes from different parts of the world. These SNPs were used to create a maximum parsimony phylogenetic tree, consisting of 9 haplotypes. No phylogeographic pattern was observed among the 9 haplotypes (paper I).

A large study of NRY SNPs, identified in Paper I, and microsatellites in 300 village dogs and Australian dingoes along with 7 wolves revealed that the village dogs cluster in two clades, one mainly consisting of Middle Eastern (Iranian) dogs and the other mainly of Southeast Asian dogs (mainly from Bali and surrounding islands) and Australian dingoes (Brown et al. 2011).

Studies of mtDNA from ancient dog material in Europe (Malmström et al. 2008, Deguilloux et al. 2009), have indicated continuity between modern and ancient mtDNA haplogroups. Haplogroups A, B and C were present in both ancient and extant dogs (Malmström et al. 2008, Deguilloux et al. 2009).

Different haplogroup frequencies in the extant compared to the ancient dog population of Europe was indicated (Malmström et al. 2008, Deguilloux et al. 2009). However, it is not known how these ancient specimens are related, especially when the samples are taken from a single locality there is a high risk of relatedness. Investigation of more ancient samples collected from well separated places over the entire European continent would be neccesary to either confirm or reject the hypothesis of a lineage replacement.

1.10.2 Dispersal

The oldest archaeological remains from dingo have been radiocarbon dated at 3,500 years ago (Milham and Thompson 1976). An early study of mtDNA sequence detected 20 haplotypes in 211 dingoes sampled from all parts of the Australian continent. Among the 20 haplotypes one was shared with domestic dog, haplotype A29, the other 19 haplotypes were dingo specific (except one A9, see below) and separated from A29 by one mutational step. The dingo specific haplotypes were arranged in a star shaped pattern surrounding the central haplotype A29 indicating a population expansion from a small set of founders possibly only a single maternal founder lineage, A29 (Savolainen et al. 2004). One of the derived haplotypes, A9, is also shared between a single dingo and domestic dog but this was interpreted as a recurrent mutation, due to the high degree of homoplasy detected in the data set (Savolainen et al. 2004). The dataset was used to calculate ρ (rho) and date the introduction of dingoes to Australia to 4,600-10,800 years ago based on a dog/wolf and coyote split 1-2 MY ago (Savolainen et al. 2004). Several studies have indicated a possible shared history for Australian dingoes

and NGSD (Savolainen et al. 2004, Runstadler, Angles and Pedersen 2006) based on sharing of mtDNA haplotypes (Savolainen et al. 2004), and DLA alleles (Runstadler et al. 2006) as well as some similarities in morphology and behavior (Koler-Matznick et al. 2003).

Knowledge about the pre-European Polynesian dog population is sparse and only 19 pre-european ancient dog remains have been investigated from Cook Island, Hawaii and New Zealand (Savolainen et al. 2004). Two haplotypes were detected in 19 samples, both represented on each of the islands (Savolainen et al. 2004).

The origin of New World dogs has been investigated in ancient material from 19 specimens from several locations of the New World (Leonard et al. 2002). All of the ancient samples carried haplotypes belonging to clade A and B already detected in the Old World (Leonard et al. 2002). A cluster within clade A consisting of 50% of the ancient Latin American samples was detected and had not been described before (Leonard et al. 2002). However, see paper VI for a discussion. In a recent study investigating mtDNA in the extant dog population none of the haplotypes belonging to the ancient subclade were detected, suggesting an almost total replacement of ancient American dogs with introduced European dogs (Castroviejo-Fisher et al. 2011). However, in paper VI we present data contradicting some of these conclusions.

2 PRESENT INVESTIGATION

The domestic dog is one of our most beloved animals. It is used both as a working companion and a life companion. The interest in dogs, their health and history is overwhelming with clubs for almost each dog breed existing. Early history of dogs has been largely unknown with a sparse archaeological record, at least for some regions, and genetic studies with conflicting results. To get more knowledge about early dog history, new genetic markers and improved samples from recently neglected areas, is important. We have used a new marker, the Y-chromosome, to analyse the paternal history of dogs and also to get a first insight into the paternal history of the Australian dingo. We have also analysed the maternally inherited mtDNA in improved samples from regions which were only sparsely sampled earlier to get further information on the early spread of dogs. Finally, we have analysed mtDNA in the Scandinavian dog population to be used in forensic investigations. This research is also interesting because it can give clues to human and dog relationships in ancient times and also on early human migrations, assuming dogs were brought along as a cultural item.

Why have we used Y-chromosome and mtDNA sequences instead of autosomal DNA markers e.g. SNPs? First of all, mtDNA and the Y-chromosome are uniparently inherited without recombination while the autosomal chromosomes are shuffled through recombination between each generation. Therefore, the autosomal regions of the genome have different genealogical histories while mtDNA and the Y-chromosome have a single and deep genealogical history. Secondly, the number of SNPs per sequenced base is higher in the mitochondrial genome and the Y-chromosome compared to the autsomal chromosomes, and thus the resolution is better using the uniparentally inherited mtDNA and Y-chromosome (Underhill and Kivisild 2007).

I hope that my research can contribute to the general knowledge of early dog history and that the dog finally gets the attention that it deserves also from researchers.

2.1 Paper I

The Y-chromosome is an important second independent marker when investigating the origin of the domestic dog since there might be a risk that the pattern revealed by mtDNA actually reflects selection, sex-bias or stochastic variation instead of dog history. Since mtDNA is maternally inherited it only reflects the female history of dogs.

The best choice would be to study the non-recombining part of the Y-chromosome since it is strictly paternally inherited and haplotypes are easy to deduce because of the haploid state of the Y-chromosome.

When we started this work only 3,200 bp of dog Y-chromosomal sequence was available in Genbank (Olivier and Lust 1998, Pujar et al. 2005). Two dog genomes had been sequenced. One was from a male Poodle (Kirkness et al. 2003) and the other from a female Boxer (Lindblad-Toh et al. 2005). However, due to the low sequence coverage in the male Poodle genome the Y-chromosome sequence could not be directly identified (Kirkness et al. 2003).

To be able to deduce which of all the fragmentary shotgun sequences stemmed from the Y-chromosome we made a comparison using BLAST between the male Poodle genome, the human Y-chromosome and the female Boxer genome and we were able to select some of the clones that had a high similarity to the human Y-chromosome but did not match sequences in the female Boxer genome. We were able to isolate 43 putative dog Y chromosomal sequences. In order to further deduce the reliability that the sequence was indeed dog Y-chromosome we tested the different sequences against a panel of male and female dogs by PCR screening. We designed primers for the different fragments and ran a first test with 2 males and 2 females and analyzed the resulting amplification products on an agarose gel electrophoresis. Thus, we could see if the male sample amplification products had distinct bands of correct size and also if the female samples were blank. For a fragment to be selected as putative Y-chromosome product it had to be amplified in both male dogs and have distinct bands and should not amplify in any of the female samples. In the first round of tests we were able to identify 35 putative Y-chromosome sequences. The sequences that failed were either not amplified at all or amplified in both males and females. We also tested sequences excluded in the Blast analysis due to a match to the female genome and all of these sequences failed in the test showing that our exclusion method was reliable. We then further tested the different sequences towards a panel of three male dogs and three female dogs, again with the same criteria and in the end we had 24,159 bp of Y chromosomal dog sequence. In the last step 3 sequences failed, two having bands only for one of the male dogs and the third not amplifying at all.

To get a first insight into the Y-chromosome phylogeny and also provide a backbone for other population studies of the Y-chromosome in dogs we sequenced a set of dogs

from different parts of the world to avoid bias in the sample. We selected 10 dogs from 6 different regions (Europe, SW Asia, Siberia, East Asia, Africa and America). We detected totally 14 SNPs and one indel in the sample. The 14 SNPs defined 9 different haplotypes which could be constructed into a maximum parsimony phylogenetic tree. Because of the low sample size and the fact that no wolves or coyotes were included in the sample we could not calculate any mutation rate and as a gross simplification we used the Y-chromosome mutation rate for humans (Schaffner 2004) to calculate the minimum number of founders. The minimum number of founders was estimated to at least 5 based on a domestication date of 15,000 years ago which was in accordance with mtDNA and DLA data (Savolainen et al. 2002, Pang et al. 2009, Vilà et al. 2005). No phylogeographic pattern could be detected with this low sample size.

In conclusion, we were able to increase the available dog Y-chromosome sequence from 3,200 bp to 24,159 bp and also constructed a first Y-chromosome phylogenetic tree for domestic dogs based on 14,437 bp of Y chromosomal sequence from ten dogs originating in different parts of the world. We calculated the minimum number of founders to at least 5 wolf lineages although the lack of a species specific mutation rate makes this calculation uncertain. We could not detect any phylogeographic pattern with this limited sample.

2.2 Paper II

The next step was to use the Y-chromosome sequence to study a larger sample covering the whole world and also include wolf and coyote samples with the main aim to study the first origins of the domestic dog and also to calculate the Y-chromosome mutation rate in domestic dogs. We also wanted to investigate if there were any signs of introgression of wolf Y-chromosomes into the dog population.

In paper II we analyzed the same Y-chromosome region as in paper I. We sequenced the 14,437 bp in 165 canids (151 dogs, 12 wolves and 2 coyotes) in order to investigate the origin and spread of the domestic dog and to confirm or refute earlier mtDNA sequence data (Savolainen et al. 2002, Pang et al. 2009).

We had a large sample bank consisting of thousands of hair samples collected globally from earlier studies (Savolainen et al. 1997, Savolainen et al. 2002, Angleby and

Savolainen 2005, Pang et al. 2009) but since we wanted to amplify nuclear DNA the hair samples were of limited use because of the low DNA content in hair roots. Therefore, we collected buccal epithelial cell samples stored on Whatman FTA cards (Whatman Inc.) The samples were collected from virtually all parts of the world to avoid geographical bias and usually only a single sample per region. In order to avoid bias due to breed specific bottlenecks in the European dog population we collected only a single sample per breed, from breeds originating in different parts of the European continent and also having different morphologies.

Since the Y-chromosome sequence was obtained from shotgun sequences (paper I) it was analyzed in 18 different amplification products. In total, we detected 49 SNPs among 165 canids defining 28 haplotypes in domestic dog, one shared with wolves, 2 haplotypes in wolves and 2 haplotypes in coyote.

We detected variation resembling diploid variation in one of the fragments (fragment G) at three different sites; the haplotypes formed partly by this variation grouped in three different positions in the phylogeny. The most parsimonious explanation for this is that the fragment has been duplicated and that one of the copies has acquired a mutation. If the diploid variation was due to recombination we would not expect the mutations to appear clustered in three parts of the tree. As observed earlier (Wayne 1993, Clutton-Brock 1995), results clearly indicate that wolves are the ancestors of dogs. Dog and wolf haplotypes differed by 0-4 substitutions while the dog and coyote haplotypes differed by at least 15 substitutions.

The Y-chromosomal dog haplotypes formed 5 different clusters in the tree which we denoted HG1, HG3, HG6, HG9 and HG23 consisting of one or two high frequency central haplotypes surrounded by less frequent haplotypes in a star shaped pattern indicating population expansion. Haplotype 2* could not be assigned to a specific haplogroup because of its position between two haplogroups separated by a single substitutional step from each haplogroup.

We calculated the mutation rate for the dog Y-chromosome by relating the average genetic distance between coyotes and dog/wolf to the time since the lineages leading to wolf/dog and coyote shared a most recent common ancestor. We used a date for the split between wolves and coyotes established from paleontological research to 1.5 – 4.5

MYA (Nowak 2003, Tedford et al. 2009), a broad range since the fossil record is incomplete, giving a mutation rate of $1.39*10^{-10}$ – $4.31*10^{-10}$ substitutions per site per year. This rate is much lower than for the human Y-chromosome which has been estimated to $1.0*10^{-9}$ substitutions per nucleotide per year (Xue et al. 2009). One possible explanation for the lower mutation rate in the dog Y-chromosome is an effect of bias introduced in the selection process of Y chromosomal DNA. We used the human Y-chromosome for comparison and therefore we might have selected regions of the Y-chromosome which are the best conserved between dog and human and consequently the slowest diverging parts of the Y-chromosome.

We used the mutation rate to estimate the most conservative number of founders assuming that all 151 dog lineages lead back to wolf founders. We made the assumption that dogs were domesticated 11,500 – 16,000 years ago based on archaeology as well as genetics (Dayan 1994, Chaix 2000, Raisor 2005, Wang and Tedford 2008, Pang et al. 2009, Napierala and Uerpmann 2010, Skoglund, Götherström and Jakobsson 2011) and calculated the number of wolf founders to at least 13. Since we detected 28 haplotypes among domestic dogs, this means that part of the Y-chromosome phylogeny is intact from the founding wolf population(s).
Among the 5 haplogroups two were universally present, HG1 and HG23, and detected in 62% of all domestic dogs analysed. The other haplogroups had restricted distributions but included dogs sampled from distant parts of the world. For example the minor haplogroup HG9 consisted of two haplotypes, haplotype 9 found in three Central African Basenjis (sampled at three different localities) and haplotype 18 found in one East Siberian Laika. Possibly this haplogroup is rare and has not been detected elsewhere due to the low sample sizes.

Generally, HG3 and HG6 were rare or absent west of the Himalayas and the Urals while HG1 and HG23 were frequent in the western part of the world but also found in moderate frequencies in the eastern part of the world.

We then compared the genetic diversity in the three regions suggested by different studies to be the domestication center(s) for the domestic dog, ASY (Pang et al. 2009), SW Asia (Dayan 1994, Vonholdt et al. 2010) and Europe (Clutton-Brock 1995). Among these areas ASY had the highest haplotype diversity and number of haplotypes (0.901, 13 haplotypes) compared to SW Asia (0.863, 9.58 haplotypes adjusted for

sample size, resampling size 23) and Europe (0.734, 6.50 haplotypes adjusted for sample size, resampling size 23). However, the sample sizes are too small to get significant results. We observed that all haplotypes throughout the world differ at most by one substitutional step from a haplotype in ASY, except for the HG9 haplotypes, and therefore can possibly have been derived from haplotypes in ASY.

The Southwestern part of ASY (SE Asia and the Chinese provinces of Guangxi and Yunnan) had the highest genetic diversity observed in this study, 0.950 and also the highest number of haplotypes 10.10 when adjusted for sample size, resampling size 14. The genetic diversity in Europe was low 5.31 haplotypes (resampling size 14), reflected by a high frequency of haplotypes H1 and H1* in the northern and southern part of the continent as well as in Britain. In SW Asia 7.35 haplotypes (resampling size 14) was detected, the most frequent haplogroup being HG23, while haplogroups HG3 and HG6 were rare. Within SW Asia, the Fertile Crescent region (Fertile Crescent, western Iran, Israel and eastern Turkey) had an especially high genetic diversity, haplotype diversity 0.923 exceeding the diversity in ASY, but a lower number of haplotypes (8) than ASY.

The European population had an extremely low haplotype diversity reflected by the high frequency of H1 (47% of all individuals) and H1* (22%). From the pattern revealed it is not possible to reject a European origin for HG1. If HG1 had a European origin we expect to find more HG1 haplotypes in Europe than in other regions. We calculated the rate of substitutions supposed to have occurred since domestication among the 26 European lineages assuming domestication between 11,500 – 16,000 years ago. This calculation revealed that 0.57 – 2.60 (95% confidence limit) substitutions would have occurred since domestication. This indicates that only two founder haplotypes, H1 and H1* were inherited from wolf and the other haplotypes were derived from substitutions within the European dog population. Therefore, an origin outside Europe for these haplotypes and a later introduction is possible. Since these haplotypes are universally present at moderate frequencies, the origin of HG1 is not possible to establish.

HG23 had the highest diversity in SW Asia (5 haplotypes) but also appears at a rather high diversity in ASY, 3 individuals having three different HG23 haplotypes, but is largely absent in Europe. We calculated the expected substitutions to have occurred since domestication to 0.37 – 1.70 (95% confidence limits) thus indicating a possible

origin for HG23 in SW Asia. However, if this haplogroup originated in SW Asia we would expect to see more HG23 haplotypes in Europe since farming and the related domestic animals were introduced from the Fertile Crescent to Europe (Bellwood 2005). The high diversity and frequency of HG23 could alternatively reflect a local wolf-dog crossbreeding but then only a low frequency of the resulting haplotypes is supposed to be detected unless the crossbreeding results in a superior phenotype (Klütsch et al. 2011). Since HG23 is present at about 50% rendering a possible crossbreeding unlikely.

The frequency of HG3 and HG6 was highest in East Asia and these two haplogroups have a clear origin in East Asia and also the origin for HG1 and HG23 in East Asia cannot be ruled out.

We could not see any signs of local wolf-dog crossbreeding in our dataset since none of the haplogroups had a local distribution. It is possible that such crossbreedings remain undetected due to the low sample sizes since the resulting haplotypes are supposed to have a low frequency in a worldwide sample (Klütsch et al. 2011). However, crossbreedings does not seem to have had a major effect on the domestic dog population. Only three cases have been reported for mtDNA adding only approximately 2% of the total worldwide mtDNA diversity (Ishiguro et al. 2009, Pang et al. 2009, Klütsch et al. 2011, Ardalan et al. 2011) and a possible fourth has been discussed (Klütsch and Savolainen 2011). There is a possibility that there were many “small” crossbreedings for both mtDNA and Y-chromosome, such events are hard to detect since the frequency will be very low and restricted to particular regions.

To conclude, we have for the first time investigated Y-chromosome sequences in a worldwide sample of domestic dogs, wolves and coyotes. We could confirm that the domestic dog has a large number of wolf founders, at least 13, in agreement with earlier studies of mtDNA (Pang et al. 2009) and MHC (Vilà et al. 2005) sequence analysis. The Y-chromosome phylogeny formed 5 different haplogroups, two of which are universally distributed, two have a restricted distribution and one is rare and was only detected in 4 samples. The difference in genetic diversity was generally low between different regions but the highest diversity was observed in Southwestern ASY. All haplotypes (except HG9 haplotypes) outside ASY were at most separated by one substitutional step from a haplotype found in ASY. Thus, all haplotypes outside ASY

might have been derived from a haplotype in ASY. In combination with the mtDNA data (Pang et al. 2009) the Y-chromosome data provides strong evidence for a single origin of the domestic dog in ASY. Excluding the Y-chromosome HG9 and the mtDNA clade D and F, ASY is the only region harbouring the full set of diversity of both the maternal and paternal gene pools. No sign of local dog-wolf crossbreeding was detected in our sample.

2.3 Paper III

Earlier studies have shown that archaeological dog samples from Polynesia had only two mtDNA haplotypes and that the dingo population possibly had a single founder haplotype A29 (Savolainen et al. 2004) but due to low sample coverage of mainland and island Southeast Asia the introduction route for these dogs is largely unknown. An introduction with the Austronesian expansion from Taiwan has been suggested (Savolainen et al. 2004). Therefore, we wanted to investigate by which route dogs were introduced to Polynesia and Australia with a new and improved sample. We also wanted to test general hypotheses about the introduction of dogs into Polynesia. An often cited model for Polynesian human and cultural origins is the express train model (Diamond 1988, Diamond 2001). This model suggests that domestic dog along with other domesticates as well as pottery and human genes were introduced to Polynesia from a Taiwanese homeland. Lastly, we also wanted to investigate the relation between NGSD and Australian dingoes. Earlier studies have detected haplotype sharing between these dogs (Savolainen et al. 2004, Runstadler et al. 2006).

The question of who the first Polynesians were and where they came from has been debated for a long time and several different models for the population expansion into Polynesia (Remote Oceania) have been suggested based on different kinds of evidence (Hurles et al. 2003); Archaeology, linguistics, genetics and recently also genetics on commensal and domestic animals used as proxies for human dispersal (Matisoo-Smith and Robins 2004, Larson et al. 2007, Diamond 1988, Oppenheimer and Richards 2001, Gray, Drummond and Greenhill 2009b).

The express train model was suggested by Diamond in 1988 and is based primarily on archaeological and linguistic evidence. The model proposes an origin for Polynesian

people on the island of Taiwan about 6,000 BP with a subsequent rapid spread into Near Oceania with minor or no admixture with local indigenous people and then a final leap out into the uninhabited Polynesia (Diamond 1988, Diamond 2001).

The entangled bank model suggests that the Polynesian culture derived from a complex network of contacts between different indigenous populations and not through a diaspora of people (Terrell 1988, Terrell et al. 2001). Several other models have also been suggested to explain how humans reached Polynesia, a slow spread from Taiwan (Kayser et al. 2006, Kayser et al. 2008), and a model suggesting an origin for the Polynesian population in the region of Wallacea (Oppenheimer and Richards 2001).

The archaeological record traces the Polynesian culture back to Lapita which is a culture that appears in Near Oceania about 3,500 BP (Kirch and Weisler 1994). The pottery and language of Lapita can be traced back to the island of Taiwan (Bellwood 2005, Blust 1995, Diamond and Bellwood 2003).

Dogs appear in the archaeological record in mainland Southeast Asia around 4,000 BP (Higham 1996) and in Island Southeast Asia about 3,000 BP (Bellwood 2005, Bellwood 1997). Also pigs appear in the archaeological record of eastern Indonesia at about the same time (Bellwood and White 2005).

However, the genetic evidence generally does not support a Taiwanese origin for Polynesian people. Most of the studies show a Melanesian origin for Polynesian Y-chromosomes and mtDNAs with only a minor contribution from Taiwan (Kayser et al. 2000, Oppenheimer and Richards 2001, Redd et al. 1995, Hagelberg et al. 1999, Soares et al. 2011, Capelli et al. 2001, Hurles et al. 2002, Kayser et al. 2006, Kayser et al. 2008).

Studies using domestic and commensal animals as proxies for human migrations agree with an origin for these animals in Near Oceania (Matisoo-Smith and Robins 2004, Larson et al. 2007). In the pig study a Pacific clade was detected and could be traced through mainland and island Southeast Asia and not found at all on Taiwan or the Philippines (Larson et al. 2007).

Earlier studies have shown that among 19 pre-European domestic dogs from all parts of Polynesia (New Zealand, Cook Island and Hawaii) two mtDNA haplotypes were detected, Arc1 and Arc2 in 100% of the samples (Savolainen et al. 2004). Both haplotypes were found in all three locations (Savolainen et al. 2004). The archaeological assemblages were dated at 400-1,000 BP and thus clearly predate any European contact (Savolainen et al. 2004).

We investigated the extant dog population with samples from Australia, New Guinea, Taiwan, the Philippines and Indonesia to trace the origin and spread of haplotypes Arc1, Arc2 and A29 in order to investigate the origin and introduction routes of early Polynesian domestic dogs and Australian dingoes. All samples were collected from rural areas with a low influx of foreign dogs and mainly from indigenous populations.

All three haplotypes were generally confined to the eastern part of the world, Arc2 and A29 found solely east of the Himalayas and Arc1 found sparsely west of the Himalayas, in a global sample of dogs (Pang et al. 2009).

We traced the Polynesian Arc haplotypes from a low frequency in mainland SE Asia (16%), through a higher frequency in ISEA (42%) to reach a maximum frequency (100%) in Polynesia. Neither of the Arc haplotypes was detected on Taiwan or the Philippines.

The data indicate strongly that Polynesian domestic dogs were introduced through Mainland SEA and ISEA and not through Taiwan and the Philippines. A similar pattern has been revealed in domestic pigs (Larson et al. 2007).

No novel haplotypes were detected in the 21 dingoes sequenced in this study. All haplotypes detected carried the dingo founder haplotype A29 (Savolainen et al. 2004).

The distribution of A29 was similar to the pattern revealed for the Polynesian Arc haplotypes. It had a low frequency in mainland SE Asia (1%) followed by a gradually higher frequency in ISEA (8%) to reach a maximum frequency (100%) in Australia, but was not detected in Taiwan or the Philippines.

The chance of not sampling any of Arc1, Arc2 or A29 in 52 samples from Taiwan assuming a similar frequency as in the source population, South China, was calculated to 0.0015.

We also analysed one New Guinea Singing Dog (NGSD) having haplotype A29. In total with the two NGSDs analyzed in Savolainen et al. (2004) two haplotypes were detected among the NGSDs, two carried haplotype A29 shared with Australian dingoes and domestic dogs and one dog carried haplotype A79 a private haplotype in NGSD and separated one mutational step from A29. Thus, indicating a common history for NGSD and Australian dingoes.

With additional dingo samples and a new mtDNA mutation rate obtained from sequencing of the entire mitochondrial genome (Pang et al. 2009) we calculated a new introduction date for the Australian dingo. Assuming that A29 is the sole founder haplotype and excluding 25 dingo samples from the Pilbarra region probably affected by genetic drift (Savolainen et al. 2004), we could calculate the mean genetic distance among dingoes to rho=0.0016 (se = 0.0005) estimating the introduction of dingoes to Australia at 4,600-18,300 years ago and thus predating the Neolithic in the region surrounding Australia.

All three haplotypes (Arc1, Arc2 and A29) were found throughout mainland and island Southeast Asia but none was present on Taiwan or the Philippines. Therefore, the dingo and the Polynesian dogs seem to have been introduced through the same route and possibly from the same dog population but due to founder bottlenecks have different haplotypes. The dingo has been isolated on the Australian continent for centuries and might actually best represent what early domestic dogs looked like. To investigate if the mtDNA studies are affected by selection, sex-bias or random variation, studies of other independent markers are necessary (see paper IV).

With these data we can show that the Express train model of Polynesian origins does not fit with the introduction of domestic dogs to Polynesia and Australia. The mtDNA data presented here indicates that Polynesian domestic dogs and Australian dingoes arrived in Island Southeast Asia before the arrival of the Neolithic.

In conclusion, both Polynesian domestic dogs and Australian dingoes trace their origin to mainland SE Asia and might have been introduced together before the arrival of the Neolithic. mtDNA data indicate a pre-Neolithic date of introduction for the Australian dingo. We could show that the Australian dingo and New Guinea Singing Dog have close ties and possibly share a common history.

2.4 Paper IV

To further investigate the ancestry of the Australian dingo population, and to get an insight into the paternal ancestry of Australian dingoes and also to investigate if earlier mtDNA data actually reflects selection, stochastic variation or sex-biased migration patterns instead of dog population history, we used the newly developed framework of Y-chromosome SNPs from paper I and paper II.

Since no dingoes were included in the test panel of paper I or paper II we started with sequencing of the previously investigated Y chromosomal DNA region in 2 dingoes and 1 NGSD to investigate if the dingo is part of the dog population. The sequencing resulted in a novel SNP leading to a new haplotype, the dingo/NGSD specific H60, providing further evidence for a connection between the two dog populations. The other haplotype H3 was shared between domestic dogs and dingoes and was primarily found among domestic dogs from the East Asian continent and in a few North European individuals, thus corroborating with mtDNA data that the dingo has an East Asian domestic dog origin and from very few founders. To save time and money, a SNP analysis of 47 Australian dingoes was performed. The SNP detection was divided into two runs where the first round sorted the samples into haplogroups (HGs) and the second round into haplotypes (HTs).

We used the PrASE SNP detection method which is based on a microarray detection system, involving tagged allele specific primers and labeled nucleotides for detection with laser (Hultin et al. 2005).

Among the 47 male dingoes analyzed we detected only two haplotypes, H3 and H60, thus suggesting that the Australian dingo population has an origin from few male founders, possibly only two. However, there is a possibility that some of the samples

have novel haplotypes not previously detected and therefore missed in the SNP analysis. The genetic diversity resembles the mtDNA data with few founders and thereafter isolation (Savolainen et al. 2004).

A weak pattern was detected in the Australian dingo population with a higher frequency of H60 in the North West and H3 in West and South. This pattern might be an effect of the low sample size in combination with these haplotypes being generally rare and not detected in all regions.

In conclusion, the paternal origin of the Australian dingo resembles that of the mtDNA, few founders from an East Asian domestic dog population and thereafter long isolation. Also, the Australian dingo and NGSD share one Y-chromosome haplotype. Earlier studies of mtDNA and DLA sequence have also revealed haplotype sharing between Australian dingoes and NGSDs (Savolainen et al. 2004, Runstadler et al. 2006). With this study we have shown that the Y-chromosome SNPs detected in paper I and paper II can be used to study paternal population histories with an easy and straightforward SNP detection method like PrASE (Hultin et al. 2005).

Note-In this thesis table S2 and S3 were not included because of the large size of the documents. These files are available upon request to the author.

2.5 Paper V

We wanted to investigate the origin of the dog population on Madagascar to detect if there was a connection to the spread of dogs to Polynesia and Australia (paper III) or if the population was mainly of mainland African origin or admixed with origins from both the African mainland and the Austronesian world.

Earlier studies have shown that the language spoken on Madagascar belongs to a subgroup of the Austronesian languages spoken in the Polynesian world (Blust 1995, Gray et al. 2009b, Diamond and Bellwood 2003, Adelaar 1995a). However, there is no continuation in archaeological sites between the Austronesian world and Madagascar (Dewar 1995).

Archaeological and paleontological evidence suggest that the island was uninhabited until approximately 2,000 BP (Burney et al. 2004, MacPhee and Burney 1991) when suddenly bones from now extinct animals bearing signs of modification by humans appear in the archaeological record. The earliest permanent settlement on Madagascar has been dated at approximately 1,300 years ago (Dewar 1995).

Human genetic studies have indicated that the Malagasy population is an admixed population of both African and Asian origin (Soodyall, Jenkins and Stoneking 1995, Hurles et al. 2005, Tofanelli et al. 2009).

Fst analysis of the Asian derived haplogroups to possible source populations in Asia has shown the population of Banjarmasin to be closest to the Asian HGs found on Madagascar (Hurles et al. 2005). Also linguistic evidence has traced the origin of the Malagasy language to the Southeast Barito languages of Borneo (Adelaar 1995a).

In order to investigate the history of the dog population on Madagascar we gathered a sample of dogs from all parts of the island both in the lowland region (n=54) and the highland region (n=38). Most of the dogs were sampled in rural areas with a low influx of foreign dogs. In total we investigated 145 samples from Madagascar, 578 samples from throughout mainland Africa and 2,076 dogs from most other regions of the world (Pires et al. 2006, Pang et al. 2009, Boyko et al. 2009).

The mtDNA gene pools of the dog populations in Africa and Indonesia are very different, with the African population carrying the same haplotypes, mostly UTs (universal haplotypes), as most other parts of the Old World while the Indonesian sample have a low frequency of UTs and a relatively high frequency of the Austronesian haplotypes Arc1, Arc2 and A29 (see paper III). This difference makes it possible to detect from which population the dog population on Madagascar was created.

We investigated the different dog populations for shared haplotypes to investigate the history of the Madagascan dog population's foundation and also try to estimate the minimum number of founders for the Madagascan dog population.

All of the dogs (100%) in the Madagascan population had haplotypes that were either shared or derived by a single substitution from a haplotype found in the mainland African dog population. While at most 63% of the dogs could be linked to the Austronesian dog population. All of the shared haplotypes between Madagascan and Austronesian dogs were UTs. Since UTs are common haplotypes found at moderate to high frequencies throughout the world they are non-informative when investigating the source population for Madagascan dogs. Ignoring the UTs, the sharing dropped from 63% to 0% between Madagascar and Austronesian dogs. However, 66% of the Madagascan dogs shared haplotypes with the mainland African population even when ignoring UTs and among these 35% were shared exclusively between Madagascan and mainland African dogs.

The Polynesian haplotypes Arc1, Arc2 and A29 which were found in 48% of the Indonesian dogs and 100% of the Polynesian dogs (paper III) was only detected at an extremely low frequency in Madagascar, Arc2 and A29 not detected at all and Arc1 found in a single individual (1%).

We also analyzed other regions to search for the source population of Madagascan dogs. Among the Malgasy dogs 93% carried either a shared or derived haplotype with the Indian and SW Asian dog populations, respectively. Most of the shared haplotypes were UTs, ignoring the UTs, the sharing dropped to 21% of the dogs sharing haplotype with the Indian dog population and 24% with the SW Asian.

We estimated the number of founders by counting of haplotypes shared between mainland Africa and Madagascar. We also calculated the mean distance from unique Madagascar haplotypes to the founder haplotypes, giving a rho value of $\rho=0.07$. The rho value and the estimated mutation rate for mtDNA (Pang et al. 2009), suggests an introduction date at 2,800-11,200 years ago.

The dog population on Madagascar is clearly different from other island populations related to the Austronesian expansion, having a much higher haplotype diversity (0.897) than the dog populations of Polynesia (0.456) and Australia (0.311), suggesting that a relatively large population of dogs were introduced to the island not experiencing any severe bottlenecks or genetic drift as in other island populations. Another

explanation for the relatively high haplotype diversity might be that there were several introductions of dogs to the island each carrying a different set of haplotypes.

In conclusion, our data indicate a predominantly mainland African origin for the dog population on Madagascar. A minimum number of 18 founders were estimated based on counting of haplotypes shared between mainland Africa and Madagascar. We calculated the introduction date at 2,800-11,200 years ago, thus predating the archaeological record of human presence. However, the calculation is highly dependent on that all founder haplotypes are being detected and there is a chance that more haplotypes might be found on the African continent if the sample size is increased and therefore the dating is only an estimation of the introduction date. We also show that the Madagascar island population has a higher genetic diversity than other island populations connected to the Austronesian expansion suggesting introduction of a large and diverse population or multiple introductions of dogs carrying different lineages.

Note-In this thesis table S1 was not included because of the large size of the document. This file is available upon request to the author.

2.6 Paper VI

Due to the large size of the supplementary material it is not included here but is available upon request.

The first humans entered the New World through the Bearing strait at least 15,000 BP (Goebel et al. 2008) when the two landmasses were connected through a land bridge.

In the first genetic study of the origin of American dogs, mtDNA variation obtained from ancient American specimens showed that the ancient New World dogs have a common origin with dogs in the Old World since both populations have the same mtDNA clades (Leonard et al. 2002). Also, a cluster of haplotypes unique to the South/Meso American dogs was detected within clade A.

In a recent study of mtDNA control region sequences from extant dogs in America, discontinuity between ancient and extant dog populations were suggested; at most 10% of indigenous mtDNA lineages was left (Castroviejo-Fisher et al. 2011). However, the samples were from free-ranging street dogs and village dogs with no indigenous breed

dogs included. Further, the results were compared to haplotypes found in only 19 ancient specimens from an earlier study (Leonard et al. 2002).

We wanted to investigate to what extent the mtDNA lineages of American breed as well as free-ranging dogs have a pre-Colombian East Asian origin or a post-Colombian European origin using an alternative approach. Previous studies (Pang et al. 2009) Have shown that the two putative source populations have very different gene pools making it easy to deduce from which population American dogs stem. Sequencing of the entire mtDNA genome in dogs revealed that the earlier detected clades consists of several subclades and in East Asia all 10 mtDNA subclades for clade A, B and C are present while only 4 are present in Europe. Therefore, if any of the 6 subclades not present in Europe is found in the American population they have most probably been introduced from East Asia.

All of the investigated putative indigenous breed dogs in America carried haplotypes derived from a probable East Asian source. In the Arctic breeds 79% of the dogs carry haplotypes not present in Europe and among the South/Meso American dogs 35% of the dogs carry haplotypes not present in the European dog population.

In two cases we detected continuity between the ancient and extant dog population in America. Firstly, haplotype A185 was detected in an ancient Mexican sample and was in the modern samples unique for Chihuahua. Secondly, a haplotype corresponding to haplotype A29 was detected in an ancient sample from Alaska and in modern samples from Alaskan Malamute. The central haplotype of the previously detected cluster of ancient South/Meso American haplotypes was detected in a modern sample from Korea and thus indicates an East Asian origin for this cluster. Four of the other haplotypes correspond to modern universal haplotypes and the rest were unique in the ancient samples. One of the unique ancient haplotypes (D40) is separated from the other haplotypes by 4 substitutional steps and could be a result of dog-wolf crossbreeding.

The free-ranging southern USA Carolina dog carried mainly haplotypes (37%) that were unique and belonged to an East Asian specific subclade. All other haplotypes were either universal or shared with Chinese non-breed dogs and a Japanese breed dog. Contrastingly, the free-ranging dogs in South America carried almost exclusively universal haplotypes and several European specific haplotypes.

In conclusion, a primarily Old World origin for the New World indigenous breed dogs investigated here was detected. Most of the indigenous American breed dogs carried universal haplotypes or haplotypes not detected in the European dog population. We also detected continuity between the ancient and extant dog population in America in two cases. We detect a large impact of European dogs only in the free-ranging dogs of South America where most of the dogs carries either UTs or haplotypes uniquely shared with European breed dogs. However, there remain pockets of populations largely unaffected by European dogs, for example the free-ranging Carolina Dog in southern USA.

Note-In this thesis table S1 was not included because of the large size of the document. This file is available upon request to the author.

3 Discussion and future perspectives

We have now come a long way from the first genetic studies on human origins using mtDNA RFLP (Cann et al. 1987). We have through the introduction of the Sanger sequencing technology (Sanger et al. 1977, Sanger 1988) and the amplification of DNA sequences (Saiki et al. 1985, Mullis and Faloona 1987) arrived at the latest technology, sequencing enormous amounts of DNA by direct sequencing of DNA libraries using for instance the Illumina GA II, in a single run (Stoneking and Krause 2011).

The most complete studies on dog origins have been investigating the mtDNA in dogs from throughout the world indicating a single origin in ASY (Pang et al. 2009) or autosomal SNPs in dogs suggesting a mainly Near Eastern origin for the domestic dog (Vonholdt et al. 2010). The detection of dog Y-chromosome sequences has been important to either corroborate or refute the different theories. The Y-chromosome data agrees most closely to mtDNA data indicating a single origin for the domestic dog in ASY (paper II). The establishment of the domestic dog Y phylogeny is also important as a backbone for future studies of Y-chromosomes in domestic dogs either by SNP analysis, microsatellite analysis or the two combined (paper I, paper II). The mtDNA and Y-chromosome studies have shown the importance of sampling dogs from ASY in studies of dog origins (Pang et al. 2009) (paper II). The diversity is higher and most or all of the mtDNA and Y-chromosome haplogroups are found there, while the diversity and representation of haplogroups are significantly lower in central and north China. Therefore, in studies of dog domestication, samples from north or central China can not compensate for samples from ASY and therefore if not including samples from ASY the origin of dogs there could be missed.

How dogs spread from their area of origin to other parts of the world is important because these dispersals are often accompanied by humans and thus can give us clues about human dispersal in ancient times. Dog mtDNA has revealed that domestic dogs were introduced to Polynesia and Australia through Mainland SEA and Indonesia (paper III) and can therefore be used as part of the puzzle to investigate how humans migrated to Polynesia and also to clarify the origin of the Australian dingo. mtDNA studies have also revealed a relationship between NGSD and Australian dingoes and a possible connection between the two landmasses (Savolainen et al. 2004, Runstadler et al. 2006)(paper III). mtDNA of domestic dogs has also been used to try to clarify how

dogs were introduced to these regions and to clarify some aspects of human migrations to Madagascar (paper V) and into the American continent (paper VI).

The Y phylogeny has been successfully tested in a study on Australian dingoes and a similar history as earlier revealed by mtDNA could be detected. This study shows that the Y phylogeny can be used to study paternal domestic dog history by using the SNPs detected in the Y phylogeny (paper IV).

The new technologies will allow a lot of DNA to be analysed in a short time and also promise good prospects for analyses of ancient material (Stoneking and Krause 2011). The next step for scientists working on dog domestication will be to pinpoint the area within ASY that harbours the highest genetic variation and to study in detail how wolves have contributed to the dog gene pool through hybridization. This can be accomplished through large scale sequencing of mtDNA, Y-chromosome and the autosomal chromosomes. Also future archaeological excavations in areas largely neglected up to now could probably reveal some interesting clues on dog origins, and sequencing of ancient material could within a few years become a relatively easy task and could potentially reveal the key to dog domestication

4 Acknowledgements

First of all I would like to thank my supervisor Peter. For accepting me as a student, for always being full of energy and always being available for questions and discussions even on a Sunday night. Thanks also to Linda.

I would also like to thank all my colleagues and collaborators. The dog group: Helen for teaching me almost everything about clean room procedures, Cornelya for all help with analysis of sequences and also for lunch company, Ai-bing for making terrific programs and Arman for all your hard work with paper writing and help with analysis of data, I am looking forward to your dissertation. Thanks also to all diploma students and project workers, for help with lab work and for being positive and nice people.

Barbara, thank you for your hard work finishing up the America paper.

Thanks to all the dog owners who contributed with samples, without you this research would not have been possible.

Afshin and Birgit Savolainen, thanks for reading an early draft of this thesis. Your comments were very valuable.

Christian N, Kicki and Bahram without you I would still be stuck in the DNA1 room.

Thanks also to the Texas Hold Em crew (Afshin, Bahram, Christian, Patrik, Savo, and all other people who have participated), the climbing crew (Anders, Erik, Mårten, Måns, Anna, Josefine, Jimmie, Sverker and all other people who have participated).

Patrik S, Pawel Z, Erik P and Christian N the Carmen crew. Why did we switch from Carmen to Jimmies? I hope that we will continue with our monthly dinner and beer routine. I could write a book with thanks to you guys, we have had so much fun during these years but I have limited space and the thesis is already long as it is.

Anna B, thank you for being such a nice and positive person. It has been fun discussing everything from shoe-gazing to the latest movies.

The HPR molbio crew both old and new members. For letting me use your computers and also sharing lab space with me. Thank you also for all the times when you have lent me a helping hand with equipment or reagents.

Emma, Francis, Josefine and everyone in the RR room. Thank you Emma for being such a nice person, I already miss our fantastic discussions about everything and nothing. Francis and Josefine deserves a special thanks, it must have been a hard time to sit next to me and all my papers and coffee cups.

Jojje and Calle H, thank you for all the dinners, beer, “Call of Duty” and movie nights. Thanks also Jojje for all the fun we had during “grundutbildningen”, I will never forget the Ingarö time.

Fredrik S, the fastest guy from Ljungsbro, or? We have to decide, I am looking forward to run with you soon again.

Everyone at Albanova floor 3 and Scilife that I have encountered with during these years.

Johan L and Lauren for nice lunch company, dinners and movie nights. Thank you for introducing me to “The Room”. Good luck in Vancouver and Sundbyberg.

Peter J for inviting me to the ” torsdags” gatherings at Indianen and for the fantastic midsummer parties that you arrange.

My loved one, Paula Borg! You have supported me throughout this process and without you it would not have been possible.

My second family, Fredrik and Veronica. You’re door is always open when I want to visit. Maybe you should close the door sometimes now that also burglars have used the opportunity to visit you.

Mom, dad, my brother Stefan and his family Karin, Benjamin and Teo! You are the best family one can wish for.

My cousins Nicklas and Richard, big thank you for taking such good care of me during these years. You are my bonus brothers. Thanks also to Britt-Marie and Terje.

Thanks to Göran, Irene, Angelica, Ellinor and Jörgen. It is always a pleasure to meet you.

Gittan and Hasse, I am looking forward to discuss this thesis and the latest horse races with you.

Thanks also to those of you that I have forgotten, I am really sorry for that but right now I am really tired and I only want to send this thesis away to print.

And finally to all the dogs that have contributed with blood, hair or buccal epithelial cells, I thank you with a big VOFF!!!

5 References

Achilli, A., A. Olivieri, P. Soares, H. Lancioni, B. Hooshiar Kashani, U. A. Perego, S. G. Nergadze, V. Carossa, M. Santagostino, S. Capomaccio, M. Felicetti, W. Al-Achkar, M. C. Penedo, A. Verini-Supplizi, M. Houshmand, S. R. Woodward, O. Semino, M. Silvestrelli, E. Giulotto, L. Pereira, H. J. Bandelt & A. Torroni (2012) Mitochondrial genomes from modern horses reveal the major haplogroups that underwent domestication. *Proc Natl Acad Sci U S A,* 109**,** 2449-54.

Adelaar, A. (1995a) Asian roots of the Malagasy - A linguistic perspective. *Bijdragen Tot De Taal- Land- En Volkenkunde,* 151**,** 325-356.

Ajmone-Marsan, P., J. F. Garcia, J. A. Lenstra & T. G. Consortium (2010) On the Origin of Cattle: How Aurochs Became Cattle and Colonized the World. *Evol Anthropol,* 19**,** 148-157.

Anderson, S., A. T. Bankier, B. G. Barrell, M. H. de Bruijn, A. R. Coulson, J. Drouin, I. C. Eperon, D. P. Nierlich, B. A. Roe, F. Sanger, P. H. Schreier, A. J. Smith, R. Staden & I. G. Young (1981) Sequence and organization of the human mitochondrial genome. *Nature,* 290**,** 457-65.

Anderson, T. M., B. M. vonHoldt, S. I. Candille, M. Musiani, C. Greco, D. R. Stahler, D. W. Smith, B. Padhukasahasram, E. Randi, J. A. Leonard, C. D. Bustamante, E. A. Ostrander, H. Tang, R. K. Wayne & G. S. Barsh (2009) Molecular and evolutionary history of melanism in North American gray wolves. *Science,* 323**,** 1339-43.

Anderung, C., A. Bouwman, P. Persson, J. M. Carretero, A. I. Ortega, R. Elburg, C. Smith, J. L. Arsuaga, H. Ellegren & A. Götherström (2005) Prehistoric contacts over the Straits of Gibraltar indicated by genetic analysis of Iberian Bronze Age cattle. *Proc Natl Acad Sci U S A,* 102**,** 8431-5.

Angleby, H. & P. Savolainen (2005) Forensic informativity of domestic dog mtDNA control region sequences. *Forensic Sci Int,* 154**,** 99-110.

Ardalan, A., C. F. Kluetsch, A. B. Zhang, M. Erdogan, M. Uhlén, M. Houshmand, C. Tepeli, S. R. Ashtiani & P. Savolainen (2011) Comprehensive study of mtDNA among Southwest Asian dogs contradicts independent domestication of wolf, but implies dog-wolf hybridization. *Ecol Evol,* 1**,** 373-85.

Bandelt, H. J. (2008) Clock debate: when times are a-changin': time dependency of molecular rate estimates: tempest in a teacup. *Heredity (Edinb),* 100**,** 1-2.

Bannasch, D. L., M. J. Bannasch, J. R. Ryun, T. R. Famula & N. C. Pedersen (2005) Y chromosome haplotype analysis in purebred dogs. *Mamm Genome,* 16**,** 273-80.

Bellwood, P. 1997. *Prehistory of the Indo-Malaysian archipelago*. Honolulu: University of Hawaii Press.

---. 2005. *First farmers: the origins of agricultural societies*. Oxford: Blackwell Publishing.

--- (2011) The Checkered Prehistory of Rice Movement Southwards as a Domesticated cereal-from the Yangzi to the Equator. *Rice,* 4**,** 93-103.

Bellwood, P. & P. White (2005) Domesticated pigs in eastern Indonesia. *Science,* 309**,** 381; author reply 381.

Bensasson, D., D. Zhang, D. L. Hartl & G. M. Hewitt (2001) Mitochondrial pseudogenes: evolution's misplaced witnesses. *Trends Ecol Evol,* 16**,** 314-321.

Binladen, J., M. T. Gilbert, J. P. Bollback, F. Panitz, C. Bendixen, R. Nielsen & E. Willerslev (2007) The use of coded PCR primers enables high-throughput sequencing of multiple homolog amplification products by 454 parallel sequencing. *PLoS One,* 2**,** e197.

Björnerfeldt, S., M. T. Webster & C. Vilà (2006) Relaxation of selective constraint on dog mitochondrial DNA following domestication. *Genome Res,* 16**,** 990-4.

Blust, R. (1995) The prehistory of the Austronesian-speaking peoples: a view from language. *J World Prehistory,* 9**,** 453-510.

Bogenhagen, D. F. (1999) Repair of mtDNA in vertebrates. *Am J Hum Genet,* 64**,** 1276-81.

Bowler, J. M., H. Johnston, J. M. Olley, J. R. Prescott, R. G. Roberts, W. Shawcross & N. A. Spooner (2003) New ages for human occupation and climatic change at Lake Mungo, Australia. *Nature,* 421**,** 837-40.

Boyko, A. R., R. H. Boyko, C. M. Boyko, H. G. Parker, M. Castelhano, L. Corey, J. D. Degenhardt, A. Auton, M. Hedimbi, R. Kityo, E. A. Ostrander, J. Schoenebeck, R. J. Todhunter, P. Jones & C. D. Bustamante (2009) Complex population structure in African village dogs and its implications for inferring dog domestication history. *Proc Natl Acad Sci U S A,* 106**,** 13903-8.

Brown, S. K., N. C. Pedersen, S. Jafarishorijeh, D. L. Bannasch, K. D. Ahrens, J. T. Wu, M. Okon & B. N. Sacks (2011) Phylogenetic distinctiveness of Middle Eastern and Southeast Asian village dog Y chromosomes illuminates dog origins. *PLoS One,* 6**,** e28496.

Burney, D. A., L. P. Burney, L. R. Godfrey, W. L. Jungers, S. M. Goodman, H. T. Wright & A. J. T. Jull (2004) A chronology for late prehistoric Madagascar. *Journal of Human Evolution,* 47**,** 25-63.

Cadieu, E., M. W. Neff, P. Quignon, K. Walsh, K. Chase, H. G. Parker, B. M. Vonholdt, A. Rhue, A. Boyko, A. Byers, A. Wong, D. S. Mosher, A. G. Elkahloun, T. C. Spady, C. André, K. G. Lark, M. Cargill, C. D. Bustamante, R. K. Wayne & E. A. Ostrander (2009) Coat variation in the domestic dog is governed by variants in three genes. *Science,* 326**,** 150-3.

Candille, S. I., C. B. Kaelin, B. M. Cattanach, B. Yu, D. A. Thompson, M. A. Nix, J. A. Kerns, S. M. Schmutz, G. L. Millhauser & G. S. Barsh (2007) A -defensin mutation causes black coat color in domestic dogs. *Science,* 318**,** 1418-23.

Cann, R. L., M. Stoneking & A. C. Wilson (1987) Mitochondrial DNA and human evolution. *Nature,* 325**,** 31-6.

Capelli, C., J. F. Wilson, M. Richards, M. P. Stumpf, F. Gratrix, S. Oppenheimer, P. Underhill, V. L. Pascali, T. M. Ko & D. B. Goldstein (2001) A predominantly indigenous paternal heritage for the Austronesian-speaking peoples of insular Southeast Asia and Oceania. *Am J Hum Genet,* 68**,** 432-43.

Castroviejo-Fisher, S., P. Skoglund, R. Valadez, C. Vilà & J. A. Leonard (2011) Vanishing native American dog lineages. *BMC Evol Biol,* 11**,** 73.

Cavalli-Sforza, L. L. & M. W. Feldman (2003) The application of molecular genetic approaches to the study of human evolution. *Nat Genet,* 33 Suppl**,** 266-75.

Chaix, L. 2000. A preboreal dog from the northern Alps (Savoie, France). In *Dogs Through Time: An Archaeological Perspective,* ed. S. Crockford, 49-59. Oxford: British archaeological reports.

Chen, S., B. Z. Lin, M. Baig, B. Mitra, R. J. Lopes, A. M. Santos, D. A. Magee, M. Azevedo, P. Tarroso, S. Sasazaki, S. Ostrowski, O. Mahgoub, T. K. Chaudhuri, Y. P. Zhang, V. Costa, L. J. Royo, F. Goyache, G. Luikart, N. Boivin, D. Q. Fuller, H. Mannen, D. G. Bradley & A. Beja-Pereira (2010) Zebu cattle are an exclusive legacy of the South Asia neolithic. *Mol Biol Evol,* 27**,** 1-6.

Clutton-Brock, J. 1995. Origins of the dog: domestication and early history. In *The Domestic Dog, its Evolution, Behavior and Interactions With People,* ed. J. Serpell, 7-20. Cambridge: Cambridge University Press.

Coelho, M., F. Sequeira, D. Luiselli, S. Beleza & J. Rocha (2009) On the edge of Bantu expansions: mtDNA, Y chromosome and lactase persistence genetic variation in southwestern Angola. *BMC Evol Biol,* 9**,** 80.

Consortium, I. H. G. S. (2004) Finishing the euchromatic sequence of the human genome. *Nature,* 431**,** 931-45.

Crockford, S. J. 2000. Dog evolution: A role for thyroid hormone physiology in domestication changes. In *Dogs through Time: An Archaeological Perspective,* ed. S. J. Crockford, 11-20. Oxford: British archaeological reports.

Davis, S. J. M. & F. R. Valla (1978) Evidence for domestication of the dog 12,000 years ago in the Natufian of Israel. *Nature,* 276**,** 608-610.

Dayan, T. (1994) Early Domesticated Dogs of the Near East. *J Archaeol Sci,* 21**,** 633-640.

Deguilloux, M. F., J. Moquel, M. H. Pemonge & G. Colombeau (2009) Ancient DNA supports lineage replacement in European dog gene pool: insight into Neolithic southeast France. *Journal of Archaeological Science,* 36**,** 513-519.

Dewar, R. E. (1995) Of Nets and Trees: Untangling the Reticulate and Dendritic in Madagascar's Prehistory. *World Archaeology,* 26**,** 301-318.

Diamond, J. (2001) Polynesian origins: slow boat to Melanesia? (Reply to: S. Oppenheimer & M. Richards). *Nature,* 410**,** 167.

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

Diamond, J. & P. Bellwood (2003) Farmers and their languages: the first expansions. *Science,* 300**,** 597-603.

Diamond, J. M. (1988) Express train to Polynesia. *Nature,* 336**,** 307-308.

Drögemüller, C., E. K. Karlsson, M. K. Hytönen, M. Perloski, G. Dolf, K. Sainio, H. Lohi, K. Lindblad-Toh & T. Leeb (2008) A mutation in hairless dogs implicates FOXI3 in ectodermal development. *Science,* 321**,** 1462.

Ellegren, H. (2000) Microsatellite mutations in the germline: implications for evolutionary inference. *Trends Genet,* 16**,** 551-8.

--- (2004) Microsatellites: simple sequences with complex evolution. *Nat Rev Genet,* 5**,** 435-45.

Fédération Cynologique Internationale, <http://www.fci.be> (Retrieved 2012-04-29).

Finlayson, C., F. G. Pacheco, J. Rodríguez-Vidal, D. A. Fa, J. M. Gutierrez López, A. Santiago Pérez, G. Finlayson, E. Allue, J. Baena Preysler, I. Cáceres, J. S. Carrión, Y. Fernández Jalvo, C. P. Gleed-Owen, F. J. Jimenez Espejo, P. López, J. A. López Sáez, J. A. Riquelme Cantal, A. Sánchez Marco, F. G. Guzman, K. Brown, N. Fuentes, C. A. Valarino, A. Villalpando, C. B. Stringer, F. Martinez Ruiz & T. Sakamoto (2006) Late survival of Neanderthals at the southernmost extreme of Europe. *Nature,* 443**,** 850-3.

Forster, P. (2004) Ice Ages and the mitochondrial DNA chronology of human dispersals: a review. *Philos Trans R Soc Lond B Biol Sci,* 359**,** 255-64; discussion 264.

Germonpré, M., M. V. Sablin, R. E. Stevens, R. E. M. Hedges, M. Hofreiter, M. Stiller & V. R. Després (2009) Fossil dogs and wolves from Palaeolithic sites in Belgium, the Ukraine and Russia: osteometry, ancient DNA and stable isotopes. *Journal of Archaeological Science,* 36**,** 473-490.

Giuffra, E., J. M. Kijas, V. Amarger, O. Carlborg, J. T. Jeon & L. Andersson (2000) The origin of the domestic pig: independent domestication and subsequent introgression. *Genetics,* 154**,** 1785-91.

Goebel, T., M. R. Waters & D. H. O'Rourke (2008) The late Pleistocene dispersal of modern humans in the Americas. *Science,* 319**,** 1497-502.

Gollan, K. 1984. The Australian dingo: in the shadow of man. In *Vertebrate zoogeography and evolution in Australasia,* eds. M. Archer & G. Clayton, 921-927. Carlisle, Australia: Hesperian Press.

Gray, M. M., J. M. Granka, C. D. Bustamante, N. B. Sutter, A. R. Boyko, L. Zhu, E. A. Ostrander & R. K. Wayne (2009a) Linkage disequilibrium and demographic history of wild and domestic canids. *Genetics,* 181**,** 1493-505.

Gray, M. W., G. Burger & B. F. Lang (1999) Mitochondrial evolution. *Science,* 283**,** 1476-81.

Gray, R. D., A. J. Drummond & S. J. Greenhill (2009b) Language phylogenies reveal expansion pulses and pauses in Pacific settlement. *Science,* 323**,** 479-83.

Green, R. E., J. Krause, A. W. Briggs, T. Maricic, U. Stenzel, M. Kircher, N. Patterson, H. Li, W. Zhai, M. H. Fritz, N. F. Hansen, E. Y. Durand, A. S. Malaspinas, J. D. Jensen, T. Marques-Bonet, C. Alkan, K. Prüfer, M. Meyer, H. A. Burbano, J. M. Good, R. Schultz, A. Aximu-Petri, A. Butthof, B. Höber, B. Höffner, M. Siegemund, A. Weihmann, C. Nusbaum, E. S. Lander, C. Russ, N. Novod, J. Affourtit, M. Egholm, C. Verna, P. Rudan, D. Brajkovic, Z. Kucan, I. Gusic, V. B. Doronichev, L. V. Golovanova, C. Lalueza-Fox, M. de la Rasilla, J. Fortea, A. Rosas, R. W. Schmitz, P. L. Johnson, E. E. Eichler, D. Falush, E. Birney, J. C. Mullikin, M. Slatkin, R. Nielsen, J. Kelso, M. Lachmann, D. Reich & S. Pääbo (2010) A draft sequence of the Neandertal genome. *Science,* 328**,** 710-22.

Hagelberg, E., M. Kayser, M. Nagy, L. Roewer, H. Zimdahl, M. Krawczak, P. Lió & W. Schiefenhövel (1999) Molecular genetic evidence for the human settlement of the Pacific: analysis of mitochondrial DNA, Y chromosome and HLA markers. *Philos Trans R Soc Lond B Biol Sci,* 354**,** 141-52.

Helena Mangs, A. & B. J. Morris (2007) The Human Pseudoautosomal Region (PAR): Origin, Function and Future. *Curr Genomics,* 8**,** 129-36.

Heptner, V. G., N. P. Naumov, P. B. Yurgenson, A. A. Sludskii, A. F. Chirkova & A. G. Bannikov. 1967. *Mammals of the Soviet Union*. Moscow: Vysshaya Shkola Publishers.

Hiendleder, S., B. Kaupe, R. Wassmuth & A. Janke (2002) Molecular analysis of wild and domestic sheep questions current nomenclature and provides evidence for domestication from two different subspecies. *Proc Biol Sci,* 269**,** 893-904.

Higham, C. F. W. (1996) A review of archaeology in Mainland Southeast Asia. *J Arch Res,* 7**,** 149-165.

Ho, S. Y. & G. Larson (2006) Molecular clocks: when times are a-changin'. *Trends Genet,* 22**,** 79-83.

Hofreiter, M., D. Serre, H. N. Poinar, M. Kuch & S. Pääbo (2001) Ancient DNA. *Nat Rev Genet,* 2**,** 353-9.

Hughes, J. F. & S. Rozen (2012) Genomics and Genetics of Human and Primate Y Chromosomes. *Annu Rev Genomics Hum Genet*.

Hughes, J. F., H. Skaletsky, L. G. Brown, T. Pyntikova, T. Graves, R. S. Fulton, S. Dugan, Y. Ding, C. J. Buhay, C. Kremitzki, Q. Wang, H. Shen, M. Holder, D. Villasana, L. V. Nazareth, A. Cree, L. Courtney, J. Veizer, H. Kotkiewicz, T. J. Cho, N. Koutseva, S. Rozen, D. M. Muzny, W. C. Warren, R. A. Gibbs, R. K. Wilson & D. C. Page (2012) Strict evolutionary conservation followed rapid gene loss on human and rhesus Y chromosomes. *Nature,* 483**,** 82-6.

Hughes, J. F., H. Skaletsky, T. Pyntikova, T. A. Graves, S. K. van Daalen, P. J. Minx, R. S. Fulton, S. D. McGrath, D. P. Locke, C. Friedman, B. J. Trask, E. R. Mardis, W. C. Warren, S. Repping, S. Rozen, R. K. Wilson & D. C. Page (2010) Chimpanzee and human Y chromosomes are remarkably divergent in structure and gene content. *Nature,* 463**,** 536-9.

Hughes, J. F., H. Skaletsky, T. Pyntikova, P. J. Minx, T. Graves, S. Rozen, R. K. Wilson & D. C. Page (2005) Conservation of Y-linked genes during human evolution revealed by comparative sequencing in chimpanzee. *Nature,* 437**,** 100-3.

Hultin, E., M. Käller, A. Ahmadian & J. Lundeberg (2005) Competitive enzymatic reaction to control allele-specific extensions. *Nucleic Acids Res,* 33**,** e48.

Hurles, M. E., E. Matisoo-Smith, R. D. Gray & D. Penny (2003) Untangling Oceanic settlement: the edge of the knowable. *Trends in Ecology & Evolution,* 18**,** 531-540.

Hurles, M. E., J. Nicholson, E. Bosch, C. Renfrew, B. C. Sykes & M. A. Jobling (2002) Y chromosomal evidence for the origins of Oceanic-speaking peoples. *Genetics,* 160**,** 289-303.

Hurles, M. E., B. C. Sykes, M. A. Jobling & P. Forster (2005) The dual origin of the Malagasy in Island Southeast Asia and East Africa: evidence from maternal and paternal lineages. *Am J Hum Genet,* 76**,** 894-901.

Iacolina, L., M. Scandura, A. Gazzola, N. Cappai, C. Capitani, L. Mattioli, F. Vercillo & M. Apollonio (2010) Y-chromosome microsatellite variation in Italian wolves: A contribution to the study of wolf-dog hybridization patterns. *Mammalian Biology - Zeitschrift für Säugetierkunde,* 75**,** 341-347.

Ingman, M., H. Kaessmann, S. Pääbo & U. Gyllensten (2000) Mitochondrial genome variation and the origin of modern humans. *Nature,* 408**,** 708-13.

Ishiguro, N., Y. Inoshima & N. Shigehara (2009) Mitochondrial DNA analysis of the Japanese wolf (Canis lupus hodophilax Temminck, 1839) and comparison with representative wolf and domestic dog haplotypes. *Zoolog Sci,* 26**,** 765-70.

Jansen, T., P. Forster, M. A. Levine, H. Oelke, M. Hurles, C. Renfrew, J. Weber & K. Olek (2002) Mitochondrial DNA and the origins of the domestic horse. *Proc Natl Acad Sci U S A,* 99**,** 10905-10.

Jobling, M. A., M. E. Hurles & C. Tyler-Smith. 2004. *Human Evolutionary Genetics - Origins, Peoples & Disease*. New York: Garland Publishing.

Jobling, M. A. & C. Tyler-Smith (2003) The human Y chromosome: an evolutionary marker comes of age. *Nat Rev Genet,* 4**,** 598-612.

Kayser, M., S. Brauer, R. Cordaux, A. Casto, O. Lao, L. A. Zhivotovsky, C. Moyse-Faurie, R. B. Rutledge, W. Schiefenhoevel, D. Gil, A. A. Lin, P. A. Underhill, P. J. Oefner, R. J. Trent & M. Stoneking (2006) Melanesian and Asian origins of Polynesians: mtDNA and Y chromosome gradients across the Pacific. *Mol Biol Evol,* 23**,** 2234-44.

Kayser, M., S. Brauer, G. Weiss, P. A. Underhill, L. Roewer, W. Schiefenhövel & M. Stoneking (2000) Melanesian origin of Polynesian Y chromosomes. *Curr Biol,* 10**,** 1237-46.

Kayser, M., Y. Choi, M. van Oven, S. Mona, S. Brauer, R. J. Trent, D. Suarkia, W. Schiefenhövel & M. Stoneking (2008) The impact of the Austronesian expansion: evidence from mtDNA and Y chromosome diversity in the Admiralty Islands of Melanesia. *Mol Biol Evol,* 25**,** 1362-74.

Keller, A., A. Graefen, M. Ball, M. Matzas, V. Boisguerin, F. Maixner, P. Leidinger, C. Backes, R. Khairat, M. Forster, B. Stade, A. Franke, J. Mayer, J. Spangler, S. McLaughlin, M. Shah, C. Lee, T. T. Harkins, A. Sartori, A. Moreno-Estrada, B. Henn, M. Sikora, O. Semino, J. Chiaroni, S. Rootsi, N. M. Myres, V. M. Cabrera, P. A. Underhill, C. D. Bustamante, E. E. Vigl, M. Samadelli, G. Cipollini, J. Haas, H. Katus, B. D. O'Connor, M. R. Carlson, B. Meder, N. Blin, E. Meese, C. M. Pusch & A. Zink (2012) New insights into the Tyrolean Iceman's origin and phenotype as inferred by whole-genome sequencing. *Nat Commun,* 3**,** 698.

Kim, K. S., S. E. Lee, H. W. Jeong & J. H. Ha (1998) The complete nucleotide sequence of the domestic dog (Canis familiaris) mitochondrial genome. *Mol Phylogenet Evol,* 10**,** 210-20.

Kim, S. & A. Misra (2007) SNP genotyping: technologies and biomedical applications. *Annu Rev Biomed Eng,* 9**,** 289-320.

Kirch, P. V. & M. I. Weisler (1994) Archaeology in the Pacific Islands: an appraisal of recent research. *J Archaeol Res,* 2**,** 285-328.

Kirkness, E. F., V. Bafna, A. L. Halpern, S. Levy, K. Remington, D. B. Rusch, A. L. Delcher, M. Pop, W. Wang, C. M. Fraser & J. C. Venter (2003) The dog genome: survey sequencing and comparative analysis. *Science,* 301**,** 1898-903.

Klütsch, C. F., E. H. Seppälä, T. Fall, M. Uhlén, A. Hedhammar, H. Lohi & P. Savolainen (2011) Regional occurrence, high frequency but low diversity of mitochondrial DNA haplogroup d1 suggests a recent dog-wolf hybridization in Scandinavia. *Anim Genet,* 42**,** 100-3.

Klütsch, C. F. C. & P. Savolainen. 2011. Geographical Origin of the Domestic Dog. In *Encyclopedia of Life Sciences (ELS)*. Chichester: John Wiley & Sons, Ltd.

Koler-Matznick, J., I. Lehr Brisbin Jr, M. Feinstein & S. Bulmer (2003) An updated description of the New Guinea Singing Dog (*Canis hallstromi,* Troughton 1957). *J Zool Lond,* 261**,** 109-118.

Koop, B. F., M. Burbidge, A. Byun, U. Rink & S. J. Crockford. 2000. Ancient DNA evidence of a separate origin for North American indigenous dogs. In *Dogs Through Time: An Archaeological Perspective,* ed. S. J. Crockford, 271-285. Oxford: British archaeological reports.

Krings, M., H. Geisert, R. W. Schmitz, H. Krainitzki & S. Pääbo (1999) DNA sequence of the mitochondrial hypervariable region II from the neandertal type specimen. *Proc Natl Acad Sci U S A,* 96**,** 5581-5.

Krings, M., A. Stone, R. W. Schmitz, H. Krainitzki, M. Stoneking & S. Pääbo (1997) Neandertal DNA sequences and the origin of modern humans. *Cell,* 90**,** 19-30.

Kruglyak, S., R. T. Durrett, M. D. Schug & C. F. Aquadro (1998) Equilibrium distributions of microsatellite repeat length resulting from a balance between slippage events and point mutations. *Proc Natl Acad Sci U S A,* 95**,** 10774-8.

Larson, G., T. Cucchi, M. Fujita, E. Matisoo-Smith, J. Robins, A. Anderson, B. Rolett, M. Spriggs, G. Dolman, T. H. Kim, N. T. Thuy, E. Randi, M. Doherty, R. A. Due, R. Bollt, T. Djubiantono, B. Griffin, M. Intoh, E. Keane, P. Kirch, K. T. Li, M. Morwood, L. M. Pedriña, P. J. Piper, R. J. Rabett, P. Shooter, G. Van den Bergh, E. West, S. Wickler, J. Yuan, A. Cooper & K. Dobney (2007) Phylogeny and ancient DNA of Sus provides insights into neolithic expansion in Island Southeast Asia and Oceania. *Proc Natl Acad Sci U S A,* 104**,** 4834-9.

Larson, G., K. Dobney, U. Albarella, M. Fang, E. Matisoo-Smith, J. Robins, S. Lowden, H. Finlayson, T. Brand, E. Willerslev, P. Rowley-Conwy, L. Andersson & A. Cooper (2005) Worldwide phylogeography of wild boar reveals multiple centers of pig domestication. *Science,* 307**,** 1618-21.

Leonard, J. A., R. K. Wayne, J. Wheeler, R. Valadez, S. Guillén & C. Vilà (2002) Ancient DNA evidence for Old World origin of New World dogs. *Science,* 298**,** 1613-6.

Li, J. Z., D. M. Absher, H. Tang, A. M. Southwick, A. M. Casto, S. Ramachandran, H. M. Cann, G. S. Barsh, M. Feldman, L. L. Cavalli-Sforza & R. M. Myers (2008) Worldwide human relationships inferred from genome-wide patterns of variation. *Science,* 319**,** 1100-4.

Lin, L., J. Faraco, R. Li, H. Kadotani, W. Rogers, X. Lin, X. Qiu, P. J. de Jong, S. Nishino & E. Mignot (1999) The sleep disorder canine narcolepsy is caused by a mutation in the hypocretin (orexin) receptor 2 gene. *Cell,* 98**,** 365-76.

Lindblad-Toh, K., C. M. Wade, T. S. Mikkelsen, E. K. Karlsson, D. B. Jaffe, M. Kamal, M. Clamp, J. L. Chang, E. J. Kulbokas, M. C. Zody, E. Mauceli, X. Xie, M. Breen, R. K. Wayne, E. A. Ostrander, C. P. Ponting, F. Galibert, D. R. Smith, P. J. DeJong, E. Kirkness, P. Alvarez, T. Biagi, W. Brockman, J. Butler, C. W. Chin, A. Cook, J. Cuff, M. J. Daly, D. DeCaprio, S. Gnerre, M. Grabherr, M. Kellis, M. Kleber, C. Bardeleben, L. Goodstadt, A. Heger, C. Hitte, L. Kim, K. P. Koepfli, H. G. Parker, J. P. Pollinger, S. M. Searle, N. B. Sutter, R. Thomas, C. Webber, J. Baldwin, A. Abebe, A. Abouelleil, L. Aftuck, M. Ait-Zahra, T. Aldredge, N. Allen, P. An, S. Anderson, C. Antoine, H. Arachchi, A. Aslam, L. Ayotte, P. Bachantsang, A. Barry, T. Bayul, M. Benamara, A. Berlin, D. Bessette, B. Blitshteyn, T. Bloom, J. Blye, L. Boguslavskiy, C. Bonnet, B. Boukhgalter, A. Brown, P. Cahill, N. Calixte, J. Camarata, Y. Cheshatsang, J. Chu, M. Citroen, A. Collymore, P. Cooke, T. Dawoe, R. Daza, K. Decktor, S. DeGray, N. Dhargay, K. Dooley, P. Dorje, K. Dorjee, L. Dorris, N. Duffey, A. Dupes, O. Egbiremolen, R. Elong, J. Falk, A. Farina, S. Faro, D. Ferguson, P. Ferreira, S. Fisher, M. FitzGerald, K. Foley, et al. (2005) Genome sequence, comparative analysis and haplotype structure of the domestic dog. *Nature,* 438**,** 803-19.

Lindgren, G., N. Backström, J. Swinburne, L. Hellborg, A. Einarsson, K. Sandberg, G. Cothran, C. Vilà, M. Binns & H. Ellegren (2004) Limited number of patrilines in horse domestication. *Nat Genet,* 36**,** 335-6.

Lippold, S., M. Knapp, T. Kuznetsova, J. A. Leonard, N. Benecke, A. Ludwig, M. Rasmussen, A. Cooper, J. Weinstock, E. Willerslev, B. Shapiro & M. Hofreiter (2011) Discovery of lost diversity of paternal horse lineages using ancient DNA. *Nat Commun,* 2**,** 450.

Liu, Y. P., G. S. Wu, Y. G. Yao, Y. W. Miao, G. Luikart, M. Baig, A. Beja-Pereira, Z. L. Ding, M. G. Palanichamy & Y. P. Zhang (2006) Multiple maternal origins of chickens: out of the Asian jungles. *Mol Phylogenet Evol,* 38**,** 12-9.

Luikart, G., L. Gielly, L. Excoffier, J. D. Vigne, J. Bouvet & P. Taberlet (2001) Multiple maternal origins and weak phylogeographic structure in domestic goats. *Proc Natl Acad Sci U S A,* 98**,** 5927-32.

Macaulay, V., C. Hill, A. Achilli, C. Rengo, D. Clarke, W. Meehan, J. Blackburn, O. Semino, R. Scozzari, F. Cruciani, A. Taha, N. K. Shaari, J. M. Raja, P. Ismail, Z. Zainuddin, W. Goodwin, D. Bulbeck, H. J. Bandelt, S. Oppenheimer, A. Torroni & M. Richards (2005) Single, rapid coastal settlement of Asia revealed by analysis of complete mitochondrial genomes. *Science,* 308**,** 1034-6.

MacPhee, R. D. E. & D. A. Burney (1991) Dating of modified femora of extinct dwarf Hippopotamus from Southern Madagascar: Implications for constraining human colonization and vertebrate extinction events. *J Archaeol Sci,* 18**,** 695-706.

Malmström, H., C. Vilà, M. T. Gilbert, J. Storå, E. Willerslev, G. Holmlund & A. Götherström (2008) Barking up the wrong tree: modern northern European dogs fail to explain their origin. *BMC Evol Biol,* 8**,** 71.

Margulis, L. 1981. *Symbiosis in Cell Evolution*. San Francisco: Freeman.

Matisoo-Smith, E. 2007. Animal translocations, genetic variation and the human settlement of the Pacific. In *Genes, language and culture history in the southwest Pacific,* ed. J. S. Friedlaender, 157-170. Oxford, UK: Oxford University Press.

Matisoo-Smith, E. & J. H. Robins (2004) Origins and dispersals of Pacific peoples: evidence from mtDNA phylogenies of the Pacific rat. *Proc Natl Acad Sci U S A,* 101**,** 9167-72.

Meadows, J. R., I. Cemal, O. Karaca, E. Gootwine & J. W. Kijas (2007) Five ovine mitochondrial lineages identified from sheep breeds of the near East. *Genetics,* 175**,** 1371-9.

Messier, W., S. H. Li & C. B. Stewart (1996) The birth of microsatellites. *Nature,* 381**,** 483.

Meyer, M., U. Stenzel & M. Hofreiter (2008) Parallel tagged sequencing on the 454 platform. *Nat Protoc,* 3**,** 267-78.

Milham, P. & P. Thompson (1976) Relative antiquity of human occupation and extinct fauna at Madura cave, Southeastern Western Australia. *Mankind,* 10**,** 175-180.

Miller, W., D. I. Drautz, A. Ratan, B. Pusey, J. Qi, A. M. Lesk, L. P. Tomsho, M. D. Packard, F. Zhao, A. Sher, A. Tikhonov, B. Raney, N. Patterson, K. Lindblad-Toh, E. S. Lander, J. R. Knight, G. P. Irzyk, K. M. Fredrikson, T. T. Harkins, S. Sheridan, T. Pringle & S. C. Schuster (2008) Sequencing the nuclear genome of the extinct woolly mammoth. *Nature,* 456**,** 387-90.

Mishmar, D., E. Ruiz-Pesini, M. Brandon & D. C. Wallace (2004) Mitochondrial DNA-like sequences in the nucleus (NUMTs): insights into our African origins and the mechanism of foreign DNA integration. *Hum Mutat,* 23**,** 125-33.

Morey, D. F. & M. D. Wiant (1992) Early Holocene Domestic Dog Burials From the North American Midwest. *Current Anthropology,* 33**,** 224-229.

Morin, P. A., G. Luikart, R. K. Wayne & S. N. P. w. g. the (2004) SNPs in ecology, evolution and conservation. *Trends in Ecology & Evolution,* 19**,** 208-216.

Mullis, K. B. & F. A. Faloona (1987) Specific synthesis of DNA in vitro via a polymerase-catalyzed chain reaction. *Methods Enzymol,* 155**,** 335-50.

Musil, R. 2000. Evidence for the domestication of wolves in central European Magdalenian sites. In *Dogs through Time: An Archaeological Perspective,* ed. S. J. Crockford, 21-28. Oxford: British archaeological reports.

Naderi, S., H. R. Rezaei, F. Pompanon, M. G. Blum, R. Negrini, H. R. Naghash, O. Balkiz, M. Mashkour, O. E. Gaggiotti, P. Ajmone-Marsan, A. Kence, J. D. Vigne & P. Taberlet (2008) The goat domestication process inferred from large-scale mitochondrial DNA analysis of wild and domestic individuals. *Proc Natl Acad Sci U S A,* 105**,** 17659-64.

Naderi, S., H. R. Rezaei, P. Taberlet, S. Zundel, S. A. Rafat, H. R. Naghash, M. A. el-Barody, O. Ertugrul, F. Pompanon & E. Consortium (2007) Large-scale mitochondrial DNA analysis of the domestic goat reveals six haplogroups with high diversity. *PLoS One,* 2**,** e1012.

Napierala, H. & H.-P. Uerpmann (2010) A ´new`palaeolithic dog from Central Europe. *International Journal of Osteoarchaeology,* 22**,** 127-137.

Neiman, M., S. Lundin, P. Savolainen & A. Ahmadian (2011) Decoding a substantial set of samples in parallel by massive sequencing. *PLoS One,* 6**,** e17785.

Nobis, G. (1979) Der älteste Haushund lebte vor 14000 Jahren. *Umschau,* 79**,** 610.

Nowak, R. M. 1996. Another look at wolf taxonomy. In *Ecology and Behavior of Wolves in a Changing World,* eds. L. N. Carbyn, S. H. Fritts & D. R. Seip. Edmonton: Canadian Circumpolar Institute.

---. 2003. Wolf evolution and taxonomy. In *Wolves: Behavior, Ecology, and Conservation,* eds. L. D. Mech & L. Boitani, 239-258. Illinois: University of Chicago Press.

Ohta, T. & M. Kimura (1973) A model of mutation appropriate to estimate the number of electrophorectically detectable alleles in a finite population. *Genet Res,* 22**,** 201-204.

Okumura, N., N. Ishiguro, M. Nakano, A. Matsui & M. Sahara (1996) Intra- and interbreed genetic variations of mitochondrial DNA major non-coding regions in Japanese native dog breeds (Canis familiaris). *Anim Genet,* 27**,** 397-405.

Olivier, M. & G. Lust (1998) Two DNA sequences specific for the canine Y chromosome. *Anim Genet,* 29**,** 146-9.

Olsen, S. J. & J. W. Olsen (1977) The Chinese wolf, ancestor of New World dogs. *Science,* 197**,** 533-535.

Olson, L. E. & A. D. Yoder (2002) Using secondary structure to identify ribosomal numts: cautionary examples from the human genome. *Mol Biol Evol,* 19**,** 93-100.

Oppenheimer, S. (2012) Out-of-Africa, the peopling of continents and islands: tracing uniparental gene trees across the map. *Philos Trans R Soc Lond B Biol Sci,* 367**,** 770-84.

Oppenheimer, S. J. & M. Richards (2001) Polynesian origins. Slow boat to Melanesia? *Nature,* 410**,** 166-7.

Ovodov, N. D., S. J. Crockford, Y. V. Kuzmin, T. F. Higham, G. W. Hodgins & J. van der Plicht (2011) A 33,000-year-old incipient dog from the Altai Mountains of Siberia: evidence of the earliest domestication disrupted by the Last Glacial Maximum. *PLoS One,* 6**,** e22821.

Page, D. C., M. E. Harper, J. Love & D. Botstein (1984) Occurrence of a transposition from the X-chromosome long arm to the Y-chromosome short arm during human evolution. *Nature,* 311**,** 119-23.

Pakendorf, B. & M. Stoneking (2005) Mitochondrial DNA and human evolution. *Annu Rev Genomics Hum Genet,* 6**,** 165-83.

Pang, J. F., C. Kluetsch, X. J. Zou, A. B. Zhang, L. Y. Luo, H. Angleby, A. Ardalan, C. Ekström, A. Sköllermo, J. Lundeberg, S. Matsumura, T. Leitner, Y. P. Zhang & P. Savolainen (2009) mtDNA data indicate a single origin for dogs south of Yangtze River, less than 16,300 years ago, from numerous wolves. *Mol Biol Evol,* 26**,** 2849-64.

Parker, H. G., A. L. Shearin & E. A. Ostrander (2010) Man's best friend becomes biology's best in show: genome analyses in the domestic dog. *Annu Rev Genet,* 44**,** 309-36.

Pedrosa, S., M. Uzun, J. J. Arranz, B. Gutiérrez-Gil, F. San Primitivo & Y. Bayón (2005) Evidence of three maternal lineages in Near Eastern sheep supporting multiple domestication events. *Proc Biol Sci,* 272**,** 2211-7.

Pionnier-Capitan, M., C. Bemilli, P. Bodu, G. Célérier, J.-G. Ferrié, P. Fosse, M. Garcià & J.-D. Vigne (2011) New evidence for Upper Palaeolithic small domestic dogs in South-Western Europe. *Journal of Archaeological Science,* 38**,** 2123-2140.

Pires, A.E., L. Ouragh, M. Kalboussi, J. Matos, F. Petrucci-Fonseca & M.W. Bruford (2006) Mitochondrial DNA Sequence Variation in Portuguese Native Dog Breeds: Diversity and Phylogenetic Affinities. *J Hered* **97**, 318-330.

Pujar, S., K. S. Kothapalli, E. Kirkness, R. H. Van Wormer & V. N. Meyers-Wallen (2005) Exclusion of Lhx9 as a candidate gene for Sry-negative XX sex reversal in the American cocker spaniel model. *J Hered,* 96**,** 452-4.

Raisor, M. J. 2005. *Determining the Antiquity of Dog Origins*. Oxford: British archaeological reports.

Redd, A. J., N. Takezaki, S. T. Sherry, S. T. McGarvey, A. S. Sofro & M. Stoneking (1995) Evolutionary history of the COII/tRNALys intergenic 9 base pair deletion in human mitochondrial DNAs from the Pacific. *Mol Biol Evol,* 12**,** 604-15.

Reed, F. A. & S. A. Tishkoff (2006) African human diversity, origins and migrations. *Curr Opin Genet Dev,* 16**,** 597-605.

Reich, D., R. E. Green, M. Kircher, J. Krause, N. Patterson, E. Y. Durand, B. Viola, A. W. Briggs, U. Stenzel, P. L. Johnson, T. Maricic, J. M. Good, T. Marques-

Bonet, C. Alkan, Q. Fu, S. Mallick, H. Li, M. Meyer, E. E. Eichler, M. Stoneking, M. Richards, S. Talamo, M. V. Shunkov, A. P. Derevianko, J. J. Hublin, J. Kelso, M. Slatkin & S. Pääbo (2010) Genetic history of an archaic hominin group from Denisova Cave in Siberia. *Nature,* 468**,** 1053-60.

Reich, D., N. Patterson, M. Kircher, F. Delfin, M. R. Nandineni, I. Pugach, A. M. Ko, Y. C. Ko, T. A. Jinam, M. E. Phipps, N. Saitou, A. Wollstein, M. Kayser, S. Pääbo & M. Stoneking (2011) Denisova admixture and the first modern human dispersals into Southeast Asia and Oceania. *Am J Hum Genet,* 89**,** 516-28.

Rozen, S., H. Skaletsky, J. D. Marszalek, P. J. Minx, H. S. Cordum, R. H. Waterston, R. K. Wilson & D. C. Page (2003) Abundant gene conversion between arms of palindromes in human and ape Y chromosomes. *Nature,* 423**,** 873-6.

Runstadler, J. A., J. M. Angles & N. C. Pedersen (2006) Dog leucocyte antigen class II diversity and relationships among indigenous dogs of the island nations of Indonesia (Bali), Australia and New Guinea. *Tissue Antigens,* 68**,** 418-426.

Ryan, M. T. & N. J. Hoogenraad (2007) Mitochondrial-nuclear communications. *Annu Rev Biochem,* 76**,** 701-22.

Sablin, M. V. & G. A. Khlopachev (2002) The earliest ice age dogs: evidence from Eliseevichi I. *Current Anthropology,* 43**,** 795-799.

Saiki, R. K., S. Scharf, F. Faloona, K. B. Mullis, G. T. Horn, H. A. Erlich & N. Arnheim (1985) Enzymatic amplification of beta-globin genomic sequences and restriction site analysis for diagnosis of sickle cell anemia. *Science,* 230**,** 1350-4.

Salmon Hillbertz, N. H., M. Isaksson, E. K. Karlsson, E. Hellmén, G. R. Pielberg, P. Savolainen, C. M. Wade, H. von Euler, U. Gustafson, A. Hedhammar, M. Nilsson, K. Lindblad-Toh, L. Andersson & G. Andersson (2007) Duplication of FGF3, FGF4, FGF19 and ORAOV1 causes hair ridge and predisposition to dermoid sinus in Ridgeback dogs. *Nat Genet,* 39**,** 1318-20.

Sanger, F. (1988) Sequences, sequences, and sequences. *Annu Rev Biochem,* 57**,** 1-28.

Sanger, F., S. Nicklen & A. R. Coulson (1977) DNA sequencing with chain-terminating inhibitors. *Proc Natl Acad Sci U S A,* 74**,** 5463-7.

Sato, M. & K. Sato (2011) Degradation of paternal mitochondria by fertilization-triggered autophagy in C. elegans embryos. *Science,* 334**,** 1141-4.

Savolainen, P., T. Leitner, A. N. Wilton, E. Matisoo-Smith & J. Lundeberg (2004) A detailed picture of the origin of the Australian dingo, obtained from the study of mitochondrial DNA. *Proc Natl Acad Sci U S A,* 101**,** 12387-90.

Savolainen, P., B. Rosén, A. Holmberg, T. Leitner, M. Uhlén & J. Lundeberg (1997) Sequence analysis of domestic dog mitochondrial DNA for forensic use. *J Forensic Sci,* 42**,** 593-600.

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

Schaffner, S. F. (2004) The X chromosome in population genetics. *Nat Rev Genet,* 5**,** 43-51.

Schlötterer, C. & D. Tautz (1992) Slippage synthesis of simple sequence DNA. *Nucleic Acids Res,* 20**,** 211-5.

Schuster, S. C., W. Miller, A. Ratan, L. P. Tomsho, B. Giardine, L. R. Kasson, R. S. Harris, D. C. Petersen, F. Zhao, J. Qi, C. Alkan, J. M. Kidd, Y. Sun, D. I. Drautz, P. Bouffard, D. M. Muzny, J. G. Reid, L. V. Nazareth, Q. Wang, R. Burhans, C. Riemer, N. E. Wittekindt, P. Moorjani, E. A. Tindall, C. G. Danko, W. S. Teo, A. M. Buboltz, Z. Zhang, Q. Ma, A. Oosthuysen, A. W. Steenkamp, H. Oostuisen, P. Venter, J. Gajewski, Y. Zhang, B. F. Pugh, K. D. Makova, A. Nekrutenko, E. R. Mardis, N. Patterson, T. H. Pringle, F. Chiaromonte, J. C. Mullikin, E. E. Eichler, R. C. Hardison, R. A. Gibbs, T. T. Harkins & V. M.

Hayes (2010) Complete Khoisan and Bantu genomes from southern Africa. *Nature,* 463**,** 943-7.

Schwartz, M. & J. Vissing (2002) Paternal inheritance of mitochondrial DNA. *N Engl J Med,* 347**,** 576-80.

Shitara, H., J. I. Hayashi, S. Takahama, H. Kaneda & H. Yonekawa (1998) Maternal inheritance of mouse mtDNA in interspecific hybrids: segregation of the leaked paternal mtDNA followed by the prevention of subsequent paternal leakage. *Genetics,* 148**,** 851-7.

Skaletsky, H., T. Kuroda-Kawaguchi, P. J. Minx, H. S. Cordum, L. Hillier, L. G. Brown, S. Repping, T. Pyntikova, J. Ali, T. Bieri, A. Chinwalla, A. Delehaunty, K. Delehaunty, H. Du, G. Fewell, L. Fulton, R. Fulton, T. Graves, S. F. Hou, P. Latrielle, S. Leonard, E. Mardis, R. Maupin, J. McPherson, T. Miner, W. Nash, C. Nguyen, P. Ozersky, K. Pepin, S. Rock, T. Rohlfing, K. Scott, B. Schultz, C. Strong, A. Tin-Wollam, S. P. Yang, R. H. Waterston, R. K. Wilson, S. Rozen & D. C. Page (2003) The male-specific region of the human Y chromosome is a mosaic of discrete sequence classes. *Nature,* 423**,** 825-37.

Skoglund, P., A. Götherström & M. Jakobsson (2011) Estimation of population divergence times from non-overlapping genomic sequences: examples from dogs and wolves. *Mol Biol Evol,* 28**,** 1505-17.

Skoglund, P. & M. Jakobsson (2011) Archaic human ancestry in East Asia. *Proc Natl Acad Sci U S A,* 108**,** 18301-6.

Soares, P., T. Rito, J. Trejaut, M. Mormina, C. Hill, E. Tinkler-Hundal, M. Braid, D. J. Clarke, J. H. Loo, N. Thomson, T. Denham, M. Donohue, V. Macaulay, M. Lin, S. Oppenheimer & M. B. Richards (2011) Ancient voyaging and Polynesian origins. *Am J Hum Genet,* 88**,** 239-47.

Soodyall, H., T. Jenkins & M. Stoneking (1995) 'Polynesian' mtDNA in the Malagasy. *Nat Genet,* 10**,** 377-8.

Stoneking, M. & J. Krause (2011) Learning about human population history from ancient and modern genomes. *Nat Rev Genet,* 12**,** 603-14.

Street, M. (2002) Ein Wiedersehen mit dem Hund von Bonn Oberkassel. *Bonner zoologische Beiträge,* 50**,** 269-290.

Sundqvist, A. K., S. Björnerfeldt, J. A. Leonard, F. Hailer, A. Hedhammar, H. Ellegren & C. Vilà (2006) Unequal contribution of sexes in the origin of dog breeds. *Genetics,* 172**,** 1121-8.

Sutter, N. B. & E. A. Ostrander (2004) Dog star rising: the canine genetic system. *Nat Rev Genet,* 5**,** 900-10.

Tapio, M., N. Marzanov, M. Ozerov, M. Cinkulov, G. Gonzarenko, T. Kiselyova, M. Murawski, H. Viinalass & J. Kantanen (2006) Sheep mitochondrial DNA variation in European, Caucasian, and Central Asian areas. *Mol Biol Evol,* 23**,** 1776-83.

Tchernov, E. & F. F. Valla (1997) Two new dogs, and other Natufian dogs, from the Southern Levant. *J Archaeol Sci,* 24**,** 65-95.

Tedford, R. H., X. Wang & B. E. Taylor (2009) Phylogenetic systematics of the North American fossil caninae (carnivore: canidae). *Bull Am Mus Nat Hist,* 325**,** 1-218.

Terrell, J. E. (1988) History as a family tree, history as an entangled bank: constructing images and interpretations of prehistory in the South Pacific. *Antiquity,* 62**,** 642-657.

Terrell, J. E., K. M. Kelly & P. Rainbird (2001) Foregone conclusions? In search of “Papuans” and “Austronesians”. *Curr Anthropol,* 42**,** 97-124.

Thomson, R., J. K. Pritchard, P. Shen, P. J. Oefner & M. W. Feldman (2000) Recent common ancestry of human Y chromosomes: evidence from DNA sequence data. *Proc Natl Acad Sci U S A,* 97**,** 7360-5.

Tixier-Boichard, M., B. Bed'hom & X. Rognon (2011) Chicken domestication: from archeology to genomics. *C R Biol,* 334**,** 197-204.

Tofanelli, S., S. Bertoncini, L. Castrì, D. Luiselli, F. Calafell, G. Donati & G. Paoli (2009) On the origins and admixture of Malagasy: new evidence from high-resolution analyses of paternal and maternal lineages. *Mol Biol Evol,* 26**,** 2109-24.

Torroni, A., R. I. Sukernik, T. G. Schurr, Y. B. Starikorskaya, M. F. Cabell, M. H. Crawford, A. G. Comuzzie & D. C. Wallace (1993) mtDNA variation of aboriginal Siberians reveals distinct genetic affinities with Native Americans. *Am J Hum Genet,* 53**,** 591-608.

Troy, C. S., D. E. MacHugh, J. F. Bailey, D. A. Magee, R. T. Loftus, P. Cunningham, A. T. Chamberlain, B. C. Sykes & D. G. Bradley (2001) Genetic evidence for Near-Eastern origins of European cattle. *Nature,* 410**,** 1088-91.

Trut, L., I. Oskina & A. Kharlamova (2009) Animal evolution during domestication: the domesticated fox as a model. *Bioessays,* 31**,** 349-360.

Trut, L. N., I. Z. Plyusnina & I. N. Oskina (2004) An experiment on fox domestication and debatable issues of evolution of the dog. *Russ J Genet,* 40**,** 644-655.

Tsuda, K., Y. Kikkawa, H. Yonekawa & Y. Tanabe (1997) Extensive interbreeding occurred among multiple matriarchal ancestors during the domestication of dogs: evidence from inter- and intraspecies polymorphisms in the D-loop region of mitochondrial DNA between dogs and wolves. *Genes Genet Syst,* 72**,** 229-38.

Underhill, A. (1997) Current Issues in Chinese Neolithic archaeology. *J World Prehistory,* 11**,** 103-160.

Underhill, P. A. & T. Kivisild (2007) Use of y chromosome and mitochondrial DNA population structure in tracing human migrations. *Annu Rev Genet,* 41**,** 539-64.

Underhill, P. A., P. Shen, A. A. Lin, L. Jin, G. Passarino, W. H. Yang, E. Kauffman, B. Bonné-Tamir, J. Bertranpetit, P. Francalacci, M. Ibrahim, T. Jenkins, J. R. Kidd, S. Q. Mehdi, M. T. Seielstad, R. S. Wells, A. Piazza, R. W. Davis, M. W. Feldman, L. L. Cavalli-Sforza & P. J. Oefner (2000) Y chromosome sequence variation and the history of human populations. *Nat Genet,* 26**,** 358-61.

Vigilant, L., M. Stoneking, H. Harpending, K. Hawkes & A. C. Wilson (1991) African populations and the evolution of human mitochondrial DNA. *Science,* 253**,** 1503-7.

Vilà, C., J. A. Leonard, A. Gotherstrom, S. Marklund, K. Sandberg, K. Liden, R. K. Wayne & H. Ellegren (2001) Widespread origins of domestic horse lineages. *Science,* 291**,** 474-7.

Vilà, C., P. Savolainen, J. E. Maldonado, I. R. Amorim, J. E. Rice, R. L. Honeycutt, K. A. Crandall, J. Lundeberg & R. K. Wayne (1997) Multiple and ancient origins of the domestic dog. *Science,* 276**,** 1687-9.

Vilà, C., J. Seddon & H. Ellegren (2005) Genes of domestic mammals augmented by backcrossing with wild ancestors. *Trends Genet,* 21**,** 214-8.

Vilà, C., C. Walker, A. K. Sundqvist, Ø. Flagstad, Z. Andersone, A. Casulli, I. Kojola, H. Valdmann, J. Halverson & H. Ellegren (2003) Combined use of maternal, paternal and bi-parental genetic markers for the identification of wolf-dog hybrids. *Heredity (Edinb),* 90**,** 17-24.

Vonholdt, B. M., J. P. Pollinger, K. E. Lohmueller, E. Han, H. G. Parker, P. Quignon, J. D. Degenhardt, A. R. Boyko, D. A. Earl, A. Auton, A. Reynolds, K. Bryc, A. Brisbin, J. C. Knowles, D. S. Mosher, T. C. Spady, A. Elkahloun, E. Geffen, M.

Pilot, W. Jedrzejewski, C. Greco, E. Randi, D. Bannasch, A. Wilton, J. Shearman, M. Musiani, M. Cargill, P. G. Jones, Z. Qian, W. Huang, Z. L. Ding, Y. P. Zhang, C. D. Bustamante, E. A. Ostrander, J. Novembre & R. K. Wayne (2010) Genome-wide SNP and haplotype analyses reveal a rich history underlying dog domestication. *Nature,* 464**,** 898-902.

Wang, X. & R. H. Tedford. 2008. *Dogs, their Fossil Relatives and Evolutionary History*. New York: Columbia University Press.

Watson, E., P. Forster, M. Richards & H. J. Bandelt (1997) Mitochondrial footprints of human expansions in Africa. *Am J Hum Genet,* 61**,** 691-704.

Wayne, R. K. (1993) Molecular evolution of the dog family. *Trends Genet,* 9**,** 218-24.

White, D. J., J. N. Wolff, M. Pierson & N. J. Gemmell (2008) Revealing the hidden complexities of mtDNA inheritance. *Mol Ecol,* 17**,** 4925-42.

Xiao, W. & P. J. Oefner (2001) Denaturing high-performance liquid chromatography: A review. *Hum Mutat,* 17**,** 439-74.

Xue, Y., Q. Wang, Q. Long, B. L. Ng, H. Swerdlow, J. Burton, C. Skuce, R. Taylor, Z. Abdellah, Y. Zhao, D. G. MacArthur, M. A. Quail, N. P. Carter, H. Yang, C. Tyler-Smith & Asan (2009) Human Y chromosome base-substitution mutation rate measured by direct sequencing in a deep-rooting pedigree. *Curr Biol,* 19**,** 1453-7.

Zeder, M. A. (2008) Domestication and early agriculture in the Mediterranean Basin: Origins, diffusion, and impact. *Proc Natl Acad Sci U S A,* 105**,** 11597-604.

Zhu, Y., J. E. Strassmann & D. C. Queller (2000) Insertions, substitutions, and the origin of microsatellites. *Genet Res,* 76**,** 227-36.