Toward understanding dog evolutionary and domestication history

Histoire de la domestication du chien

Francis Galibert *, Pascale Quignon, Christophe Hitte, Catherine André

UMR 6061 CNRS, institut de génétique et développement de Rennes (IGDR), faculté de médecine, université de Rennes 1, 35000 Rennes, France

1. Introduction

Among the many species of plants and animals that have been domesticated, the dog holds the distinction of being the first. Domestication is a long process during which humans, by selecting animals for future crosses, are able to modify a number of physiological and behavioral traits, generations after generations. As such, domestication implies a vision or a plan from humans and differs from training or taming. Here we will describe the process of dog domestication and its main consequences, which differ by several aspects to other animal domestications.

ARTICLE INFO

Article history:
Available online 1 February 2011

Keywords:
Dog
Domestication
Breed
Genome
Evolution

ABSTRACT

Dog domestication has probably started very early during the Upper paleolithic period (~35,000 BP), thus well before any other animal or plant domestication. This early process, probably unconscious, is called proto-domestication to distinguish it from the real domestication process that has been dated around 14,000 BC. Genomic DNA analyses have shown recently that domestication started in the Middle East and rapidly expanded into all human populations. Nowadays, the dog population is fragmented in several hundreds of breeds well characterized by their phenotypes that offer a unique spectrum of polymorphism. More recent studies detect genetic signatures that will be useful to highlight breed history as well as the impact of domestication at the DNA level.

© 2011 Académie des sciences. Published by Elsevier Masson SAS. All rights reserved.

R É S U M É

Le chien est de loin la première espèce à avoir été domestiquée. Cette domestication a eu lieu de façon probablement inconsciente au Paléolithique supérieur (~35,000 BP) par des populations humaines vivant en contact plus ou moins rapproché avec des loups. Par référence à l’effort de domestication dont les premières traces apparaissent aux alentours de 14,000 ans avant notre ère, cette première période est qualifiée de proto-domestication. Les sélections d’individus ayant des traits particuliers se sont amplifiées au cours des millénaires et plus particulièrement depuis le Moyen Âge. De nos jours, la population canine, très abondante dans toutes les sociétés humaines, est fragmentée en de très nombreuses races fortement individualisées par leurs phénotypes très caractéristiques. Les analyses moléculaires d’abord réalisées sur l’ADN mitochondrial, puis sur l’ADN génomique ont permis de retracer l’origine de la domestication au Moyen-Orient. Ces mêmes analyses permettent actuellement de rechercher des signatures génétiques qui permettront de retracer l’histoire complexe des races et l’impact moléculaire de la domestication.

2. Dog history and archaeological data

The dog family or canidae is a group of carnivore mammals divided into 38 species [1]. Among them the dog, Canis familiaris, is the only member to have been domesticated, although the red fox and the raccoon have been bred for their fur. Until rather recently, there have been many debates concerning dog ancestor(s). Considering the huge phenotypic diversity, the main point of discussion was whether a single wild species or several were at the origin of today’s dogs, some believing the wolf was the unique ancestor and others the jackal or an unknown extinct species were also ancestors. Darwin himself suggested that dogs were derived from several species [2].

Today, the combined studies on behavior, vocalizations, morphology and more particularly molecular biology clearly show that wolf, Canis lupus is the principal if not the unique dog ancestor. One study showed that dog and wolf share 98% of their mitochondrial (mt) DNA, which contrasts to the 7.5% difference between wolf and coyotes (Canis latrans), the species that is their closest wild relative [3].

Bones of wolves and of early hominids have been found together at several locations such as Zhoukoundian in North China dated 300,000 years BP (before present) or the cave of Lazaret in the south of France dated 150,000 years BP. These associations do not demonstrate domestication was already en route but suggest that humans and wolves probably shared the same territories and lived in close contact. The oldest dog skull showing clear marks of differentiation from wolves was discovered in the Goyet cave, Belgium, during the mid-19th century. It was recently re-examined among a number of other fossil canids and dated by Accelerator Mass Spectrometry (AMS) at 31,700 years BP [4]. By all criteria analyzed, the large canid fossil is clearly different from the recent wolves, presumably resembling the prehistoric dogs. It is thus identified as a Paleolithic dog. The study examining the Goyet cave also identified what appear to be prehistoric dogs at the Chauvet Cave (~26,000 BP) and Mezirich in the Ukraine (~15,000 years BP). These results suggest that dog domestication has already started during the Aurignacian period. However, not everyone is convinced by this suggestion and some authors argued that, in these early days, hominids might not have the brain ability to project themselves in a complex program of domestication. Rather, as well explained by Clutton-Brock [1], one may consider that wolves in search of food were attracted by human camps where they could find some leftovers and thus became progressively accustomed to the human presence. Human hunters, also, probably killed wolves looking for food around their camps and used their skins for clothing. Sometimes they would carry around live pup that would occasionally become used to the family and be tamed. These tamed wolves were many generations away from the true domesticated dogs, but were undoubtedly their precursors. This theory of proto-domestication, as opposed to actual domestication as it happens in many other occasions, is also shared by Coppinger and Coppinger [5].

Stronger evidence of domestication could be found at a burial site in Germany called Bonn-Oberkassel, dated at 14,000 years BP. This site belongs to a cultural period named Epipaleolithic. Until that period, animals were killed by direct impact but in the Epipaleolithic, hunting strategies changed as humans started to use arrows armed with tiny stone blades. The success of this new strategy would have enhanced a new partnership between hunters and predomesticated dogs, which would help track and bring back wounded animals. In the context of this stronger human–dog relationship, the archaeological remains found at the site of Ein Mallaha dated at 12,000 years BP is of interest. Inhabitants of this site were hunter-gatherers, who buried their dead in stone-covered tombs. In one of these tombs, a human skeleton laying on its right side in a flexed position has its left hand on the thorax of a puppy. From the following prehistoric period (11,000–7000 years BP), a large number of dog remains were found in many parts of the world, among which the Danger Cave in Utah, the earliest case of dog burial in the Americas, at about 11,000 years ago. In Western Europe, remains dated at least 8500 years BP bring information about wolf domestication in Europe [6,7].

Thus, even if no precise date can be ascertained mainly because of the time length of the whole domestication process, the dog is by far the first domesticated animal. Its domestication predated the advent of agriculture at around 10,000 years BP and the domestication of other animals of agricultural values. One can also suggest that the domestication process of the dog served as a model for other animal domestication. As we reported, the beginning of domestication during Aurignacian was probably not an intentional action from humans, which explained this process might has covered a period lasting 15,000 years, between 30,000 and 15,000 years BP. Completely different situations and events characterize the domestication of sheep, horses or cattle. The biological process of domestication appears like natural evolution in that the parent animals become reproductively isolated from the wild population and constitute a small founder group that progressively increases in number and diverges genetically under a process of genetic drift. Over successive generations, this group of animals accumulates genetic and phenotypic differences in response to its novel human environment.

So far, no genomic data related to any dog remains have been produced. But with the improvement and cost decrease that characterize the recent genome sequencing techniques known under the generic term of next generation sequencing, it is reasonable to think that such data will soon become available. Already numerous archaeological samples have been collected by Vigne et al. (this issue) and are processed at the PalGene laboratory in ENS Lyon, France (http://igfl.ens-lyon.fr/PALGENE/plateforme-nationale-de-paleogenetique-pal-gene/). The sequencing of regions known to have been under positive selection during domestication would bring new pieces to the puzzle [8].

3. Phenotypic and behavioral diversity

Canid bones retrieved from archaeological sites have revealed very large differences in size in the prehistoric
dog population, which already demonstrated some phenotypic variability. However, distinctive breeds, as we intend this word nowadays, do not seem to have existed for very long. Breed creation is consubstantial with domestication and nowadays more than 400 dog breeds are recognized worldwide by the different dog societies like the American Kennel Club (AKC; http://www.akc.org/) in the United States, the Société Centrale Canine (SCC; http://www.scc.asso.fr/) in France or the Fédération Cynologique Internationale (FCI; http://www.fci.be/). These many different dog breeds were created along the centuries (Fig. 1). The Saluki is certainly one of the oldest breeds. They seem to have originated in the Fertile Crescent area of Mesopotamia from the Sumerian period around 7000–6000 years BP. Carvings from this period show what appear to be images of Saluki dogs working alongside humans. Many breeds were famous in Ancient Egypt (3,000 years BP) where they already had various functions and served for hunting, as police guards, in military actions or simply as household pets. The extraordinary increase of the number of breeds took place during the Medieval and Renaissance periods and the more recent centuries with the creation of hundreds of breeds, each of them responding to precise standards in term of coat color, body shape, size, stature. Humans have exerted a strong selective pressure on dogs. This led to the creation of breeds better adapted to purposes such as herding, guarding, hunting or just companion pets to please. All these breeds show a wild range of phenotypes unmatched by any other species. For example, St. Bernard and Great Dane are up to 100 times heavier than Chihuahua, and greyhound is eight times larger than Pekinese. Behavior and aptitude to perform various tasks were also a strong driving force in the creation of many breeds.

The ability to perform specific tasks was also a strong impetus in the creation of breeds and one is just fascinated by the differences that were developed by the different breeds. Keeshonds are excellent watchdogs, they bark loudly to alert their owners of intruders, but would not attack, contrary to German shepherds, which are excellent guard dogs. If one looks at herding dogs, large differences exist in the way they perform their tasks. Some breeds called heelers like the Australian Cattle dog typically nip at the heels of animals and keep pushing the animals forward, whereas other breeds, collectively named headers like Border Collie, get in front of the animals and use what is called strong eye to stare down the animals to turn or stop the animal’s movement [9]. Training cannot abolish or profoundly change these traits. It can enhance or decrease behavioral differences between breeds but one cannot train a retriever dog to behave like a livestock guard. The social character of the Labrador, which is an ideal

---

**Fig. 1.** Illustrations of dogs throughout Human history. Upper left corner: Renegade Canyon, California, about 4000 BP. Upper right corner: Dog of Fo, Chinese polychrome statuette, about 960–1000. Burchard gallery. Lower left corner: Artemis goddess, greek oenochoe. From Athens, about 450 BC, Louvres Museum. Lower right corner: “Très riches heures du duc de Berry, janvier”, details. Limbourg brothers, 1416.
companions for children, and the attacking and guarding attitude of the pit bull and the Rottweiler (ROT) are examples. This selective pressure has also profoundly affected the social-cognitive capability of dogs. A series of experiments showed that adult dogs, as well as puppies raised with little human contact, could read and interpret human signals significantly better than wolves or even chimpanzees [10,11]. These differences are deeply inscribed in the genetic make-up of the breeds and result from the selection, generations after generations, of many alleles through the selection of dogs having the most suitable or desired phenotype [5].

4. Dog evolutionary history and molecular data

Until recently, dog history has been studied at the molecular level mostly using mitochondrial DNA [12,13]. Initially, these data have suggested that wolves and dogs split into different species around 100,000 years ago [3]; but whether humans had been involved with that, remains unknown. Later studies have suggested that the canine species have been domesticated for ~15,000 years [14]. A recent mtDNA and nuclear DNA analysis suggests that the origin and location of dog domestication, long thought to be in East Asia, is in some doubt [15]. The authors showed that mitochondrial sequencing alone does not seem well suited to determine the timing and location of domestication. In their work, the authors analyzed modern day “village dogs” that almost certainly lived as human commensals that were not subjected to the same degree of intense artificial selection and closed breeding practices that characterize modern dog breeds. These ‘ancient’ dog populations developed genetic signatures characteristic of their geographic localizations. These signatures would persist in both modern day village dog populations that descend from these ancient populations. However, when dogs from modern breeds cross with village dogs, the observed genetic diversity is greater than one could expect. In their work, Boyko et al. [15] show that African village dogs contain mostly native genetic diversity in contrast to Asian dogs that include a large part of diversity coming from modern breeds. From this study, although the authors do not suggest that Africa is actually the site of dog domestication, they do suggest for the first time that the East Asian origin of dogs should be further scrutinized.

A very recent study led by Robert Wayne used genome-wide SNP and haplotype analyses to better understand dog diversification and domestication [8]. They identified the primary source of genetic diversity for domestic dogs using multiple approaches based on genetic patterns sharing between domestic dog and wolf populations sampled from 11 grey wolf populations globally distributed. Haplotype sharing between modern and ancient dog breeds with specific wolf populations from Europe, Middle East and China for 500-kilobases (kb) haplotype windows has been carried out. The Middle East and China have been previously implicated as centers for dog origin based on the archaeological record or mtDNA diversity [14,16]. Absence of haplotype sharing between dog breeds and North American wolves has been assessed as a negative control, as dogs did not originate there but crossed the Bering Strait with humans [13].

Across all breeds, levels of sharing between dogs and North American wolves were substantially lower than levels of sharing between dogs and Old World wolves, as expected. The majority of breeds, including some dog breeds of diverse geographic origins (for example, basenji, Chihuahua, basset hound and borzoi) has the highest level of sharing with Middle Eastern wolves. The basenji, a breed of Middle Eastern origin, had a greater proportion of shared haplotypes with Middle Eastern wolves than any other domestic dog. This result suggests that basenjis had a larger effective population size early in domestication or that they have more recently backcrossed with wolves. In addition, significant sharing with European wolves was found in miniature pinschers, Staffordshire bull terriers, greyhounds and whippets. Overall, these results implicate the Middle East as a primary source of genetic variations in the dog, and with potential secondary origins from Europe and East Asia. It is important to remind that these results are contrasted with the mtDNA results, for which East Asian wolves are a predominant source of haplotype diversity for only a few East Asian dog breeds that have a long history in that region.

5. Genetic signatures of recent selection between canine breeds

Beyond the initial event of domestication, modern dog breed radiation is another event that is difficult to resolve because most have originated recently and still lack deep, detailed molecular analyses. The canine species is today composed of ~400 distinct breeds that result from intense artificial population statistics that aim to detect differentiation of allelic frequencies, selective sweeps and extended haplotypes in the vicinity of the allele selected by the domestication process. It thus leads to a distorted pattern of genetic variation that indicate such recent selection events and the underlying functional mutations mostly remain unknown. The evolutionary process underlying the genesis of functional, behavioral or fixed phenotypes within breeds is obscure although parallel evolution and genetic homogeneity within phenotypic, functional or behavioral breed groupings is expected. We, and others, are analyzing the distinct genetic signatures within modern dogs that reflect artificial selection, mostly occurring from breeding practices. Identifying patterns of genetic variation that indicate recent selection events of the canine species is based on genetic population statistics that aim to detect differentiation of allelic frequencies, selective sweeps and extended haplotypes in the vicinity of the allele selected by the domestication process. It thus leads to a distorted pattern of genetic variation that can be useful for detecting selection. The genome-wide scans used to detect selection signals have largely been based on searching for a distortion in the allele frequency spectrum or haplotype structure in a single population [17]. It is now feasible to apply such statistics to genome-wide scans using genotyped SNPs from 50 to 170K available SNPs arrays. In the latest Illumina Infinium CanineSNP170 array, SNPs are distributed evenly throughout the genome, at a mean density of one SNP per 15 kb. Sets of unrelated purebred dogs from as many as possible distinct breeds can be
genotyped on these arrays and the data analyzed to detect selection signatures.

The strategy developed is based on a population genetic statistical approach that measures the allelic differentiation level between populations in order to detect polymorphism pattern specific to one population. The method is based on the variance of allelic frequency using Wright’s Fixation index ($F_{st}$) statistic. For each SNP, the $F_{st}$ index is calculated for all pairwise breed combinations. To compare each $F_{st}$ value, a statistical value ‘di’ can be derived [16]. It is a function of standard variation of Wright’s Fixation index ($F$) method is based on the variance of allelic frequency using polymorphism pattern specific to one population. The statistical approach that measures the allelic differentiation signatures.

The strategy developed is based on a population genetic statistical approach that measures the allelic differentiation level between populations in order to detect polymorphism pattern specific to one population. The method is based on the variance of allelic frequency using Wright’s Fixation index ($F_{st}$) statistic. For each SNP, the $F_{st}$ index is calculated for all pairwise breed combinations. To compare each $F_{st}$ value, a statistical value ‘di’ can be derived [16]. It is a function of values of $F_{st}$ by pair between a breed ‘i’ and the totality of the other breeds as: $di = \sum_{j} (\sigma_{ij}^{2} - E(\sigma_{ij}^{2})) / \sigma_{ij}^{2}$ where $E(\sigma_{ij}^{2})$ and $\sigma_{ij}^{2}$ represent the expected value and the standard variation of $F_{st}$ between breeds $i$ and $j$ computed from the SNPs, respectively.

We have the goal to establish a highly resolutive map of polymorphism patterns which could represent recent signatures of selection of the canine genome that pinpoint new functional candidates to account for most differentiated phenotypes in purebred dogs. Loci previously published that defined canine breed differentiation will be used as a set of internal controls. As an example, we detected a short region of 187 kb on chromosome 13 that has been previously found to differentiate specifically the Shar-pei breed from all others [18]. This region contains a single gene – HAS2 – that is associated to the ‘skin wrinkling’ phenotype of Shar-pees. We also pinpointed short regions that both include the RSPO2 and KRT77 genes that are associated with the furnished and curly phenotypes of dog breeds [19]. Other loci can also be used to validate our findings such as the IGFI gene that is associated with the ‘small size’ phenotype or the FGFR retrogene associated to chondrodysplasia [20]. The functional analysis of the artificial selection signatures will be analyzed to pinpoint whether specifically pathways and genes are targeted by selection. To test the underlying assumption that genes targeted by selection may be implicated in particular biological pathways, we will examine their functional annotation in the context of Gene Ontology (GO). To focus on genes that are most likely to be targeted by artificial selection, a possibility will be to consider regions that pinpoint one coding gene only.

6. Dog breed population structure

In addition to detecting selection signatures, analyses were performed to unravel the genetic relationship between all the breeds. Indeed, before these analyses, dog breeds were generally grouped on the basis of their roles that human created them for. The first study by Parker et al. used 96 microsatellites that were genotyped on a collection of purebred dogs representing 85 breeds [21]. They showed that breed membership could be determined from individual dog genotypes, confirming the intra-breed homogeneity. Using a clustering method, they were able to determine clusters of breeds that are related, even if a precise phylogeny of all these breeds cannot be reliably determined. The first group of breeds that separate from the others is from ancient origins and contains breeds like the Shiba Inu, Chow chow, Afghan hound, or the Saluki. The 14 breeds composing this first group all date from antiquity. The breeds composing the second group that separates from the others all have the Mastiff appearance. The third group includes herding breeds like Collie or Shetland sheepdog. The remaining breeds of the last group are primarily of recent European origins and are mainly hunting dogs like scent hounds, terriers or pointers. This classification was later confirmed and narrowed using the same method and 132 breeds [22].

In addition, a fifth group was identified and contains large mountain dogs like the Bernese Mountain dog (BMD) and some spaniels. The most recent study used SNPs that were genotyped using the Affymetrix SNPs array [8]. Thus, 48,000 SNPs were genotyped on 912 dogs from 85 breeds. Phylogenetic trees based on these data are consistent with the previous findings, i.e. one group of ancient breeds that can be divided in three: an Asian group, a Middle Eastern group and a northern group. However, even with this large amount of data, the radiation of the different modern breeds is still difficult to resolve. In addition, using an analysis of molecular variation, they showed that 65% of the variation is due to variation within dog breeds, and 31% is due to variation within breed groups. Genetic variations between populations can be important in the design of genome wide association studies (GWAS), where population structure can increase the number of false positive. In these past years, the dog proved to be a powerful model using GWAS to highlight genes involved in morphological phenotypes [19,20,23–25]. Mapping by association requires a population-based sample as opposed to linkage mapping approaches that need multi-generational family samples. Since only a limited number of markers such as SNPs can be genotyped, the feasibility and the efficiency of GWAS are dependent on the level of linkage disequilibrium (LD) that exists in the studied species. SNPs are then used to detect correlation between surrounding regions of a locus through the non-random association between markers because of LD. In the canine species LD extends over relatively long distances [26,27] and comparatively to human, fewer markers have to be typed, as each marker allows to infer about relatively large segments of flanking DNA. Indeed, LD in dog breeds ranges from 0.4 to over 3 Mb, which is dramatically longer than the ~100 kb range reported for human populations [28]. Such LD pattern makes GWAS more straightforward in dog than in human genetics, using array with only 50–150K informative SNPs.

In addition to LD pattern, many dog breeds are characterized by reduced genetic diversity related to small numbers of founders, whose allelic pool is over represented in all subsequent generations. As a result, while phenotypic variation across breeds is enormous, within breed phenotypic variation, and thus variation at the DNA level, is considerably more limited than in human populations. Using a single breed to investigate a trait or a disease is based on the assumption that all dogs of a single breed share the same level and type of genetic variation. However, the accuracy of association studies can be jeopardized if the over representation of a rare allele can lead to the conclusion that a linked marker has been found when, in fact, the frequency of the allele reflects the relatedness of cases in the population. It is therefore
important to determine the extent to which dog breeds may be subdivided into smaller genetically differentiated entities, especially for cohorts sampled from different countries. Toward that end, we have evaluated the genetic relatedness of independently bred lines of European and American dogs from the BMD, flat-coated retriever (FCR), golden retriever (GR) and ROT breeds [29]. Our results have demonstrated that, as expected, dog breeds do not constitute homogeneous entities, and population stratification needs to be systematically assessed in association studies. As an example, the GR has the highest rate of polymorphic SNPs (76.5%) and also has the greatest number of haplotypes within an LD block. These results correlate with the popularity of the breed in both Europe and the US. In the US, the GR rose in popularity in the 1930s and was ranked 4th among nearly 155 breeds in 2006 with more than 42,000 American Kennel Club (AKC) registrations (http://www.akc.org/breeds/golden_retriever/history.cfm). In France, GR was ranked 2nd in popularity in 2006. Those results were supported by the fact that the GR analyzed in this study separate significantly in clustering from other breeds. Toward that end, we have evaluated the genetic relatedness of independently bred lines of European and US is rare. The breed creates a large breeding pool on both continents and geographical origins. While the US and the European GR populations have common ancestors, the popularity of this breed creates a large breeding pool on both continents and mixing between European populations and US is rare. Thus, clear population differences and associated differences in allele frequency and distribution exist in the GR that can be traced to sampling location.

7. Conclusion

Dog domestication process started several thousand years earlier than any other mammal domestication, even without taking into consideration the very long phase that some call proto-domestication. This is not the only characteristic of the dog domestication. Indeed, the presence of dog is universal, dogs were present and are present in all cultures and occupy a very special place in our life within the modern societies and have an ambivalent status mid-domesticated animal mid-person as explained by Dominique Guillo in his book [30]. Thus, English-speaking people tend to use “he, she, him and her” pronouns when they speak about their dogs and not the neutral form “it or its”. According to a survey conducted by the American Animal Hospital Association, 94% of owners consider their pet to have human-like personality traits and 93% say that they would risk their own life for their pet [31]. This is a rather unique situation, not well understood if one considers the cost of their presence among us. Up to two percent of the population (4.7 millions) in the US are bitten every years and 30 (absolute number) of them deceased following dog injuries with a total cost for the American society of one billion dollars per year. Sanitary problems and dejection management in large modern cities are a nightmare. For the city of Paris, dejection management alone costs 1.5% of the city budget. There have been long debates about the role that dogs can play in term of affectivity and substitute for the absence of person-to-person communication in our modern societies. However, if one considers that there are as many dogs owned by persons living alone and dogs living in families with several children, this explanation is thus partly realistic. Whatever the actual explanation from a Darwinian point of view, domestication was a great success. While the wolf population is declining worldwide, the dog population is increasing: 75 millions of purebred dogs are registered in the US and in Beijing the number of dogs is presently exploding. From a dog point of view, the success of domestication might appear not so evident, as purebred dogs have to pay a big toll to enjoy their somewhat enviable status. The practices used for the creation and development of many breeds were not well mastered and their consequences understood. Due to the excessive use of some male champions, dissemination and frequency of some disabled alleles have increased dramatically. Nowadays all purebred dogs suffer from a large number of genetic diseases, i.e. up to 30% of Dalmatian become deaf, 25% of BMD suffer from cancer, especially histiocytic sarcoma and 50% of large size dogs are affected by hip dysplasia to cite a few. Obviously, this calls us for a change in our practice and breed management. The good side of this dreadful situation is that the dog population, structured in hundreds of breeds well characterized by their phenotypes – with many specific traits being fixed – constitutes a unique and very efficient model to unravel physiological and pathological phenotype genotype relationships [19,32,33].

Conflict of interest statement

There is no conflict of interest.

Acknowledgements

We want to thank the Centre National de la Recherche Scientifique (CNRS) and the University of Rennes 1 for their continuous support and encouragements. We greatly acknowledge Sophie Licari for the illustration.

References
