



The canine genome

Elaine A. Ostrander and Robert K. Wayne

Genome Res. 2005 15: 1706-1716

Access the most recent version at doi:[10.1101/gr.3736605](https://doi.org/10.1101/gr.3736605)

References This article cites 96 articles, 23 of which can be accessed free at:
<http://genome.cshlp.org/content/15/12/1706.full.html#ref-list-1>

License

Email Alerting Service Receive free email alerts when new articles cite this article - sign up in the box at the top right corner of the article or [click here](#).



To subscribe to *Genome Research* go to:
<https://genome.cshlp.org/subscriptions>

Cold Spring Harbor Laboratory Press

The canine genome

Elaine A. Ostrander^{1,3} and Robert K. Wayne²

¹Cancer Genetics Branch, National Human Genome Research Institute, National Institutes of Health, Bethesda, Maryland 20892, USA; ²Department of Ecology and Evolutionary Biology, University of California at Los Angeles, Los Angeles, California 90095, USA

The dog has emerged as a premier species for the study of morphology, behavior, and disease. The recent availability of a high-quality draft sequence lifts the dog system to a new threshold. We provide a primer to use the dog genome by first focusing on its evolutionary history. We overview the relationship of dogs to wild canids and discuss their origin and domestication. Dogs clearly originated from a substantial number of gray wolves and dog breeds define distinct genetic units that can be divided into at least four hierarchical groupings. We review evidence showing that dogs have high levels of linkage disequilibrium. Consequently, given that dog breeds express specific phenotypic traits and vary in behavior and the incidence of genetic disease, genomic-wide scans for linkage disequilibrium may allow the discovery of genes influencing breed-specific characteristics. Finally, we review studies that have utilized the dog to understand the genetic underpinning of several traits, and we summarize genomic resources that can be used to advance such studies. We suggest that given these resources and the unique characteristics of breeds, that the dog is a uniquely valuable resource for studying the genetic basis of complex traits.

As one of the premier journals in genome biology celebrates its' 10th anniversary, the scientific community studying dogs also enjoys a year of major advances and milestones, particularly with regard to canine genomics and comparative genetics. In July of 2004, the first high-quality draft (7.5×) sequence of the Boxer dog was made publicly available (Lindblad-Toh et al. 2005). This advance followed on the heels of other major milestones in the past several months, including the availability of a 1.5× Poodle sequence (Kirkness et al. 2003), a dense high quality radiation hybrid (RH) map (Breen et al. 2004), a detailed comparative map (Hitte et al. 2005), the localization and cloning of several disease genes, the successful application of dogs for gene therapy studies (Howell et al. 1997; Acland et al. 2001; Mount et al. 2002; Ponder et al. 2002), and new insights into the evolution of dogs and dog breeds (Parker et al. 2004).

As a result, the genome community is well poised to take advantage of the canine system and begin to fulfill some of the expectations advanced nearly 15 yr ago. First, with the development of appropriate molecular resources, the canine system was proposed to hold the power to map and clone disease genes that had proven intractable through studies of human families. Second, the variation in size and skeletal proportions that are segregated into distinct breeds of dog was hypothesized to provide a unique resource for dissecting genetic pathways underlying skeletal development. Finally, the range of behavioral traits that appeared strongly associated with individual breeds suggested a mechanism to decipher the basic genetic vocabulary of behavior (Patterson et al. 1982; Ostrander et al. 1993, 2000; Galibert et al. 1998; Patterson 2000). At the heart of these questions lies a fundamental conundrum. Why has the wolf genome, from which the dog is recently evolved, retained alleles controlling such a

large amount of genetic variability, particularly as regards morphology? Is the dog genome somehow unique from other genomes? Or would strong selective pressures applied to any mammalian genome result in a range of species with a level of phenotypic variation that rivals the dog? Research done to date cannot readily answer these questions. However, we are beginning to understand how to localize the genes that regulate morphology (Chase et al. 2002). In so doing, we can begin to understand how genetic variation leads to major phenotypic changes. With the sequencing of the dog genome, it may be within our grasp to localize genes that cause the difference between Giant Mastiffs and Pekingese, Pointer and Terrier, and sight and scent hounds.

In this celebratory review, we first discuss the evolutionary framework and domestication of dogs. We then consider the recent accomplishments of the canine genome community. Finally, we highlight ongoing studies aimed at addressing some of the questions above.

The evolutionary framework

The domestic dog is the most recently evolved species in the dog family Canidae, a group that has a long history spanning the last 50 million years (Myr). This history can be portrayed as a succession of phylogenetic hierarchies defined by DNA sequence information (Fig. 1) and is a necessary structure for understanding molecular data. Of note is that dogs are the earliest divergence in the superfamily Canoidae that includes bears, weasels, skunks, raccoons, and the pinnipeds (seals, sea lions, and walrus) (Fig. 1A). This kinship predicts dogs will share more molecular similarities with these taxa than with cats, mongooses, civets, or hyenas. However, because of the early divergence of dogs from all other carnivores, only slowly evolving regions will show substantial sequence similarities. A second important point is that the 35 species of extant canids are genetically very similar, having radiated from a common ancestor less than about 10 Mya. The recent

³Corresponding author.

E-mail eostrand@mail.nih.gov; fax (301) 480-0472.

Article and publication are at <http://www.genome.org/cgi/doi/10.1101/gr.3736605>.



Figure 1. Evolutionary relationships of the dog. (A) The evolutionary relationships of carnivores based on DNA hybridization data. (Wayne et al. 1989). (B) A neighbor-joining tree of canids based on 2001 bp of mitochondrial DNA sequence (cytochrome b, cytochrome c oxidase I, and cytochrome c oxidase II) (Wayne et al. 1997a). (C) A neighbor-joining tree of wolf (W) and dog (D) haplotypes based on 261 bp of control region I sequences (Vila et al. 1997). Dog haplotypes are grouped in four sequence clades, numbered I to IV.

radiation in a family that otherwise has a long evolutionary history suggests that genetic comparisons among extant canids will highlight rapidly evolving sequences and that they all may share uniquely evolved molecular structures such as SINE elements inherited from their recent common ancestor (Fanning et al. 1988; Kirkness et al. 2003) or rapidly evolving genes such as olfactory receptors, immune related genes, or reproductive proteins (e.g., Clark et al. 2003). In fact, although the dog family has a diverse chromosome complement ranging from 36 to 78 chromosomes, they all can be reconstructed through simple chromosome rearrangement from a common ancestral karyotype (Nash et al. 2001).

Within the Canidae, three distinct phylogenetic groupings are apparent (Fig. 1B) (Wayne et al. 1987a,b, 1997b) as follows: (1) the fox-like canids, which include species closely related to the red fox (genus *Vulpes*), as well as the arctic and fennec fox (genus *Alopex* and *Fennecus*, respectively); (2) the wolf-like canids including dog, wolf, coyote, Ethiopian wolf or Simien jackal, and three other species of jackals (genus *Canis*), as well as the African hunting dog (genus *Lycaon*) and the dhole (genus *Cuon*); and (3) the South American canids including fox-sized canids such as the pampus fox, crab-eating fox, and small-eared dog (genus *Pseudolopex*, *Lycolopex*, *Atelocynus*) and the maned wolf (genus *Chrysocyon*) and bushdog (genus *Speothos*). Additionally, there are several canids that have no close living relatives and define distinct

evolutionary lineages such as the gray fox (genus *Urocyon*), the bat-eared fox (genus *Otocyon*), and the raccoon dog (genus *Nyctereutes*).

These phylogenetic relationships imply that the dog has several close relatives within its genus, in fact, all members of *Canis* can produce fertile hybrids and several species may have genomes that reflect hybridization in the wild (Wayne and Jenks 1991; Gottelli et al. 1994; Roy et al. 1996; Wilson et al. 2000; Adams et al. 2003). Furthermore, the wolf-like canids are grouped more closely with the South American canids and the red and gray fox are very distinct groups whose common ancestry with dogs extends to the beginning of the modern radiation. Consequently, molecular tools developed from the dog genome sequencing project are likely to be most applicable to the wolf-like canids. For instance, fewer than half of microsatellite primers developed in the dog amplify DNA in the gray fox (Goldstein et al. 1999).

The domestication of the dog

The essential questions about dog domestication concern the species from which the dog originated and the location, number, and timing of domestication or interbreeding events. Molecular data has shed some light on all of these questions. First, with regard to species origins, Charles Darwin and others such as Konrad Lorenz, the renowned behavioral biologist, speculated that

given the great diversity in form and behavior of dogs, they might share ancestry with wolves and other canids, such as any one of the three species of jackals. However, extensive genetic analyses of the dog and other wolf-like canids clearly show that the dog is derived from gray wolves only, rather than jackals, coyotes, or Ethiopian wolves (Fig. 1C; Wayne et al. 1987a,b; Vila et al. 1997, 2005; Leonard et al. 2002; Savolainen et al. 2002). Consequently, the immense phenotypic diversity in the dog owes its origin to primarily the standing genetic variation existing in the ancestral population of gray wolves and any subsequent mutations that occurred during the brief history of domestication. At least for structural genes, such mutations are expected to be few since their mutation rate is so low, on the order of 10^{-5} mutations per gene per generation (Hartl and Clark 1997).

Mitochondrial DNA (mtDNA) sequence analysis has shed some light on the location of dog domestication as well as the number of founding matrilineal lines. MtDNA analysis offers a unique perspective on evolutionary history because the mitochondrial genome is maternally inherited, and hence, only females leave a genetic legacy. Moreover, because the mitochondrial genome does not recombine, phylogenetic analysis of mtDNA sequence data defines a uniquely bifurcating haplotype tree (Fig. 1A,B,C). Phylogenetic analysis of dog and gray wolf mitochondrial sequences clearly show that dog sequences are found in at least four distinct clades, implying a single origination event and at least three other origination or interbreeding events. The latter are difficult to distinguish once the first domestication had occurred, although extensive marker analysis of the nuclear genome might be able to discriminate the two alternatives. A striking finding of the mtDNA analysis is that one sequence clade (clade I, Fig. 1C) contains the majority of dog sequences and that the nucleotide diversity of this clade is high, implying an origin of the clade from 40 to 135 thousand years ago (Vila et al. 1997; Savolainen et al. 2002). This date exceeds the 15,000-yr-old archeological record of dogs and suggests that dogs may have had a long prehistory when they were not phenotypically distinct from wolf progenitors. These early dogs may not have been recognized as domesticated by study of the archeological record before 15,000 yr ago because of their physical similarity to gray wolves. The initial change to the diagnostic phenotype of domestic dogs beginning about 15,000 yr ago may have instead indicated a change in the selection pressures associated with the transition from hunter gatherer to more sedentary lifestyles (Wayne et al. 2006).

Conceivably, a more recent date can be made consistent with the archeological record if it is assumed that dogs were founded from multiple matrilineal lines in clade one (Savolainen et al. 2002). To determine whether such a diverse founding is likely, analysis of nuclear genes sequence data is needed (e.g., Parker et al. 2004). In fact, recent analysis of major histocompatibility (MHC) genes in dogs and wolves suggest that the origin of dogs involved several populations and hundreds of individuals (Vila et al. 2005). Consequently, the model emerging from mitochondrial DNA, MHC analysis, and microsatellite loci is that the dogs had a diverse origin in East Asia that likely involved multiple contributions from several populations, and thereafter, they may have been other origins of domestication and backcrossing (Vila et al. 1997, 2005; Leonard et al. 2002; Savolainen et al. 2002; Parker et al. 2004). A multiple and diverse origin model describes domestication in other domestic animals such as cattle, sheep, and goats (Bruford et al. 2003). Furthermore, once domesticated,

dogs rapidly spread around the earth and as a result, genetically divergent populations and breeds are found in Africa, Asia, the Arctic, Australia, the Middle East, and historically, the New World (Leonard et al. 2002; Parker et al. 2004; Savolainen et al. 2004).

Breed diversity and genetic structure

The explosion of dog breeds over the past two centuries represents perhaps one of the greatest genetic experiments ever conducted by humans. Distilled from the genome of the wild wolf are animals that differ by more than 40-fold in size with the ability to herd, guard, hunt, and guide (American Kennel Club 1998). Behavioral variation is surpassed by morphologic variation, with individual breeds represented by dogs of every imaginable size and proportion. Coats alone can be described by color, texture, length, thickness, and curl. Tails can be described as plumed, curled, double curled, gay (upright), sickled (arching), otter (down and flat), whipped, ringed, screwed, or snapped (American Kennel Club 1998). The diversity in skeletal size and proportion of dogs is greater than any mammalian species and even exceeds that of the entire canid family (Wayne 1986a,c). Such variation may reflect simple modifications of post-natal development (Wayne 1986a,c), but the specific genetic mechanisms are not well known (see below).

Much of the morphologic variation in dogs is partitioned into over 350 distinct breeds worldwide as a result of the development of breed standards and controlled breeding. In general, in order to register a dog in the American Kennel Club at least both parents must have been registered in the same breed. Consequently, purebred dogs are members of closed breeding populations, which receive little genetic variation beyond that existing in the original founders (Ostrander and Giniger 1997; Galibert et al. 1998; Ostrander et al. 2000; Sutter and Ostrander 2004).

Common to the origin and development of many breeds is a founder event involving only a few dogs and, thereafter, reproductive dominance by popular sires that conform most closely to the breed standard. These restrictive breeding practices reduce effective population size and increase genetic drift, resulting in the loss of genetic diversity within breeds and allele frequency divergence among them. For example, in a genetic study of 85 breeds, Parker et al. (2004) showed that humans and dogs have similar levels of overall nucleotide diversity, 8×10^{-4} , which represent the overall number of nucleotide substitutions per base/pair. However, the variation between dog breeds is much greater than the variation between human populations (27.5% versus 5.4%). Conversely, the degree of genetic homogeneity is much greater within individual dog breeds than within distinct human populations (94.6% versus 72.5%). Furthermore, in some breeds, genetic variation has been additionally reduced by bottlenecks associated with catastrophic events such as war and economic depression, making them analogous to human populations of limited genetic variation used for disease-mapping studies such as the Finns, Icelanders, and Bedouins. As a result, the unique pattern of LD in dogs provides an exceptional opportunity to study complex traits that are relevant to human biology using robust approaches that would not be possible in human populations.

Because many breeds represent closed gene pools, they may define distinct genetic clusters. Analysis of microsatellite loci have strongly supported this notion (Koskinen 2003). For example, in

the Parker et al. (2004) study, 96 microsatellite markers were genotyped that spanned all dog autosomes at approximately a 30-Mb resolution (Parker et al. 2004). Excluding data from the highly related Belgian Sheepdog and Belgian Tervuren breeds, they observed that 99% of 414 dogs were correctly assigned to breed. Consequently, a “breed” can be defined at the molecular level and dogs can be correctly assigned to their breed with small amounts of data. These results strongly imply that breeds are distinct genetic units and even closely related breeds do not represent genetic replicates.

Breed origin and relationship

Mitochondrial DNA studies have not been useful for the reconstruction of breed origins or relationships because the origin of the vast majority of sequence polymorphisms found in dogs preceded the development of modern breeds. Therefore, phylogenetic hierarchies based on DNA sequences reveal the history of mutations that occurred before dogs were domesticated (e.g., Fig. 1C). However, many breeds contain several mitochondrial DNA haplotypes, suggesting that multiple matrines were involved in the founding of a dog breed. To assess the recent evolution and relationships of breeds, microsatellite loci provide a better tool, as their high variability insures allele frequency divergence through drift. Genetic distance trees based on the microsatellite dataset from Parker et al. (2004) revealed several distinct breed clusters. The most divergent grouping presumably contained the most ancient breeds, but none of these nine ancient breeds were of European origin. The ancient breeds included dogs from a wide geographic area including the Arctic, Asia, Africa, and the Middle East. By comparison, the majority of breeds, including European breeds, appeared to stem from a single node without

significant phylogenetic structure, which has been termed a “hedge,” indicating a recent origin and extensive hybridization between the breeds (Parker et al. 2004; Fig. 2). The focus on breeds belonging to this hedge in past studies probably explains the observed lack of phylogenetic resolution (Zajc et al. 1997; Koskinen and Bredbacka 2000; Irion et al. 2003).

This evolutionary hierarchy suggests breeds should cluster genetically into groups sharing recent common ancestry. A genetic clustering algorithm, deployed in the computer program “structure” was used to explore the possible groupings within dogs (Pritchard et al. 2000). Structure assigned 335 dogs correctly to 69 unique breed specific clusters that represented either single breeds or sets of very closely related breeds. However, the program could not easily distinguish a half-dozen obviously related pairs such as the Bernese Mountain Dog and Greater Swiss Mountain Dog or Mastiff and Bullmastiff. This lack of resolution in these few breeds is predicted based on breed history. For instance, the Bullmastiff is reported to be 60% Mastiff and 40% Bulldog and was created by crossing the two breeds in the mid-1800s (Rogers and Brace 1995).

Individual breeds represented the smallest definable cluster; however, higher order clusters are expected given the origins of many dogs breeds. Consequently, the number of groups (K) was set to two, three, and finally, four. The first distinct cluster to be defined at $K = 2$ included nearly all breeds of Asian origin (Akita, Shiba Inu, Shar Pei, Lhasa Apso, etc.), some sled dogs, and some known ancient hounds such as the Saluki (Fig. 2). When added to the analysis, gray wolves from eight countries all grouped in the first cluster as well. The early divergence of the Asian breeds on the phylogenetic tree and their association with the wolves in clustering analysis (Fig. 2) supports the conclusions of mitochon-

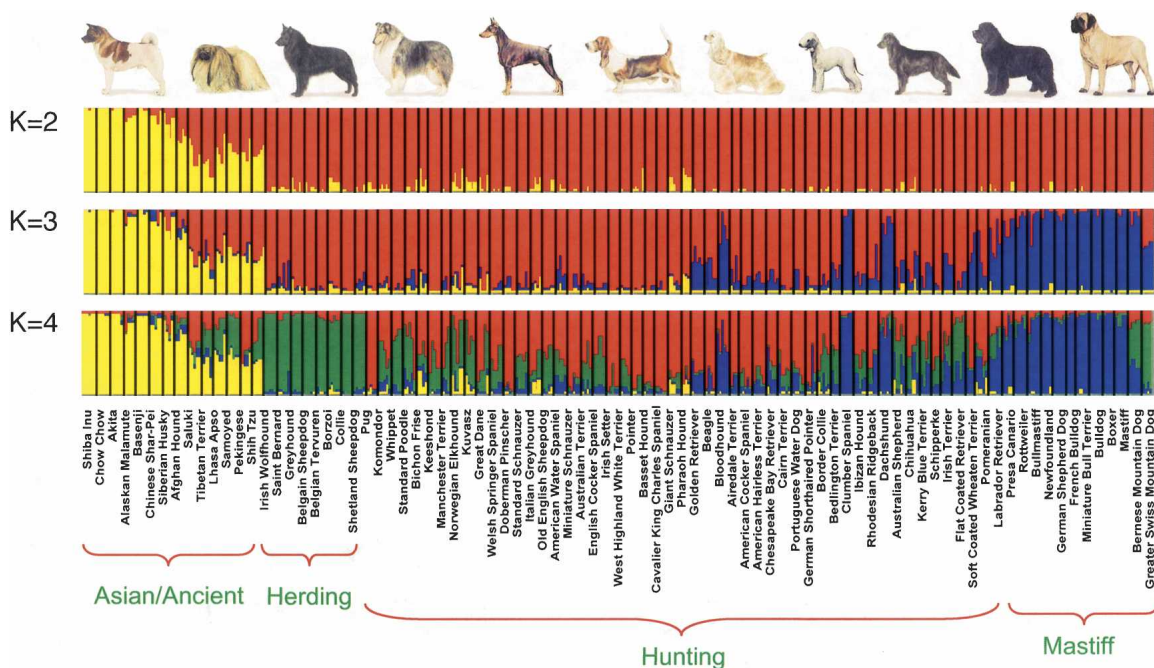


Figure 2. Structure analysis of 85 dog breeds. Cluster results from a structure analysis of 414 dogs from 69 breeds and based on 96 microsatellite markers. Each breed was usually represented by five dogs, and all dogs were unrelated to one another at the grandparent level. Structure implements a Bayesian model-based clustering algorithm that attempts to identify genetically distinct subpopulations based on patterns of allele frequencies (Pritchard et al. 2000). Each genotyped dog is represented by a single vertical line divided into K colors, where K is the number of clusters assumed in each structure analysis. The length of the colored segment represents the individual’s estimated proportion of membership in that cluster (Parker et al. 2004). At $K = 4$, four clusters are clearly defined representing genetically distinct breed grouping within the domestic dog (see text).

drial DNA analysis that domestication first took place in East Asia (Savolainen et al. 2002). The next cluster to be defined at $K = 3$ was comprised of mastiff-type dogs including the Mastiff, Bullmastiff, Bulldog, Boxer, etc. Finally, at $K = 4$, the third cluster to be defined included working dogs such as the Collie and Shetland Sheepdog, together with a subset of the sight hounds, such as the Greyhound. The final cluster comprised mostly modern breeds used in hunting and included gun dogs, hounds, and terriers. On-going analysis is focusing on defining clusters within this hedge group, using more highly mutable tetranucleotide-based microsatellite markers (Francisco et al. 1996) and less mutable markers based on single nucleotide polymorphisms (SNPs). However, the structure analysis for the first time defined groups based on common ancestry and genetic similarity rather than function (e.g., hunting or herding breeds) and provides a genetic guide to the design of whole-genomic scans (see below).

Another promising approach toward reconstructing breed history utilizes single gene histories. For example, study of the multidrug resistance gene (*MDR1*) and four closely linked microsatellite markers was used to reconstruct the history of a group of related breeds (Neff et al. 2004). A single *MDR1* mutation was found to segregate in nine breeds that included seven herding breeds and two sight hound subgroups, which were likely related to at least one of the herding breeds. Haplotype analysis confirmed this relationship by revealing that the region around *MDR1* was identical by descent in all nine breeds, suggesting that they inherited this haplotype from an exclusive common ancestor. Additional study of single gene mutations in dogs will help dissect the branching structure of "twigs" in the phylogenetic tree of dogs.

Mapping and sequencing the dog genome

The success of disease-mapping studies and those unraveling the mysteries of canine evolution were clearly dependent on the prior development of key resources. Meiotic linkage maps and RH maps based on family studies (Mellersh et al. 1997) and a 5000 rad panel (Vignaux et al. 1999) were first made available in the late 1990s and were essential to subsequent map-building efforts (Mellersh et al. 1997, 2000; Priat et al. 1998; Neff et al. 1999). The first comparative maps and later dense RH maps that followed allowed researchers to take full advantage of the much more well-developed human and mouse genome mapping resources (Breen et al. 2001; Guyon et al. 2003, 2004). A recent integrated RH map of the dog, including microsatellites, genes, and BAC ends (Breen et al. 2004), has proven invaluable in allowing investigators to do positional cloning experiments following initial findings of linkage. Most recent mapping efforts focused on developing a high-resolution 9000 rad comparative map (Hitte et al. 2004), which includes 10,348 canine markers, 9850 corresponding to canine orthologs of human genes derived from a $1.5 \times$ poodle shotgun sequence (Kirkness et al. 2003). For online information, see <http://sun-recomgen.med.univ-rennes1.fr/Dogs/> and http://research.nhgri.nih.gov/dog_genome/.

Very recently, the landscape for canine genome studies has been changed by the availability of a $7.5 \times$ assembled sequence of the Boxer genome (<http://www.genome.ucsc.edu>), completed by investigators at the Broad Institute (CanFam1.0 and CanFam2.0) (Lindblad-Toh et al. 2005). These data suggest that the euchromatic portion of the dog genome is $\sim 18\%$ smaller than the human genome and 6% smaller than the mouse ge-

nome. The size difference is explained by a lower rate of repeat insertions in the dog genome relative to both human and mouse, while the deletion rate of ancestral bases has been approximately equal between the dog and human lineages. The relatively low level of recent repeats in the dog genome contributes, together with high quality data and improved assembly algorithms, to the high connectivity and quality of the dog genome assembly. This is well supported by the above-mentioned RH gene map of the dog, which shows high concordance with the assembled sequence as well as a set of several hundred BAC ends previously localized by FISH (Hitte et al. 2005).

The assembled sequence demonstrates that $\sim 94\%$ of the dog genome is contained in clear segments of conserved synteny relative to the human and mouse genomes. The gene count of $\sim 19,000$ canine genes is slightly lower than that currently considered for human, which is somewhat surprising. The accuracy of these data, however, is high; of the 19,000 reported canine genes, 14,200 represent 1-1-1 orthologs between dog, human, and mouse. Approximately 5.4% of the orthologous nucleotides between human and dog appears to be under purifying selection. The purifying selection acting on conserved orthologous genes appears significantly higher in the lineage leading to dog than in that leading to human, but lower than in the lineage leading to mouse. However, the relative constraints between orthologs with different functions have been highly correlated between the three lineages. Only genes involved in nervous system function have diverged faster in both dog and human relative to mouse, but not relative to each other, consistent with similar selection pressures, and possibly, convergent evolution. Finally, gene family expansions are less common in dog than in human, suggesting that the dog has the most primitive gene content of the currently sequenced placental mammals.

Linkage disequilibrium across and between dog breeds

To fully exploit the unique genetic characteristics of the dog, the architecture of linkage disequilibrium (LD) in the canine genome needs to be understood. This knowledge would facilitate the mapping and cloning of genes important to canine health, as well as the discovery of loci regulating phenotypic traits. The importance of this knowledge is demonstrated in human studies where LD mapping in well-defined populations has simplified locus heterogeneity problems associated with complex traits (Kruglyak 1999a; Sundin et al. 2000; Ophoff et al. 2002; Friedrichsen et al. 2004). Three fundamental questions have been addressed. First, how does the extent of LD compare to that which has been reported in humans? Second, how does LD differ between breeds, and finally, how well does breed history predict the extent of LD?

These issues have been addressed in two major studies (Sutter et al. 2004; Wade et al. 2005). Sutter et al. (2004) examined 189 SNPs from five unlinked loci in five breeds using 20 unrelated dogs from each breed (Fig. 3). They found that in the Golden Retriever, LD falls to half of its maximum value at about 0.48 Mb. However, in the other breeds, LD is more extensive, increasing to about 0.9 in the Pekingese and Labrador Retriever and to 2.2 Mb in the Bernese Mountain Dog. Finally, at 3.8 Mb, LD in the Akita is nearly $10 \times$ greater than that observed in the Golden Retriever. In some cases, these observations agree well with recorded breed history (Fogel 1995; Wilcox and Walkowicz 1995; American Kennel Club 1998;). For instance, the Golden and Labrador Retriever are among the most popular breeds and

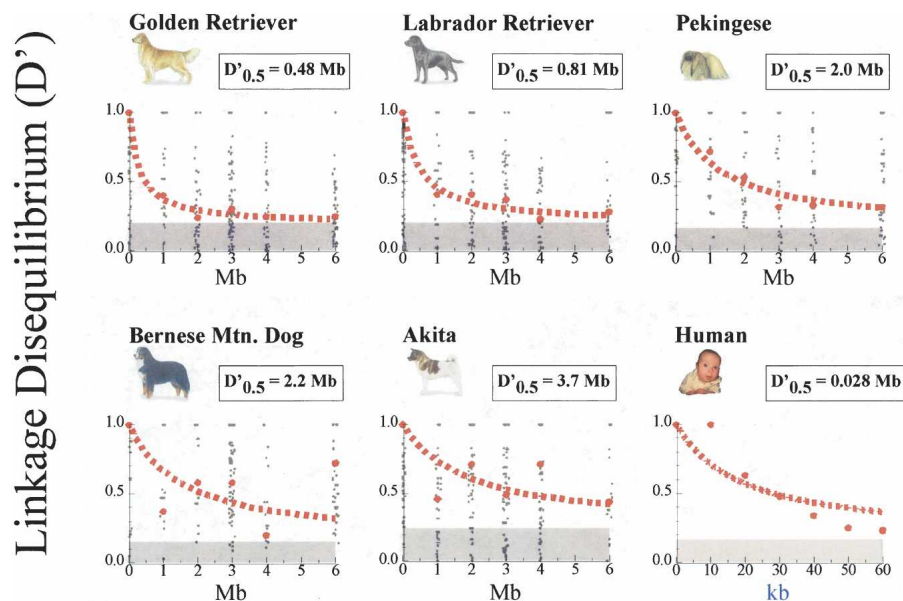


Figure 3. LD in five breeds of dog. LD in 20 unrelated dogs from each of the five breeds scanned for a total of 51 Kb in five unlinked regions on chromosomes 1, 2, 3, 34, and 37. The scan revealed 189 SNPs and those with a minor allele frequency greater than 0.2 in each breed were used on LD calculations. Data were averaged across the five sites and the D' statistic used to indicate the level of linkage disequilibrium. $D'_{0.5}$ indicates the point at which the D' statistic decays by 50%. Data are given in Mb for dog and Kb for human.

neither breed has experienced significant population bottlenecks (Fogel 1995; Wilcox and Walkowicz 1995). By comparison, LD is expected to be greater in the Pekingese, as these dogs are derived from a small number of founders that came to the U.S. from China (Fogel 1995; Wilcox and Walkowicz 1995). LD is predicted to be most extreme in the Akita, a relatively rare breed with a restricted gene pool.

These results suggest two important considerations for the design of mapping and cloning studies. First, as there is at least a 10-fold difference in the extent of LD between dog breeds, breed selection deserves careful consideration. Second, LD in dogs is 20–50 times more extensive than that found in humans, where LD is typically reported to be about 0.28 Mb (Reich et al. 2001; Weiss and Clark 2002). More than 500,000 SNPs must be genotyped for whole-genome association studies in humans (Kruglyak 1999b; The International HapMap Consortium 2003). In contrast, only about 10,000 SNPs are hypothesized to be needed for the comparable dog study (Sutter et al. 2004). Thus, the mapping of common and complex diseases such as epilepsy, cancer, autoimmune disease, deafness, and heart disease in dogs may be more economical than similar efforts in humans.

The canine genome sequencing effort has made 2.1 million SNPs publicly available (<http://www.broad.mit.edu/mammals/dog/snp/>) (Lindblad-Toh et al. 2005). To determine how to best use this resource, Sutter et al. (2004) examined the extent of haplotype sharing for the five breeds described above. For any one breed, 80% of chromosomes examined had, on average, just 2.7 haplotypes. For all 100 dogs examined, 80% of chromosomes carried just 4.5 haplotypes. The overall degree of haplotype sharing, measured as the proportion of a breed's chromosomes carrying haplotypes shared with another breed, ranged from 46% to 84%. These findings of low haplotype diversity and high haplotype sharing, albeit with great variability, suggest that a universal

SNP set of modest size will be sufficient to successfully accomplish whole-genome association studies in most breeds.

A more in-depth analysis of the same general questions, as well as issues regarding the overall haplotype structure of the dog were examined using ~1300 SNPs plus resequencing data drawn from 10 random regions covering 6% of the genome. The study was undertaken as part of the canine genome sequencing effort (Lindblad-Toh et al. 2005) and the conclusions largely agree with those of Sutter et al. (2004). In addition to the 7.5× Boxer sequence, the genome sequencing effort generated 100,000 sequence reads from each of nine diverse breeds representing all seven AKC groups, and 20,000 reads from each of five wild canids (four wolves and one coyote). The resulting SNP frequencies of 1/900 bp between breeds, 1/580 bp between dogs and wolves, and 1/420 bp between dogs and coyote, emphasizes that all three species are more closely related than human and chimpanzee. The resulting set of 2.1 million SNPs have a polymorphism rate across breeds of ~72% within any given

breed, suggesting that most SNPs discovered as part of the sequencing effort will be useful for mapping in any breed.

Comparison of the two boxer haplotypes, as well as extensive resequencing and genotyping in 10 breeds by the sequencing group has been illustrative for understanding the detailed haplotype structure of the dog. Such analyses demonstrate megabase sized portions of the genome that are alternatively homozygous and heterozygous exist both for the sequenced boxer, as well as for 24 dogs from different breeds and 20 dogs from each of 10 breeds. Thus, megabase-sized haplotypes will be common within virtually any purebred dog.

Lindblad-Toh and collaborators conclude that LD within any breed is actually dependent on the intensity and duration of two bottlenecks. The first is an ancient bottleneck occurring at the time of canine domestication that is common to all dogs. The second likely occurred during breed formation. In combination, these bottlenecks resulted in LD that extends for megabases in most breeds and limited haplotype diversity. Indeed, across the dog population as a whole, ancestral haplotype blocks are roughly 5–10 kb long with approximately five alleles in each block. Thus, when LD is examined carefully across many breeds, typically, five haplotypes are observed across each 10–500-kb window, with one or two being common and the rest rare. The recent ancestry of these haplotypes supports the idea that a modest number of SNPs, perhaps as few as 5000, will be sufficient for genome-wide association mapping. However, the underlying ancestral haplotype block structure implies that the false-positive rate will be high if only single SNP association is used. Consequently, haplotype-based association should be used instead for most mapping studies.

Canine disease gene mapping

Billions of dollars are spent on canine health in the United States each year (Association 2002) and much of it is focused on a

limited number of diseases including cancer, epilepsy, blindness, cataracts, autoimmune disease, and heart disease. Over 360 genetic disorders found in humans have also been described in the dog (Patterson 2000; Sargan 2004), and about 46% of these genetic diseases occur predominantly or exclusively in one or a few breeds. A detailed listing of over 1000 canine diseases, and descriptions of each, appears in the database of inherited diseases in dogs (IDID, <http://www.vet.cam.ac.uk/idid>) (Sargan 2004).

To date, the location of many canine disease loci has been determined, and in some cases the underlying gene has been cloned (for review, see Patterson et al. 1982; Ostrander and Giniger 1997; Galibert et al. 1998; Ostrander et al. 2000; Sutter and Ostrander 2004; Switonski et al. 2004).

In some cases, identification of canine disease genes has opened new avenues of research for human biologists. For instance, the identification of a mutation in the hypocretin 2 receptor gene (Lin et al. 1999) in Doberman Pinschers with inherited narcolepsy has proven key to understanding the molecular mechanisms which regulate sleep (Nishino et al. 2000; Thanickal et al. 2000). In humans, the disease is associated with a progressive loss of hypocretin-expressing neurons and is a non-Mendelian trait mediated by a unique mechanism different from that causing the disease in Dobermans. However, study of the simpler etiology in dogs provided the requisite tools for understanding the more complex disease in humans.

In other cases, study of canine disease genes has increased our understanding of the interaction between genes and how such interactions affect disease. Such interactions have proven difficult to study in human populations, where the size of even the largest case-control studies is simply too small to identify anything but major effects. The identification of the *MURR1* gene associated with copper toxicosis in Bedlington Terriers (van De Sluis et al. 2002) provides an excellent example. Contrary to expectation, this disease did not map to the portion of the canine genome analogous to the Wilson's disease locus in humans (Yuzbasiyan-Gurkan et al. 1997; van de Sluis et al. 1999). Analysis of the human homolog of *MURR1* in Wilson's disease patients has subsequently proven provocative, as those who carry particular sequence variants appear to present with earlier onset disease (Stuehler et al. 2004), suggesting that the two genes or their products interact to accelerate disease.

Another significant advance concerns the identification of novel disease mechanisms through the study of dog genetics. Lohi et al. (2005) recently identified a gene for progressive myoclonic epilepsy (PME) in a population of purebred miniature wirehaired dachshunds. About 5% of the breed suffers from this autosomal recessive disease, which was shown to be analogous to the human disorder, Lafora disease. As in the human disease, affected individuals carry mutations in the *NHLRC1* gene. However, in contrast to the human disease, the disease in dogs is due exclusively to bi-allelic expansion of a dodecamer repeat found within the 5' end of the genes' single large exon. Affected individuals carry 19 to 26 copies of the repeat sequence rather than the expected two copies. This is the first example of a dodecamer repeat expansion associated with disease in any mammalian system and suggests a potential novel mechanism for human disorders.

Currently, perhaps the greatest concentrated collaborative efforts are focused on the study of canine cancer (Chun and de Lorimier 2003; Ettinger 2003; Fan 2003; London and Seguin 2003; Porrello et al. 2004; Modiano et al. 2005). Dogs develop cancer about twice as frequently as humans and the disease pre-

sentation and pathology of canine cancers is similar to analogous human tumors. Genetic studies are ongoing to find susceptibility genes for canine osteosarcoma, lymphoma, mast cell tumors, malignant histiocytosis, and kidney cancer, and a BAC CGH array resource is in development to better understand somatic events leading to tumor growth and metastasis (Thomas et al. 2003a,b). Of primary interest is determining whether different types of tumors have unique or shared origins. If a common origin of a particular canine cancer is established, then considering data from several breeds simultaneously can facilitate the localization of the susceptibility gene. Breeds of similar appearance and sharing common ancestry as suggested by historical record may often share variants for disease phenotypes (e.g., Neff et al. 2004). However, in most cases, rigorous studies such as those described below are needed to address the issue.

Genetics of morphology

The genetic basis for differences in size and proportion among dogs has yet to be revealed. However, both candidate gene and association studies are beginning to provide insight into the complexity underlying morphological differentiation. For example, two potential candidate genes, *MSX2* and *TCOF1*, which are expressed during cranial facial development, were sequenced in 10 different dog breeds that varied in cranial and face shape (Haworth et al. 2001a,b). However, only a single amino acid change in the *TCOF1* protein showed an association with short and broad skulls. Nonetheless, greatly expanded surveys of candidate genes may prove more fruitful; for example, variation in the production of insulin-like growth factor 1 (IGF-1) was shown to correlate with differences in the body size of poodles, suggesting it may be a candidate gene for size variation in dogs (Eigenmann et al. 1984).

More definitive associations have been demonstrated through quantitative analysis of morphologic measurements combined with genome marker scans. For example, Chase et al. (2002) analyzed data from nearly 700 Portuguese Water Dogs genotyped with ~500 markers and <http://www.georgieproject.com/>. For 460 dogs, they recorded 91 measurements from a set of five x-rays taken on each dog. The data were analyzed using principal component analysis, which defines independent component axes based on linear combinations of variables. Each axis is ordered by a decreasing fraction of the total variation in the data set. The first four axes explained 61% of the variation in the data set and represented different components of size and shape. For example, the first principal component axis reflected overall size variation of the skeleton, whereas the second reflected the relationship between the pelvis, head, and neck, such that the size and strength of the pelvis and head-neck musculoskeletal systems are inversely related. Quantitative Trait Loci (QTLs) have been localized that are related to variation on each of the above four principal components. Moreover, using a data set of 286 phenotyped dogs, Chase et al. (2004) defined two loci on chromosome one spaced 95 Mb apart that appear to account for a modest percentage of hip dysplasia, as defined by Norberg angle in the Portuguese Water Dog.

Nonclassical genetic variation may also be an important source of phenotypic variation in dogs. Fondon III and Garner (2004) suggested that highly mutable simple tandem repeats imbedded in genes may be the source of new variation in recent developed lines and may explain their high rate of morphologic change. To test this hypothesis, these investigators analyzed

three-dimensional models of dog skulls from 20 breeds and seven mongrels. In representatives of 92 different breeds, they also sequenced 37 repeat-containing regions from 17 genes known or thought to be involved in craniofacial development. In general, they found that dogs had more perfect repeats than humans and may be changing faster in length. Additionally, they found that the size and the ratio of lengths of two tandem repeats in the *Rumx-2* gene correlated with the degree of dorsoventral nose bend (clinorhynch) and mid-face length in a variety of breeds. Although this evidence is suggestive, clearly more detailed studies are needed associating repeat change with specific phenotypic traits (Pennisi 2000). If such genetic mechanisms are unique to the dog, they may explain, in part, the apparent phenotypic plasticity of dogs. However, dogs also have a unique skeletal development whose alterations may more readily result in novel phenotypes (Wayne 1986a,b,c; Morey 1992, 1994).

One area of morphology we do not discuss in detail is that of canine coat color, which has been written about extensively in the past. More recently, progress on dissecting coat color genetics in the dog has been done by two groups (Kerns et al. 2003; Berryere et al. 2005). Particular progress has been made in understanding the interactions between the Agouti protein and the Melanocortin 1 receptor, which control the type of pigment synthesized in mammalian hair (Berryere et al. 2005). Additional recent work has focused on black color in dogs, which appears to be independent of the above interactions (Schmutz et al. 2002; Kerns et al. 2003). Very interesting work that is just beginning focuses on the role of polymorphisms in coat color affecting genes, such as the melanophilin gene (Philipp et al. 2005). With the availability of the canine genome sequence, this is an area that will surely expand in the coming years.

Genetics of behavior

Dog breeds have distinct behaviors, and dogs as a whole have unique behaviors not found in gray wolves (Hare et al. 2002). However, the genetic basis of behavior is less well understood than morphology. In general, the greatest need remains the development of assays to reproducibly score specific behaviors. However, some understanding is likely to come from the study of pedigrees of dogs displaying aberrant behaviors. For example, Moon-Fanelli et al. (1998) have characterized pedigrees of Bull Terriers displaying obsessive compulsive disease (OCD) phenotypes, such as tail chasing, which in other respects is similar to human OCD. As genome scans of affected pedigrees are completed, they may shed light on both the human and canine disease conditions.

Expression patterns may also provide clues to the genetic basis of behavior. Saetre et al. (2004) surveyed the expression pattern of 7762 genes in three different regions in the brains of domestic dogs and in gray wolves and coyotes. They found that the pattern of gene expression in the hypothalamus of domestic dogs was different from that in gray wolves and coyotes, whereas patterns of gene expression in the amygdala and frontal cortex were less differentiated. The hypothalamus controls specific emotional, endocrinological, and autonomic responses of dogs and is highly conserved throughout mammals. The results of Saetre et al. (2004) suggest that behavioral selection in dogs may have affected this central part of the brain, initiating a cascade of effects that result in some of the unique behaviors found in dogs.

Conclusions

The domestic dog has long fascinated evolutionary biologists and geneticists because of the extreme phenotypic diversity exhibited by the species and the short time frame over which this diversity has evolved. Molecular genetic evidence suggests that dogs are indeed the oldest domesticated species and their origin may have even well preceded their first appearance in the archeological record about 15,000 yr ago. The dog has a diverse genetic origin that likely involved multiple gray wolf populations and subsequently was enriched by backcrossing with wolves throughout their history. This substantial input of variation from wild ancestors has provided the raw material for phenotypic change, but unique development and genetic mechanisms may also have assisted the course of artificial selection. Dogs clearly have behaviors, phenotypes, and diseases that are not evident in their wild progenitors. Finally, in the more recent evolution of dog breeds, limited interbreeding has imposed a remarkable genetic structure such that nearly all breeds represent distinct genetic pools that can be divided into at least four distinct genetic groupings.

Understanding the genetic mechanisms that have given rise to the unique attributes of domestic dogs may finally be within reach. A complete and a partial genome sequence are available from a boxer and a poodle, respectively, and mapping resources are well developed and increasing in sophistication. The dog genome in general has high levels of LD, such that whole-genome association studies will be facilitated and genomic scans of specific breeds segregating traits of interest may readily be found through patterns of LD or reductions in heterozygosity due to selective sweeps (Weiss and Clark 2002; Bamshad and Wooding 2003; Luikart et al. 2003; Pollinger et al. 2005). In this review, we have provided the evolutionary and empirical framework for understanding the molecular diversity of dogs with the aim of taking the first step toward answering the questions posed in the introduction. The primary intent of this article was to help generate the enthusiasm that will lead to realizing the promise of the dog genome for solving significant problems in evolution, genetics, and human health.

Acknowledgments

We thank two anonymous reviewers, Kerstin Lindblad-Toh, Heidi Parker, Nate Sutter, Ed Giniger, and Francis Galibert for thoughtful comments and helpful suggestions on this manuscript. We also thank Kerstin Lindblad-Toh for sharing data in advance of publication. Finally, we thank the many colleagues, dog owners, and breeders who have generously shared samples and made much of the work reviewed here possible.

References

- Acland, G.M., Aguirre, G.D., Ray, J., Zhang, Q., Aleman, T.S., Cideciyan, A.V., Pearce-Kelling, S.E., Anand, V., Zeng, Y., Maguire, A.M., et al. 2001. Gene therapy restores vision in a canine model of childhood blindness. *Nat. Genet.* **28**: 92–95.
- Adams, J.R., Leonard, J.A., and Waits, L.P. 2003. Widespread occurrence of a domestic dog mitochondrial DNA haplotype in southeastern US coyotes. *Mol. Ecol.* **12**: 541–546.
- American Kennel Club 1998. *The complete dog book*. Howell Book House, New York.
- American Veterinary Medical Association 2002. *U.S. pet ownership and demographics sourcebook*. American Veterinary Medical Association, Schaumburg, IL.
- Bamshad, M. and Wooding, S.P. 2003. Signatures of natural selection in the human genome. *Nat. Rev. Genet.* **4**: 99–111.

- Berryere, T.G., Kerns, J.A., Barsh, G.S., and Schmutz, S.M. 2005. Association of an Agouti allele with fawn or sable coat color in domestic dogs. *Mamm. Genome* **16**: 262–272.
- Breen, M., Jouquand, S., Renier, C., Mellersh, C.S., Hitte, C., Holmes, N.G., Cheron, A., Suter, N., Vignaux, F., Bristow, A.E., et al. 2001. Chromosome-specific single-locus FISH probes allow anchorage of an 1800-marker integrated radiation-hybrid/linkage map of the domestic dog genome to all chromosomes. *Genome Res.* **11**: 1784–1795.
- Breen, M., Hitte, C., Lorentzen, T.D., Thomas, R., Cadieu, E., Sabacan, L., Scott, A., Evanno, G., Parker, H.G., Kirkness, E., et al. 2004. An integrated 4249 marker FISH/RH map of the canine genome. *BMC Genomics* **5**: 1–11.
- Bruford, M.W., Bradley, D.G., and Luikart, G. 2003. DNA markers reveal the complexity of livestock domestication. *Nat. Rev. Genet.* **4**: 900–910.
- Chase, K., Carrier, D.R., Adler, F.R., Jarvik, T., Ostrander, E.A., Lorentzen, T.D., and Lark, K.G. 2002. Genetic basis for systems of skeletal quantitative traits: Principal component analysis of the canid skeleton. *Proc. Natl. Acad. Sci.* **99**: 9930–9935.
- Chase, K., Lawler, D.F., Adler, F.R., Ostrander, E.A., and Lark, K.G. 2004. Bilaterally asymmetric effects of quantitative trait loci (QTLs) that affect laxity in the right versus left coxofemoral (hip) joints of the dog (*Canis familiaris*). *Am. J. Med. Genet. A* **124**: 239–247.
- Chun, R. and de Lorimier, L.P. 2003. Update on the biology and management of canine osteosarcoma. *Vet. Clin. North Am. Small Anim. Pract.* **33**: 491–516, vi.
- Clark, A.G., Glanowski, S., Nielsen, R., Thomas, P.D., Kejarwal, A., Todd, M.A., Tanenbaum, D.M., Civello, D., Lu, F., Murphy, B., et al. 2003. Inferring nonneutral evolution from human–chimpanzee–mouse orthologous gene trios. *Science* **302**: 1960–1963.
- Eigenmann, J.E., Patterson, D.F., and Froesch, E.R. 1984. Body size parallels insulin-like growth factor I levels but not growth hormone secretory capacity. *Acta Endocrinol.* **106**: 448–453.
- Ettinger, S.N. 2003. Principles of treatment for soft-tissue sarcomas in the dog. *Clin. Tech. Small Anim. Pract.* **18**: 118–122.
- Fan, T.M. 2003. Lymphoma updates. *Vet. Clin. North Am. Small Anim. Pract.* **33**: 455–471.
- Fanning, T.G., Modi, W.S., Wayne, R.K., and O'Brien, S.J. 1988. Evolution of heterochromatin-associated satellite DNA loci in felids and canids (Carnivora). *Cytogenet. Cell Genet.* **48**: 214–219.
- Fogel, B. 1995. *The encyclopedia of the dog*. DK Publishing, Inc., New York.
- Fondon III, J.W. and Garner, H.R. 2004. Molecular origins of rapid and continuous morphological evolution. *Proc. Natl. Acad. Sci.* **101**: 18058–18063.
- Francisco, L.V., Langston, A.A., Mellersh, C.S., Neal, C.L., and Ostrander, E.A. 1996. A class of highly polymorphic tetranucleotide repeats for canine genetic mapping. *Mamm. Genome* **7**: 359–362.
- Friedrichsen, D.M., Stanford, J.L., Isaacs, S.D., Janer, M., Chang, B.L., Deutsch, K., Gillanders, E., Kolb, S., Wiley, K.E., Badzioch, M.D., et al. 2004. Identification of a prostate cancer susceptibility locus on chromosome 7q11–21 in Jewish families. *Proc. Natl. Acad. Sci.* **101**: 1939–1944.
- Galibert, F., Andre, C., Cheron, A., Chuat, J.C., Hitte, C., Jiang, Z., Jouquand, S., Priat, C., Renier, C., and Vignaux, F. 1998. The importance of the canine model in medical genetics. *Bull. Acad. Natl. Med.* **182**: 811–821.
- Goldstein, D.B., Roemer, G.W., Smith, D.A., Reich, D.E., Bergman, A., and Wayne, R.K. 1999. The use of microsatellite variation to infer population structure and demographic history in a natural model system. *Genetics* **151**: 797–801.
- Gottelli, D., Sillero-Zubiri, C., Applebaum, G.D., Roy, M.S., Girman, D.J., Garcia-Moreno, J., Ostrander, E.A., and Wayne, R.K. 1994. Molecular genetics of the most endangered canid: The Ethiopian wolf *Canis simensis*. *Mol. Ecol.* **3**: 301–312.
- Guyon, R., Lorentzen, T.D., Hitte, C., Kim, L., Cadieu, E., Parker, H.G., Quignon, P., Lowe, J.K., Renier, C., Gelfenbeyn, B., et al. 2003. A 1-Mb resolution radiation hybrid map of the canine genome. *Proc. Natl. Acad. Sci.* **100**: 5296–5301.
- Guyon, R., Kirkness, E.F., Lorentzen, T.D., Hitte, C., Comstock, K.E., Quignon, P., Derrien, T., Andréa, C., Fraser, C.M., Galibert, F., et al. 2004. Comparative mapping of human chromosome 1p and the canine genome. In *The genome of Homo sapiens. The 68th Cold Spring Harbor Symposium*, pp. 171–177. Cold Spring Harbor Press, Cold Spring Harbor, NY.
- Hare, B., Brown, M., Williamson, C., and Tomasello, M. 2002. The domestication of social cognition in dogs. *Science* **298**: 1634–1636.
- Hartl, D. and Clark, A. 1997. *Principles of population genetics*. Sinauer Associates, Inc., Sunderland, MA.
- Haworth, K., Breen, M., Binns, M., Hopkinson, D.A., and Edwards, Y.H. 2001a. The canine homeobox gene MSX2: Sequence, chromosome assignment and genetic analysis in dogs of different breeds. *Anim. Genet.* **32**: 32–36.
- Haworth, K.E., Islam, I., Breen, M., Putt, W., Makrinou, E., Binns, M., Hopkinson, D., and Edwards, Y. 2001b. Canine TCOF1: cloning, chromosome assignment and genetic analysis in dogs with different head types. *Mamm. Genome* **12**: 622–629.
- Hitte, C., Derrien, T., Andre, C., Ostrander, E.A., and Galibert, E. 2004. CRH_Server: An online comparative and radiation hybrid mapping server for the canine genome. *Bioinformatics* **20**: 3665–3667.
- Hitte, C., Madeoy, J., Kirkness, E.F., Priat, C., Lorentzen, T.D., Senger, F., Thomas, D., Derrien, T., Ramirez, C., Scott, C., et al. 2005. Survey sequencing combined with dense radiation hybrid gene mapping facilitates genome navigation. *Nat. Rev. Genet.* **6**: 643–648.
- Howell, J.M., Fletcher, S., Kakulas, B.A., O'Hara, M., Lochmuller, H., and Karpati, G. 1997. Use of the dog model for Duchenne muscular dystrophy in gene therapy trials. *Neuromus. Disord.* **7**: 325–328.
- The International HapMap Consortium 2003. The International HapMap Project. *Nature* **426**: 789–796.
- Irion, D.N., Schaffer, A.L., Famula, T.R., Eggleston, M.L., Hughes, S.S., and Pedersen, N.C. 2003. Analysis of genetic variation in 28 dog breed populations with 100 microsatellite markers. *J. Hered.* **94**: 81–87.
- Kerns, J.A., Olivier, M., Lust, G., and Barsh, G.S. 2003. Exclusion of melanocortin-1 receptor (mc1r) and agouti as candidates for dominant black in dogs. *J. Hered.* **94**: 75–79.
- Kirkness, E.F., Bafna, V., Halpern, A.L., Levy, S., Remington, K., Rusch, D.B., Delcher, A.L., Pop, M., Wang, W., Fraser, C.M., et al. 2003. The dog genome: Survey sequencing and comparative analysis. *Science* **301**: 1898–1903.
- Koskinen, M.T. 2003. Individual assignment using microsatellite DNA reveals unambiguous breed identification in the domestic dog. *Anim. Genet.* **34**: 297–301.
- Koskinen, M.T. and Bredbacka, P. 2000. Assessment of the population structure of five Finnish dog breeds with microsatellites. *Anim. Genet.* **31**: 310–317.
- Kruglyak, L. 1999a. Genetic isolates: Separate but equal? *Proc. Natl. Acad. Sci.* **96**: 1170–1172.
- . 1999b. Prospects for whole-genome linkage disequilibrium mapping of common disease genes. *Nat. Genet.* **22**: 139–144.
- Leonard, J.A., Wayne, R.K., Wheeler, J., Valadez, R., Guillen, S., and Vila, C. 2002. Ancient DNA evidence for Old World origin of New World dogs. *Science* **298**: 1613–1616.
- Lin, L., Faraco, J., Li, R., Kadotani, H., Rogers, W., Lin, X., Qiu, X., Jong, P.J.D., Nishino, S., and Mignot, E. 1999. The sleep disorder canine narcolepsy is caused by a mutation in the hypocretin (orexin) receptor 2 gene. *Cell* **98**: 365–376.
- Lindblad-Toh, K., Wade, C.M., Mikkelsen, T.S., Karlsson, E.K., Jaffe, D.B., Kamal, M., Clamp, M., Chang, J.L., Kulbokas III, E.J., Zody, M.C., et al. 2005. Genome sequence, comparative analysis and haplotype structure of the domestic dog. *Nature* (in press).
- Lohi, H., Young, E.J., Fitzmaurice, S.N., Rusbridge, C., Chan, E.M., Vervoort, M., Turnbull, J., Zhao, X.C., Ianzano, L., Paterson, A.D., et al. 2005. Expanded repeat in canine epilepsy. *Science* **307**: 81.
- London, C.A. and Seguin, B. 2003. Mast cell tumors in the dog. *Vet. Clin. North Am. Small Anim. Pract.* **33**: 473–489, v.
- Luikart, G., England, P.R., Tallmon, D., Jordan, S., and Taberlet, P. 2003. The power and promise of population genomics: From genotyping to genome typing. *Nat. Rev. Genet.* **4**: 981–994.
- Mellersh, C.S., Langston, A.A., Acland, G.M., Fleming, M.A., Ray, K., Wiegand, N.A., Francisco, L.V., Gibbs, M., Aguirre, G.D., and Ostrander, E.A. 1997. A linkage map of the canine genome. *Genomics* **46**: 326–336.
- Mellersh, C.S., Hitte, C., Richman, M., Vignaux, F., Priat, C., Jouquand, S., Werner, P., Andre, C., DeRose, S., Patterson, D.F., et al. 2000. An integrated linkage-radiation hybrid map of the canine genome. *Mamm. Genome* **11**: 120–130.
- Modiano, J.F., Breen, M., Burnett, R.C., Parker, H.G., Inusah, S., Thomas, R., Avery, P.R., Lindblad-Toh, K., Ostrander, E.A., Cutter, G.C., et al. 2005. Distinct B and T cell lymphoproliferative disease prevalence among dog breeds indicates heritable risk. *Cancer Res.* **65**: 5654–5661.
- Moon-Fanelli, A.A. and Dodman, N.H. 1998. Description and development of compulsive tail chasing in terriers and response to clomipramine treatment. *J. Am. Vet. Med. Assoc.* **212**: 1252–1257.
- Morey, D. 1994. The early evolution of the domestic dog. *Am. Scientist* **82**: 336–347.
- Morey, D.F. 1992. Size, shape, and development in the evolution of the domestic dog. *J. Archaeol. Sci.* **19**: 181–204.
- Mount, J.D., Herzog, R.W., Tillson, D.M., Goodman, S.A., Robinson, N.,

- McClelland, M.L., Bellinger, D., Nichols, T.C., Arruda, V.R., Lothrop Jr., C.D., et al. 2002. Sustained phenotypic correction of hemophilia B dogs with a factor IX null mutation by liver-directed gene therapy. *Blood* **99**: 2670–2676.
- Nash, W.G., Menninger, J.C., Wienberg, J., Padilla-Nash, H.M., and O'Brien, S.J. 2001. The pattern of phylogenomic evolution of the Canidae. *Cytogenet. Cell Genet.* **95**: 210–224.
- Neff, M.W., Broman, K.W., Mellers, C.S., Ray, K., Acland, G.M., Aguirre, G.D., Ziegler, J.S., Ostrander, E.A., and Rine, J. 1999. A second-generation genetic linkage map of the domestic dog, *Canis familiaris*. *Genetics* **151**: 803–820.
- Neff, M.W., Robertson, K.R., Wong, A.K., Safra, N., Broman, K.W., Slatkin, M., Mealey, K.L., and Pedersen, N.C. 2004. Breed distribution and history of canine *mdr1-1Delta*, a pharmacogenetic mutation that marks the emergence of breeds from the collie lineage. *Proc. Natl. Acad. Sci.* **101**: 11725–11730.
- Nishino, S., Ripley, B., Overeem, S., Lammers, G.J., and Mignot, E. 2000. Hypocretin (orexin) deficiency in human narcolepsy. *Lancet* **355**: 39–40.
- Ophoff, R.A., Escamilla, M.A., Service, S.K., Spesny, M., Meshi, D.B., Poon, W., Molina, J., Fournier, E., Gallegos, A., Mathews, C., et al. 2002. Genomewide linkage disequilibrium mapping of severe bipolar disorder in a population isolate. *Am. J. Hum. Genet.* **71**: 565–574.
- Ostrander, E.A. and Giniger, E. 1997. Semper fidelis: What man's best friend can teach us about human biology and disease. *Am. J. Hum. Genet.* **61**: 475–480.
- Ostrander, E.A., Rine, J., Sack Jr., G.H., and Cork, L.C. 1993. What is the role of molecular genetics in modern veterinary practice? *J. Am. Vet. Med. Assoc.* **203**: 1259–1262.
- Ostrander, E.A., Galibert, F., and Patterson, D.F. 2000. Canine genetics comes of age. *Trends Genet.* **16**: 117–124.
- Parker, H.G., Kim, L.V., Sutter, N.B., Carlson, S., Lorentzen, T.D., Malek, T.B., Johnson, G.S., DeFrance, H.B., Ostrander, E.A., and Kruglyak, L. 2004. Genetic structure of the purebred domestic dog. *Science* **304**: 1160–1164.
- Patterson, D. 2000. Companion animal medicine in the age of medical genetics. *J. Vet. Internal Med.* **14**: 1–9.
- Patterson, D.F., Haskins, M.E., and Jezyk, P.F. 1982. Models of human genetic disease in domestic animals. *Adv. Hum. Genet.* **12**: 263–339.
- Pennisi, E. 2000. Human genome. Finally, the book of life and instructions for navigating it. *Science* **288**: 2304–2307.
- Philipp, U., Hamann, H., Mecklenburg, L., Nishino, S., Mignot, E., Gunzel-Apel, A.R., Schmutz, S.M., and Leeb, T. 2005. Polymorphisms within the canine *MLPH* gene are associated with dilute coat color in dogs. *BMC Genet.* **6**: 34.
- Pollinger, J.P., Bustamante, C.D., Fledel-Alon, A., Schmutz, S., Gray, M.M., and Wayne, R.K. 2005. Selective sweep mapping of genes with large phenotype effects. *Genome Res.* (this issue).
- Ponder, K.P., Melniczek, J.R., Xu, L., Weil, M.A., O'Malley, T.M., O'Donnell, P.A., Knox, V.W., Aguirre, G.D., Mazrier, H., Ellinwood, N.M., et al. 2002. Therapeutic neonatal hepatic gene therapy in mucopolysaccharidosis VII dogs. *Proc. Natl. Acad. Sci.* **99**: 13102–13107.
- Porrello, A., Cardelli, P., and Spugnini, E.P. 2004. Pet models in cancer research: General principles. *J. Exp. Clin. Cancer Res.* **23**: 181–193.
- Priat, C., Hitte, C., Vignaux, F., Renier, C., Jiang, Z., Jouquand, S., Cheron, A., Andre, C., and Galibert, F. 1998. A whole-genome radiation hybrid map of the dog genome. *Genomics* **54**: 361–378.
- Pritchard, J.K., Stephens, M., Rosenberg, N.A., and Donnelly, P. 2000. Association mapping in structured populations. *Am. J. Hum. Genet.* **67**: 170–181.
- Reich, D.E., Cargill, M., Bolk, S., Ireland, J., Sabeti, P.C., Richter, D.J., Lavery, T., Kouyoumjian, R., Farhadian, S.F., Ward, R., et al. 2001. Linkage disequilibrium in the human genome. *Nature* **411**: 199–204.
- Rogers, C.A. and Brace, A.H. 1995. *The International Encyclopedia of Dogs*. Howell Book House, New York.
- Roy, M.S., Geffen, E., Smith, D., and Wayne, R.K. 1996. Molecular genetics of pre-1940 red wolves. *Conservation Biol.* **10**: 1413–1424.
- Saetre, P., Lindberg, J., Leonard, J.A., Olsson, K., Pettersson, U., Ellegren, H., Bergstrom, T.F., Vila, C., and Jazin, E. 2004. From wild wolf to domestic dog: Gene expression changes in the brain. *Brain Res. Mol. Brain Res.* **126**: 198–206.
- Sargan, D. 2004. IDID: Inherited diseases in dogs: Web-based information for canine inherited disease genetics. *Mamm. Genome* **15**: 503–506.
- Savolainen, P., Zhang, Y.P., Luo, J., Lundeberg, J., and Leitner, T. 2002. Genetic evidence for an East Asian origin of domestic dogs. *Science* **298**: 1610–1613.
- Savolainen, P., Leitner, T., Wilton, A.N., Matisoo-Smith, E., and Lundeberg, J. 2004. A detailed picture of the origin of the Australian dingo, obtained from the study of mitochondrial DNA. *Proc. Natl. Acad. Sci.* **101**: 12387–12390.
- Schmutz, S.M., Berryere, T.G., and Goldfinch, A.D. 2002. TYRP1 and MC1R genotypes and their effects on coat color in dogs. *Mamm. Genome* **13**: 380–387.
- Stuehler, B., Reichert, J., Stremmel, W., and Schaefer, M. 2004. Analysis of the human homologue of the canine copper toxicosis gene *MURR1* in Wilson disease patients. *J. Mol. Med.* **82**: 629–634.
- Sundin, O.H., Yang, J.M., Li, Y., Zhu, D., Hurd, J.N., Mitchell, T.N., Silva, E.D., and Maumenee, I.H. 2000. Genetic basis of total colourblindness among the Pingelapese islanders. *Nat. Genet.* **25**: 289–293.
- Sutter, N.B. and Ostrander, E.A. 2004. Dog star rising: The canine genetic system. *Nat. Rev. Genet.* **5**: 900–910.
- Sutter, N.B., Eberle, M.A., Parker, H.G., Pullar, B.J., Kirkness, E.F., Kruglyak, L., and Ostrander, E.A. 2004. Extensive and breed-specific linkage disequilibrium in *Canis familiaris*. *Genome Res.* **14**: 2388–2396.
- Switonski, M., Szczerbal, I., and Nowacka, J. 2004. The dog genome map and its use in mammalian comparative genomics. *J. Appl. Genet.* **45**: 195–214.
- Thannickal, T.C., Moore, R.Y., Nienhuis, R., Ramanathan, L., Gulyani, S., Aldrich, M., Cornford, M., and Siegel, J. 2000. Reduced number of hypocretin neurons in human narcolepsy. *Neuron* **27**: 469–474.
- Thomas, R., Bridge, W., Benke, K., and Breen, M. 2003a. Isolation and chromosomal assignment of canine genomic BAC clones representing 25 cancer-related genes. *Cytogenet. Genome Res.* **102**: 249–253.
- Thomas, R., Fiegler, H., Ostrander, E.A., Galibert, F., Carter, N.P., and Breen, M. 2003b. A canine cancer-gene microarray for CGH analysis of tumors. *Cytogenet. Genome Res.* **102**: 254–260.
- van de Sluis, B.J., Breen, M., Nanji, M., van Wolferen, M., de Jong, P., Binns, M.M., Pearson, P.L., Kuipers, J., Rothuizen, J., Cox, D.W., et al. 1999. Genetic mapping of the copper toxicosis locus in Bedlington terriers to dog chromosome 10, in a region syntenic to human chromosome region 2p13-p16. *Hum. Mol. Genet.* **8**: 501–507.
- van de Sluis, B., Rothuizen, J., Pearson, P.L., van Oost, B.A., and Wijmenga, C. 2002. Identification of a new copper metabolism gene by positional cloning in a purebred dog population. *Hum. Mol. Genet.* **11**: 165–173.
- Vignaux, F., Hitte, C., Priat, C., Chuat, J.C., Andre, C., and Galibert, F. 1999. Construction and optimization of a dog whole-genome radiation hybrid panel. *Mamm. Genome* **10**: 888–894.
- Vila, C., Savolainen, P., Maldonado, J.E., Amorim, I.R., Rice, J.E., Honeycutt, R.L., Crandall, K.A., Lundeberg, J., and Wayne, R.K. 1997. Multiple and ancient origins of the domestic dog. *Science* **276**: 1687–1689.
- Vila, C., Seddon, J., and Ellegren, H. 2005. Genes of domestic mammals augmented by backcrossing with wild ancestors. *Trends Genet.* **21**: 214–218.
- Wade, C., Karlsson, E.K., Mikkelsen, J., Kulbokas, E.J.I., Zody, M.C., and Lindblad-Toh, K. 2005. The dog genome: Sequence, evolution and haplotype structure. In *The genome of the domestic dog* (eds. E.A. Ostrander, K. Lindblad-Toh, and U. Giger) Cold Spring Harbor Press, Cold Spring Harbor, New York (in press).
- Wayne, R.K. 1986a. Cranial morphology of domestic and wild canids: The influence of development on morphological change. *Evolution* **40**: 243–261.
- . 1986b. Developmental constraints on limb growth in domestic and some wild canids. *J. Zool.* **210**: 381–399.
- . 1986c. Limb morphology of domestic and wild canids: The influence of development on morphologic change. *J. Morphol.* **187**: 301–319.
- Wayne, R.K. and Jenks, S.M. 1991. Mitochondrial DNA analysis implying extensive hybridization of the endangered red wolf *Canis rufus*. *Nature* **351**: 565–568.
- Wayne, R.K., Nash, W.G., and O'Brien, S.J. 1987a. Chromosomal evolution of the Canidae. I. Species with high diploid numbers. *Cytogenet. Cell Genet.* **44**: 123–133.
- . 1987b. Chromosomal evolution of the Canidae. II. Divergence from the primitive carnivore karyotype. *Cytogenet. Cell Genet.* **44**: 134–141.
- Wayne, R.K., Van Valkenburgh, B., Kat, P.W., Fuller, T.K., Johnson, W.E., and O'Brien, S.J. 1989. Genetic and morphological divergence among sympatric canids. *J. Hered.* **80**: 447–454.

- Wayne, R.K., Geffen, E., Girman, D.J., Koepfli, K.P., Lau, L.M., and Marshall, C.R. 1997a. Molecular systematics of the Canidae. *Syst. Biol.* **46**: 622–653.
- Wayne, R.K., Geffen, E., Girman, D.J., Koepfli, K.P., Lau, L.M., and Marshall, C.R. 1997b. Molecular systematics of the Canidae. *System. Biol.* **46**: 622–653.
- Wayne, R.K., Leonard, J.A., and Vila, C. 2006. Genetic analysis of dog domestication. In *Documenting domestication: New genetic and archeological paradigms* (ed. M.E. Zeder). Smithsonian Institution Press, Washington, DC. (in press).
- Weiss, K.M. and Clark, A.G. 2002. Linkage disequilibrium and the mapping of complex human traits. *Trends Genet.* **18**: 19–24.
- Wilcox, B. and Walkowicz, C. 1995. *Atlas of dog breeds of the world*. T.F.H. Publications, Neptune City, NJ.
- Wilson, P.J., Grewal, S., Lawford, I.D., Heal, J.N.M., Granacki, A.G., Pennock, D., Theberge, J.B., Theberge, M.T., Voigt, D.R., Waddell, W., et al. 2000. DNA profiles of the eastern Canadian wolf and the red wolf provide evidence for a common evolutionary history independent of the gray wolf. *Canadian J. Zool.* **78**: 2156–2166.
- Yuzbasiyan-Gurkan, V., Blanton, S.H., Cao, V., Ferguson, P., Li, J., Venta, P.J., and Brewer, G.J. 1997. Linkage of a microsatellite marker to the canine copper toxicosis locus in Bedlington terriers. *Am. J. Vet. Res.* **58**: 23–27.
- Zajc, I., Mellersh, C.S., and Sampson, J. 1997. Variability of canine microsatellites within and between different dog breeds. *Mamm. Genome* **8**: 182–185.

Web site references

- <http://www.broad.mit.edu/mammals/dog/snp/>; Broad Institute's Dog SNP library.
- http://research.nhgri.nih.gov/dog_genome/; The NHGRI Dog Genome Project.
- <http://www.genome.ucsc.edu/>; UCSC Genome Bioinformatics.
- <http://www.georgieproject.com/>; The Georgie Project home page.
- <http://sun-recomgen.med.univ-rennes1.fr/Dogs/>; Canine Radiation Hybrid Project.
- <http://www.vet.cam.ac.uk/idid/>; IDID; Database of inherited diseases in dogs.