DOG STAR RISING: THE CANINE GENETIC SYSTEM

Nathan B. Sutter and Elaine A. Ostrander

Abstract | Purebred dogs are providing invaluable information about morphology, behaviour and complex diseases, both of themselves and humans, by supplying tractable populations in which to map genes that control those processes. The diversification of dog breeds has led to the development of breeds enriched for particular genetic disorders, the mapping and cloning of which have been facilitated by the availability of the canine genome map and sequence. These tools have aided our understanding of canine population genetics, linkage disequilibrium and haplotype sharing in the dog, and have informed ongoing efforts of the need to identify quantitative trait loci that are important in complex traits.

COMPLEX TRAIT
A measured phenotype, such as disease status or a quantitative character, that is influenced by many environmental and genetic factors, and potentially by interactions in and between them.

National Human Genome Research Institute, National Institutes of Health, 50 South Drive, MSC8002, Building 50, Room 5222, Bethesda, Maryland 20892, USA. Correspondence to E.A.O. e-mail: eostrand@mail.nih.gov
doi:10.1038/nrg1492

Almost 38 million households in the United States (36%) own at least one dog[1], and it is estimated that 55% of these dogs are purebred. The American Kennel Club (AKC), the most recognizable purebred registry in the United States, registers 154 breeds of dog, with the 20 most popular breeds making up over 70% of registrations (FIGS 1,2). The total health-care expenditure on dogs is nearly 20 billion US$ (REF. 1); the level of medical surveillance and care that pet dogs receive is second only to that to which we treat ourselves[2]. Nearly half of genetic diseases reported in dogs occur predominantly or exclusively in one or a few breeds[2,3]. Susceptibility of some breeds to particular diseases such as types of cancer, deafness, forms of blindness, cataracts or metabolic disorders, coupled with a near absence in other breeds indicates that a subset of dog breeds are strongly enriched for particular disease alleles (BOX 1). Such an enrichment can be caused by origination from a small group of founders, population bottlenecks and popular-sire effects[4]. But this enrichment will only occur when the number of risk alleles is small, and they are relatively rare in the overall population. Dog breeds are therefore similar to geographically isolated human populations, such as those from Finland or Iceland, except that the isolation is more extreme. This offers an enormous advantage in the search for genes associated with complex diseases, which, in theory, can be more easily mapped using dog families than human families.

Here we describe the growing role of modern dog genetics in improving human health through the use of purebred dogs to uncover genes that are important in development, behaviour and disease susceptibility. First, we summarize what is known about the domestication of the dog from the wolf. We then discuss the canine genome map and sequence, focusing on comparative aspects, and how our understanding of dog population genetics can inform the mapping of human disease genes, including COMPLEX TRAITS. Each of these advances is tied to recent and rapid progress by the research community in mapping, and ultimately sequencing, the canine genome. As information about dog health and biology grows, knowledge about our own species also advances. In the following discussion we describe work that, because of its timeliness or innovation, has moved dog genetics forwards and has contributed to the development of purebred dogs as a new and unique genetic system.

Origin and domestication of the dog

In ancient Egypt, the appearance of Sirius, the dog star, heralded the onset of the Nile flood that brought prosperity and life. The dog is the oldest domesticated species and has served man by herding flocks, guarding homes, hunting by sight and scent, retrieving wild animals, pulling sledges and simply by providing companionship. The dog was the only domesticated

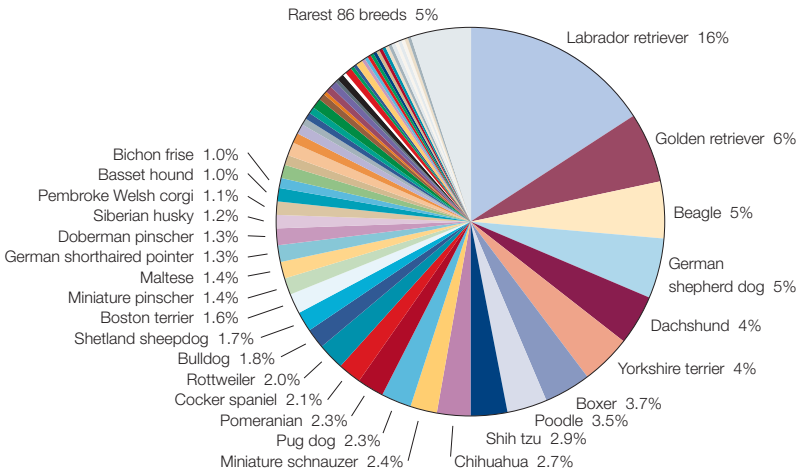

Figure 1 | **Proportion of purebred dog registrations by the American Kennel Club (AKC).** A total of 916,000 purebred dogs were newly registered in 2003 in the United States by the AKC, which is the largest purebred dog registry in the United States and that recognizes over 150 distinct breeds. The most popular 20 breeds account for 70% of all registrations. The rarest 86 breeds cumulatively account for just 5% of registrations. Many rare breeds newly register fewer than 100 animals per year.

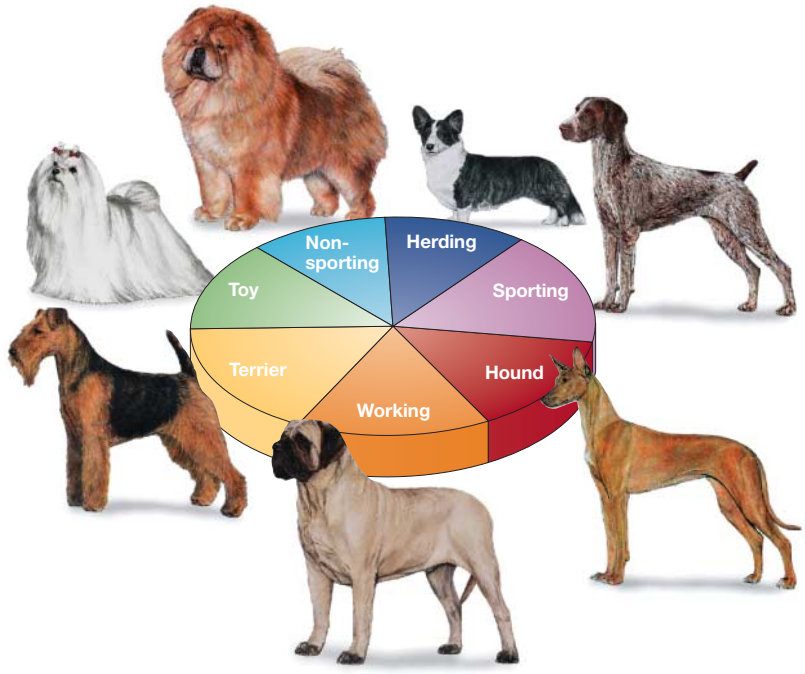

Figure 2 | **Grouping of dog breeds by the American Kennel Club (AKC).** The American Kennel Club categorizes dog breeds into seven groups based on historical development, morphology and behaviour. Representative breeds from each of the AKC groupings are shown. The German shorthaired pointer is a member of the sporting group, which includes breeds used for hunting and retrieval of wild animals. The Pharaoh hound represents the hound group, whose members hunt independently using sight and speed (greyhound) or in packs (beagle) or by scent (bloodhound). The mastiff is a member of the working group, a set of breeds used for guarding, fighting and drafting. Many of these breeds are heavily muscled, sharing the 'molosser' morphology. The Welsh terrier depicts the terrier group, a set of breeds used primarily for vermin extermination. The Maltese represents the toy group, which features breeds of small body size. Many share morphology with larger breeds in other groups (for example, the Italian greyhound is a toy with the sighthound morphology). The chow chow is a member of the non-sporting group, a miscellaneous cluster of breeds with little shared morphology or history. The cardigan Welsh corgi is a herding dog, a group defined functionally as breeds used to herd and guard animals.

species populating both the Old and New World five centuries ago, when Europeans found the Americas[5]. So, dogs either came to the Americas with humans from the Old World or they were domesticated independently in the Old and the New World. The phenotype variation in dogs led Charles Darwin to argue that different breeds had been independently domesticated from wild CANID species on different continents, perhaps descending from wolf, coyote and several jackal species[6]. Mitochondrial sequence data, however, show that dogs descend solely from wolves. In a key experiment, Vila *et al.*[7] analysed 261 bp of mitochondrial sequence from dogs, wolves and other wild canids, and showed that dog and wolf mitochondrial HAPLOTYPES were similar. Indeed, the most dissimilar haplotypes between the two species differed by no more than 12 substitutions. By contrast, the minimum number of differences between dog and other wild canid haplotypes was 20 substitutions[7]. Subsequent phylogenetic analysis of mitochondrial haplotypes identified four canid CLADES, including one containing 19 of the 26 dog haplotypes, but no wolf haplotypes. The existence of a dog-specific clade with high haplotype diversity supports a relatively ancient time frame for the domestication of dogs, perhaps >100,000 years before the present[7].

Mitochondrial sequence data have also been used to refute the New World domestication hypothesis. Mitochondrial DNA was sequenced from archaeological dog specimens collected throughout the Americas and dating before the arrival of Europeans. The haplotypes fit with those from modern breeds, which are known to have come from Old World domestication. They also included a large number of haplotypes assigned to one of the dog-specific clades[5]. Therefore, the haplotypes do not provide any evidence for an independent domestication in the New World. Savolainen and colleagues refined this picture by examining mitochondrial haplotypes from 654 dogs representing main populations around the world[8]. They report a similar phylogeny in every geographic region, suggesting a single domestication event in one geographic region. The greatest haplotype diversity occurs in East Asian dog populations, consistent with the notion that the key domestication event occurred in Asia[8].

Beyond the time and place of dog domestication, there is interest in how domestication occurred. In what ways do all domestic dogs differ from wolves? This has been addressed recently by Hare and colleagues, who studied how dogs and wolves differ in their ability to relate to humans[9]. They devised a study in which dogs were challenged to find a hidden food reward, aided by a human who stared and pointed at the treat's hiding place. Adult dogs and puppies, even those raised with little human interaction, did significantly better than wolves. Dogs have therefore acquired a social-cognitive capability to 'read' humans, which is lacking in wolves[9]. Identifying the genes that regulate such evolutionary events will be both challenging and exciting.

The canine genome

Early support from a number of non-traditional sources provided the resources needed to develop meiotic linkage[10–12], RADIATION HYBRID maps[13] and integrated maps of the canine genome[14,15] (TABLE 1). Analysis of genotyping data from several disease-mapping efforts[16,17] indicated that in the dog, 1 cM corresponds to about 800 kb in physical distance[18].

A particular challenge has been the establishment of a nomenclature for dog chromosomes, most of which are small and ACROCENTRIC. In 1996, Switonski and colleagues initially named the larger chromosomes of the dog[19], but to unambiguously order dog chromosomes it was necessary to use reciprocal chromosome paint studies[20–22] using FLOW-SORTED canine chromosomes[23]. For this reason, the first complete DAPI-BANDED karyotype,

Box 1 | Top diseases of domestic dogs and breeds at risk*

Cancers
Airedale terrier | akita | American eskimo dog | Belgian malinois | bloodhound | boxer | briard | canaan dog | curley coated retriever | dandie dinmont terrier | English foxhound | English setter | flax coated | retriever | German wirehaired pointer | great dane | greyhound | Irish water spaniel | Irish wolfhound | Japanese chin | kuvasz | otterhound | pembroke Welsh corgi | Portuguese water dog | Rhodesian ridgeback | rottweiler | Scottish deerhound | Scottish terrier | Skye terrier | soft coated wheaten terrier | Staffordshire terrier | Tibetan terrier | vizsla

Epilepsy
Australian terrier | Belgian malinois | Belgian tervuren | Boston terrier | canaan dog | Chesapeake Bay retriever | clumber spaniel | collie | curley coated retriever | dachshund | dalmatian | English foxhound | English springer spaniel | English toy spaniel | Field spaniel | giant schnauzer | great pyrenees | greater Swiss mountain dog | harrier | Irish setter | Irish water spaniel | Italian greyhound | Japanese chin | miniature pincher | Newfoundland | otterhound | petit basset griffon vendeen | poodle | Portuguese water dog | pug | schipperke | Sealyham terrier | Shetland sheepdog | Siberian husky | vizsla | Welsh springer spaniel | Welsh terrier

Hip dysplasia
Airedale terrier | American eskimo dog | American water spaniel | basset hound | Belgian malinois | black and tan coonhound | bloodhound | Chesapeake Bay retriever | clumber spaniel | curly coated retriever | English foxhound | English setter | field spaniel | flat coated retriever | German wirehaired pointer | giant schnauzer | great dane | greyhound | harrier | mastiff | Newfoundland | otterhound | pembroke Welsh corgi | pug | Rhodesian ridgeback | rottweiler | Shetland sheepdog | Staffordshire bull terrier | Staffordshire terrier | Tibetan terrier | weimaraner | Welsh springer spaniel

Thyroid disease
Akita | American eskimo dog | Australian terrier | basset hound | Belgian tervuren | canaan dog | Chesapeake Bay retriever | dachshund | dandie dinmont terrier | English foxhound | English terrier | English springer spaniel | field spaniel | German wirehaired pointer | giant schnauzer | greyhound | Irish water spaniel | Italian greyhound | kuvasz | Maltese | miniature pincher | Norwegian elkhound | Rhodesian ridgeback | Scottish deerhound | Shetland sheepdog | Siberian husky | Sussex spaniel | Tibetan terrier | vizsla | Welsh springer spaniel

Allergies
Airedale terrier | American water spaniel | Australian terrier | bichon frise | black and tan coonhound | Boston terrier | bull terrier | Chinese shar-pei | dalmatian | English springer spaniel | French bulldog | Irish setter | Irish water spaniel | kuvasz | otterhound | Scottish deerhound | Staffordshire bull terrier

Bloat
Akita | bloodhound | briard | collie | curly coated retriever | English foxhound | great dane | greater Swiss mountain dog | greyhound | Irish setter | Irish wolfhound | komondor | poodle | Scottish deerhound | Sussex spaniel | weimaraner

Heart disease
American water spaniel | Belgian malinois | bloodhound | briard | bull terrier | English toy spaniel | field spaniel | Irish wolfhound | Japanese chin | Maltese | mastiff | miniature bull terrier | Staffordshire terrier | Sussex spaniel | West Highland white terrier

Autoimmune diseases
Airedale terrier | akita | Belgian tervuren | briard | canaan dog | English cocker spaniel | German wirehaired pointer | Italian Greyhound | maltese | pembroke Welsh corgi | petit basset griffon vendeen | Skye terrier | Staffordshire terrier

Progressive retinal atrophy
Airedale terrier | American eskimo dog | basenji | Belgian tervuren | Chesapeake Bay retriever | dachshund | English springer spaniel | Italian greyhound | mastiff | papillon | Portuguese water dog | Tibetan terrier

Cataracts
Bedlington terrier | bichon frise | black and tan coonhound | Boston terrier | Cairn terrier | Chesapeake Bay retriever | Havanese | Norwegian elkhound | Siberian husky | Staffordshire bull terrier | Tibetan terrier

*Information provided by the American Kennel Club Canine Health Foundation based on surveys of purebred dog owners and breeders.

CANID
A member of the Canidae family of carnivorous mammals that include the wolves, jackals, foxes, coyotes and the domestic dog.

HAPLOTYPE
An experimentally determined profile of genetic markers that is present on a single chromosome of any given individual.

CLADE
A taxon or other grouping of organisms consisting of a single species and its descendents.

RADIATION HYBRID (MAPPING)
A determination of marker order along chromosomes. This is done by assessing the presence or absence of alleles associated with markers in a set of hybrid cell lines, each of which carries a distinct portion of the genome in a rodent background.

ACROCENTRIC
A chromosome in which the centromere lies near to one end, such that one arm of the chromosome is much larger than the other.

FLOW SORTING
The analysis of single cells or subcellular particles by the detection of their light absorption, scattering and/or fluorescence properties as they pass through a laser beam in a directed fluid stream.

DAPI-BANDING
The pattern created from treatment with the sensitive fluorescent probe for DNA, 4′6-diamidino-2-phenylindole-2HCl, used in fluorescence microscopy.

Table 1 | **Dog genetic resources**

Description	Website/reference
Genome sequence	
NCBI's dog resources page. Links to most of the resources below	http://www.ncbi.nlm.nih.gov/genome/guide/dog
White paper arguing for full sequencing of the dog genome	http://www.genome.gov/Pages/Research/Sequencing/SeqProposals/CanineSEQedited.pdf
UCSC genome browser. Viewing the boxer 7x assembled sequence	http://www.genome.ucsc.edu
Dog gene index at The Institute for Genomic Research	http://www.tigr.org/tigr-scripts/tgi/T_index.cgi?species=dog
Olfactory receptor gene family in canines	http://bip.weizmann.ac.il/HORDE; http://www-recomgen.univ-rennes1.fr/Dogs/OR.html
Mapping	
Radiation hybrid map	http://www-recomgen.univ-rennes1.fr/ Dogs/maquette-pnas.html; http://www.fhcrc.org/science/dog_genome
Meiotic linkage map	http://www.fhcrc.org/science/dog_genome
Cytogenetic map	http://www2.ncsu.edu/unity/lockers/project/cvmaprhome/breen_matthew.htm
Comparative and radiation hybrid mapping server	http://idefix.univ-rennes1.fr:8080/Dogs/rh-server.html
Radiation hybrid panel of cell lines	http://www-recomgen.univ-rennes1.fr/Dogs/outils.html
Multiplexed microsatellites for genome scans	http://www.cvm.tamu.edu/cgr/multiplex.html
Canine reference families at Purina	http://www.purina.com/us/institute/whoweare.asp?article=209
BAC library made from a doberman pinscher	http://bacpac.chori.org/mcanine81.htm
Inherited diseases	
Canine Inherited Disorders Database	http://www.upei.ca/~cidd/intro.htm
Online Mendelian Inheritance in Animals database	http://www.angis.org.au/Databases/BIRX/omia
Inherited Diseases in Dogs Database	http://server.vet.cam.ac.uk:591/index.html
Canine Genetic Disease Information System	100
American Kennel Club — Canine Health Foundation. A list of supported grants for canine diseases	http://www.akcchf.org
Background on breeds	
American Kennel Club	http://www.akc.org

BAC, bacterial artificial chromosome; NCBI, The National Center for Biotechnology Information; UCSC, University of California, Santa Cruz.

enumerating all 38 autosomes and sex chromosomes, was not standardized until 1999 (REF. 20).

In 2003, using a 5,000-rad panel[24], we published a comprehensive radiation hybrid map of the dog. The map was composed of 3,270 markers and an average inter-marker distance of 1 Mb (REF. 25). With recent additions, the map now stands at 4,249 markers, including 1,760 BAC (bacterial artificial chromosome) ends, 1,589 MICROSATELLITE markers, and featuring one marker every 900 kb (REF. 26). A subset of 804 BACs have also been localized by fluorescence *in situ* hybridization (FISH), providing an unusually high level of confidence in map order. Canine chromosome 1 and segments from human chromosomes 18, 6, 9 and 19 are compared in FIG. 3. A multiplexed set of microsatellite markers at a density of 9 Mb is also available for genome-wide scans[27].

Sequencing the dog genome

Although the radiation hybrid map has proven invaluable for identification of preliminary loci associated with various diseases[28], the complete sequence of the canine genome was needed to propel the field forward. In 2003, a 1.5x sequence of the dog generated from a male standard poodle was produced[29]. The finished sequence consisted of 6.22 million reads (mean read length 576 bp), representing about 1.5x coverage of the 2.5-Gb haploid canine genome[30]. When assembled, the final output consisted of 522,101 SCAFFOLDS with a mean length of 3.8 kb and a mean span of 8.6 kb (REF. 29).

Although it was only 1.5x, the sequence demonstrated several interesting features of the canine genome. First, nucleotide conservation between the dog and other mammalian genomes is high. Despite the fact that the dog lineage was the first to diverge from the common ancestor of human–dog–mouse[31,32], the genomes of dog and human show a higher level of nucleotide conservation than those of human and mouse. So, whereas the 1.5x coverage represents only ~78% of the dog genome, the BLASTN best-hit for these sequences provides 649-Mb coverage of unique human sequence, which is nearly twice that observed for the near-complete mouse genome.

Segments of conserved synteny are those that contain adjacent markers in the same order and orientation, whereas blocks of synteny contain one or more segments that are adjacent in one genome, but might be shuffled in order or orientation. Comparison of the human and dog genomes by RECIPROCAL ZOO-FISH[20–22] revealed ~70 syntenic blocks, slightly less than the 85 blocks indicated by the 1-Mb radiation hybrid map[25].

MICROSATELLITE
A class of repetitive DNA that is made up of repeats that are 2–8 nucleotides in length. They can be highly polymorphic and are frequently used as molecular markers in population genetics studies.

SCAFFOLD
A portion of the genome sequence composed of contigs and gaps and reconstructed from end-sequenced whole-genome shotgun clones.

BLASTN
Basic local alignment search tool that compares a nucleotide query sequence against a nucleotide sequence database.

RECIPROCAL ZOO-FISH
Bidirectional heterologous chromosome painting using fluorescence *in situ* hybrization.

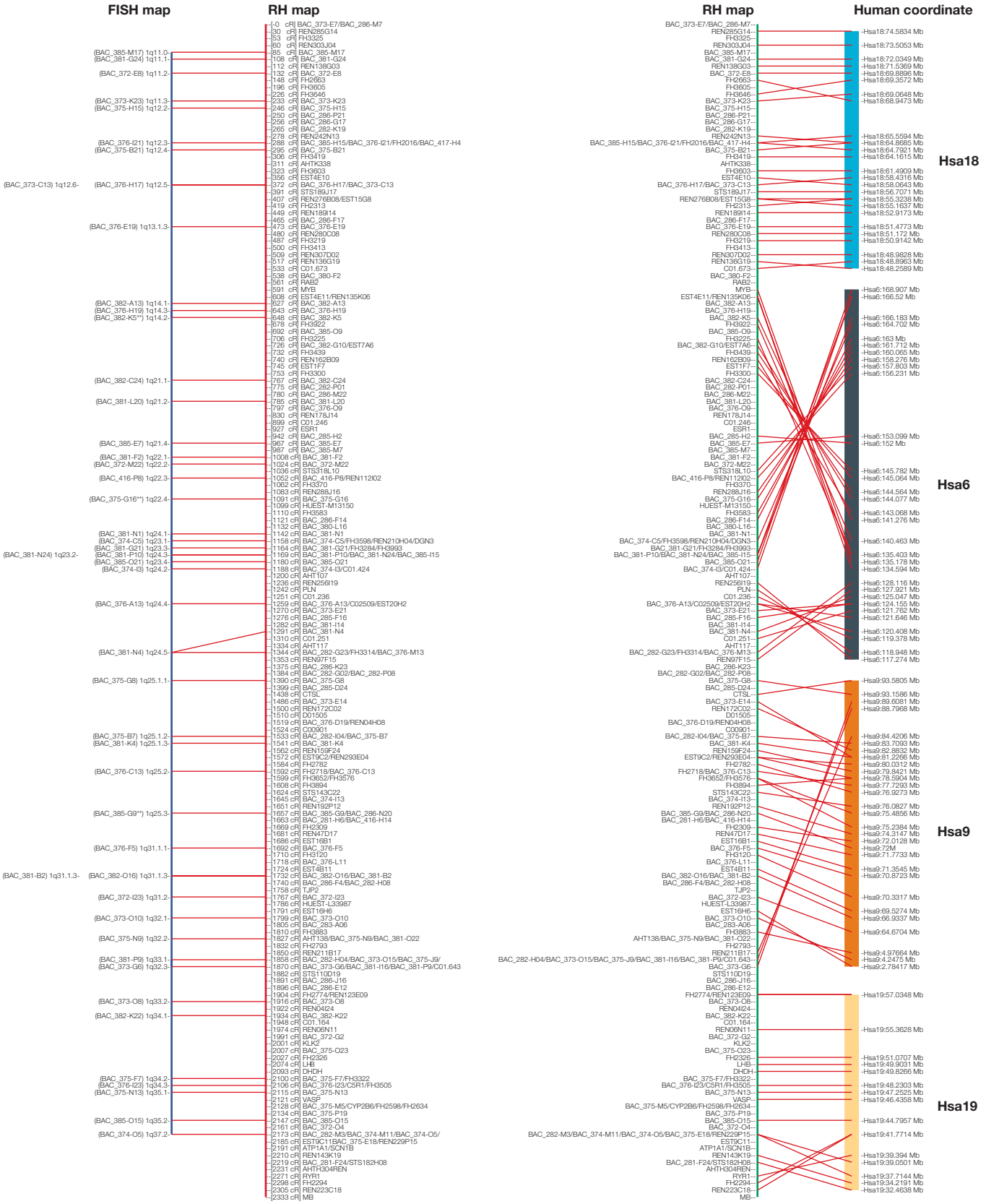

Figure 3 | **Comparative map of dog chromosome 1 and orthologous segments of human chromosomes (Hsa) 18, 6, 9 and 19.** On the left panel, BACs (bacterial artificial chromosomes) mapped by both fluorescence *in situ* hybridization (FISH) and radiation hybrid (RH) mapping are shown, demonstrating high concordance between the two mapping methods. On the right panel, canine BACs mapped by RH (left) are compared with the human coordinates for orthologous sequences (right, coloured bars). Dense RH mapping demonstrates that there are multiple rearrangements within larger conserved segments. Reproduced, with permission, from REF. 26 © (2004) BioMed Central Ltd. Asterisks (**) represent BAC clones that did not have a unique cytogenic location.

Analysis of the 1.5x sequence, however, indicates that the number is closer to 200, with large conserved segments containing many minor rearrangements that are not found by radiation hybrid mapping.

To develop the much-needed high resolution comparative map of the dog, Francis Galibert at the University of Rennes, Panos Deloukas at the Sanger Center, Ewen Kirkness at the Institute for Genomic Research, and our own laboratory have collaborated on the construction of a high-density radiation hybrid map, which places over 10,000 individual genes selected from the 1.5x sequence on a 9,000-rad dog radiation hybrid panel (F. Galibert, P. Deloukas, E. Kirkness & E.A.O., unpublished data). Although it is still under analysis, it is clear that this approach, encompassing what is likely to be over one-third of the individual genes in the genome, will delineate the breakpoints of nearly all syntenic blocks.

Finally, the 1.5x sequence has provided a great deal of information about repeat elements in the canine genome. Approximately 31% of the dog genome consists of repetitive sequences, less than the 46% and 38% that are found in the human and mouse genomes, respectively. An abundant repeat, a SINE element termed SINEC_CF, seems to be derived from tRNA sequences and many copies are not yet fixed in the dog genome. An estimated 7% (16,000 of 230,000 copies) of the SINEC_CF elements represent bimorphic insertions in the sequence that was derived from the standard poodle. By comparison, the number of bimorphic *ALU* INSERTIONS in the human genome is estimated to be only 1,200.

Two years ago, the canine genome community submitted a White Paper to request 6x sequencing of the dog genome (see links to this proposal in the Online links box). Sequencing of an anonymous female boxer dog began at the Broad Institute in June 2003. The recently completed effort (see links to Genbank and UCSC Genome Bioinformatics Site in the Online links box), released in July of 2004, generated 35 million sequence reads and a 7.8x sequence of the dog. The N50 CONTIG size is 121 kb and the N50 supercontig size is 42 Mb. Accompanying the primary sequencing effort are one million additional reads from nine breeds of dogs, which are currently being annotated. These reads will be compared with the reference sequence for developing a SNP resource.

Genetics of canine and human diseases

In the interest of space we can do little more than survey the extensive literature on common canine diseases. Of greatest relevance, however, is that the top ten diseases in purebred dogs include several that are of concern to human health, such as cancer, epilepsy, autoimmune diseases, blindness, cataracts and heart disease (see links to online disease databases in TABLE 1). To date, loci have been mapped, and in some cases identified, for several of these diseases, including vision disorders[16,17,33,34], kidney cancer[35], NARCOLEPSY[36], rheumatoid arthritis[37], SCID[38], keratin-associated diseases[39], CYSTINURIA[40], bleeding disorders[41,42], CEROID LIPOFUSCINOSIS[43] and copper toxicosis[44,45].

Molecular cytogenetics is an important development in canine cancer studies. Comparative genomic hybridization (CGH) studies, using either a chromosome-based approach[46] or an array of canine BAC clones (array-CGH)[47] can now be used to assay the patterns of chromosomal changes that occur in tumours. The aim of such studies is to improve both the diagnosis and therapeutic intervention for the most common cancers that afflict dogs. Such detailed genetic dissection will also inform us of the link between canine and human tumours.

Three particular success stories document the importance of the canine model in facilitating our understanding of human health, with each one illustrating distinct strengths of the canine system. The three disorders that these are centred on — narcolepsy, hereditary kidney cancer and blindness — are described below.

Narcolepsy. The first example, the cloning of the canine narcolepsy gene, demonstrates how the study of rare canine disorders can lead to an understanding of the molecular biology of common cellular processes[36]. Although narcolepsy is a relatively rare disease in humans, sleep disorders in general are common in the human population. In doberman pinschers, a highly penetrant and rare form of narcolepsy was found to be caused by a mutation in the hypocretin-2-receptor gene *HCRTR2* (REF. 36). Shortly thereafter, clinical studies established that hypocretin deficiency is associated with most cases of narcolepsy in humans[48–50]. Ongoing studies indicate that hypocretin might have a key role in circadian clock-dependent alertness as well as integrating hypothalamic signals for neuroendocrine release and for regulating metabolic rate, appetite, mood and sleep. These findings indicate that the hypocretin system might be a therapeutic target not only in the treatment of narcolepsy, but also for more common sleep disturbances. This study therefore provides a clear example of how the understanding of common human diseases can be facilitated by studies of rare genetic diseases in dogs.

Hereditary kidney cancer. A second success story, the mapping and cloning of a gene for hereditary kidney cancer in the German shepherd dog, is an example of the use of disease gene mapping in the dog to define new animal models for human disease. Canine hereditary multifocal renal cystadenocarcinoma and nodular dermatofibrosis (RCND) is a naturally occurring inherited cancer syndrome in German shepherd dogs and is characterized by multifocal tumours in kidneys and skin. RCND has an autosomal dominant pattern of inheritance, and about 50% of affected dogs experience metastasis[51,52]. Using a large family of Norwegian dogs, we mapped RCND to canine chromosome 5 with a LOD SCORE of 16.7 (θ, the recombination fraction between a trait locus and a marker locus = 0.016) (REF. 35) in a region that overlapped the recently mapped, but as yet uncloned, human Birt–Hogg–Dubé (BHD) disease locus[53,54]. BHD is a multisystem disorder in humans that bears strong similarity to RCND,

ALU INSERTION
A dispersed, intermediately repetitive 300-bp DNA sequence, found in the human genome in 300,000 copies, that is named after the restriction endonuclease (*AluI*) that cleaves it.

CONTIG
(and supercontig). A contiguous region of DNA sequence constructed by aligning many sequence reads.

NARCOLEPSY
A sleep disorder characterized by excessive sleepiness, cataplexy, sleep paralysis, hypnologic hallucinations and an abnormal tendency to pass directly from wakefulness into REM sleep.

CYSTINURIA
An inherited abnormality of renal tubular transport of dibasic amino-acids leading to massive urinary excretion of cystine, lysine, arginine and ornithine.

CEROID LIPOFUSCINOSIS
An inherited neurodegenerative disorder associated with the accumulation of an abnormal pigment in the brain called lipofuscin.

LOD SCORE
A method of hypothesis testing that uses the logarithm of the ratio between likelihoods under the null and alternative hypotheses.

indicating that the same gene might be responsible for both the human and canine disease. This was indeed shown to be true: a single base change leads to alteration of a highly conserved amino acid, resulting in a disease-associated mutation in the canine-encoded protein folliculin[55,56]. We observed an absence of recombinants between the disease locus and the mutation in multiple US and Norwegian dogs that were separated by as many as 20 generations, arguing strongly that the missense change is the disease-causing mutation. Although little is known about the protein, crosses between heterozygote RCND dogs have so far failed to produce any homozygote dogs, indicating that the mutation is homozygous lethal ($P < 0.01$) (REF. 56). We now have in hand a superb, naturally occurring animal model for a rare and intriguing human disorder. In achieving this we have, appropriately, demonstrated the usefulness of large canine pedigrees for finding relatively subtle disease mutations.

Blindness. One final story of interest that is likely to direct future scientific directions is the mapping of genes for canine blindness. Retinitis pigmentosa (RP) is a human retinal degeneration that results in blindness, affecting approximately 1 in 4,000 people. Autosomal dominant, autosomal recessive, X-linked, and digenically inherited forms of RP are recognized, and linkage studies have identified at least 37 RP loci in the human genome, for which at least 22 genes have been cloned to date. Many genes, however, remain to be identified. Investigators at Cornell University and the University of Pennsylvania Veterinary School have studied naturally occurring hereditary retinal degeneration in dogs, termed progressive retinal atrophies (PRAs), for over two decades[17,57–60]. So far, loci have been identified for several forms of canine PRA[16,17,33,61–63]. Mapping of the underlying loci provides valuable opportunities to investigate not only the cell biology and pathogenesis of these diseases, but also to begin therapeutic interventions[64].

In the latter instance, gene therapy has been successfully applied in the canine form of Leber congenital amaurosis (LCA). LCA causes near-total blindness in infancy and results from mutations in the *RPE65* gene. A naturally occurring animal model, the *RPE65*$^{-/-}$ dog, suffers from early and severe visual impairment that is similar to that seen in humans with LCA. Using a recombinant adeno-associated virus carrying the wild-type gene, Acland and colleagues restored full visual function to affected dogs[64]. A similar success story has been reported by Mark Haskins and colleagues in the treatment of dogs with mucopolysaccharidosis VII, using a retroviral vector expressing wild-type canine β-glucuronidase[65].

What remains to be done? In the case of PRA, the mapping and cloning of the genes that are associated with PRA for dozens of different breeds of dog is yet to be completed. This task would be much easier if more was known about the ancestral relationships between dog breeds that manifest seemingly similar forms of any given disease[4]. Simultaneously, mapping studies from multiple breeds could be considered, thereby providing more statistical power. Initial studies accomplishing this are described below.

Population structure of the domestic dog

Although mitochondrial DNA analyses have been used successfully to elucidate the relationship between the domestic dog and the wolf[5,7,8], the evolution of mitochondrial DNA is too slow to allow inference of relationships among modern dog breeds, most of which have existed as closed breeding populations for <400 years[66–68]. One previous study showed that nuclear microsatellite loci could be used to assign dogs from five breeds to their breed of origin, demonstrating large genetic distances among these breeds[69]. In a larger effort, involving many more breeds, we recently showed that microsatellite typing, combined with phylogenetic analysis and genetic clustering methods, could be used to define relatedness among groups of breeds, and that genetic relatedness among breeds often correlates with morphological similarity and geographic origins[70].

These results provide information about dogs that is key to the design of mapping and cloning studies. For instance, we found that when all dogs are considered as a single population, the observed nucleotide heterozygosity is 8×10^{-4} — essentially the same as that found for the human population. However, variation among breeds accounts for more than 27% of the total genetic variation observed in dogs, and the average genetic distance between breeds, as calculated from SNP data, is $F_{ST} = 0.36$. These observations demonstrate the reality of the breed barrier and are in striking contrast to the much lower levels of genetic differentiation (5–10%) that are reported for human populations.

We then used cluster-based methods to determine the relationships among 85 breeds, and identified 4 genetic clusters, each of which has breeds with similar geographic origin, behaviour or morphology (FIG. 4). Expansion of this data set to include more breeds and distinct lines within breeds (such as 'show' versus 'field' lines) is currently under way.

A comprehensive definition of breed relationships such as this results in many hypotheses about disease alleles and haplotype sharing that can subsequently be tested within small clusters of related breeds. Recent work exemplifies this: seven herding breeds and two sighthound breeds were all found to carry the *mdr1-1*Δ allele identical by descent[71]. Locus-by-locus evidence of breed history can complement genome-wide studies.

Linkage disequilibrium and haplotype sharing

To best use this breed-relationship data to map genes important in morphology, disease susceptibility and behaviour, a clear understanding of the background levels of LINKAGE DISEQUILIBRIUM (LD) in the dog genome is needed, particularly as the extent of LD determines the feasibility of whole-genome ASSOCIATION MAPPING STUDIES (FIG. 5). To address this issue, we measured LD at five loci, none of which is under obvious selective pressure, in five breeds: akita, Bernese mountain dog, golden retriever, Labrador retriever and Pekingese[72]. We found

F_{ST}
A measure of population subdivision that indicates the proportion of genetic diversity found between populations relative to the amount within populations.

LINKAGE DISEQUILIBRIUM
This occurs when the frequency of a particular haplotype for two or more loci deviates significantly from that expected from the product of the observed allelic frequencies at each locus.

ASSOCIATION MAPPING STUDIES
A set of methods used to correlate polymorphisms in genotype to polymorphisms in phenotype in populations.

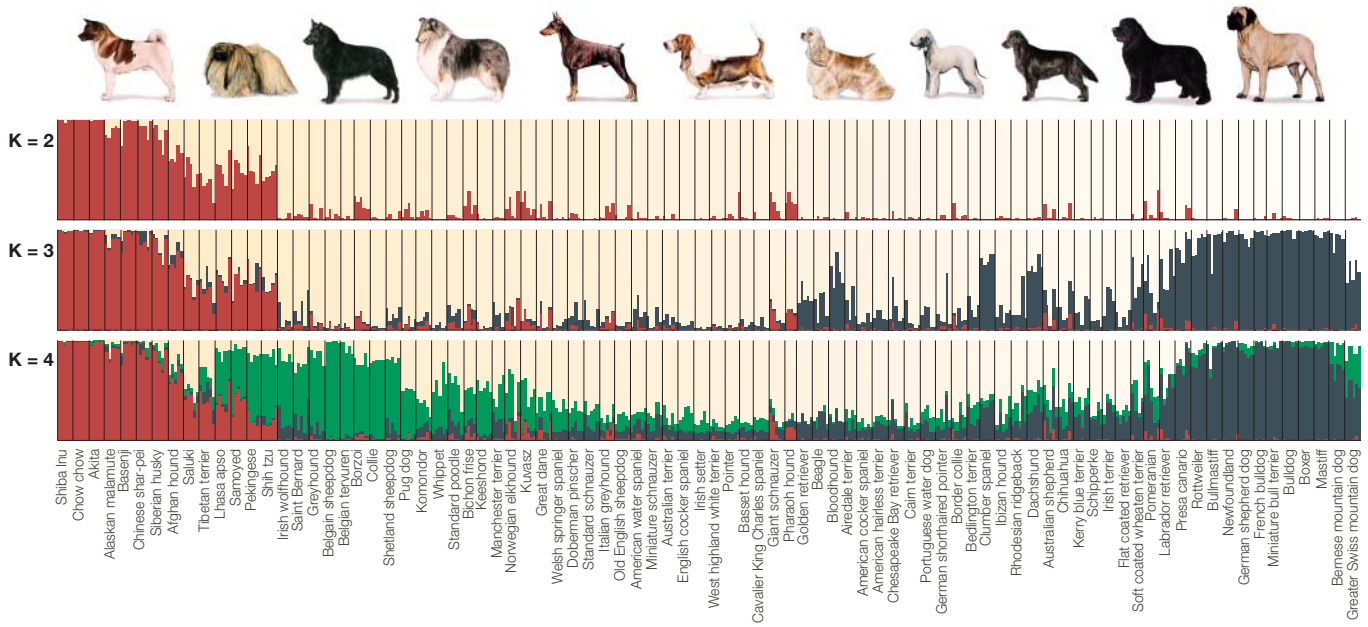

Figure 4 | **Population structure of 85 domestic dog breeds.** Each individual dog is represented by a single vertical line divided into 1–4 colours, specified as K, where K is the number of clusters assumed. Each colour represents one cluster, and the length of the coloured segment in each column shows the individual's estimated fractional membership for that cluster. Black vertical lines separate the breeds (five unrelated dogs per breed), which are labelled below the figure. Representative breeds are pictured above the graph. Results shown are averages over 15 structure runs for each value of K. The first cluster (red) was composed of Asian breeds (Chinese shar-pei, akita and shiba-inu), ancient hounds (Afghan and saluki) and spitz-type dogs (Alaskan malamute and Siberian husky). Dogs in group two (blue) shared a common morphological appearance, having broad frames and large square shaped heads and included, among others, the mastiff, bullmastiff, bulldog, boxer, Bernese mountain dog and greater Swiss mountain dog. The third group (green) shared behaviour patterns and included many working breeds such as the Belgian sheepdog, border collie and Australian shepherd. The fourth and final group (yellow) included many dogs of recent European descent, largely terriers and hounds. Reproduced, with permission, from REF 70 © (2004) American Association for the Advancement of Science.

that LD in domestic dogs is much more extensive than in humans, ranging from <1 Mb in golden retriever and Labrador retriever to 3.2 Mb in the Pekingese. Interestingly, these findings correlate with reported breed history: golden retriever and Labrador retriever are the most popular US breeds (FIG. 1) and, more importantly, their populations have never experienced bottlenecks, unlike the akita, Bernese mountain dog and Pekingese, all of which reportedly underwent tight population bottlenecks in the last 100 years. The above numbers are in striking contrast to the <100-kb extent of LD reported for human populations[73–82]. As LD in dog breeds is 20–100 times more extensive than in humans, a much smaller number of markers will be required for whole-genome association mapping studies in dogs compared with humans.

Haplotype diversity in dogs is relatively low in regions of extensive LD, and dog breeds, although highly differentiated genetically, share haplotypes with each other to a high degree. Using the above data set, we found that most chromosomes within a breed carry just a few haplotypes: 80% of a breed's chromosomes were found to carry an average of 2.1–3.4 haplotypes at any one locus. When all dogs were considered together, without regard to breed, an average of 4.5 haplotypes explained 80% of the total chromosomes at a locus, representing a relatively low level of diversity and high level of haplotype sharing[72]. This indicates that standardized marker sets can be constructed for use in mapping traits in a large number of distinct breeds. A modest level of haplotype diversity, like extensive LD, lowers the number of markers that will be required for whole-genome association mapping studies in dogs.

Mapping complex traits

The notion that breeds of domestic dog might provide an optimal system for mapping COMPLEX TRAITS has been suggested for over a decade, but it is the recent analysis of canid morphology in the Portuguese water dog that has clearly demonstrated the value of dog breeds for studies in quantitative genetics[83]. The Portuguese water dog population is derived from a small founder population that has expanded rapidly in recent years. Investigators at the University of Utah phenotyped 330 dogs for skeletal traits using 90 metrics derived from X-ray images (The Georgie Project, see Online links). Simultaneously, 550 microsatellite markers were analysed using DNA isolated from each dog. Principal-component analysis (PCA) was then used to define the genetic networks regulating the skeleton shape and size of Portuguese water dogs. PCA classifies phenotypic variation into independent systems of correlated components. Because principal components are phenotypes, QUANTITATIVE TRAIT LOCI (QTLs) that control them

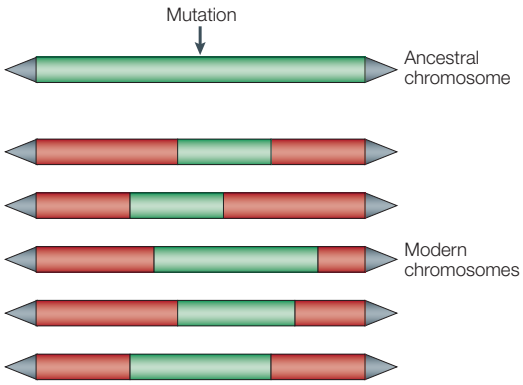

Figure 5 | **Linkage disequilibrium around a mutation is variable in the population of modern chromosomes that carry the mutation.** A mutation arises against the backdrop of a particular haplotype (shown in green) on a chromosome that is ancestral to a population of modern chromosomes. Modern-day chromosomes carry particular fragments of the original haplotype owing to their individual histories of recombination. Breeds in which the mutation occurs are expected to carry distinct fragments of the ancestral haplotype, which will help to narrow the interval in which the mutation must be found. Reproduced, with permission, from *Nature Reviews Genetics* (REF. 99) © (2002) Macmillan Magazines Ltd.

can then be localized[83]. In this particular study, many statistically significant QTLs were identified for each of the four principal components. The first component (PC1) represents the overall size of the skeleton. PC2 regulates metrics of the pelvis, head and neck such that the size and strength of the pelvic and head–neck musculoskeletal systems are inversely related. PC3 regulates an inverse relationship between the metrics of cranial volume and the length of the skull and limbs. Last, PC4 controls the skull and axial skeletons, representing a trade-off between strength and speed[83]. What remains now is to identify the precise variants that control the dramatic differences seen between breeds, as well as the genetic and environmental interactions that affect expression of those variants. We can speculate that perhaps these same genes will be shown to be important in human evolution as well.

Using the same population of Portuguese water dogs, Lark and colleagues have also begun mapping genes for complex diseases such as hip dysplasia[84]. They have identified two QTLs, both on CFA1, that are associated with SUBLUXATION of the hip joint, as measured by the Norberg angle — a quantitative radiographic measure of laxity. A complementary approach to the one described above has come from Todhunter *et al.* who have developed a four-generation pedigree of dogs that segregates hip dysplasia[85]. The cross is between Labrador retrievers and greyhounds, the latter of which do not get hip dysplasia. Preliminary data indicate that a limited number of QTLs control this complex phenotype[86,87]. The successful mapping of many other complex traits, such as rheumatoid arthritis and allergies, will benefit both humans and their companion animals (BOX 1).

QUANTITATIVE TRAIT LOCI
Genetic loci or chromosomal regions that contribute to the variability in complex quantitative traits (such as body weight), as identified by statistical analysis. Quantitative traits are typically affected by several genes and can be affected by the environment.

SUBLUXATION
A slight dislocation or disfunction of vertebrae or other joints.

Limitations of the dog

What, then, are the limitations of the dog and its genome for a study subject? Of note is the nature of dog populations themselves. As already discussed, these populations are invaluable and, we believe, uniquely powerful for tackling complex traits. That said, it is also true that we have only a limited knowledge of the recorded histories of dog breeds and, more importantly, exercise no control over their maintenance. Dog breeds are not mouse or rat strains, and the geneticist is somewhat at the mercy of the breed club for access to phenotypes and DNA samples. This lack of direct control is ameliorated in the following ways. First, the large number of breeds and canine diseases ensures that at least some breed clubs — in fact, the majority of them — are more than willing to assist in research to aid their breed. One outstanding example is the dedicated assistance that Gordon Lark and his team have received from members of the Portuguese water dog club of America. Another challenge is the common difficulty of collecting samples from extended dog families for linkage mapping studies, as progeny are often dispersed throughout the country. However, these disadvantages are far outweighed by the many advantages detailed above, including the enormous enthusiasm expressed by both individual breed clubs and the American and European kennel clubs for finding disease genes in dogs that will lead to the eventual development of genetic tests and the production of healthier and more long-lived breeds.

Future directions

As molecular information for the dog becomes increasingly available, and the number of success stories like that of Chase *et al.* increase[83], canine geneticists will almost certainly turn their attention to the most challenging of mapping studies, the identification of genes associated with behaviour[88]. Initial studies will focus on the identification of genes important in aberrant behaviour patterns such as the obsessive–compulsive tail-chasing behaviour seen in bull terriers[89,90]. Beyond that, Overall has suggested that dogs present naturally occurring models for many of man's psychiatric illnesses including separation anxiety, impulse control disorders, panic disorder and cognitive dysfunction[91]. Understanding the contribution of genes to such behavioural disturbances in the dog will certainly have strong implications for human psychiatric illnesses. Among the most difficult, and certainly most controversial, behaviours to study will be aggression[92]. The legal and social implications are significant, and there are strong arguments both for and against whether such studies should even be undertaken[93].

Of equal interest will be the identification of genes associated with domestication. One approach to understanding the genetics of this process derives from a 40-year-long study at the Russian Academy of Sciences in Novosibirsk, where a strain of silver fox has been selectively bred for tameability[94–96]. Foxes selected for tameness for over 30 generations are friendly, and even compete for attention from humans. They have also assumed a more dog-like appearance. Understanding

the genes that control this process will be highly informative for many species[97].

Last, and of greatest interest to naturalists, will be efforts to identify genes associated with breed-specific behaviours. In an interesting study, Quignon *et al.* have estimated that there are 1,300 olfactory receptor genes in the canine genome, 20% of which are pseudogenes[98]. This proportion is similar to that found for rodents, but much lower than what is observed in the human genome. Different dog breeds have different abilities to smell, but the genetic, physiological, morphological and developmental bases of these differences are unknown. The same can be said of other breed-specific behaviours: why do herding dogs herd, pointers point and retrievers retrieve? Such behaviours are clearly 'hard-wired', and understanding the genetics of these behaviours will provide tantalizing keys as to which types of human behaviours are genetically controlled.

Outlook and conclusions

Dogs have been our closest companions and greatest protectors for thousands of years. More than simply our servants, they are a species that has evolved in harmony with us, so that encoded within their genome is the ability to read both our hearts and minds, humbling us, but never diminishing us. Now it seems that our own genetic health might also ultimately owe much to the population structure we have created in dogs. As the dog star continues to rise, it will be incumbent on us to use the information from canine genetic studies to improve not only our own health and well being, but that of our closest companion as well.

1. Association, A. V. M. *U. S. Pet Ownership and Demographics Sourcebook*, 126 (American Veterinary Medical Association, Schaumburg, Illinois, 2002).
2. Patterson, D. Companion animal medicine in the age of medical genetics. *J. Vet. Internal Med.* **14**, 1–9 (2000).
3. Patterson, D. F., Haskins, M. E. & Jezyk, P. F. Models of human genetic disease in domestic animals. *Adv. Hum. Genet.* **12**, 263–339 (1982).
4. Ostrander, E. A. & Kruglyak, L. Unleashing the canine genome. *Genome Res.* **10**, 1271–1274 (2000).
5. Leonard, J. A. *et al.* Ancient DNA evidence for Old World origin of New World dogs. *Science* **298**, 1613–1616 (2002).
6. Darwin, C. Variation in animals and plants under domestication (Appelton and Co, New York, 1883).
7. Vila, C. *et al.* Multiple and ancient origins of the domestic dog. *Science* **276**, 1687–1689 (1997).
Provides strong evidence that dog domestication derives solely from the grey wolf and supports an ancient time frame for domestication.
8. Savolainen, P., Zhang, Y. P., Luo, J., Lundeberg, J. & Leitner, T. Genetic evidence for an East Asian origin of domestic dogs. *Science* **298**, 1610–1613 (2002).
9. Hare, B., Brown, M., Williamson, C. & Tomasello, M. The domestication of social cognition in dogs. *Science* **298**, 1634–1636 (2002).
First attempts to understand theory of mind in the domestic dog.
10. Mellersh, C. S. *et al.* A linkage map of the canine genome. *Genomics* **46**, 326–336 (1997).
11. Neff, M. W. *et al.* A second-generation genetic linkage map of the domestic dog, *canis familiaris*. *Genetics* **151**, 803–820 (1999).
12. Werner, P. *et al.* Anchoring of canine linkage groups with chromosome specific markers. *Mamm. Genome* **10**, 812–823 (1999).
13. Priat, C. *et al.* A whole-genome radiation hybrid map of the dog genome. *Genomics* **54**, 361–378 (1998).
14. Mellersh, C. S. *et al.* An integrated linkage-radiation hybrid map of the canine genome. *Mamm. Genome* **11**, 120–130 (2000).
15. Breen, M. *et al.* 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 (2001).
16. Acland, G. M. *et al.* A novel retinal degeneration locus identified by linkage and comparative mapping of canine early retinal degeneration. *Genomics* **59**, 134–142 (1999).
17. Acland, G. M. *et al.* Linkage analysis and comparative mapping of canine progressive rod-cone degeneration (prcd) establishes potential locus homology with retinitis pigmentosa (RP17) in humans. *Proc. Natl Acad. Sci. USA* **96**, 3048–3053 (1998).
18. Gordon, D., Corwin, M. B., Mellersh, C. S., Ostrander, E. A. & Ott, J. Establishing appropriate genome-wide significance levels for canine linkage analyses. *J. Hered.* **94**, 1–7 (2003).
Correlates physical and genetic distance in the dog genome.
19. Switonski, M. *et al.* Report on the progress of standardization of the G-banded canine (*Canis familiaris*) karyotype. *Chromosome Res.* **4**, 306–309 (1996).
20. Breen, M., Bullerdiek, J. & Langford, C. F. The DAPI banded karyotype of the domestic dog (*Canis familiaris*) generated using chromosome-specific paint probes. *Chromosome Res.* **7**, 401–406 (1999).
21. Yang, F. *et al.* A complete comparative chromosome map for the dog, red fox, and human and its integration with canine genetic maps. *Genomics* **62**, 189–202 (1999).
22. Sargan, D. R. *et al.* Use of flow-sorted canine chromosomes in the assignment of canine linkage, radiation hybrid, and syntenic groups to chromosomes: refinement and verification of the comparative chromosome map for dog and human. *Genomics* **69**, 182–195 (2000).
23. Langford, C. F., Fischer, P. E., Binns, M. M., Holmes, N. G. & Carter, N. P. Chromosome-specific paints from a high-resolution flow karyotype of the dog. *Chromosome Res.* **4**, 115–123 (1996).
24. Vignaux, F. *et al.* Construction and optimization of a dog whole-genome radiation hybrid panel. *Mamm. Genome* **10**, 888–894 (1999).
25. Guyon, R. *et al.* A 1 Mb resolution radiation hybrid map of the canine genome. *Proc. Natl Acad. Sci. USA* **100**, 5296–5301 (2003).
26. Breen, M. *et al.* An integrated 4249 marker FISH/RH map of the canine genome. *BMC Genomics* **5**, 65 (2004).
Fully integrated FISH/radiation hybrid map of the dog demonstrating high fidelity in BAC positioning and most complete map published to date.
27. Clark, L. A. *et al.* Chromosome-specific microsatellite multiplex sets for linkage studies in the domestic dog. *Genomics* (in the press).
28. Ostrander, E. A. & Comstock, K. E. The domestic dog genome. *Curr. Biol.* **14**, R98–99 (2004).
29. Kirkness, E. F. *et al.* The dog genome: survey sequencing and comparative analysis. *Science* **301**, 1898–1903 (2003).
First whole-genome shotgun sequence in the dog providing 1.5x coverage of the standard poodle genome.
30. Vinogradov, A. E. Genome size and GC-percent in vertebrates as determined by flow cytometry: the triangular relationship. *Cytometry* **31**, 100–109 (1998).
31. Madsen, O. *et al.* Parallel adaptive radiations in two major clades of placental mammals. *Nature* **409**, 610–614 (2001).
32. Murphy, W. J. *et al.* Molecular phylogenetics and the origins of placental mammals. *Nature* 409, 614–618 (2001).
33. Sidjanin, D. J. *et al.* Canine *CNGB3* mutations establish cone degeneration as orthologous to the human achromatopsia locus *ACHM3*. *Hum. Mol. Genet.* **11**, 1823–1833 (2002).
34. Lowe, J. K. *et al.* Linkage mapping of the primary disease locus for collie eye anomaly. *Genomics* **82**, 86–95 (2003).
35. Jonasdottir, T. J. *et al.* Genetic mapping of a naturally occurring hereditary renal cancer syndrome in dogs. *Proc. Natl Acad. Sci. USA* **97**, 4132–4137 (2000).
First mapping of a cancer susceptibility gene in the dog following a genome-wide scan.
36. Lin, L. *et al.* The sleep disorder canine narcolepsy is caused by a mutation in the hypocretin (orexin) receptor 2 gene. *Cell* **98**, 365–376 (1999).
37. Ollier, W. E. *et al.* Dog MHC alleles containing the human RA shared epitope confer susceptibility to canine rheumatoid arthritis. *Immunogenetics* **53**, 669–673 (2001).
38. Henthorn, P. S. *et al.* IL-2Rγ gene microdeletion demonstrates that canine X-linked severe combined immunodeficiency is a homologue of the human disease. *Genomics* **23**, 69–74 (1994).
39. Keller, R. C. *et al.* Chromosomal assignment of two putative canine keratin gene clusters. *Anim. Genet.* **29**, 141–143 (1998).
40. Henthorn, P. S. *et al.* Canine cystinuria: polymorphism in the canine *SLC3A1* gene and identification of a nonsense mutation in cystinuric Newfoundland dogs. *Hum. Genet.* **107**, 295–303 (2000).
41. Chao, H. *et al.* Persistent expression of canine factor IX in hemophilia B canines. *Gene Ther.* **6**, 1695–1704 (1999).
42. Venta, P. J., Li, J., Yuzbasiyan-Gurkan, V., Brewer, G. J. & Schall, W. D. Mutation causing von Willebrand's disease in Scottish Terriers. *J. Vet. Intern. Med.* **14**, 10–19 (2000).
43. Lingaas, F. *et al.* Genetic markers linked to neuronal ceroid lipofuscinosis in English setter dogs. *Anim. Genet* **29**, 371–376 (1998).
44. Yuzbasiyan-Gurkan, V. *et al.* Linkage of a microsatellite marker to the canine copper toxicosis locus in Bedlington terriers. *Am. J. Vet. Res.* **58**, 23–27 (1997).
45. van de Sluis, B. J. *et al.* 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 (1999).
46. Thomas, R., Smith, K. C., Ostrander, E. A., Galibert, F. & Breen, M. Chromosome aberrations in canine multicentric lymphomas detected with comparative genomic hybridisation and a panel of single locus probes. *Br. J. Cancer* **89**, 1530–1537 (2003).
47. Thomas, R. *et al.* A canine cancer-gene microarray for CGH analysis of tumors. *Cytogenet Genome Res.* **102**, 254–260 (2003).
48. Nishino, S., Ripley, B., Overeem, S., Lammers, G. J. & Mignot, E. Hypocretin (orexin) deficiency in human narcolepsy. *Lancet* **355**, 39–40 (2000).
49. Peto, J. *et al.* Prevalence of *BRCA1* and *BRCA2* gene mutations in patients with early-onset breast cancer. *J. Natl Cancer Inst.* **91**, 943–949 (1999).
50. Thannickal, T. C. *et al.* Reduced number of hypocretin neurons in human narcolepsy. *Neuron* **27**, 469–474 (2000).
51. Lium, B. & Moe, L. Hereditary multifocal renal cystadenocarcinomas and nodular dermatofibrosis in the German shepherd dog: macroscopic and histopathologic changes. *Vet. Pathol.* **22**, 447–455 (1985).
52. Moe, L. & Lium, B. Hereditary multifocal renal cystadenocarcinomas and nodular dermatofibrosis in 51 German shepherd dogs. *J. Small Anim. Pract.* **38**, 498–505 (1997).
53. Schmidt, L. S. *et al.* Birt–Hogg–Dube syndrome, a genodermatosis associated with spontaneous pneumothorax and kidney neoplasia, maps to chromosome 17p11.2. *Am. J. Hum. Genet.* **69**, 876–882 (2001).
54. Comstock, K. E. *et al.* A high-resolution comparative map of canine Chromosome 5q14.3–q33 constructed utilizing the 1.5x canine genome sequence. *Mamm. Genome* **15**, 544–551 (2004).
55. Nickerson, M. *et al.* Mutations in a novel gene lead to kidney tumors, lung wall defects, and benign tumors of the hair follicle in patients with the Birt–Hogg–Dube syndrome. *Cancer Cell* **2**, 157 (2002).
56. Lingaas, F. *et al.* A mutation in the canine BHD gene is associated with hereditary multifocal renal cystadenocarcinoma and nodular dermatofibrosis in the German Shepherd dog. *Hum. Mol. Genet.* **12**, 3043–3053 (2003).

57. Aguirre, G. D. *et al.* Congenital stationary night blindness in the dog: common mutation in the *RPE65* gene indicates founder effect. *Mol. Vis.* **4**, 23 (1998).
58. Aguirre, G. D., Baldwin, V., Weeks, K. M., Acland, G. M. & Ray, K. Frequency of the codon 807 mutation in the cGMP phosphodiesterase β-subunit gene in Irish setters and other dog breeds with hereditary retinal degeneration. *J. Hered.* **90**, 143–147 (1999).
59. Zhang, Q. *et al.* Fine mapping of canine XLPRA establishes homology of the human and canine RP3 intervals. *Invest. Ophthalmol. Vis. Sci.* **42**, 2466–2471 (2001).
60. Kijas, J. *et al.* Naturally-occurring *rhodopsin* mutation in the dog causes retinal dysfunction and degeneration mimicking human dominant retinitis pigmentosa. *Proc. Natl Acad. Sci. USA* **99**, 6328–6333 (2002).
61. Petersen-Jones, S. M., Entz, D. D. & Sargan, D. R. cGMP phosphodiesterase-alpha mutation causes progressive retinal atrophy in the Cardigan Welsh corgi dog. *Invest. Ophthalmol. Vis. Sci.* **40**, 1637–1644 (1999).
62. Aguirre, G., Lolley, R., Farber, D., Fletcher, T. & Chader, G. Rod-cone dysplasia in Irish Setter dogs: a defect in cyclic GMP metabolism in visual cells. *Science* **201**, 1133 (1978).
63. Zeiss, C. J., Ray, K., Acland, G. M. & Aguirre, G. D. Mapping of X-linked progressive retinal atrophy (XLPRA), the canine homolog of retinitis pigmentosa 3 (RP3). *Hum. Mol. Genet.* **9**, 531–537 (2000).
64. Acland, G. M. *et al.* Gene therapy restores vision in a canine model of childhood blindness. *Nature Genet.* **28**, 92–95 (2001).
First successful gene therapy in the dog, demonstrating recovery of sight following introduction of vector carrying wild-type gene.
65. Ponder, K. P. *et al.* Therapeutic neonatal hepatic gene therapy in mucopolysaccharidosis VII dogs. *Proc. Natl Acad. Sci. USA* **99**, 13102–13107 (2002).
66. Rogers, C. A. & Brace, A. H. *The International Encyclopedia of Dogs*, 496 (Howell Book House, New York, 1995).
67. Fogel, B. *The Encyclopedia of the Dog*, (DK Publishing, Inc., New York, 1995).
68. American, K. C. *The Complete Dog Book*, 790 (Howell Book House, New York, New York, 1998).
69. Koskinen, M. T. Individual assignment using microsatellite DNA reveals unambiguous breed identification in the domestic dog. *Anim. Genet.* **34**, 297–301 (2003).
70. Parker, H. G. *et al.* Genetic structure of the purebred domestic dog. *Science* **304**, 1160–1164 (2004).
71. Neff, M. W. *et al.* Breed distribution and history of canine *mdr1-1*Δ, a pharmacogenetic mutation that marks the emergence of breeds from the collie lineage. *Proc. Natl Acad. Sci. USA* **101**, 11725–11730 (2004).
72. Sutter, N. B., Eberle, M., H. G., P., Kruglyak, L. & E. A., O. Extensive and breed specific linkage disequilibrium in *Canis familiaris*. *Genome Res.* (in the press).
73. Daly, M. J., Rioux, J. D., Schaffner, S. F., Hudson, T. J. & Lander, E. S. High-resolution haplotype structure in the human genome. *Nature Genet.* **29**, 229–232 (2001).
74. Dawson, E. *et al.* A first-generation linkage disequilibrium map of human chromosome 22. *Nature* **418**, 544–548 (2002).
75. Gabriel, S. B. *et al.* The structure of haplotype blocks in the human genome. *Science* **296**, 2225–2229 (2002).
76. Ke, X. *et al.* The impact of SNP density on fine-scale patterns of linkage disequilibrium. *Hum. Mol. Genet.* **13**, 577–588 (2004).
77. Moffatt, M. F., Traherne, J. A., Abecasis, G. R. & Cookson, W. O. Single nucleotide polymorphism and linkage disequilibrium within the TCR α/δ locus. *Hum. Mol. Genet.* **9**, 1011–1019 (2000).
78. Patil, N. *et al.* Blocks of limited haplotype diversity revealed by high-resolution scanning of human chromosome 21. *Science* **294**, 1719–1723 (2001).
79. Reich, D. E. *et al.* Linkage disequilibrium in the human genome. *Nature* **411**, 199–204 (2001).
80. Stephens, J. C. *et al.* Haplotype variation and linkage disequilibrium in 313 human genes. *Science* **293**, 489–493. (2001).
81. Varilo, T. *et al.* The interval of linkage disequilibrium (LD) detected with microsatellite and SNP markers in chromosomes of Finnish populations with different histories. *Hum. Mol. Genet.* **12**, 51–59 (2003).
82. Weiss, K. M. & Clark, A. G. Linkage disequilibrium and the mapping of complex human traits. *Trends Genet.* **18**, 19–24 (2002).
83. Chase, K. *et al.* Genetic basis for systems of skeletal quantitative traits: principal component analysis of the canid skeleton. *Proc. Natl Acad. Sc.i USA* **99**, 9930–9935 (2002).
First use of PCA to identify QTLs important in morphological variation in the dog.
84. Chase, K., Lawler, D. F., Adler, F. R., Ostrander, E. A. & Lark, K. G. Bilaterally asymmetric effects of quantitative trait loci (QTLs): QTLs that affect laxity in the right versus left coxofemoral (hip) joints of the dog (*Canis familiaris*). *Am. J. Med. Genet.* **124A**, 239–247 (2004).
85. Todhunter, R. J. *et al.* An outcrossed canine pedigree for linkage analysis of hip dysplasia. *J. Hered.* **90**, 83–92 (1999).
86. Todhunter, R. J. *et al.* Genetic structure of susceptibility traits for hip dysplasia and microsatellite informativeness of an outcrossed canine pedigree. *J. Hered.* **94**, 39–48 (2003).
87. Todhunter, R. J. *et al.* Power of a Labrador Retriever–Greyhound pedigree for linkage analysis of hip dysplasia and osteoarthritis. *Am. J. Vet. Res.* **64**, 418–424 (2003).
88. Ostrander, E. A. & Giniger, E. *Semper fidelis*: what man's best friend can teach us about human biology and disease. *Am. J. Hum. Genet.* **61**, 475–480 (1997).
89. Dodman, N. H. *et al.* Behavioral changes associated with suspected complex partial seizures in bull terriers. *J. Am. Vet. Med. Assoc.* **208**, 688–691 (1996).
90. Moon-Fanelli, A. A. & Dodman, N. H. Description and development of compulsive tail chasing in terriers and response to clomipramine treatment. *J. Am. Vet. Med. Assoc.* **212**, 1252–1257 (1998).
91. Overall, K. L. Natural animal models of human psychiatric conditions: assessment of mechanism and validity. *Prog. Neuropsychopharmacol. Biol. Psychiatry* **24**, 727–776 (2000).
92. van den Berg, L., Schilder, M. B. & Knol, B. W. Behavior genetics of canine aggression: behavioral phenotyping of golden retrievers by means of an aggression test. *Behav. Genet.* **33**, 469–483 (2003).
93. Feddersen-Petersen, D. U. Biology of aggression in dogs. *Dtsch. Tierarztl. Wochenschr.* **108**, 94–101 (2001) (in German).
94. Kolesnikova, L. A., Trut, L. N. & Beliaev, D. K. Changes in the morphology of the epiphysis of silver foxes during domestication. *Zh. Obshch. Biol.* **49**, 487–492 (1988) (in German).
95. Vasil"eva, L. L. & Trut, L. N. The use of the method of principal components for phenogenetic analysis of the integral domestication trait. *Genetika* **26**, 516–524 (1990) (in Russian).
96. Prasolova, L. A., Trut, L. N., Vsevolodov, E. B. & Latyshov, I. F. Phenogenetic analysis of various changes in the color of fur in silver foxes originated during domestication. *Genetika* **25**, 1626–1635 (1989) (in Russian).
97. Kukekova, A. V. *et al.* A marker set for construction of a genetic map of the silver fox (*Vulpes vulpes*). *J. Hered.* **95**, 185–194 (2004).
98. Quignon, P. *et al.* Comparison of the canine and human olfactory receptor gene repertoires. *Genome Biol* .**4**, R80 (2003).
Characterizes the canine odorant receptor gene family, demonstrating a low proportion of pseudogenes and describing two large families on canine chromosomes 18 and 21.
99. Ardlie, K. G. Kruglyak, L. & Seielstad, M. Patterns of linkage disequilibrium in the human genome. *Nature Rev. Genet.* **3**, 299–309 (2002).
99. Patterson, D. F. *Canine Genetic Disease Information System — A Computerized Knowledgebase of Genetic Diseases in Dogs* (Elsevier, in the press).
100. Patterson, D. F. *Genetic Diseases of the Dog* (Elsevier, in the press).

Acknowledgements

N.B.S. is a Waltham Foundation Fellow. E.A.O. gratefully acknowledges support from the NIH and the Burroughs Wellcome Foundation. We thank Ed Giniger, Kerstin Lindblad-Toh, Ewen Kirkness and Francis Galibert for their thoughtful comments and helpful suggestions on this manuscript, and the AKC and Chet Jezierski for use of the copyrighted canine artwork included in figures 2 and 4.

Competing Interests Statement

The authors declare no competing financial interests.

Online links

FURTHER INFORMATION

American Kennel Club: http://www.akc.org
Canine Resources at the University of Rennes: http://www-recomgen.univ-rennes1.fr/Dogs/paper/FISH-RHmap.html
Elaine Ostrander's homepage: http://myprofile.cos.com/ostrander1
Genbank: http://www.ncbi.nih.gov/Genbank
Matthew Breen's homepage: http://www.cvm.ncsu.edu/mbs/breen_matthew.htm
Proposal to sequence the dog genome: http://www.genome.gov/Pages/Research/Sequencing/SeqProposals/CanineSEQedited.pdf
The FHCRC Dog Genome Project: http://www.fhcrc.org/science/dog_genome/dog.html.
The Portuguese Water Dog Georgie Project: http://www.georgieproject.com/
UCSC Genome Bioinformatics Site: http://www.genome.ucsc.edu
Access to this interactive links box is free online.