Both Ends of the Leash — The Human Links to Good Dogs with Bad Genes

Elaine A. Ostrander, Ph.D.

For nearly 350 years, veterinary medicine and human medicine have been separate entities, with one geared toward the diagnosis and treatment in animals and the other toward parallel goals in the owners. However, that model no longer fits, since research on diseases of humans and companion animals has coalesced.1-4 The catalyst for this union has been the completion of the human genome sequence, coupled with draft sequence assemblies of genomes for companion animals.5,6 Here, we summarize the critical events in canine genetics and genomics that have led to this development, review major applications in canine health that will be of interest to human caregivers, and discuss expectations for the future.

Human and Canine Genomics

In 2001, two independent draft versions of the human genome sequence and the concomitant identification of approximately 30,000 genes were the seminal events that defined completion of the Human Genome Project.7-8 The genome was officially declared to be finished in 2004, with sequencing reported to include 99% of transcribing DNA.9 By comparison, the genome of the domestic dog, Canis lupus familiaris, was sequenced twice, once to 1.5× density (i.e., covering the genome, in theory, 1.5 times) and once to 7.8× density (providing sequencing for more than 95% of base pairs) in the standard poodle and boxer, respectively.5,10 Subsequent contributions to the canine genome have focused on better annotation to locate missing genes,11 understanding chromosome structure,12 studying linkage disequilibrium,5,13 identifying copy-number variants,14-16 and mapping the transcriptome.17

The use of the canine genome to understand the genetic underpinning of disorders that are difficult to disentangle in humans has been on the rise for nearly two decades,1,2,18 The reason relates back to the domestication of dogs from gray wolves (C. lupus), an event that began at least 30,000 years ago.19-21 Since their domestication, dogs have undergone continual artificial selection at varying levels of intensity, leading to the development of isolated populations or breeds5,22,23 (Fig. 1). Many breeds were developed during Victorian times24 and have been in existence for only a few hundred years, a drop in the evolutionary bucket.25 Most breeds are descended from small numbers of founders and feature so-called popular sires (dogs that have performed well at dog shows and therefore sire a large number of litters). Thus, the genetic character of such founders is overrepresented in the population.25,26 These facts, coupled with breeding programs that exert strong selection for particular...
physical traits, mean that recessive diseases are common in purebred dogs, and many breeds are at increased risk for specific disorders. We, and others, have chosen to take advantage of this fact in order to identify genes of interest for human and canine health.

The Genetic Power of Canine Families

One of the most striking features of canine families is their large size, which makes them amenable to conventional linkage mapping. This fact...
was particularly well illustrated in the search for the canine gene for hereditary multifocal renal cystadenocarcinoma and nodular dermatofibrosis (RCND) in German shepherds. Although rare, RCND is a naturally occurring inherited cancer syndrome that includes bilateral, multifocal tumors in kidneys and numerous, dense collagen-based nodules in the skin, a disorder that is similar to the Birt–Hogg–Dubé syndrome (BHD) in humans. In dogs, the disease allele is highly penetrant and transmitted in an autosomal dominant fashion. The dog pedigree that was used for mapping the disease included one affected founder male who sired several litters (Fig. 2). With DNA available from nearly all dogs, this single pedigree had sufficient power to localize the disease gene to canine chromosome 5q12 with a logarithm of odds (LOD) score of 4.6, giving odds of more than 10,000 to 1 that the mapping was correct.

After the localization of RCND, the human BHD locus was mapped to human chromosome 17p12q11, which corresponds to canine chromosome 5q12. Both affected dogs and humans were found to carry mutations in the same gene encoding tumor-suppressor protein folliculin, which is hypothesized to interact with the energy and nutrient-sensing signaling pathway consisting of AMP-activated protein kinase (AMPK) and mammalian target of rapamycin (mTOR).

Three issues about this example are striking. First, the single, large dog pedigree was collected and genotyped in a fraction of the time it took to collect and characterize the many necessary human pedigrees. Second, BHD is associated with substantial variability in disease presentation in humans and may be hard to distinguish from similar disorders. In the case of the large extended dog family, phenotyping was easy, since every dog had the same genetic background and the disease presentation was highly uniform. Also, the dog locus was found before the human locus. Other disease genes that were first mapped in dogs for which there is a close human proxy include narcolepsy, copper toxicity, neuronal ceroid lipofuscinosis, and ichthyosis, to name a few. Each of such stories is illuminating in its own way. In the case of narcolepsy in the Doberman pinscher, the identification of a mutation in the gene encoding hypocretin receptor 2 suggested a newly recognized pathway that is involved in the molecular biology of sleep. Another example is canine neuronal ceroid lipofuscinosis, a late-onset disorder of American Staffordshire terriers with symptoms that are similar to a human adult-onset form of the disorder known as Kuf’s disease. In American Staffordshire terriers, neuronal ceroid lipofuscinosis is caused by an R99H mutation in exon 2 of the gene encoding arylsulfatase G (ARSG), leading to a 75% decrease in sulfatase activity. This study, therefore, both identified a new gene for consideration in human neuronal ceroid lipofuscinosis and provided new information regarding sulfatase deficiency and pathogenesis of the disease.

A recurring theme in the gene mapping of canine diseases is the power of the breed structure. To be a registered member of a breed, the dog’s ancestors must have been registered members as well. In 2011, the American Kennel Club (www.akc.org) recognized 173 distinct dog breeds, with European clubs taking the number of established breeds to more than 400. Dog breeds offer the same advantage of reducing locus heterogeneity that is gained by studying humans from geographically isolated countries such as Finland or Iceland. For any given complex disease, a small number of genes and deleterious alleles will dominate the breed, much as the 999de15 BRCA2 mutation does in Icelandic women with hereditary breast cancer.

Epilepsy is a good example, since this disease has been difficult to disentangle genetically in humans because of indistinct clinical phenotypes and a high degree of locus heterogeneity. The disease affects 5% of dogs and is reported in dozens of breeds. Remitting focal epilepsy in the Lagotto Romagnolo breed is caused by variants in LGI2, a homologue of the human epilepsy LGI1 gene. In contrast, miniature wire-haired dachshunds have a form of epilepsy reminiscent of the progressive myoclonic disease known as Lafora’s disease, which in humans is the most severe form of teenage-onset epilepsy. The similar disease in dachshunds is caused by an unusual expansion of a dodecamer repeat within the gene encoding malin (EPM2B) that modulates gene expression by a factor of nearly 900. The presentation of epilepsy is expectedly unique in other breeds. Thus, one way to disentangle complex diseases like epilepsy is to study the disorder in different dog breeds.
Figure 2. Mapping Pedigree for Canine Renal Cystadenocarcinoma and Nodular Dermatofibrosis (RCND).
A single affected male dog carrying an autosomal dominant allele for RCND sired five litters of pups with five unique and unaffected females. Affected dogs are shown in black, and unaffected dogs in white. Squares indicate males, circles females, and lines relationships. The portion of canine chromosome 5q14 showing linkage is indicated as a rectangle below each square or circle. Black bars indicate the portion of the affected parental chromosome inherited by each offspring from the affected father, and white bars indicate the portion inherited from the normal chromosome of the father. Alleles for each marker are indicated as numbers. Breakpoints allow the disease gene to be localized to a region adjacent to marker ZuBeCa6. Reprinted from Jónasdóttir et al., with the permission of the publisher.
within dog breeds it can extend for megabases. The second way in which breed structure offers unique advantages to genetic mapping is that when used judiciously, it allows researchers to move quickly from linked or associated markers to causative genes. Although progressive rod–cone degeneration was initially mapped in miniature and toy poodles, the disorder appears in more than a dozen breeds and is phenotypically similar to one form of human adult-onset, autosomal recessive retinitis pigmentosa. Analysis of additional SNPs allowed the investigators to reduce the disease locus to a 106-kb haplotype that is shared by affected dogs from 14 breeds. A mutation in a novel gene was ultimately determined to cause the disease.50 Had there not been 14 affected breeds sharing the founder mutation, which allowed the haplotype to be significantly reduced, only next-generation sequencing could have ultimately localized the disease gene.

Although researchers could have correctly guessed a subset of the breeds that shared the same mutation at the causative locus for progressive rod–cone degeneration by knowing about their shared heritage, common geographic origin, or shared morphologic features, in many cases the relationship among the breeds is too ancient to be obvious. With the use of both cluster analysis51,52 and neighbor-joining trees,23 a clear picture is emerging regarding how breeds are related to one another genetically (Fig. 3). This type of information highlights groups of breeds that probably share common founders (and hence the same disease alleles) and facilitates experimental design.

The examples discussed thus far have focused on disease phenotypes. However, canine morphologic studies have been informative for both discovering new ways of perturbing the genome and suggesting candidate genes for related diseases. For instance, chondrodysplasia is a fixed trait for more than 20 breeds with disproportionately short legs recognized by the American Kennel Club, including the dachshund, corgi, and basset hound (Fig. 5).53 A genomewide association study comparing 95 dogs from eight chondrodysplastic breeds with 702 dogs from 64 breeds lacking the trait identified a single strong association (P=1.0×10−102) with canine chromosome 18. Although this very low P value is probably exaggerated because of the population structure, such a strong association is not unusual when breeds sharing a trait
from a common founder are compared with a large number of unrelated control breeds. In this case, the trait is caused by expression of an fgf4 retrogene. This retrogene encodes fibroblast growth factor 4 in which all fgf4 exons are present, but introns and regulatory signals are missing (Fig. 5). The spliced copy of the gene is located a large distance away from the source gene. Although such an arrangement is common in insects, this was the first report of an expressed retrogene that alters a mammalian trait. Expression studies showed that the fgf4 retrogene was expressed in the long bones of 4-week-old puppies, suggesting that mistimed expression, incorrect RNA levels, or mislocalization of the retrogene product caused premature closure of the growth plates in the long bones of the carrier breeds. It will be interesting to see whether this gene, or this method of mutating mammalian genomes, turns out to be important in similar human diseases.

Other canine morphologic traits that include such characteristics as body size, leg width, and coat color have been mapped. Not surprisingly, loci that control both a morphologic trait and a disease have been identified. This may be a result of strong selection by breeders to propagate dogs of a certain appearance, which results in piggybacking of disease alleles, or in some cases, diseases are associated with the same genetic variants that create a morphologic effect. This is best illustrated by dermoid sinus, a neural-tube defect in the ridgeback breed that is caused by the same copy-number variant that produces the hair ridge characteristic of the Rhodesian ridgeback.

MAPping MULTIGENIC TRAITS

When the dog genome sequence was published in 2005, Lindblad-Toh et al. hypothesized that breed structure would enable mapping of simple recessive traits in dogs with a genomewide association study of no more than 20 cases and controls each. They further reasoned that complex traits that are controlled by, for instance, five genes could be mapped with 97% certainty on the basis of just 100 cases and 100 controls. This was a bold prediction, since most genomewide association studies of complex human disorders require thousands of samples. But the investigators’ prediction proved to be correct, and many genomewide association studies in dogs have successfully mapped complex traits on the basis of no more than 50,000 SNPs and fewer than 200 dogs.

Recent work by Wilbe et al. that identifies genes for systemic lupus erythematosus (SLE)–related disease complex illustrates this point. Nova Scotia duck-tolling retrievers have an abnormally high rate of autoimmune diseases, including SLE. The breed is descended from a small number of founders that survived two major outbreaks of canine distemper virus in the early 1900s. It has been hypothesized that autoimmune disorders develop in these dogs because they have a particularly strong or reactive immune system, which helped them to survive the distemper outbreaks. In an analysis of 81 cases and 57 controls in a genomewide association study of 22,000 SNPs, investigators found five associated loci, three of which have already been validated. Candidate genes of particular interest in-
include those associated with T-cell activation such as PPP3CA, BANK1, and DAPPI.

**DOGS AND CANCER**

Of all the disorders for which dogs are likely to inform human health, canine cancer is likely to have the greatest effect.63 Cancers are the most frequent cause of disease-associated death in dogs, and naturally occurring cancers are well described in several breeds.3,64,65 Although considerable effort has gone into the study of common cancers, the dog has also served as a model for studies of rare tumors, including histiocytic sarcomas, which are highly aggressive, lethal, dendritic-cell neoplasms.66 In dogs, two forms exist: a localized variant, in which skin and subcutaneous tumors develop in a leg and metastasize to lymph nodes and blood vessels, and a dissem-
Histiocytic sarcomas are among the most common malignant tumors of dogs and humans that is caused by a loss of the RPE65 protein owing to mutations in RPE65, causing blindness shortly after birth. In a landmark study in 2001, Acland et al. used a recombinant adeno-associated virus carrying wild-type RPE65 to restore vision in a dog that was homozygous for the RPE65 mutation. Replication was successful, and treated dogs maintained stable vision for at least 3 years. Humans with Leber’s congenital amaurosis are now being successfully treated for the disorder. Progressive retinal atrophy occurs in more than 100 breeds of dogs, suggesting dozens of naturally occurring models for additional study. So far, 18 genes for canine retinal diseases have been found.

The canine system is valuable for mapping behaviors that are specific to both breed and species. Abnormal behaviors, including separation anxiety, dominance aggression, and obsessive–compulsive disorder, are most amenable to genetic studies. Partial success has been achieved with obsessive–compulsive disorder in bull terriers and Doberman pinschers. In Dobermans, the disease presents as flank or blanket sucking and was recently mapped to a 1.7-Mb region of chromosome 7 near the CDH2 gene. CDH2 mediates synaptic activity-regulated neuronal adhesion, but to date no functional studies have illuminated these findings and no mutation has been reported.

What we most wish to understand about dog health is the very same thing we wish to know about ourselves. When will we, or they, get sick? How is the illness best treated? And what is the likely outcome? Each half of a pet–human pair wants to know what to expect from the other end of the leash and how to prolong the relationship. Finally, as the end of life approaches, we seek to make both our canine companions and ourselves comfortable, settled in the knowledge that a full life has been achieved. When considered in that frame, we are not so different from our canine companions. As the scientific advances coalesce, joining us ever closer to the one family member we actually get to choose, it is worth bearing in mind the idea that these independently propagating species that are specific to both breed and species.
mind that though our methods may be different, our goals are the same: a healthy life well spent in the best of company.

Disclosure forms provided by the author are available with the full text of this article at NEJM.org.