*Am. J. Hum. Genet. 61:475–480, 1997*

# INSIGHTS FROM MODEL SYSTEMS
# Semper Fidelis: What Man's Best Friend Can Teach Us about Human Biology and Disease

Elaine A. Ostrander[1] and Edward Giniger[2]

[1]Clinical Research Division and [2]Basic Science Division, Fred Hutchinson Cancer Research Center, Seattle

> Many millions of years ago the country which is now called Spain seethed uneasily in the ferment of creation. Ages passed; vegetation appeared; where there is vegetation the law of Nature has decreed that there shall be rabbits; where there are rabbits, Providence has ordained there shall be dogs. There is nothing in this that calls for question or comment. But when we ask why the dog that caught rabbits was called a Spaniel, then doubts and difficulties begin (Woolf 1933, p. 1).

If genetics is the study of biology by analysis of heritable variation in a species, then the raw material of genetics is the natural variation in the properties of a species. It follows that the most efficacious way to study mammalian genetics is by exploiting a species that combines the maximum range of natural variation with the capacity to minimize that variation within a defined pedigree. In this review, we will present the virtues and challenges of using the domestic dog to study mammalian genetics.

The dog arose from the domestication of wild canids, probably with additional infusions of genes from the wild at multiple times since the initial domestication events (Vila et al. 1997). A modest number of relatively ancient founder breeds are still in existence today. Among the oldest of these are the Maltese, which is believed to have originated some 28 centuries ago; the Pharaoh Hound, which is depicted on the tomb of Tutankhamen (1339 B.C.); the Mastiff, whose great loyalty in fighting alongside its masters was noted by Caesar during the invasion of Britain in 55 B.C.; and the Pekinese, which is traceable to the Tang Dynasty of the eighth century.

The vast multiplicity of contemporary breeds has largely been generated over just the past 250 years. Despite the short time involved, each of the modern breeds displays a phenotypic uniformity that is suggestive of a high degree of genetic homogeneity. The very fact that such a diverse array of phenotypically stable breeds has been created in so few generations may imply that a relatively small number of key loci are responsible for the characteristic features that define particular breeds. It is the dog's unique combination of interbreed diversity with intrabreed uniformity that makes it the organism of choice for dissecting complex mammalian traits.

Received June 16, 1997; accepted for publication June 27, 1997.

Address for correspondence and reprints: Dr. Elaine A. Ostrander, Clinical Research Division M318, Fred Hutchinson Cancer Research Center, 1124 Columbia Street, Seattle, WA 98104. E-mail: eostrand@fhcrc.org

This article represents the opinion of the authors and has not been peer reviewed.



0002-9297/97/6103-0003$02.00

## Simple Traits

Most of classical genetics has been concerned with traits that can be changed by modifying single genes. The morphology of peas, eye color of fruit flies, coat color of mice, and some relatively rare diseases of humans, such as Huntington, are familiar examples. In dogs, the most thoroughly studied single gene traits have been genetic diseases.

It is not surprising that the same mating programs that have resulted in breeds with distinctive size, shape, and abilities have produced animals with a variety of genetic disorders. Genetic diseases are predicted to occur with high frequency in populations with closed gene pools and in which breeding of close relatives is used to propagate desired traits. Breeds established from a small number of founders and expanded rapidly to meet breeders' and consumers' demands suffer the most. Autosomal recessive traits present the biggest problem, since the status of asymptomatic carriers may not be suspected until several litters have been produced. This includes diseases such as cancer, heart disease, skin disorders, and a host of autoimmune disorders. If a popular stud dog turns out to be a carrier of a recessive trait, the effect on the breed can be devastating, because some stud dogs sire dozens of litters in a year. For breeders and owners (not to mention dogs!), the proliferation of genetic diseases has been an unhappy byproduct of breeding practices; for the biomedical researcher, it is a unique resource.

Even in the absence of an ordered genetic map, significant efforts have been made to localize genes of interest in dogs, particularly disease genes. In part, this is because the development of markers linked to disease

genes will lead to commercially available diagnostics, with consequent improvements in canine health. Toward that end, a number of canine disease genes have already been identified, based largely on their similarity to known human disease genes. These include genes for von Willebrand disease (Brooks et al. 1996), severe combined immunodeficiency disorder (Somberg et al. 1995), X-linked Duchenne type muscular dystrophy (Bartlett et al. 1996), pyruvate kinase deficiency (Whitney et al. 1995), muscle type phosphofructokinase deficiency (Smith et al. 1996), and hemophilia (Mauser et al. 1996). In one case, an anonymous marker has been found that is linked to a disease: Yuzbasiyan-Gurkan et al. (1997) demonstrated linkage of a microsatellite marker to copper toxicosis, an autosomal recessive disease of copper accumulation that affects 25% of Bedlington Terriers in the United States. The marker can now be used by breeders to identify carriers of the disease, which can then be selectively removed from breeding programs.

The second motivation for mapping disease genes in dogs is that many human and canine diseases are thought to have related causes, and localizing disease genes in dogs is likely to be substantially easier than in humans. First, large purpose-bred canine pedigrees with dozens of offspring can readily be developed. By comparison, human families are naturally smaller, and the number of informative meioses is therefore restricted. This is particularly true in families with inherited disease, because couples often elect not to have additional children once they produce a single affected child. Second, selective inbreedings can be conducted in canine pedigrees to eliminate heterogeneous genetic backgrounds that may complicate phenotypic assignment. Thus, the problem of phenocopy, which often confounds human studies, is greatly reduced. Examples of purpose-bred pedigrees that are currently the focus of genetic mapping efforts include those with spinal muscular atrophy, cleft palate, narcolepsy, cardiac malformation, and retinal disease.

## Genetic Pathways

No gene acts in isolation. The execution of even the simplest biological process requires the sequential action of a pathway of genes, as in the familiar metabolic pathways of *Escherichia coli.* One of the key insights in modern genetics was the realization that inactivation of many of the genes in a dependent pathway can lead to essentially the same phenotype, inactivation of the pathway. Conversely, isolation of multiple unlinked mutations with a common phenotype may identify many pieces of the molecular machinery responsible for carrying out a complex biological task, as has been exemplified in studies of developmental genetics (Nusslein-Volhard and Wieschaus 1980; Lehmann et al. 1983).

The notion that a complex process can be dissected molecularly by the identification of mutations with a common phenotype has also proven its value in the study of human disease. Genetic study of a group of phenotypically identical lethal skin blistering diseases, for example, identified the genes for each of the two chains of a single receptor, $\alpha_6\beta_4$ integrin, as well as its ligand, laminin V. Furthermore, analysis of diseases with related but distinguishable phenotypes identified the cytokeratins and accessory proteins that link to the intracellular domain of the integrin receptor. Thus, the approach of investigating the biological pathway of a human disease by genetic isolation of its constituents is tremendously powerful, in principle. It suffers from a serious limitation, however: human genetic disorders are inherently rare, far too rare to be useful for defining multiple steps in most biochemical pathways.

The dog provides an opportunity for deducing the genetic pathways of disease in an animal whose physiology is closely matched to that of humans. For example, generalized progressive retinal atrophy (PRA) in dogs encompasses a heterogeneous group of genetic disorders that show similarity to retinitis pigmentosa in humans. PRA has been reported in a number of breeds. The clinical phenotype varies in different breeds primarily in the age at onset and rate of progression, but various studies clearly demonstrate that mutations in different genes are responsible for the disease. For example, breeding studies show that three of the early-onset autosomal disorders are not allelic, while a fourth form of the disease (in Siberian Huskies) is X-linked (Acland et al. 1989, 1994). Moreover, the PRA mutation found in Irish Setters (*rcd1*) has been shown to be a nonsense mutation of the canine PDEβ gene (Suber et al. 1993), but this gene is wild type in PRA-affected Tibetan Terriers and Miniature Longhaired Dachsunds (Clements et al. 1993). In contrast, later onset autosomal PRA in several breeds is allelic with *prcd,* a mutation characterized in Miniature Poodles (Aguirre and Acland 1988). Mapping and cloning the mutant PRA loci from a variety of breeds will test the hypothesis that in mammals, as in invertebrates, genetic pathways can be constructed by systematic analysis of mutations that display similar phenotypes.

A second way to define gene pathways is to screen for genetic modifiers, extragenic mutations that enhance or suppress the phenotype of a starting mutation. This approach has been a remarkably effective for finding the missing pieces of complex genetic networks in simple model systems such as *Drosophila* (Simon et al. 1991). In mammalian systems, the equivalent experiment can be done simply and economically, by outbreeding to

uncover the effects of genetic background on the expression of a phenotype. The interacting loci so identified may provide unique insights into the variability of phenotypes seen among affected individuals in normal outbred populations, such as human societies. In mice, for example, strain differences in the number of tumors developed by animals that are heterozygous for a mutation in adenomatous polyposis coli were used to map a second gene, *Mom1,* involved in the development of intestinal adenomas (Dietrich et al. 1993). The power of the approach depends directly on the amount of natural genetic variation that can be brought to bear on a problem. It follows that the diversity among modern dog breeds should provide an enormous reservoir of allelic differences, which can be exploited to discover new elements of many genetic pathways.

## Complex Traits

Almost all of classical genetics has been built on the strength of single gene traits. It is clear, however, that many of the properties of an animal reflect the concerted action of arrays of genes. Historically, the study of traits that derive irreducibly from the interactions of multiple genes was not practical. Technical advances of the past decade, however, have made study of complex traits feasible, and they promise to be a major focus of genetic study in the coming decade. Quantitative trait loci (QTL) analysis of complex traits requires the ability to mate parental individuals who display the two extremes, respectively, of a quantifiable phenotype; to perform directed backcrosses that generate large pedigrees; and to correlate the segregation of polymorphic markers in the $F_2$ generation with the segregation of the trait of interest.

It is in the study of complex traits that the unique advantages of canine genetics truly come to the fore. Dogs are unique among mammalian species in the extent of variation they show in morphological traits such as height, weight, mass, and shape, yet within a breed each of these traits is inherited within extremely narrow limits. The Chihuahua is <6 inches high at the shoulder; the Irish wolfhound close to 3 ft. The Pomeranian weighs between 4 and 5 lb.; the St. Bernard may weigh 150 lb. The Collie has a small, narrow head like that of a fox; the Pug has a massive head with a short, blunt muzzle. No other mammalian species presents natural variation on a scale to rival these, yet individuals from nearly any breeds can be mated to yield fertile offspring. It would appear that application of QTL methodology to appropriately chosen and bred dogs should be suitable for localizing the major loci responsible for the morphological differences that distinguish different breeds.

The notion of mapping genes for morphological traits in dogs is a straightforward extrapolation from mapping loci responsible for external characteristics in other organisms. What is perhaps less transparent is the applicability of this formalism to mapping the genes that underlie behavior. The relative importance of genetics and environment in controlling behavior has been argued for at least 2,500 years (Aristotle [c. 335–322 B.C.] 1974, chap. 5, sect. 10). Animal breeding studies and twin/adoption studies of humans in this century provide evidence for a genetic component to behavior. Nevertheless, the variability of even relatively stereotyped behavior, the clear influence of learning, and the challenge of quantifying behavior have all conspired to confound the debate. How can the study of dogs serve to clarify this contentious issue?

It is apparent from even casual observation that the behaviors and much of the personality of a dog can be characteristic of its breed, rather than being strictly individual (Darwin [1859] 1983, p. 239). This in itself hints at a genetic contribution to canine behavior. Consider the Pointer that points, the draft dog that pulls, and the Bloodhound that tracks. A Border Collie raised away from a farm will still demonstrate the classic herding behaviors of circling, crouching, and eye when appropriately prompted. Few would contest that Dalmatians are high-strung, Border Collies obsessive, and Doberman Pinschers protective. Simple breeding studies have suggested that aggression, herding, and, perhaps, loyalty are among the canine behavioral traits that are likely to be controlled, in part, by genetic factors (for review, see Willis 1989). More extensive behavioral and genetic analysis has been done to analyze a "human aversion" trait, which has been maintained by selective mating and inbreeding of two lines of Pointers (Murphree and Newton 1971; Murphree 1973). Affected lines are nervous, fearful, and timid. They avoid human contact and will become catatonic and rigid when humans approach. Dogs tested from a control line, bred under the same conditions, behave normally. Attempts to condition friendliness into dogs from a fearful line through a daily program of friendly human approach are of little success, suggesting that some aspects of fearfulness or shyness may have strong underlying genetic components.

If a trait, even a behavioral trait, can be measured reliably and can be bred, its genes can be mapped; if it can be mapped, it can be cloned. Thus, the dog is uniquely suited to the molecular study of complex behaviors. What is required at the outset is a genetic map of the dog genome.

## Organization of the Canine Genome

One challenge to the use of dogs as a genetic system has been the difficulty in characterizing the organization

of the canine genome. Cytogenetic analysis of dog chromosomes historically has been extremely difficult, since the dog has 38 pairs of autosomes, most of which are small and acrocentric. New cytological techniques have recently permitted dog chromosomes to be defined to the 400-band level (Graphodatsky et al. 1995). Chromosome numbers have now been assigned to the 21 largest chromosomes (Switonski 1995). The most convenient resource for deciphering the organization of the canine genome, however, is a set of chromosome-specific "paints," which were recently developed from flow-sorted canine chromosomes (Langford et al. 1995). These paints will provide a resource for investigators to tag and number the smaller chromosomes unambiguously.

## The Development of a Canine Map

The key resource necessary for mapping traits of interest in dogs is a canine genetic map. To be useful in the dog, such a map should have markers placed at least every 10 cM. The development of microsatellite-based markers that can be used for the mapping of traits within purebred dogs is challenging. Because of significant inbreeding, microsatellite-based markers based on $(CA)_n$ repeats often show inadequate heterogeneity between dogs of the same breed (Ostrander et al. 1995; Zajc et al. 1997). Recently, however, Francisco et al. (1996) described a class of tetranucleotide repeats in dogs that is based on the $(GAAA)_n$ repeat sequence and that have an average PIC of 0.75. This class of markers appears to be ideally suited for mapping, even within pedigrees of purebred dogs.

Preliminary linkage groups have been reported previously for a small number of $(CA)_n$ repeat–based markers (Lingaas et al. 1997). More recently, a genetic linkage map spanning a significant portion of the canine genome has been developed (C. S. Mellersh, A. A. Langston, G. M. Acland, M. A. Fleming, K. Ray, N. A. Weigand, L. V. Francisco, M. Gibbs, G. D. Aguirre, and E. A. Ostrander, unpublished data). Two factors were key to the success of this latter effort. First, the mapping panel consisted of 26 individual three-generation pedigrees containing 351 individuals with 282 $F_2$ offspring. This panel contained sufficient phase-known meioses to detect linkage with a LOD score of 3.0 between maximally informative markers spaced $\leq$40 cM apart. Second, a large number of markers placed on the map were based on tetranucleotide repeats. The combination of highly informative markers and large, purpose-bred pedigrees allowed investigators to overcome the difficulties faced by previous efforts.

If the size of the dog genome is assumed to be similar to that of humans, a framework map will require $\sim$350 well-spaced and highly informative markers to cover the entire genome. Increasing map density further will, of course, improve the precision with which a trait can be mapped. Because of practical limitations, however, it is unlikely that the canine map will ever approach the density of the human map, which currently has $>$6,000 markers. The most efficient approach for extending the utility of a map made with a limited number of markers is to exploit the evolutionary relationship that exists between all mammalian genomes.

Comparative genome analyses require that a common set of loci be mapped in all genomes, to serve as landmarks for conserved segments of chromosomes. Lyons et al. (1997) described a set of 410 primer pair sequences that were designed to amplify specific, highly conserved genes from any mammalian species. Since these primer pairs were designed to span introns, where a maximum of polymorphism is expected to occur, their placement on the canine map will allow speedy integration of the linkage map with the gene map. For the short term, establishing regions of synteny will allow investigators to take advantage of the human and mouse maps to identify candidate genes for linkages established between anonymous microsatellite-based markers and traits of interest in the dog.

The quickest way to bridge the gap between anonymous genetic markers, genes, and chromosomes is through the use of hybrid cell lines. Ideally, such lines should each contain a small number of canine chromosomes in some other genetic background, such as a rodent cell, making it possible to use PCR to assign a particular gene or DNA segment to a given canine chromosome. A panel of such hybrids that appears to span most of the canine genome has now been constructed (A. A. Langston, C. S. Mellersh, C. Neal, G. M. Acland, K. Ray, G. D. Aguirre, R. E. K. Fournier, and E.A. Ostrander, unpublished data). Testing of primer pairs defining anchor genes, as well as anonymous microsatellite-based markers, quickly determines which genes and markers are likely to be on the same canine chromosome. If one of the markers is found to be linked to a trait of interest, examination of the cosegregating genes will often suggest a region of the human or canine genome to search for candidate genes.

The value of the comparative approach is illustrated by Werner et al. (1997), who showed that a 38-cM region of canine chromosome 9 has syntenic homology to human chromosome 17q. The organization of genes is maintained from *P4HB* to *NF1,* including *GALK, TK1, GH1, BRCA1, RARA,* and *MPO,* although the entire region is inverted in relation to the centromere. Similarly, two genes located on human 17p are located on canine chromosome 5 (Werner et al. 1997). Obviously, the next step will be integration of these gene

maps with the linkage map of anonymous microsatellite markers. Once the syntenic relationships of mammalian genomes are fully characterized, the mapping of a gene in any one mammalian species will be tantamount to its mapping in all mammals. This will permit any gene of interest to be studied in the animal model that is most convenient for its analysis.

### Conclusion

The dog has been man's faithful servant for 1,000 centuries. From the dog's wolf ancestors, man has created companions, shepherds, hunters, guides and guards, each with a distinctive size, shape and coloration appropriate to its function. From within the wolf personality, breeders have purified affection and aggression, loyalty, industry, and playfulness. Each of these various and disparate traits was already present as a separable element of the dog's genetic heritage, as were the disease alleles that continue to plague modern dogs. The opportunity and challenge for contemporary genetics is to exploit that heritage to understand the genetic regulation of morphology, disease, and behavior in the mammals.

## Acknowledgments

E.A.O. is the recipient of an American Cancer Society Junior Faculty Award and grants from the Canine Health Foundation of the American Kennel Club, Muscular Dystrophy Association, and American Cancer Society Grant VM-165. E.G. is supported, in part, by funds from the Fred Hutchinson Cancer Research Center and is a Basil O'Connor Starter Scholar of the March of Dimes Birth Defects Foundation. We thank Drs. A. Langston, M. Gibbs, P. Welch, G. Lark, and G. Acland for careful reading of this manuscript and helpful comments.

## References

Acland GM, Blanton SH, Hershfield B, Aguire GD (1994) XLPRA: a canine retinal degeneration inherited as an X-linked trait. Am J Med Genet 52:27–33

Acland G, Fletcher RT, Gentleman S, Chader G, Aguirre G (1989) Non-allelism of three genes (rcd1, rcd 2, and erd) for early-onset hereditary retinal degeneration. Exp Eye Res 49:983–998

Aguirre GD, Acland GM (1988) Variation in retinal degeneration phenotype inherited at the prcd locus. Exp Eye Res 46: 663-687.

Aristotle (1974) The politics. Barker E (ed and trans). Oxford University Press, New York

Bartlett RJ, Winand NJ, Secore SL, Singer JT, Fletcher S, Wilton S, Bogan DJ, et al (1996) Mutation segregation and rapid carrier detection of X-linked muscular dystrophy in dogs. Am J Vet Res 57:650–654

Brooks M, Raymond S, Catalfamo J (1996) Severe, recessive von Willebrand's disease in German Wirehaired Pointers. J Am Vet Med Assoc 209:926–929

Clements PJM, Gregory CY, Peterson-Jones SM, Sargan DR, Bhattacharya SS (1993) Confirmation of the rod cGMP phosphodiesterase β subunit (PDEβ) nonsense mutation in affected rcd-1 Irish setters in the UK and development of a diagnostic test. Curr Eye Res 9:861–866

Darwin C ([1859] 1983) The origin of species. Burrow JW (ed). Penguin, New York

Dietrich WF, Lander ES, Smith JS, Moser AR, Gould KA, Luongo C, Borenstein N, et al (1993) Genetic identification of Mom-1, a major modifier locus affecting Min-induced intestinal neoplasia in the mouse. Cell 75:631–639

Francisco LV, Langston AA, Mellersh CS, Neal CL, Ostrander EA (1996) A class of highly polymorphic tetranucleotide repeats for canine genetic mapping. Mamm Genome 7:359–362

Graphodatsky AS, Beklemisheva VR, Dolf G (1995) High-resolution GTG-banding patterns of dog and silver fox: description and comparative analysis. Cytogenet Cell Genet 69:226–231

Langford CF, Fischer PE, Binns MM, Holmes NG, Carter NP (1995) Chromosome-specific paints from a high–resolution flow karyotype of the dog. Chrom Res 3:1–10

Lehman R, Jimenez F, Dietrich U, Campos-Ortega JA (1983) On the phenotype and development of mutants of early neurogenesis in *Drosophila melanogaster:* Wilhelm Roux's arch. Dev Biol 192:62–74

Lingaas F, Sorensen A, Juneja RK, Johansson S, Fredholm F, Wintero AK, Sampson J, et al (1997) Towards construction of a canine linkage map: establishment of 16 linkage groups. Mamm Genome 8:218–221

Lyons LA, Laughlin TF, Copeland NG, Jenkins NA, Womack JE, O'Brien SJ (1997) Comparative anchor tagged sequences (CATS) for integrative mapping of mammalian genomes. Nat Genet 15:47–56

Mauser AE, Whitlark J, Whitney KM, Lothrop CD (1996) A deletion mutation causes hemophilia B in Lhasa Apso dogs. Blood 88:3451–3455

Murphree OD (1973) Inheritance of human aversion and inactivity in two strains of the pointer dog. Biol Psych 7:23–29

Murphree OD, Newton JEO (1971) Crossbreeding and special handling of genetically nervous dogs. Cond Reflex 6:129

Nusslein-Volhard C, Wieschaus E (1980) Mutations affecting segment number and polarity in *Drosophila*. Nature 287: 795–801

Ostrander EA, Mapa F, Yee M, Rine J (1995) One hundred and one new simple sequence repeat-based markers for the canine genome. Mamm Genome 6:192–195

Simon MA, Bowtell DDL, Dodson GS, Laverty TR, Rubin GM (1991) Ras1 and a putative guanine nucleotide exchange factor perform crucial steps in signaling by the SEVENLESS protein tyrosine kinase. Cell 67:701–716

Smith BF, Stedman H, Rajpurohit Y, Henthorn PS, Wolfe JH, Patterson DF, Giger U (1996) Molecular basis of canine muscle type phosphofructokinase deficiency. J Biol Chem 271:20070–20074

Somberg RL, Pullen RP, Casal ML, Patterson DF, Felsburg

PJ, Henthorn PS (1995) A single nucleotide insertion in the canine interleukin-2 receptor gamma chain results in X-linked severe combined immunodeficiency disease. Vet Immunol Immunopathol 47:203–213

Suber ML, Pittler SJ, Qin N, Wright GC, Holcombe V, Lee RH, Craft CM, et al (1993) Irish setter dogs affected with rod/cone dysplasia contain a nonsense mutation in the rod cGMP phosphodiesterase β subunit gene. Proc Natl Acad Sci USA 90:3968–3972

Switonski M (1995) Committee for the Standardization of the Canine Karyotype. Second International DOGMAP Meeting, Cambridge, United Kingdom

Vila C, Savolainen P, Maldonado JE, Amonm IR, Rice JE, Honeycutt RL, Crandall KA, et al (1997) Multiple and ancient origins of the domestic dog. Science 276:1687–1689

Whitney KM, Lothrop CD (1995) Genetic test for pyruvate kinase deficiency of Basenjis. J Am Vet Med Assoc 207: 918–921

Woolf V (1933). Flush: a biography. Harcourt, Brace, New York

Werner P, Raducha MG, Prociuk U, Henthorn PS, Patterson DF (1997) Physical and linkage mapping of human chromosome 17 loci to dog chromosomes 9 and 5. Genomics 42: 74–82

Willis MW (1989) Genetics of the dog. Howell, New York

Yuzbasiyan-Gurkan V, Blanton SH, Cao Y, Ferguson P, Li J, Vental PJ, Brewer GJ (1997) Linkage of a microsatellite marker to the canine copper toxicosis locus in the Bedlington terrier. Am J Vet Res 58:23–27

Zajc I, Mellersh CS, Sampson J (1997) Variability of canine microsatellites within and between different dog breeds. Mamm Genome 8:182–185