Evaluation of Candidate Genes in the Absence of Positional Information: A Poor Bet on a Blind Dog!

J. Aguirre-Hernández and D. R. Sargan

From the Centre for Veterinary Science, University of Cambridge, Cambridge CB3 0ES, UK (Aguirre-Hernández and Sargan), and Hospital Infantil de México Federico Gómez, Dr. Márquez 162, México, D.F. 06720, México (Aguirre-Hernández).

Address correspondence to J. Aguirre-Hernández, Centre for Veterinary Science, University of Cambridge, Cambridge CB3 0ES, UK, or e-mail: ja248@cam.ac.uk.

Abstract

More than 350 inherited diseases have been reported in dogs and at least 50% of them have human counterparts. To remove the diseases from dog breeds and to identify canine models for human diseases, it is necessary to find the mutations underlying them. To this end, two methods have been used: the functional candidate gene approach and linkage analysis. Here we present an evaluation of these in canine retinal diseases, which have been the subject of a large number of molecular genetic studies, and we show the contrasting outcomes of these approaches when dealing with genetically heterogeneous diseases. The candidate gene approach has led to 377 published results with 23 genes. Most of the results (66.6%) excluded the presence of a mutation in a gene or its coding region, while only 3.4% of the results identified the mutation causing the disease. On the other hand, five linkage analysis studies have been done on retinal diseases, resulting in three identified mutations and two mapped disease loci. Mapping studies have relied on dog research colonies. If this favorable application of linkage analysis can be extended to dogs in the pet population, success in identifying canine mutations could increase, with advantages to veterinary and human medicine.

Dogs were domesticated more than 15,000 years ago and subgroups have been selected for particular traits ever since (Savolainen et al. 2002). More than 350 inherited diseases have been reported in this species (Nicholas 2003), and a recent survey of the literature suggests that the number of diseases with major inherited components may considerably exceed this figure (Sargan 2004). This places dogs as the species with the second largest number of known genetic diseases, surpassed only by humans. Particular diseases occur with a high incidence in specific breeds, which may be thought of as genetically isolated and inbred subgroups within the entire species. This large number of reported inherited diseases has two main causes: the close surveillance to which dogs are subjected, and the founder effect and inbreeding practiced in pure-bred dogs that uncovers recessive disease alleles. The frequency of these alleles may increase by “popular sire” effects, whereby a very small proportion of the available males (those successful in dog shows or in competition events for working breeds) contribute disproportionately to the next generation. Thus over the last 30 years, only 3%-5% of registered dogs were used to produce the current Dutch pure-bred dog population (Ubbink et al. 1998).

To reduce the incidence of canine recessive diseases, which account for about 70% of all inherited diseases with a known mode of inheritance (Brooks and Sargan 2001), carriers of the mutant allele, as well as affecteds, have to be identified before they reach sexual maturity. Carriers and affecteds are then either removed from breeding or mated only to individuals with the wild-type allele (Petersen-Jones et al. 1995).

In addition, more than 50% of canine inherited diseases are shared with humans (Nicholas 2001; Ostrander et al. 2000; Sargan 2004) and the coding sequences of dogs and humans show an overall greater similarity to each other than to mouse coding sequences (Kirkness et al. 2003), even though mouse and human share a more recent common ancestor. Thus dogs can be used as models to understand many diseases and to develop therapies before they are tested in humans, as is being done with muscular dystrophy (Bartlett et al. 2000), Leber congenital amaurosis (Acland et al. 2001; Narfstrom et al. 2003), clotting disorders (Mount et al. 2002), lysosomal storage diseases (Ponder et al. 2002), dystrophic epidermolysis bullosa (Balducci et al. 2003), hemophilia A (Chuah et al. 2003), and hemophilia B (Mount et al. 2002),
among others. For this to be done, the first step is to find the mutations.

Two approaches have predominated in the search for mutations involved in dog diseases: the functional candidate gene (FCG) approach and linkage analysis (LA). In the former, the mutation is pursued in genes that have been implicated in a similar disease in another species or in genes implicated by their known normal activity or by their distribution among cells or tissues. In LA, on the other hand, the procedure consists of three steps. First, the segregation of genetic markers and the disease are analyzed in families with affected individuals to identify those markers that cosegregate with the disease. This indicates the region where the disease locus is found. Then, fine mapping with additional markers in the region of interest is performed to increase the resolution. Finally, candidate genes within the region of interest, chosen by their role, structure, or transcript distribution, are analyzed, hence the terms positional cloning and positional candidate genes used in this context (Collins 1992).

A large number of genetic studies have been done on canine retinal diseases, so in this article we use those results to show the contrasting outcomes of using the FCG approach and LA to find genes involved in canine diseases, particularly those with genetic heterogeneity.

**Canine Retinal Diseases**

These comprise, among others, progressive retinal atrophy (PRA), cone dystrophy (cd), and retinal dystrophy, diseases that impair the normal functioning of dog photoreceptors or the retinal pigment epithelium, leading to partial or total blindness (Clements et al. 1996; Lin et al. 2002; Petersen-Jones 1998b). The age of onset varies for these diseases, depending on the nature of the mutation and, to a lesser extent, on the genetic background of affected individuals. These diseases have human counterparts (Petersen-Jones 1998a). For example, PRA, cd (Sidjanin et al. 2002), and retinal dystrophy (Aguirre et al. 1998; Veske et al. 1998, 1999) are homologous to retinitis pigmentosa (RP), achromatopsia (Sundin et al. 2000), and Leber congenital amaurosis (Gu et al. 1997; Marhens et al. 1997; Morimura et al. 1998), respectively.

Studies to identify the mutations underlying dog retinal diseases have been numerous for three reasons. First, these diseases affect many breeds (e.g., PRA has been reported for more than 100 breeds) (Clements et al. 1996; Petersen-Jones 1998b). Second, they are comparatively easy to detect. Third, they show genetic heterogeneity, which means that the identification of a mutation or the exclusion of a gene in one breed may tell nothing about that gene in other breeds. In addition to this genetic heterogeneity among breeds, there is also genetic heterogeneity within some breeds (e.g., Briard and Labrador retriever; see supplementary information in Table S1). The opposite situation may also exist: progressive rod-cone degeneration (prcd) is allelic in at least six breeds (Narfstrom and Wrigstad 1999; Ray et al. 1999) and cd is allelic in two (Sidjanin et al. 2002), and two breeds with X-linked PRA (XLPRA) share the same mutation (Zhang et al. 2002).

To date, 10 mutations causing canine retinal diseases have been found, affecting 11 breeds plus a single individual of mixed breed (Table 1). In addition, two PRA loci have been mapped, but the genes have not yet been identified (Table 2).

**FCG Studies**

Using the candidate gene approach, 377 results have been published on 23 genes addressing 72 breed diseases in 64 breeds plus a mongrel individual (Table S1). We define a breed disease as each disease present in one breed. For example, prcd and retinal degeneration, both seen in Labrador retrievers, are two different breed diseases (Kommonen and Karhunen 1990); likewise, prcd in the poodle and the English cocker spaniel are considered different breed diseases inasmuch as they occur in different breeds, even though it is known that they are allelic (Aguirre and Acland 1988).

Most of the results (60.2%) have excluded the presence of mutations in particular genes (Table S1). The exclusions have been derived from the distribution of intragenic polymorphisms in functional candidate genes (FCGs) among affected and nonaffected individuals. In this approach, it is assumed that the mutation underlying each disease occurred only once in the breed, and it appeared in the context of a particular haplotype, so all affected individuals in a breed disease must share that haplotype, while nonaffecteds may have several different haplotypes. However, instead of analyzing complete haplotypes, usually individual intragenic polymorphisms are studied separately. If the recessive mutation lies in the gene being studied, all affected individuals are expected to be homozygous for the same allele, while nonaffecteds may have any of a number of different combinations of alleles in homozygous and heterozygous genotypes. Thus heterozygosity in intragenic polymorphisms of a particular gene in affected dogs, or the presence of two or more different homozygous genotypes, excludes that gene as the cause of the disease. For this procedure to be applied, it is important to have clear evidence regarding the recessive inheritance of the disease.

In 6.4% of the results, exon or cDNA sequencing excluded the presence of a mutation in the coding region of candidate genes or their transcripts. The sequences were compared in affected and nonaffected dogs, and no differences were found. Nevertheless, strictly speaking, the possibility remains for a mutation to be present in noncoding regions of the gene. For this reason, the analysis of intragenic polymorphisms is more powerful than merely sequencing the coding regions.

In a third category representing 8.22% of the results, instead of sequencing coding regions, intragenic polymorphisms in coding or noncoding regions were searched for. However, results were uninformative since no intragenic polymorphisms were found among the dogs studied—affected and nonaffected—so no information could be obtained regarding the segregation of these genes and their possible involvement in causing the disease.
<table>
<thead>
<tr>
<th>Breed</th>
<th>Disease</th>
<th>Inheritance and age of onset</th>
<th>Gene</th>
<th>Mutation</th>
<th>Protein</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>PRA</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>English mastiff, bull mastiff</td>
<td>PRA</td>
<td>AD, EO</td>
<td>RHO</td>
<td>Exon 1, missense, 147C&gt;G, T4R (X*1380)</td>
<td>Mutation located in N-terminal portion of the protein, forming part of the extracellular domain</td>
<td>Kijas et al. 2002, 2003</td>
</tr>
<tr>
<td>Irish setter</td>
<td>rcd1</td>
<td>AR, EO</td>
<td>PDE6B</td>
<td>Exon 21, nonsense, 2426G&gt;A, W807X (Z23014)</td>
<td>Truncated by 49 residues</td>
<td>Clements et al. 1993; Ray et al. 1994; Suber et al. 1993</td>
</tr>
<tr>
<td>Miniature schnauzer</td>
<td>pd</td>
<td>AR', EO</td>
<td>PDC</td>
<td>Exon 4, missense, 326C&gt;G, R82G (Y17697)</td>
<td>Non-conservative substitution close to the residue (Glu 85) interacting with the beta and gamma-subunits of transducin</td>
<td>Zhang et al. 1998</td>
</tr>
<tr>
<td>Mongrel, Sloughi</td>
<td>XLPRA2</td>
<td>X-linked</td>
<td>RPGR</td>
<td>Exon ORF15, frameshift, 2 bp deletion, 1084_1085delGA (AF388629)</td>
<td>Premature termination, protein lacks C-terminal tail</td>
<td>Zhang et al. 2002</td>
</tr>
<tr>
<td>Samoyed, Siberian</td>
<td>XLPRA1</td>
<td>X-linked</td>
<td>RPGR</td>
<td>Exon ORF15, frameshift, 5 bp deletion, 1028_1032delGAGAA (AF388629)</td>
<td>Premature termination, protein lacks 230 C-terminal tail residues</td>
<td>Zhang et al. 2002</td>
</tr>
<tr>
<td>Welsh corgi (Cardigan)</td>
<td>rcd3</td>
<td>AR, EO</td>
<td>PDE6A</td>
<td>Exon 15, frameshift, 1 bp deletion, 1940delA, N616fs (Z68340)</td>
<td>Truncated by 218 residues</td>
<td>Petersen-Jones et al. 1999</td>
</tr>
<tr>
<td>Alaskan malamute</td>
<td>cd</td>
<td>AR, EO</td>
<td>CNGB3</td>
<td>Deletion of region comprising entire gene and adjacent sequences</td>
<td>Absent</td>
<td>Sidjanin et al. 2002</td>
</tr>
<tr>
<td>German shorthaired pointer</td>
<td>cd</td>
<td>AR, EO</td>
<td>CNGB3</td>
<td>Exon 6, missense, 808G&gt;A, D262N (AF490511)</td>
<td></td>
<td>Sidjanin et al. 2002</td>
</tr>
<tr>
<td>Retinal dystrophy</td>
<td>Briard</td>
<td>AR, EO</td>
<td>RPE65</td>
<td>Exon 5, frameshift, 4-bp deletion, 487_490delAAGA, K154fs (Y16567)</td>
<td>Premature termination, protein truncated by more than two-thirds</td>
<td>Aguirre et al. 1998; Veske et al. 1998, 1999</td>
</tr>
</tbody>
</table>

*AD, autosomal dominant; ADPRA, autosomal dominant progressive retinal atrophy; AR, autosomal recessive; cd, cone degeneration; csnb, congenital stationary night blindness; EO, early age of onset; ORF, open reading frame; pd, photoreceptor dysplasia; PRA, progressive retinal atrophy; rcd1, rod-cone dysplasia 1; rcd3, rod-cone dysplasia 3; XLPRA1, X-linked progressive retinal atrophy 1; XLPRA2, X-linked progressive retinal atrophy 2.

* The position of the mutation is given according to the numbering in the sequence in parentheses.

* The suggestion has been made that in a number of cases the disease could be digenic (Zhang et al. 1998).

* Several splice variants are known for RPGR; the nucleotide position of the mutation corresponds to the numbering in ORF15.
### Table 2. Mapping studies on canine disease loci (retinal or otherwise)*

<table>
<thead>
<tr>
<th>Disease</th>
<th>Inheritance</th>
<th>Gene</th>
<th>Mutation</th>
<th>Breeds</th>
<th>Resources employed</th>
<th>Markers used</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Narcolepsy</td>
<td>AR</td>
<td>Hctr2</td>
<td>SINE insertion in intron 3 with exon skipping Transition leading to exon 6 skipping p.E54K</td>
<td>Doberman pinscher</td>
<td>Colony (15 affected, 9 carrier) and general population Colony (7 affected, 9 carrier)</td>
<td>RFLP in canine Sf and MHC loci</td>
<td>Cederberg et al. 1998; Huns et al. 2001; Lin et al. 1999; Mignot et al. 1991</td>
</tr>
<tr>
<td>SCID</td>
<td>X linked</td>
<td>IL2RG</td>
<td>exon 1, 4 bp deletion</td>
<td>Dachshund</td>
<td>General population</td>
<td>Four markers on canine CFAX WGS with 213 microsatellites</td>
<td>Deschenes et al. 1994; Henthorn et al. 1994 van De Sluis et al. 2002; Yueh and Gruenen 2007</td>
</tr>
<tr>
<td>CT</td>
<td>AR</td>
<td>MMUR1</td>
<td>exon 2 deleted</td>
<td>Bedlington terrier</td>
<td>General population (75 individuals): 25 affected</td>
<td>WGS with 213 microsatellites</td>
<td>Zeiss et al. 2000; Zhang et al. 2001, 2002</td>
</tr>
<tr>
<td>XLPRA1</td>
<td>X linked</td>
<td>RPGR</td>
<td>exon ORF15, 5 bp deletion</td>
<td>Samoyed and Siberian husky</td>
<td>Colony; Siberian husky male bred to beagles (eight three-generation families with 77 backcross informative individuals); further mapping with 150 dogs</td>
<td>Five intragenic markers on CFAX, followed by 11 microsatellites on CFAX</td>
<td></td>
</tr>
<tr>
<td>MH</td>
<td>AD</td>
<td>RYR1</td>
<td>exon 15, p.V547A</td>
<td>Mixed breed (Doberman/German shepherd/collie)’</td>
<td>Colony; mixed breed male bred to Labrador retriever</td>
<td>Nine microsatellite markers in the vicinity of RYR1 on CFA1</td>
<td>Roberts et al. 2001</td>
</tr>
<tr>
<td>cd</td>
<td>AR</td>
<td>CNGB3</td>
<td>Deletion of region comprising entire gene and adjacent sequences p.D26SN</td>
<td>Alaskan malamute</td>
<td>Colony; Alaskan malamute male bred to beagles (11 families with 135 individuals; 88 informative progeny)’</td>
<td>WGS with microsatellites</td>
<td>Sidjian et al. 2002</td>
</tr>
<tr>
<td>RCND</td>
<td>AD</td>
<td>BD</td>
<td>exon 7, p.H225R</td>
<td>German shepherd</td>
<td>Colony; German shepherd/flat-coated retriever bred to German shepherd and English setters (72 dogs: 27 affected)</td>
<td>WGS with microsatellites at approximately 10 cM resolution</td>
<td>Jonasdottir et al. 2000; Lingaas et al. 2003</td>
</tr>
</tbody>
</table>

### Mapped loci

<table>
<thead>
<tr>
<th>Disease</th>
<th>Inheritance</th>
<th>Gene</th>
<th>Mutation</th>
<th>Breeds</th>
<th>Resources employed</th>
<th>Markers used</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>prcd</td>
<td>AR</td>
<td>CFA9</td>
<td>American cocker spaniel, English cocker spaniel, Labrador retriever, papillon, poodle (miniature), Portuguese water dog’</td>
<td>Colony; Miniature poodles bred to beagles (nine three-generation families with 70 informative progeny)</td>
<td>WGS, approximately 100 microsatellites</td>
<td>Acland et al. 1998; Narfstrom and Wrigstad 1999; Ray et al. 1999</td>
<td></td>
</tr>
<tr>
<td>HCSMA</td>
<td>AD</td>
<td>CFA1</td>
<td>Brittany spaniel</td>
<td>Colony; Brittany spaniels mated to two beagles (16 dogs: 11 affected)</td>
<td>WGS, approximately 150 microsatellites</td>
<td>Blazek et al. 1998</td>
<td></td>
</tr>
<tr>
<td>CNCL</td>
<td>AR</td>
<td>CFA37</td>
<td>English setter</td>
<td>Colony; (133 individuals)</td>
<td>WGS with 103 microsatellites</td>
<td>Lingaas et al. 1998</td>
<td></td>
</tr>
<tr>
<td>erd</td>
<td>AR</td>
<td>CFA27</td>
<td>Norwegian elkhound</td>
<td>Colony; Norwegian elkhound bred to beagles</td>
<td>WGS with more than 150 microsatellites</td>
<td>Acland et al. 1999</td>
<td></td>
</tr>
<tr>
<td>CTVM</td>
<td>AD, reduced penetrance</td>
<td>CFA9</td>
<td>Labrador retriever</td>
<td>General population (37 dogs: 18 affected)</td>
<td>WGS with 172 microsatellites</td>
<td>Andelfinger et al. 2003</td>
<td></td>
</tr>
<tr>
<td>MGA</td>
<td>AR</td>
<td>CFA8</td>
<td>Giant schnauzer</td>
<td>Colony; two affected giant schnauzers bred to three unaffected dogs from other breeds (initially 88 dogs; fine mapping with 128 dogs)</td>
<td>WGS with 172 microsatellites; fine mapping with 6 microsatellites and 3 SNPs</td>
<td>He et al. 2003</td>
<td></td>
</tr>
<tr>
<td>Disease</td>
<td>Type</td>
<td>Breed</td>
<td>Colony</td>
<td>Method</td>
<td>References</td>
<td></td>
<td></td>
</tr>
<tr>
<td>---------</td>
<td>------</td>
<td>-------</td>
<td>--------</td>
<td>--------</td>
<td>------------</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Epilepsy</td>
<td>Polygenic with a single AR locus of large effect</td>
<td>Border collies, Australian shepherds&lt;sup&gt;a&lt;/sup&gt;</td>
<td>Colony; rough, smooth, and border collies, and Australian shepherds bred to mixed-breed dogs (104 dogs from 15 litters)</td>
<td>WGS with 172 microsatellites</td>
<td>Lowe et al. 2003</td>
<td></td>
<td></td>
</tr>
<tr>
<td>cnm</td>
<td>AR</td>
<td>Belgian tervuren and sheepdogs</td>
<td>General population (230 dogs: 42 affected)</td>
<td>WGS with 100 microsatellites</td>
<td>Oberbauer et al. 2003</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Labrador retriever</td>
<td>Colony (40 dogs in four generations; 20 affected)</td>
<td>WGS with 66 microsatellites and pooled DNA; individual genotyping of selected regions</td>
<td>Tiet et al. 2003</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CTDs</td>
<td>AR</td>
<td>Keeshond</td>
<td>Colony; keeshonds bred to beagles (83–88 individuals)</td>
<td>RFLP in four loci on telomeric end of CFA26</td>
<td>Werner et al. 1999</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DCM</td>
<td>AD?, incomplete penetrance, age-related penetrance</td>
<td>Newfoundland</td>
<td>General population (48 individuals)</td>
<td>WGS with 200 markers</td>
<td>Dukes-McEwan and Jackson, 2002</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Other</td>
<td>CHD</td>
<td>Labrador retriever</td>
<td>Colony, Labrador retriever bred to greyhounds&lt;sup&gt;b&lt;/sup&gt;</td>
<td>WGS with 1 marker every 10 cM&lt;sup&gt;c&lt;/sup&gt;</td>
<td>Bliss et al. 2002; Todhunter et al. 1999, 2003a,b</td>
<td></td>
<td></td>
</tr>
<tr>
<td>IE</td>
<td>QTL</td>
<td>Labrador retriever</td>
<td>Simulation study; general population (29 affected, plus siblings and parents from 13 three-generation pedigrees)</td>
<td>WGS with 1 marker</td>
<td>Patterson et al. 2003</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>AR</td>
<td>Vizslas</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<sup>a</sup> AD, autosomal dominant; AR, autosomal recessive; cd, cone degeneration; cea, collie eye anomaly; CHD, canine hip dysplasia; cM, centiMorgan; CNCL, neuronal ceroid lipofuscinosis; cnm, centronuclear myopathy; CT, copper toxicosis; CTDs, conotruncal defects; CTVM, canine tricuspid valve malformation; DCM, dilated cardiomyopathy; erd, early retinal degeneration; HCSMA, hereditary canine spinal muscular atrophy; IE, idiopathic epilepsy; MGA, megaloblastic anemia or Iverslund-Grasbeck syndrome; MH, canine malignant hyperthermia; MHC, major histocompatibility complex; PRA, progressive retinal atrophy; pred, progressive rod-cone degeneration; QTL, quantitative trait loci; RCND, hereditary multifocal renal cystadenocarcinoma and nodular dermatofibrosis; RFLP, restriction fragment length polymorphism; SCID, X-linked severe combined immunodeficiency; SINE, short interspersed nucleotide element; SNP, single-nucleotide polymorphism; S<sub>l</sub>, switch region of the immunoglobulin heavy-chain gene; WGS, whole genome scan; XLPRA1, X-linked progressive retinal atrophy 1.

<sup>b</sup> After the disease locus was mapped by linkage analysis in Doberman and Labrador colonies, linkage analysis was also performed in dachshunds of the general population, and linkage to the same locus was observed (Hungs et al. 2001; Lin et al. 1999).

<sup>c</sup> Within each category, studies are listed by year of publication.

<sup>d</sup> Within each category, studies are listed by year of publication.

<sup>e</sup> A disease known as XLPRA2 was caused, in a mongrel, by a different RPGR mutation, a 2 bp deletion leading to a truncated protein. It was identified by analyzing this candidate gene (Zhang et al. 2002).

<sup>f</sup> Allelism has been established for the diseases in these breeds (Lowe et al. 2003).

<sup>g</sup> After the disease locus was mapped by linkage analysis in Doberman and Labrador colonies, linkage analysis was also performed in dachshunds of the general population, and linkage to the same locus was observed (Hungs et al. 2001; Lin et al. 1999).

<sup>h</sup> A research colony was been developed, but mapping studies have not been published.

<sup>i</sup> Only simulation studies have been performed and they indicate that the disease locus can be mapped (average LOD score 3.23, with a maximum of 6.56).
In 14.6% of the results, the exclusion of a previously known mutation or any other mutation in that same exon was derived. For these results, the exclusion of a known mutation was obtained by restriction fragment length polymorphism (RFLP) analysis or by sequencing a small region covering the site of a known mutation; this does not tell if a mutation exists in that gene outside the small region studied.

Some results (7.2%) have been equivocal, meaning that the distribution of intragenic polymorphisms among affected and nonaffected dogs was consistent with the presence of a mutation in the gene. However, a very small sample of affected dogs was studied (one to three), so the observed distribution of polymorphic alleles could be due to chance, thus a larger sample needs to be studied or the gene must be screened for mutations.

Finally, only 3.4% of the results identified the mutation. This strikingly low success suggests that the FCG approach tends to be very inefficient.

Most of the 23 FCGs studied to date are involved in the phototransduction cascade or have mutations causing human retinal diseases. Surprisingly, fulfillment of these criteria does not often lead to the identification of the mutations causing canine retinal diseases. It could be argued that many other genes fulfilling the criteria remain to be studied and that mutations may be found in them. Unfortunately, however, including more FCGs would only perpetuate, and in fact exacerbate, the problem of having a low success rate with the FCG approach.

Half (50.4%) of the results correspond to studies on just five genes (Table S2)—PDE6B, RHO, PDC, RPE65, RDS—and mutations have been found in four of them (Table 1): a nonsense mutation in PDE6B in Irish setters with PRA (Clements et al. 1993; Dekomien et al. 2000; Ray et al. 1994; Suber et al. 1993), a missense mutation in RHO in English mastiffs and some Bull mastiffs, a missense mutation in PDC responsible for some cases of photoreceptor dysplasia (PD) in miniature schnauzers (Zhang et al. 1998), and a 4 bp deletion in RPE65 in Briards with retinal dystrophy (Aguirre et al. 1998; Veske et al. 1998, 1999). RDS is the other most intensively studied gene, and it codes for a structural protein on the photoreceptor membrane; it is involved in a digenic retinal disease in humans (Kajiwara et al. 1994), but no mutations have been found in dogs.

To date, half (50.1%) of all the FCG results correspond to only 13 breed diseases (19% of all breed diseases; Table S3), but interestingly, in only 3 of them has the mutation been found (retinal dystrophy in the Briard, PD in some miniature schnauzers, and early onset PRA in the Irish setter), again underscoring the difficulties encountered with the FCG approach. For the six most intensively studied breed diseases, mutations have not yet been found.

Pooling together the results on prcd, which has been shown to be allelic among six breeds, they represent one-fifth (19.6%) of all the results.

As shown in the section on LA, the lack of success with the FCG approach in some of these most intensively studied diseases is due to the fact that they seem to be caused by genes that have not yet been associated with retinal diseases in humans or any other species.

Finally, the level of redundancy in the FCG studies (Table S1) constitutes a strong argument in favor of publishing in a rapid manner negative results obtained with this approach in order to avoid further redundant experiments.

The analysis of intragenic polymorphisms or the comparison of the sequence of the gene in affected and nonaffected individuals generates very detailed information on the genes. For example, all FCG studies have reported gene or cDNA sequences, except TTR, and all have been mapped to particular canine chromosomes, except GCE and TTR.

Intragenic polymorphisms have been observed in the laboratory for all FCGs except four: CNGB3, CRX, GNB1, and GNGT2; polymorphisms have also been reported for two positional candidate genes—APOH (Gu et al. 1999) and TIMP1 (Zeiss et al. 1998)—giving a total of 154 intragenic polymorphisms. In addition, for some genes and cDNAs, or parts of them, more than one sequence has been deposited in the databases. By comparing them, 104 additional polymorphisms have been found, corresponding to 13 genes. These in silico polymorphisms still need to be validated.

Of the polymorphisms observed in the laboratory, 132 correspond to single nucleotide polymorphisms (SNPs) belonging to 19 FCGs, resulting in 1.19 SNP/kb. Here, SNPs are defined in the narrow sense, as a change of one base for another, thus excluding indels (Brookes 1999). Coding SNPs (cSNPs) constitute 37.1% of the 132 observed SNPs, with missense cSNPs representing 13.6%. On the other hand, in silico polymorphisms contain 78 additional SNPs in 13 genes (37.2% cSNPs, with 15.4% missense cSNPs). If these in silico SNPs turn out to be real, there would be 1.89 SNP/kb. Previously a SNP density of 2.56 SNP/kb was obtained by a pooling and sequencing approach from 12 canine genes from 10 different breeds (Brouillette et al. 2000), while the analysis of the canine genome sequence of a single male poodle resulted in a value of 0.67 SNP/kb (Kirkness et al. 2003). The average density of SNPs in the human genome has been estimated at 0.53 SNP/kb (Sachidanandam et al. 2001) and 0.83 SNP/kb (Zhao et al. 2003).

**Linkage Analysis**

Only four canine retinal diseases have been studied with this technique (Table 2). Three mutations have been found and two additional disease loci have been mapped.

One of the advantages of LA is that it determines the region containing the disease locus, thus limiting the study of candidate genes to those within that region. Once the locus has been mapped, and fine mapped, three outcomes are possible.

The first possible outcome is that the region where the canine disease locus has been mapped may contain a known gene causing a similar disease in another species, making it the obvious positional candidate and leading to the identification of the mutation, as happened with XLPR in the Siberian husky (Zeiss et al. 2000) and cd in Alaskan malamutes. The former
was mapped to an X chromosome region homologous to that containing\textit{RPGR} in humans (OMIM 300389), which is a known X-linked RP gene (Meindl et al. 1996; Roepman et al. 1996; Zhang et al. 2001). When this gene was studied in affected dogs, a deletion was found (Zhang et al. 2002). For\textit{cd}, a whole-genome scan (WGS) was undertaken to map the disease locus (Sidjanin et al. 2002). The result pointed to a region homologous to a human chromosome location harboring the\textit{CNGB3} gene, which is involved in achromatopsia (OMIM 262300), the equivalent human disease (Sundin et al. 2000). This gene was shown to be absent in affected malmutes, pointing to a microdeletion as the cause of the disease.

The second possible outcome of LA mapping is that it may point to a region containing a disease locus that has already been mapped, but not yet identified, in another species. Suitable candidate genes have to be studied within that region in both species until the mutations are found. Canine\textit{prcd} and human\textit{RP17} loci (OMIM 600852) may be an example of this situation. With a WGS, the\textit{prcd} locus was mapped to the centromeric portion of CFA9 (Acland et al. 1998), a region homologous to part of the short arm of human chromosome 17, where the\textit{RP17} locus had been mapped, but whose identity was unknown (Bardien et al. 1995, 1997; Bardien-Kruger et al. 1999; den Hollander et al. 1999). It was suggested that the\textit{prcd} and\textit{RP17} loci might be homologous, although the canine disease is recessive, while human\textit{RP17} is dominant (Acland et al. 1998). Since the gene involved in\textit{RP17} has recently been identified as\textit{CA4} (Rebello et al. 2004), it should now be possible to determine if mutations in this gene also account for canine\textit{prcd}.

The third possible outcome when mapping canine disease loci is that the locus mapped may be in a region where no genes implicated in a similar disease are known in humans or other species, so the search for suitable candidates has to proceed without reference to known disease genes or mapped disease loci. This seems to be the case with early retinal degeneration (erd) in the Norwegian elkhound. The locus has been mapped to a region of CFA27 with homology to human chromosome 12p13-12q13 and to mouse chromosome 6 (Acland et al. 1999). Two human retinal disease loci (\textit{COL2A1}, OMIM 120140;\textit{RDH5}, OMIM 136880) have been identified on HSA12 (Gonzalez-Fernandez et al. 1999; Yamamoto et al. 1999), but the characteristics of the diseases differ from erd, so they are unlikely candidates.

If this review of LA studies is expanded to include not only retinal diseases, but any canine disease, it emerges that only two LA studies have failed to map the disease locus: canine conotruncal heart defect and dilated cardiomyopathy (Table 2). In the first case, the failure may be attributed to the fact that only the telomeric end of CFA26 was studied instead of performing a WGS (Werner et al. 1999). In the second case, there seems to have been difficulties in assigning the correct phenotypes to some individuals, and a number of markers were uninformative, thus leaving large gaps of the genome inadequately covered (Dukes-McEwan and Jackson 2002); it is also possible that this disease may have a complex genetic basis. An interesting characteristic of canine LA studies is that almost all of them have relied on research colonies of affected dogs, and in some cases the structure of the pedigrees has been tailored to maximize their usefulness for LA. In total, LA has led to the identification of seven disease loci and an additional nine have been mapped. This is still a relatively small number, but the success rate is high. The number of success stories is expected to increase dramatically because of the availability of both a high-density map (Guyon et al. 2003) and a rough draft sequence of the canine genome (Kirkness et al. 2003).

### Conclusion

A large number of FCG studies have been done to identify the mutations causing canine retinal diseases, but only a minimal proportion of them (3.4%) have been successful. This problem with the FCG approach is likely to be most acute in diseases with heterogeneous genetic etiologies.

On the other hand, few studies have used LA, but the results obtained are encouraging. Assuming an adequate number of cases, the best strategy to find canine disease loci seems to consist of performing a WGS to map the disease locus. Then, if the region points to a previously known FCG, the information already amassed for these genes will lead rapidly to the mutation. If no obvious candidate disease genes are located in that region, LA nevertheless provides information on where to search for them and what remains to be done; for example, study positional candidates with appropriate expression patterns that have been identified in homologous chromosome regions in other species.

The immediate task for LA studies on canine disease loci is to extend its application to individuals in the general population (as opposed to breeding research colonies). The difficulty here lies in collecting complete multigenerational families, particularly for diseases of late onset. A partial solution to this problem may reside in using nonparametric forms of linkage analysis. Another possibility is to use association studies not requiring pedigree information (Hyun et al. 2003), and which would take advantage of the genetic homogeneity of many breed-specific diseases. Success in developing such methods will have a great impact on reducing the incidence of genetic diseases in affected breeds and will also open up a large pool of models to study human diseases and for testing therapies against them.

### Supplementary Data

Supplementary tables are available at \textit{Journal of Heredity} online (www.jhered.oxfordjournals.org).

### Acknowledgments

J. Aguirre-Hernández was supported by grants from the Consejo Nacional de Ciencia y Tecnologìa de México (CONACYT), and Pembroke College, Cambridge; this work was supported by grants from the Kennel Club and The Wellcome Trust for progressive retinal atrophy studies (Sargan).

### References


Received October 26, 2004
Accepted May 12, 2005

Corresponding Editor: Stephen J. O’Brien