A mutation in the canine *BHD* gene is associated with hereditary multifocal renal cystadenocarcinoma and nodular dermatofibrosis in the German Shepherd dog

Frode Lingaas[1,†], Kenine E. Comstock[2,†], Ewen F. Kirkness[3], Anita Sørensen[1],
Tone Aarskaug[1], Christophe Hitte[4], Michael L. Nickerson[5], Lars Moe[1],
Laura S. Schmidt[6], Rachael Thomas[7], Matthew Breen[8], Francis Galibert[4],
Berton Zbar[5] and Elaine A. Ostrander[2,*]

[1]Norwegian School of Veterinary Science, PO Box 8146 N-0033 Oslo, Norway, [2]Clinical and Human Biology
Divisions, Fred Hutchinson Cancer Research Center, PO Box 19024, D4-100, Seattle, WA 98109-1024, USA,
[3]The Institute for Genomic Research, 9712 Medical Center Drive, Rockville, MD 20850, USA, [4]UMR 6061 CNRS,
Génétique et Développement, Faculté de Médecine, 35043 Rennes Cédex, France, [5]Laboratory of Immunobiology,
Center for Cancer Research, National Cancer Institute, Frederick, MD 21702, USA, [6]Basic Research Program,
SAIC-Frederick, Inc., National Cancer Institute Frederick, Frederick, MD 21702, USA, [7]Oncology Research,
Animal Health Trust, Lanwades Park, Kentford, Newmarket, Suffolk CB8 7UU, UK and [8]Department of Molecular
Biomedical Sciences, College of Veterinary Medicine, North Carolina State University, Raleigh, NC 27606, USA

**Hereditary multifocal renal cystadenocarcinoma and nodular dermatofibrosis (RCND) is a naturally occurring
canine kidney cancer syndrome that was originally described in German Shepherd dogs. The disease
is characterized by bilateral, multifocal tumors in the kidneys, uterine leiomyomas and nodules in the skin
consisting of dense collagen fibers. We previously mapped RCND to canine chromosome 5 (CFA5) with
a highly significant LOD score of 16.7 ($\theta = 0.016$). We have since narrowed the RCND interval following
selection and RH mapping of canine genes from the $1.3\times$ canine genome sequence. These sequences also
allowed for the isolation of gene-associated BACs and the characterization of new microsatellite markers.
Ordering of newly defined markers and genes with regard to recombinants localizes RCND to a small
chromosomal region that overlaps the human Birt–Hogg–Dubé locus, suggesting the same gene may be
responsible for both the dog and the phenotypically similar human disease. We herein describe a disease-
associated mutation in exon 7 of canine *BHD* that leads to the mutation of a highly conserved amino acid of
the encoded protein. The absence of recombinants between the disease locus and the mutation in US and
Norwegian dogs separated by several generations is consistent with this mutation being the disease-causing
mutation. Strong evidence is provided that the RCND mutation may have a homozygous lethal effect
($P < 0.01$).**

INTRODUCTION

Canine hereditary multifocal renal cystadenocarcinoma and
nodular dermatofibrosis (RCND) is a naturally occurring

inherited cancer syndrome in German Shepherd dogs that
was first described in 1985 (1,2). The syndrome is character-
ized by bilateral, multifocal tumors in kidneys and numerous
firm nodules, consisting of dense collagen fibers in the skin and

---

subcutis. In addition, all females examined at an appropriate age demonstrate uterine leiomyomas and ∼50% of dogs experience metastasis (2). Analysis of canine families with RCND strongly indicates an autosomal dominant pattern of inheritance (1,2). Using a large resource family of Norwegian dogs, we previously mapped RCND to canine chromosome 5 (CFA5) with a highly significant LOD score of 16.7 ($\theta = 0.016$) (3). Using whole-chromosome paint probes, evolutionarily conserved chromosome segments between the canine and the human genomes were identified, suggesting that CFA5 contains several conserved segments corresponding to portions of HSA 11q, 17p, 1p and 16q (4–7). By low-density radiation hybrid (RH) mapping, C02608 originally appeared to lie in the region close to the boundary between HSA 17p and 1p (3).

RCND has some similarities to several human cancer syndromes. A number of provocative genes based upon their phenotype were investigated as possible candidates including the *TSC1*, *TSC2*, *TP53*, *PDK1*, *KRT9*, *WT1*, *FH* and *NF1* genes. Nevertheless, all of these genes have been eliminated based upon their location in the canine map (3,8–12).

We mapped the RCND locus to a region of CFA5 corresponding predominantly to human chromosome (HSA) 17p11.2. During the course of this work, a human renal cancer syndrome called Birt–Hogg–Dubé (BHD) that shows some similarity to RCND was mapped to 17p11.2 and the disease-associated gene, termed *BHD*, was subsequently cloned (13–15) [#2443; Nickerson *et al*. (19)]. In addition, the Nihon rat model for hereditary renal cell carcinoma was described and the gene responsible was mapped to a portion of rat chromosome 10 that also corresponds to HSA 17p11.2 (16,17). The function of the protein folliculin, encoded by the *BHD* gene, is unknown. Because of the similarity in phenotype and the corresponding locations in the human and canine genomes, we cloned and then searched for disease-associated mutations in the canine ortholog of the *BHD* gene.

RESULTS

The minimum RCND recombinant interval includes the canine *BHD* gene

A high-density RH map of CFA5, including 41 microsatellite markers, 10 BACs and 59 genes, and an integrated linkage map including 18 markers were constructed as a first step to narrowing the critical region (K. Comstock *et al.*, personal communication). Microsatellites were identified using the following procedure: the human genome sequence assembly was scanned using the University of California Santa Cruz Human Genome Project Working Draft (http://genome.ucsc.edu/) for genes located on HSA 1p and HSA 17p. These sequences were used to scan the canine 1.3× sequence for orthologous sequences. The resulting sequence reads represented partial sequence of genes from within the region of interest and were used as probes to screen an 8× canine BAC library for large genomic clones from the region of interest (18). BACs were pooled and used to construct mini-libraries that were then screened for microsatellites that were subsequently ordered on the RH map.

Twenty-six markers were found to be polymorphic in the founder dog and were used to analyze the pedigree for additional recombinants (Fig. 1). Further genotyping and haplotype analysis of the Norwegian RCND family identified a recombination in the proximal marker FH4160 that eliminated all genes centromeric to this marker as candidates. We also identified a recombination in the distal marker FH4442 by genotyping and subsequent haplotype analysis that eliminated all genes telomeric to this marker as candidates. The corresponding interval on the human map was inclusive of the genes *GLP2R* (NM_004246, 10.8 Mb) and *MAP2K3* (NM_002756, 22.8 Mb) and spans 12 Mb on HSA 17 p. This interval contains the human *BHD* gene (NM_144606, 18.5 Mb) and 85 other genes (RefSeq genes, UCSC Genome Browser on the Human April 2003 Freeze).

Sequence analysis of the canine *BHD* gene identified a disease-associated mutation

Portions of the orthologous canine *BHD* gene were obtained by screening the canine 1.3× sequence database with the sequence of the human gene. This strategy yielded 44 sequences that either encompassed or were within 1 kb of all *BHD* exons. The intron–exon structure of the canine *BHD* gene was deduced from the structure of human *BHD* (19). Intronic primers were designed to amplify all exons, except exon 1, using DNA from a healthy male Standard Poodle. For exon 1, an untranslated exon, cDNA was isolated from an unaffected Beagle kidney to obtain the sequence near the 5′ end of the mRNA. All the canine sequence obtained was compared with the human *BHD* sequence.

All exons were sequenced in three affected dogs and three unaffected dogs from the Norwegian family. In all affected dogs from the family and none of the unaffected dogs, an adenine to guanine mutation in exon 7 was detected (Fig. 2). This nucleotide change confers a histidine-to-arginine mutation in the expressed protein. No missense, nonsense or deletion mutations were found in any other exon segregating with affected dogs.

We next tested 12 RCND-affected German Shepherds from Norway and three from the USA, none of which were descendants of the founder of the Norwegian pedigree. Significantly, the mutation in exon 7 was detected in all 15 RCND-affected dogs. The exon 7 mutation was not detected in 264 unaffected dogs including 63 unrelated, unaffected German Shepherds, 28 Labrador Retrievers, 13 English Setters, 18 Golden Retrievers, 23 Norwegian Elkhounds, 10 Flat-coated Retrievers, 15 Pitbull Terriers, 20 Rottweilers, 16 Boxers, eight Newfoundlands, three Bernese Mountain Dogs and a single dog from each of 47 other breeds. Exon 7 was also examined in a single wolf, revealing a sequence identical to the unaffected dogs.

Expression patterns of canine *BHD* were investigated by northern blot experiments. We saw expression of an ∼3.8 kb transcript and no smaller transcripts in unaffected adult canine lung, muscle, skin, kidney, heart, colon, brain and uterus when northern blots were probed with *BHD* exon 5 (data not shown).

In addition, expression patterns of canine *BHD* in five affected and eight unaffected dogs from the Norwegian pedigree were compared by northern blot experiments. We
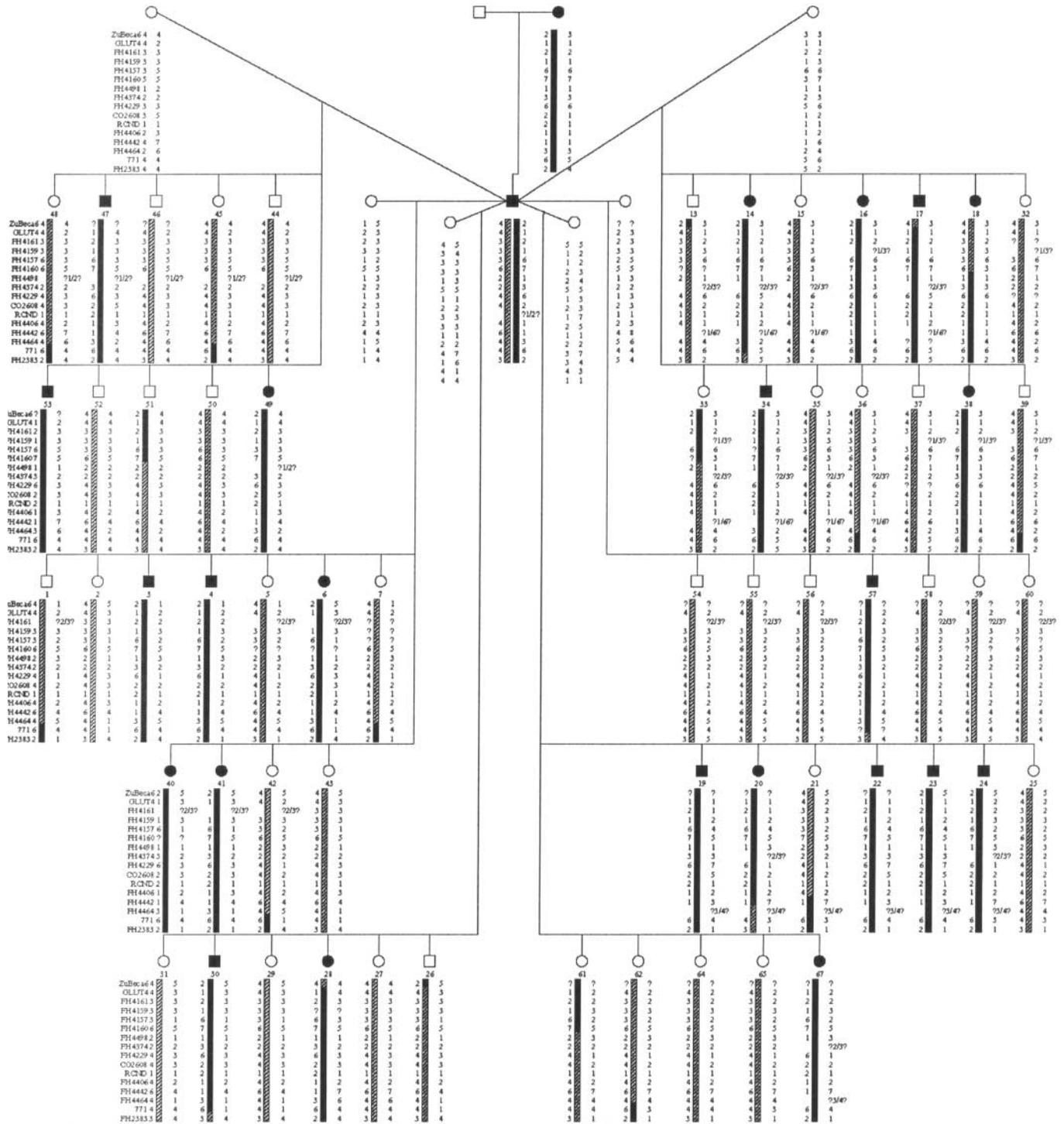
Figure 1. The canine pedigrees segregating RCND. Affected dogs are represented with black shading and unaffected dogs are unshaded. Marker names, ZuBeCa6, GLUT4, FH4161, FH4159, FH4157, FH4160, FH4498, FH4374, FHFH4229, CO2608, RCND, FH4406, FH4442, FH4464, CO5.771, FH2383, are indicated to the left of each row of genotypes. The genotypes of all markers are shown, but the vertical bar representing the haplotypes in the offspring is only shown for the affected proband's side. For the RCND locus, a '1' indicates the wild-type allele (unaffected) and a '2' indicates the mutant allele (affected).

saw equivalent levels of expression of an ~3.8 kb transcript and no smaller transcripts in kidneys of both affected and unaffected dogs when Northern blots were probed with *BHD* exons 6–9 (Fig. 3).

Comparison of RCND and BHD tumor histologic types

For a comparison between the human and canine diseases, slides stained by standard methods using hematoxyllin/eosin

A

GTATTTGAGGCAGAGCAATTTGGTTGCCCACAGCGTGCCCAGAGAATGAACACGGGCTTCACACCAT
TCCTGCACCAACGCAATGGGAACGCAGCTCGTTCACTGACCTCCTTGACAAGCGATGACAACTTGTG
GGCATGCCTTC(A/G)TACCTCCTTTGCTTG

B
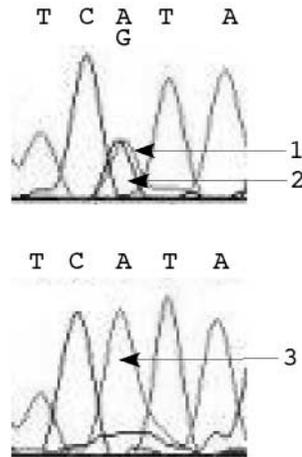
C

```
Dog:       VFEAEQFGCPQRAQRMNTGFTPFLHQRNGNAARSLTSLTSDDNLWACLHTSFA
RCND dog:  VFEAEQFGCPQRAQRMNTGFTPFLHQRNGNAARSLTSLTSDDNLWACLRTSFA
```

Figure 2. (**A**) The canine *BHD* nucleotide sequence of exon 7 is shown with the mutation indicated in parentheses. (**B**) Chromatographs showing the nucleotide sequences surrounding the mutation. Arrows 1 and 2 indicate the sequence of a heterozygous affected dog. Arrow 3 indicates the sequence of a homozygous unaffected dog. (**C**) The canine folliculin amino acid sequence showing the H255R mutation.

and Van Gieson that represented typical tumors from canine kidneys and skin were sent to a pathologist at NCI for evaluation. The histologic appearance of the dog renal tumors was found to be similar in appearance to the predominant type of renal tumor found in patients affected with the Birt–Hogg–Dubé syndrome (M. Merino, personal communication), which has elements of chromophobe renal carcinomas and renal oncocytoma. This oncocytic hybrid tumor, thought to arise from the distal tubule of the kidney, was observed in >50% of renal tumors from BHD patients (20).

Conservation of the *BHD* amino acid sequence

The folliculin protein is highly conserved across species. Full-length homologs of the human protein (NP_659434) are encoded by the genomes of mouse (NP_666130), rat (XP_220518), *Drosophila melanogaster* (NP_648090), *Caenorhabditis elegans* (NP_495422), and *Schizosaccharomyces pombe* (NP_595962). In addition, gene fragments that are homologous to exons 7 and 8 of the human folliculin gene have been obtained from another mammal (*Bos taurus*; BE481158), a bird (*Gallus gallus*, BG712454), two fish (*Danio rerio*, AL923165; *Oryzias latipes*, BJ487768), a sea squirt (*Molgula tectiformis*, AU281864) and another insect (*Anopheles gambia*, EAA04758). For each of these species, the predicted protein-coding sequence in the region

of the canine mutation was aligned (Fig. 4). Without exception, all genes and gene fragments encode a His residue at the location of the canine mutation.

A shared haplotype is present in affected Norwegian dogs and distantly related American dogs

Haplotypes were determined in a subset of the Norwegian dogs from the family and the two dogs from the USA with available pedigrees that were diagnosed with RCND. All of the RCND-affected dogs tested shared the haplotype and had the exon 7 mutation. These dogs were genotyped with markers surrounding the RCND locus (FH4229, FH4406, FH4442, FH4464). All the affected dogs share a four marker haplotype spanning a distance of ~3.0 Mb (Fig. 5).

The number of generations between the Norwegian proband and two of the American dogs for whom pedigrees were available, through a common affected ancestor, can be predicted. However, some uncertainty remains due to the high number of common ancestors, some missing pedigree information, and the lack of disease records in the population of German Shepherds. The shortest possible distance between the Norwegian proband and one of the American dogs is eight generations. However, following the pedigrees through the most likely common ancestors due to accumulation of a
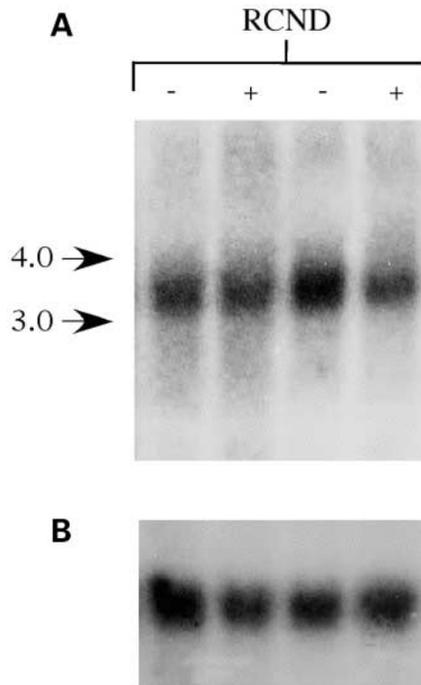
Figure 3. Analysis of *BHD* mRNA levels in RCND-affected and unaffected dogs. (**A**) Northern blot probed with *BHD* exons 6–9. (**B**) Probed with the *GAPDH* gene. The arrows indicate size in kb.

number of other affected Norwegian dogs in these lines, they are separated by ∼12–14 generations. The distance between the Norwegian proband and the other American dog is ∼22 generations. The two American dogs are separated by at least 10 generations (Fig. 5).

In addition, most of 85 other dogs diagnosed by us as having RCND by pathology (2) and where suitable pedigrees were available, could be traced back to the same pedigrees. Unfortunately, DNA samples are not available to test these dogs for the H255R mutation described here.

A homozygous A-to-G mutation in *BHD* exon 7 confers a loss of fetal viability

Nineteen offspring comprising four litters from matings between confirmed RCND-affected dogs (AG*AG) heterozygous for the mutation were analyzed. All the offspring were genotyped for the mutation in exon 7 by direct sequencing, as described earlier. The expected segregation of 1 : 2 : 1 was not observed; three offspring had the genotype AA, 16 offspring had the genotype AG, and none of the puppies had GG. This provides strong evidence that the RCND mutation may have a homozygous lethal effect ($P < 0.01$).

DISCUSSION

We have identified a canine gene, *BHD*, which substantial evidence suggests may play a critical role in the pathology of an inherited cancer syndrome in German Shepherd dogs. Identification of gene mutations in human families with BHD disease, together with the high level of identity observed

between the *BHD* homologs in divergent species, implies a critical functional role for the folliculin protein. In the German Shepherd Dog, we observed a disease-associated mutation (H255R) in the canine *BHD* gene that confers an amino acid change in a highly conserved region of the protein. It is often difficult to determine if a given missense change is actually disease-causing rather than simply disease-associated in the absence of detailed functional information about the protein. Indeed, while many disease-associated mutations have been reported for cancer susceptibility genes, such as *ATM*, *BRCA1*, and *BRCA2* (21,22), only a subset are confirmed as being disease-causing (23,24).

In the case of canine *BHD*, three lines of reasoning suggest that the H255R mutation is responsible for RCND. First, evolutionary analysis demonstrates a high level of amino acid sequence conservation between multiple species across exon 7, which contains the H255R mutation. This indicates that this region of the protein is likely to be of functional significance. Specifically, we observed no amino acid differences in H255 in any of 12 species ranging from *H. sapiens* to *S. pombe*. Significantly, by examining the offspring of a cross between two dogs who are heterozygous for the mutation, there is strong evidence that a homozygous mutation confers a loss of viability of the fetus. The future availability of functional assays, such as a binding assay to show interactions with other proteins, would allow us to definitively test the biological implications of the H255R mutation.

Secondly, while affected Norwegian and US dogs are separated by at least eight generations, we did not observe the H255R mutation in 264 unaffected dogs of 58 breeds originating from both Norway and the USA. Significantly, the H255R mutation was not observed in 63 unaffected German Shepherd dogs, including animals from both the USA and Norway.

Thirdly and perhaps most compellingly, we found the same H255R mutation in RCND-affected dogs in both the USA and Norway and showed that all affected dogs share a common haplotype of four markers spanning ∼3 Mb. The presence of a shared haplotype among all RCND-affected dogs from USA and Norway, known to be separated by several generations, makes a strong argument for a founder event. Founder effects are common in dog breeds, resulting when popular sires carrying undetected disease alleles are repeatedly bred into multiple lines within the breed (25). At the very least, the presence of a shared haplotype among affected individuals argues that if the H255R mutation is not responsible for the disease, another mutation in the shared haplotype that is in linkage disequilibrium with the H255R mutation is. Given that consideration, we cannot formally rule out the possibility that there are additional disease-associated mutations in a very closely linked gene or a *BHD* intron or regulatory region. However northern blot analyses using total RNA from affected and unaffected dogs revealed no apparent differences in expression levels, which argues that message levels and stability are unaffected in RCND dogs. This eliminates the second, but not the first possibility.

We focused our search for candidate genes utilizing both map position and predicted phenotype data. Our ability to map the gene associated with RCND to a small interval was of great importance when selecting *BHD* as the most likely gene. While
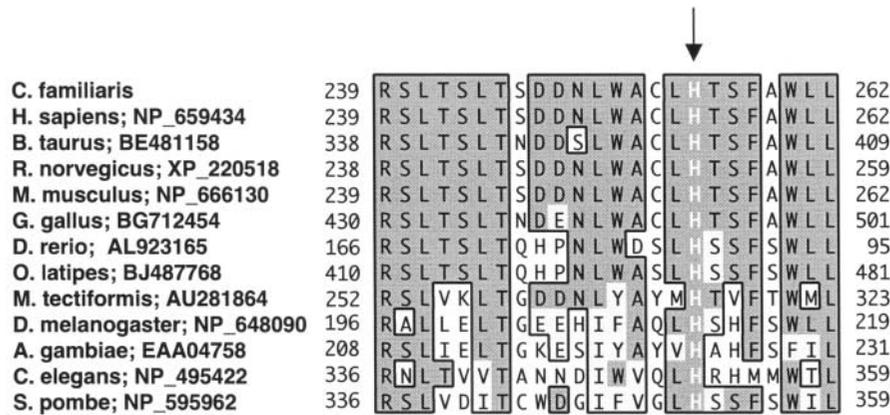
Figure 4. Alignment of folliculin homologues. The arrow indicates the location of the amino acid mutation in RCND-affected dogs. Identical amino acids are in dark gray, conservative differences are in light gray, and nonconservative differences are unshaded.

the region of minimal recombination contains 85 predicted genes based upon comparison with the human sequence, the fact that RCND shares common features with several related syndromes allowed us to limit our search for candidate genes significantly. For instance, human tuberous sclerosis complex (TSC), is similar to RCND except that the latter includes skin tumors and lacks vascular neoplasms (26,27). Mutations in the gene encoding fumarate hydratase (*FH*) cause a predisposion to uterine leiomyomas, benign tumors of the skin and papillary renal cell carcinoma, a phenotype that is quite similar to RCND (28,29). However, both *TSC1* and *FH* were definitively eliminated as candidate genes by their canine map position (3).

While the phenotypes of RCND and BHD syndrome are quite similar, they are not precisely identical. Human BHD syndrome shows similarity to RCND in that affected individuals experience firm nodules in the skin and subcutis and kidney tumors. Unlike RCND-affected dogs, however, BHD-affected humans frequently experience pneumothorax and do not experience uterine leiomyomas. In addition, there are distinct differences in the types of skin tumors that occur in the two hereditary syndromes. In BHD, the skin tumors are hamartomas of the hair follicle termed fibrofolliculomas, composed of elongated, delicate epithelial strands in a dense stroma. RCND-affected dogs do not present with hamaratomas, do not show the strands of epithelial cells, and the hair follicles are generally not involved. One striking feature of RCND is that renal tumors were observed in 100% of 51 autopsied RCND-affected dogs in the Norwegian pedigree. However for 19 of these dogs, the tumors were not detected until an age of 9–11 years (30,31). By comparison, renal tumors are reported in about 15% of BHD-affected humans (32), although differences between the occurrence of renal tumors in BHD-affected humans and RCND-affected dogs could be due to differences between diagnostic methods or age of diagnosis.

Nonetheless, the histologic appearance of the canine renal tumors is similar to the appearance of the predominant type of renal tumor found in patients affected with BHD (M. Merino, personal communication). The predominant type of tumor found in humans with BHD has elements of chromophobe renal carcinomas and renal oncocytoma. These similarities suggest that RCND may be an appropriate model for human BHD-induced renal tumors, while also providing insights into

species-specific tumorigenesis. One of the primary advantages of an RCND model is that the dogs' disease presentation is much earlier and disease progression is greatly accelerated compared to humans with BHD.

Finally, we found none of the mutations in RCND-affected dogs that have been observed in the BHD-affected human families described previously (14,19,33). Interestingly, the hypermutable C_8 tract in human *BHD* is interrupted in the center by an 'AT' dinucleotide pair in dogs, possibly explaining why the insertion/deletion mutations in this tract that comprised 44% of the *BHD* mutations observed in humans were not seen in any of the RCND-affected dogs. Likewise, in the mouse, the hypermutable C_8 tract is interrupted in the center by a 'TG'. As additional data becomes available from human families, it will be of interest to see if the differences observed in phenotype between BHD-affected humans and RCND-affected dogs can be correlated with specific genotypes.

Our work described here constitutes the first example of human and canine inherited cancer syndromes with similar phenotypes displaying disease-associated mutations in a both a human gene and its canine ortholog. The work has implications for analysis of human pedigrees, segregating disease linked to the BHD region, but for who obvious disease associated changes, such as large genomic deletions or insertions, have not been found. This particular example focused on kidney cancer, but we hypothesize that similarly structured studies could be used to map other cancer susceptibility genes as well. Many of the same cancers that occur in humans are observed at a very high frequency in certain dog breeds (34,35). Breed-associated cancers are observed for Boxers and Pointers (lymphoma) (36), Airedale Terriers and Golden Retrievers (soft tissue tumors) (37), Scottish Terriers (melanoma) (38), Scottish Deerhounds and Rotweillers (osteosarcoma) and Sky Terriers (breast cancer). By utilizing the advantages of canine families and homogenous breed structure, together with the now well developed canine genome map (9), we hypothesize that genes involved in both human and canine cancer biology can be mapped. Canine models can provide important additional information as well, as demonstrated by the likely lethal effect of the mutation in the present study. This sets the stage for future studies involving canine pedigrees aimed at mapping and cloning genes for complex human

Figure 5. Haplotype sharing analysis. Affected dogs are black, unaffected dogs are white. The 'FH' prefix on the marker names was omitted. Shared haplotypes are boxed and dotted lines indicate more distantly related dog.

diseases that have not been tractable through the study of large numbers of small human families.

MATERIALS AND METHODS

Canine pedigree development, phenotypic assessment and sample collection

A Norwegian canine colony segregating RCND was established by breeding a single affected male German Shepherd/Flat-coated Retriever to one unaffected female German Shepherd and five unaffected female English Setters (3). All females were unrelated to the male. The five female English setters were related to each other. In addition, three confirmed RCND-affected female dogs heterozygous for the mutation described here were bred to a single affected male German Shepherd/Flat-coated Retriever to generate 19 offspring in a total of four litters. Offspring were examined for the presence of multiple microscopic renal cysts by exploratory laparotomy, kidney biopsy or necropsy and subsequent histologic examination as described (31). Blood samples were also drawn and DNA prepared and sequenced, as described previously, from all dogs in the pedigree.

Samples from German Shepherd dogs affected with RCND residing in the USA were obtained by sending requests to veterinarians throughout the country. Once dogs were identified, their owners were contacted asking for participation in the study. Blood samples were drawn from the dogs by their own veterinarians and sent for DNA isolation. Blood samples from control dogs of all breeds were donated by their owners and collected by their own veterinarians either in the USA or Norway.

The Norwegian Animal Research Authority approved colony development, maintenance and sample collection in Norway. Canine blood and DNA samples in the USA were handled as specified by the Fred Hutchinson Cancer Research Center Institutional Animal Care and Use Committee.

Genomic DNA isolation

Genomic DNA was isolated from EDTA-anticoagulated whole blood using standard procedures (39). All DNA samples were resuspended in 10 mM Tris–HCl (pH 8.0), then quantitated by spectrophotometry.

Partial canine gene sequences were obtained

The 1.3× canine genome sequence used for this study was originally obtained from plasmid libraries of small- (2 kb) and medium-sized (10 kb) inserts, prepared and sequenced at Celera Genomics, as described previously for the human genome (40). The finished sequence data consists of 6.2 million reads (average read length, 576 bases), representing ∼1.3× coverage of the haploid canine genome (2.8 Gb) (41).

The human genome sequence assembly was scanned to identify genes located on HSA1p and 17 p using the University of California Santa Cruz Human Genome Project Working Draft (http://genome.ucsc.edu/). To obtain the corresponding partial canine sequence, the associated human peptide sequence was searched against the complete collection of dog reads

using tblastn ($W = 12$). Rarely (∼3% of searches) putative dog orthologs were detected using less stringent parameters. For each peptide, all homologous dog reads that were identified by the blast searches were assembled at high stringency (99% nucleotide identity) using TIGR Assembler (www.tigr.org/softlab/assembler/). Each assembly, or unassembled read, was then searched back against the Ensembl (release 1.1) collection of confirmed cDNAs and peptides (using blastn and blastx, respectively). If the assembly was most similar (at both the DNA and protein levels) to the gene that was used originally for searching, the assembly was considered a fragment of a putative ortholog.

Construction of DNA minilibraries and screening for microsatellites

The partial canine gene sequences were also used to probe a canine BAC library with a mean insert size of 155 kb and a 8.1-fold predicted coverage of the canine genome (18). For each gene, BAC filters were probed with PCR products of ≥400 bp, labeled by random primer incorporation and used at a concentration of 10^6 cpm/ml hybridization solution. Up to five filters were hybridized with 10–15 labeled probes simultaneously in a 7.5 cm diameter bottle, washed and then exposed to autoradiography film using standard techniques (42). The resulting positive clones were picked from the primary BAC library plates.

To construct mini-libraries, the isolated BAC clones were grown in LB with antibiotics using standard protocols and then pooled (42). BAC DNA was isolated using a Qiagen Large Construct kit and established procedures (43). The BAC DNA was partially digested with two 4-bp cutters, *Bfa*I and *Mse*I, and the resulting fragments were purified and cloned into the unique *Nde*I site in pGEM-5fZ(+/−). The libraries were transformed into DH5-alpha cells, then screened for common canine microsatellites using $(CA)_{15}$, $(GAAA)_{10}$, $(GTAT)_{10}$ and $(CCTT)_{10}$ oligonucleotides as described previously (44). The resulting clones were sequenced using the pUC/M13 forward and reverse primers using an ABI3700 automated sequencer. After BLAST analysis to eliminate clones containing LINE or SINE elements and identify any gene sequence, primers that bracket microsatellite repeats of 12 or greater were selected using the web-based program Primer3 (www-genome.wi.mit.edu/cgi-bin/primer/primer3_www.cgi). Primer sequences and product sizes are reported in Table 1.

Linkage and recombinant mapping

Microsatellites were typed in both members of the RCND pedigree as well as in unrelated affected and unaffected dogs. PCR was performed using 5′-Cy5 labelled primers as reported previously (12). The PCR products was analysed on an ALFexpress® sequencher (Amersham) with software for Fragment analysis (Allelinks®).

RCND is assumed to be inherited in an autosomal dominant manner and is fully penetrant. Using the PREPARE option of the Multimap program, each marker was checked for Mendelian inheritance. Two-point linkage analysis was carried out between RCND and each marker and between each pair of markers using the MultiMap software package and markers

were ordered by multipoint analyses. The most likely order and spacing of the markers within the linkage group were calculated using multipoint analysis and the GET-LIKELIHOODS function of Multimap. Maps were constructed with framework markers ordered at odds greater than 1000 : 1 and all remaining markers ordered at odds greater than 10 : 1 using MultiMap (45).

Haplotype sharing analysis

One primer of each primer pair was end-labeled using standard conditions (46). Amplification was carried out with 5 ng genomic DNA using previously published conditions (3). Primer sequences and product sizes are shown in Table 1. PCR products were separated on 4–6% polyacrylamide gels under denaturing conditions at 55°C, visualized by autoradiography and scored manually.

Cloning and sequencing of the canine *BHD* gene

To obtain the partial canine sequence corresponding to the human *BHD* gene, the associated human gene sequence was searched against the complete 1.3× collection of dog reads. Canine sequence was found within 1 kb of all the corresponding human exons. BLAST (www.ncbi.nlm.nih.gov:80/BLAST/) and Repeat Masker (www.repeatmasker.genome.washington.edu/cgi-bin/RepeatMasker) were used to identify any repeated elements in the sequence. Primers were designed using Primer3 (www-genome.wi.mit.edu/cgi-bin/primer/primer3_www.cgi) to flank all exons except exon 1 by at least 40 bp. PCR product sizes ranged from 304 to 1755 bp. Treatment of the PCR products with exonuclease I and shrimp alkaline phosphatase was done prior to sequencing. Sequencing was done using the BigDye kit (Applied Biosystems Inc., Foster City, CA, USA) and an ABI3700 or 3730 automated sequencer. The sequence of canine *BHD* exon 7 has been submitted to Genbank (Accession no. AY326427). Alignment and comparison of sequences from affected and unaffected dogs was done using the Phred/Phap/Consed software packages (47–49).

Mutation detection

The sequence of exon 7 was identified by sequencing of a PCR-product from cDNA using primers from exon 6 to exon 9. After initial identification of a mutation in the end of exon 7 in an affected dog, new primers were designed from the start of exon 7 (Ex7F) to intron 7 (In7R) for the purpose of genomic PCR and mutation detection. The sequencing reaction was performed with a nested primer (Ex7FS) 18 bp downstream of the forward primer. A PCR product of ∼1500 bp was generated using primers Ex7F (5′-GAGGCAGAGCAATTTGGTT-3′) and In7R (5′-TGTTGGATGATTTTGTGTTTGA-3′) and standard protocols for PCR under the following cycling conditions: 95°C for 3 min, followed by 35 cycles each of 95°C for 30 s, 60°C for 45 s and 72°C for 90 s. The sequencing reaction was performed with the ET-terminator kit for MegaBACE (Amersham) in accordance with recommendations from the manufacturer using the sequencing primer Ex7FS 5′-GAGA-ATGAACACGGCCTTC-3′ in a 20 µl reaction mixture containing 8 µl 'ET-mix', 2.5 µl PCR product, 1.5 µl sequencing

Table 1. Primer sequences and product sizes for canine microsatellites

Microsatellite	Repeat	Forward primer (5′–3′)	Reverse primer (5′–3′)	Annealing temperature (°C)	Size (bp)
FH2383	Tetra	GACCTGTCTTCTCCTGAGTCTACC	TACCAGAAATTACCTGCCCG	58	500
FH3278	Tetra	CTGCTCTTTGTAACCCATGC	AATGCCTACCAGGTGAAGG	58	324
FH3978	Tetra	ACCATAGAAGGAATGGTCAGTG	TCAGAATCTCTGGGGTCATTAG	58	331
FH4157	Tetra	AATCAAACATAGGCAGTGTGG	ACGAATCAGCCAGGAGAAGG	58	452
FH4160	Di	ACCACAAACACAAGCCTACAG	GTTCTCACGCTAGAGAAGGAAG	55	132
FH4166	Di	TATGTTTCTTCTTTCCCACCAG	CAGGACCTTTATTTCTCATTGG	58	151
FH4167	Di	GAAGATCATCGTGGGAGATG	TATAGGATGGAGTCTACGGGTG	57/32cyc	340
FH4168	Di	AGGACCCTTCTCTTATGGAGTC	ACACATGCAGAATGTATCGAAG	55	479
FH4169	Di	ATTCTGGACAAGTTACTGTGGG	ATTTCCCTGGCCTATAGTTTTC	58	272
FH4171	Di	AGGAGATGCTACAGGCAGG	CTTTGTGGAATGAAATGTAGGG	58	215
FH4229	Di	CTCGTGGAGCTTACCATCC	CTGAGGGAGCCCTCTACC	60	374
FH4241	Di	ATGGACCCAGGTTATCTCAGC	ATATACGGACTGGGACACTGG	58	203
FH4367	Di	GCTGGGTATCCACGACTGG	AGTGGGGAGACCCTGACC	58	357
FH4374	Di	AGTGGGAGAGTCTCAGTGTCC	GTGCTTTCAAGTGTCCTGACC	60	241
FH4379	Di	GGCTTCAAGCAGATAAAGGAC	GAGCATGGAGCTTGCTTG	60	314
FH4381	Di	GCATGAACTTTGTGGAACTGC	GCTCTCTGTTCCTGAGTGTCC	60	197
FH4404	Di	GGACCGTCAGATTACATGAGC	ATATACGGACTGGGACACTGG	62	246
FH4406	Di	CTCTCATCTATGAAGCATTGTCC	ATGGCACTTTTCTGCTTACG	62	168
FH4422	Di	TTCTAAAGGGTAGGAATTGAAGG	GGAATAGTCTATGTAATCTCAATGTGC	58	203
FH4442	Tetra	GGTTTAGTTTGGTTTTGTTTGG	CATTCTCAGCCAGGTTTGG	58	359
FH4464	Tetra	CACCTGCCTGGCTTAACA	CTGCCTGATGTTCAGTGTCTT	62	247
FH4487	Tetra	AACCACAAGTTTGCTTTTAGC	ATCTGATTTTCCCATCTCAGG	62	332
FH4496	Tetra	GTCTCTGCCTCTGTGTCTCTAT	CTCCTCAAAGCTTACCCTCA	60	197
FH4498	Tetra	GCATGGATGATAAAAGCAACC	TGTGAGTTCTCTATGGCAAAGC	58	284
FH4509	Di	CCAGTCCACTTGAGTTGCTT	CCGCCATCTTGAGGAGTT	td 61–51	176
FH4512	Di	TTAGGATATGGAACACCGTGAAC	AAGCAGGGTTTGTGTTGTCTG	td 61–51	185
FH4517	Tetra	GTTCAACACTACAATGATCAAAAGG	CTAGAGCCTTTCTCAGGTTTGC	60	475
FH4526	Tetra	AGTCAGGTGTGAGATCCAGTAGC	GTGTTTGCTTCATAATCAACAAGG	60	262
FH4532	Tetra	CGCAGGTACACCTTCCTAAACC	AGTTTCTGATTTCAGAGCTCAAGG	60	477
CPH18	Tetra	CAGAGATACGTCTTGACACTAGCAGA	AGCAGACAGTGGGCCATGTT	58	237
ZuBeCa6	Tetra	AGGAGTTACATGCCATAAGCC	CCAGTAAGGATTTTACCAGCC	58	100

primer (5 pmol/µl) and 8 µl H$_2$O. The sequencing reaction was performed by cycle sequencing with the following protocol: initially 95°C for 1 min then 29 cycles each of 95°C for 20 s, 59°C for 15 s and 60°C for 60 s. Sequencing was done using an automated sequencer (Molecular Dynamics MegaBACE 1000).

mRNA isolation, northern analysis and 5-RACE

Tissues were collected from all dogs shortly post-mortem. All tissues were then immediately immersed in liquid nitrogen. RNA was isolated from 50–100 mg canine tissues using TRIZOL reagent (Invitrogen Inc., Carlsbad, CA, USA) as recommended by the manufacturer. Total RNA was isolated from tissues of two unaffected adult dogs. In addition, total RNA was isolated from kidneys from three unaffected and three dogs affected with RCND.

A kit was used for the Northern blot procedure (NorthernMax, Ambion Inc., Austin TX, USA). Ten micrograms of total RNA for each sample were loaded onto 1.0% agarose gels containing formaldehyde and electrophoresed at 5 V/cm for ~2 h. Ribosomal RNA was visualized under UV light after ethidium bromide staining to check for possible degradation. RNA was transferred to a nylon membrane (Hybond N+, Amersham, Inc. Piscataway, NJ, USA) by capillary transfer, then UV crosslinked.

The blot was prehybridized in ULTRAhyb solution (Ambion Inc., Austin, TX USA) for 30 min at 68°C. A PCR product from exon 6–9 or exon 5 was radioactively labeled and used as probe. The probe was labeled in a reaction mixture containing 0.5 µl PCR product, 1.0 µl cold dNTP, 1.0 µl 20 pmol primer, 0.2 µl *Taq* polymerase, 2.6 µl 10× PCR buffer, 15.7 µl H$_2$O (total 22 µl). Four microlitres [α^{32}P]dCTP were added and the following protocol was used: initial denaturation at 95°C for 3 min followed by 30 cycles each of 95°C for 30 s, 58°C for 20 s, 72°C for 60 s and 72°C for 5 min. The radiolabeled DNA probe was added at a concentration of 10^6 cpm per ml to the ULTRAhyb solution. The blot was incubated overnight at either 42 or 68°C in a roller bottle hybridization oven.

The blot was washed in 2× SSC at room temperature two times for 5 min each, then washed in 0.1× SSC at 42°C (exon 5 probe) or 65°C (exon 6–9 probe) two times for 15 min each. The blot was then exposed to film at −80°C with an intensifying screen overnight for autoradiography.

For 5′ amplification of BHD cDNA, the SMART RACE cDNA amplification kit (Clontech, Palo Alto, CA, USA) was used. One microgram of total RNA isolated from a Beagle kidney was used. First-strand cDNA was made according to the manufacturer's instructions. The cDNA was then specifically amplified according to the manufacturer's instructions using a canine *BHD*-specific primer (5′-CGATGGCATTCATGGTGT-CCTGGAG-3′). The resulting PCR product was sequenced as described above.

Comparison of kidney tumor histologic types

Samples were taken from kidney tumors arising in RCND-affected dogs. Staining was done by standard techniques using HE (hematoxyllin–eosin) and VG (Van Gieson). Slides were then digitally scanned at high resolution and sent to NCI for comparison.

REFERENCES

1. Lium, B. and Moe, L. (1985) Hereditary multifocal renal cystadenocarcinomas and nodular dermatofibrosis in the German shepherd dog: macroscopic and histopathologic changes. *Vet. Pathol.*, **22**, 447–455.
2. Moe, L. and Lium, B. (1997) Hereditary multifocal renal cystadenocarcinomas and nodular dermatofibrosis in 51 German shepherd dogs. *J. Small Anim. Pract.*, **38**, 498–505.
3. Jonasdottir, T.J., Mellersh, C.S., Moe, L., Heggebo, R., Gamlem, H., Ostrander, E.A. and Lingaas, F. (2000) Genetic mapping of a naturally occurring hereditary renal cancer syndrome in dogs. *Proc. Natl Acad. Sci. USA*, **97**, 4132–4137.
4. Thomas, R., Breen, M., Langford, C.F. and Binns, M.M. (1999) Zoo-FISH analysis of dog chromosome 5: identification of conserved synteny with human and cat chromosomes. *Cytogenet. Cell Genet.*, **87**, 4–10.
5. Breen, M., Thomas, R., Binns, M.M., Carter, N.P. and Langford, C.F. (1999) Reciprocal chromosome painting reveals detailed regions of conserved synteny between the karyotypes of the domestic dog (*Canis familiaris*) and human. *Genomics*, **61**, 145–155.
6. Yang, F., O'Brien, P.C., Milne, B.S., Graphodatsky, A.S., Solanky, N., Trifonov, V., Rens, W., Sargan, D. and Ferguson-Smith, M.A. (1999) A complete comparative chromosome map for the dog, red fox, and human and its integration with canine genetic maps. *Genomics*, **62**, 189–202.
7. Sargan, D.R., Yang, F., Squire, M., Milne, B.S., O'Brien, P.C. and Ferguson-Smith, M.A. (2000) 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.
8. Breen, M., Jouquand, S., Renier, C., Mellersh, C.S., Hitte, C., Holmes, N.G., Cheron, A., Suter, N., Vignaux, F., Bristow, A.E. *et al.* (2001) Chromosome-specific single-locus FISH probes allow anchorage of an 1800-marker integrated radiation-hybrid/linkage map of the domestic dog genome to all chromosomes. *Genome Res.*, **11**, 1784–1795.
9. Guyon, R., Lorentzen, T.D., Hitte, C., Kim, L., Cadieu, E., Parker, H.G., Quignon, P., Lowe, J.K., Renier, C., Gelfenbeyn, B. *et al.* (2003) A 1 Mb resolution radiation hybrid map of the canine genome. *Proc. Natl Acad. Sci. USA*, **100**, 5296–5301.
10. Werner, P., Mellersh, C.S., Raducha, G.M., DeRose, S., Acland, G.M., Prociuk, U., Wiegand, N., Aguirre, G.D., Henthorn, P.S., Patterson, D.F. *et al.* (1999) Anchoring of canine linkage groups with chromosome specific markers. *Mammal. Genome*, **10**, 812–823.
11. Priat, C., Hitte, C., Vignaux, F., Renier, C., Jiang, Z., Jouquand, S., Cheron, A., Andre, C. and Galibert, F. (1998) A whole-genome radiation hybrid map of the dog genome. *Genomics*, **54**, 361–378.
12. Jonasdottir, T.J., Mellersh, C.S., Moe, L., Vignaux, F., Ostrander, E.A. and Lingaas, F. (2000) Chromosomal assignment of canine TSC2, PKD1 and CLN3 genes by radiation hybrid- and linkage analyses. *Anim. Genet*, **31**, 123–126.
13. Birt, A.R., Hogg, G.R. and Dube, W.J. (1977) Hereditary multiple fibrofolliculomas with trichodiscomas and acrochordons. *Arch. Dermatol.*, **113**, 1674–1677.
14. Khoo, S.K., Bradley, M., Wong, F.K., Hedblad, M.A., Nordenskjold, M. and Teh, B.T. (2001) Birt-Hogg-Dube syndrome: mapping of a novel hereditary neoplasia gene to chromosome 17p12–q11.2. *Oncogene*, **20**, 5239–5242.
15. Schmidt, L.S., Warren, M.B., Nickerson, M.L., Weirich, G., Matrosova, V., Toro, J.R., Turner, M.L., Duray, P., Merino, M., Hewitt, S. *et al.* (2001) 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.
16. Hino, O., Mitani, H., Nishizawa, M., Katsuyama, H., Kobayashi, E. and Hirayama, Y. (1993) A novel renal cell carcinoma susceptibility gene maps on chromosome 10 in the Eker rat. *Jpn. J. Cancer Res.*, **84**, 1106–1109.
17. Hino, O., Kobayashi, T., Tsuchiya, H., Kikuchi, Y., Kobayashi, E., Mitani, H. and Hirayama, Y. (1994) The predisposing gene of the Eker rat inherited cancer syndrome is tightly linked to the tuberous sclerosis (TSC2) gene. *Biochem. Biophys. Res. Commun.*, **203**, 1302–1308.
18. Li, R., Mignot, E., Faraco, J., Kadotani, H., Cantanese, J., Zhao, B., Lin, X., Hinton, L., Ostrander, E.A., Patterson, D.F. *et al.* (1999) Construction and characterization of an eightfold redundant dog genomic bacterial artificial chromosome library. *Genomics*, **58**, 9–17.
19. Nickerson, M., Warren, M., Toro, J., Matrosova, V., Glenn, G., Turner, M., Duray, P., Merino, M., Choyke, P., Pavlovich, C. *et al.* (2002) 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.
20. Pavlovich, C.P., McClellan, M.W., Eyler, R.A., Hewitt, S.M., Zbar, B., Linehan, W.M. and Merino, M.J. (2002) Renal tumors in the Birt-Hogg-Dube syndrome. *Am. J. Surg. Pathol.*, **26**, 1542–1552.
21. Deffenbaugh, A.M., Frank, T.S., Hoffman, M., Cannon-Albright, L. and Neuhausen, S.L. (2002) Characterization of common BRCA1 and BRCA2 variants. *Genet. Test.*, **6**, 119–121.
22. Boultwood, J. (2001) Ataxia telangiectasia gene mutations in leukaemia and lymphoma. *J Clin. Pathol.*, **54**, 512–516.
23. Hayes, F., Cayanan, C., Barilla, D. and Monteiro, A.N. (2000) Functional assay for BRCA1: mutagenesis of the COOH-terminal region reveals critical residues for transcription activation. *Cancer Res.*, **60**, 2411–2418.
24. Lavin, M.F. and Shiloh, Y. (1997) The genetic defect in ataxia-telangiectasia. *Annu. Rev. Immunol.*, **15**, 177–202.
25. Ostrander, E.A. and Kruglyak, L. (2000) Unleashing the canine genome. *Genome Res.*, **10**, 1271–1274.
26. Roach, E.S., Gomez, M.R. and Northrup, H. (1998) Tuberous sclerosis complex consensus conference: revised clinical diagnostic criteria. *J. Child Neurol.*, **13**, 624–628.
27. Franz, D.N. (1998) Diagnosis and management of tuberous sclerosis complex. *Semin. Pediatr. Neurol.*, **5**, 253–268.
28. Tomlinson, I.P.M., Alam, N.A., Rowen, A.J., Barclay, E., Jaeger, E.E., Kelsell, D., Leigh, I., Gorman, P., Lamlum, H., Rahman, S. *et al.* (2002) Germline mutations in FH preidspose to dominantely inherited uterine fibroids, skin leiomyomata and papillary renal cell cancer. *Nat. Genet.*, **30**, 406–410.
29. Toro, J.R., Nickerson, M.L., Wei, M.H., Warren, M.B., Glenn, G.M., Turner, M.L., Stewart, L., Duray, P., Tourre, O., Sharma, N. *et al.* (2003) Mutations in the fumarate hydratase gene cause hereditary leiomyomatosis and renal cell cancer in families in north america. *Am. J. Hum. Genet.*, **73**, 95–106.
30. Moe, L. and Lium, B. (1997) Hereditary multifocal renal cystadenocarcinomas and nodular dermatofibrosis in 51 German shepherd dogs. *J. Small Anim. Pract.*, **38**, 498–505.

31. Moe, L., Gamlem, H., Jonasdottir, T.J. and Lingaas, F. (2000) Renal microcystic tubular lesions in two 1-year-old-dogs—an early sign on hereditary renal cystadenocarcinomas? *J. Comp. Pathol.*, **123**, 218–221.

32. ZBar, B., Alvord, W.G., Glenn, G., Turner, M., Pavlovich, C.P., Schmidt, L., Walther, M., Choyke, P., Weirich, G., Hewitt, S.M. *et al.* (2002) Risk of renal and colonic neoplasms and spontaneous pneumothorax in the Birt–Hogg–Dube syndrome. *Can. Epidemrbl. Biol. Prev.*, **11**, 393–400.

33. Khoo, S., Giraud, S., Kahnoski, Chen, J., Motorna, O., Nickolov, R., Binet, O., Lambert, D., Friedel, J., Lévy, R. *et al.* (2002) Clinical and genetic studies of Birt–Hogg–Dubé syndrome. *J. Med. Genet.*, **39**, 906–912.

34. Ostrander, E.A., Galibert, F. and Patterson, D.F. (2000) Canine genetics comes of age. *Trends Genet.*, **16**, 117–123.

35. Arnesen, K., Gamlem, H., Glattre, E., Grøndalen, J., Moe, L. and Nordstoga, K. (2001) The Norwegian canine cancer register 1990–1998. Report from the project 'Cancer in dogs'. *Eur. J. Compan. Anim. Pract.*, **11**, 159–169.

36. Dorn, R.C. and Priester, W.A. (1987) Epidemiology. In Theilen, G.H. and Madewell, B.R. (eds), *Veterinary Cancer Medicine*. Lea and Febiger, Philadelphia, PA, pp. 27–52.

37. Priester, W.A. and Mantel, N. (1971) Occurrence of tumors in domestic animals. Data from 12 United States and Canadian colleges of veterinary medicine. *J. Natl Cancer Inst.*, **47**, 1333–1344.

38. Theilen, G. and Madewell, B. (1987) *Veterinary Cancer Medicine*. 2nd edn. Lea and Febiger, Philadelphia, PA.

39. Bell, G.I., Karam, J.H. and Rutter, W.J. (1981) Polymorphic DNA region adjacent to the 5′ end of the human insulin gene. *Proc. Natl Acad. Sci. USA*, **78**, 5759–5763.

40. Venter, J.C., Adams, M.D., Myers, E.W., Li, P.W., Mural, R.J., Sutton, G.G., Smith, H.O., Yandell, M., Evans, C.A., Holt, R.A. *et al.* (2001) The sequence of the human genome. *Science*, **291**, 1304–13051.

41. Vinogradov, A.E. (1998) Genome size and GC-percent in vertebrates as determined by flow cytometry: the triangular relationship. *Cytometry*, **31**, 100–109.

42. Ausebel, F.M., Brent, R., Kingston, R.E., Moore, D.D., Seidman, J.G., Smith, J.A. and Struhl, K. (1987) *Current Protocols in Molecular Biology*. Wiley Interscience, Boston, MA.

43. Kelley, J.M., Field, C.E., Craven, M.B., Bocskai, D., Kim, U.J., Rounsley, S.D. and Adams, M.D. (1999) High throughput direct end sequencing of BAC clones. *Nucl. Acids Res.*, **27**, 1539–1546.

44. Francisco, L.V., Langston, A.A., Mellersh, C.S., Neal, C.L. and Ostrander, E.A. (1996) A class of highly polymorphic tetranucleotide repeats for canine genetic mapping. *Mammal. Genome*, **7**, 359–362.

45. Matise, T.C., Perlin, M. and Chakravarti, A. (1994) Automated construction of genetic linkage maps using an expert system (MultiMap): a human genome linkage map. *Nat. Genet.*, **6**, 384–390.

46. Maniatis, T., Fritsch, E.F. and Sambrook, J. (1982) *Molecular Cloning: a Laboratory Manual*, 1st edn. Cold Spring Harbor Laboratory Press, New York.

47. Ewing, B., Hillier, L., Wendl, M.C. and Green, P. (1998) Base-calling of automated sequencer traces using phred. I. Accuracy assessment. *Genome Res.*, **8**, 175–185.

48. Ewing, B. and Green, P. (1998) Base-calling of automated sequencer traces using phred. II. Error probabilities. *Genome Res.*, **8**, 186–194.

49. Gordon, D., Abajian, C. and Green, P. (1998) Consed: a graphical tool for sequence finishing. *Genome Res.*, **8**, 195–202.