Biochemical Genetics, Vol. 34, Nos. 7/8, 1996

Gene-Specific Universal Mammalian Sequence-Tagged Sites: Application to the Canine Genome

Patrick J. Venta,[1-3,5] James A. Brouillette,[1,3] Vilma Yuzbasiyan-Gurkan,[2] and George J. Brewer[4]

Received 19 Feb. 1996—Final 4 June 1996

We are developing a genetic map of the dog based partly upon markers contained within known genes. In order to facilitate the development of these markers, we have used polymerase chain reaction (PCR) primers designed to conserved regions of genes that have been sequenced in at least two species. We have refined the method for designing primers to maximize the number that produce successful amplifications across as many mammalian species as possible. We report the development of primer sets for 11 loci in detail: CFTR, COL10A1, CSFIR, CYP1A1, DCN1, FES, GHR, GLB1, PKLR, PVALB, *and* RB1. *We also report an additional 75 primer sets in the appendices. The PCR products were sequenced to show that the primers amplify the expected canine genes. These primer sets thus define a class of gene-specific sequence-tagged sites (STSs). There are a number of uses for these STSs, including the rapid development of various linkage tools and the rapid testing of genomic and cDNA libraries for the presence of their corresponding genes. Six of the eleven gene targets reported in detail have been proposed to serve as "anchored reference loci" for the development of mammalian genetic maps [O'Brien, S. J.,* et al., Nat. Genet. *3:103, 1993]. The primer sets should cover a significant portion of the canine genome*

---

[1] Department of Microbiology, College of Veterinary Medicine, Michigan State University, East Lansing, Michigan 48824-1314.
[2] Department of Small Animal Clinical Sciences, College of Veterinary Medicine, Michigan State University, East Lansing, Michigan 48824-1314.
[3] Genetics Program, Michigan State University, East Lansing, Michigan 48824-1314.
[4] Department of Human Genetics, The University of Michigan Medical School, Ann Arbor, Michigan 48109-0618.
[5] To whom correspondence should be addressed at Department of Microbiology, College of Veterinary Medicine, Michigan State University, East Lansing, Michigan 48824-1314.

0006-2928/96/0800-0321$09.50/0 © 1996 Plenum Publishing Corporation

for the development of a linkage map. In order to determine how useful these primer sets would be for the other genome projects, we tested the 11 primer sets on the DNA from species representing five mammalian orders. Eighty-four percent of the gene–species combinations amplified successfully. We have named these primer sets "universal mammalian sequence-tagged sites" because they should be useful for many mammalian genome projects.

KEY WORDS: genome mapping; evolution; homology; polymerase chain reaction.

INTRODUCTION

Efforts have intensified in recent years to develop comprehensive genomic maps for many eukaryotic species using molecular techniques. Many of these efforts have focused on mammalian species, including human, mouse, rat, ox, sheep, pig, horse, cat, and dog (e.g., Buchanan *et al.,* 1994; Dietrich *et al.,* 1992; Ellegren *et al.,* 1992; O'Brien, 1986; Serikawa *et al.,* 1992; Weissbach *et al.,* 1992; Winterø *et al.,* 1991; Barendse *et al.,* 1994; present report). For the nonhuman species, these projects should lead to more successful breeding strategies, both for selecting desirable characteristics and for removing genes that lead to various genetic diseases. Comparisons made between these genome maps should also lead to new insights on the mechanisms of chromosomal evolution (e.g., see O'Brien *et al.,* 1993).

We are developing a comprehensive map of the canine genome, with our ultimate aim being to reduce the incidence of canine genetic diseases. In addition to developing random, highly polymorphic genetic markers (Type 2 markers), we are also developing markers for specific genes (Type 1 markers). An appropriate mix of these two types of markers should maximize our ability to map disease genes.

The traditional method for developing gene-specific markers, Southern blotting and cross-species hybridization, is very time-consuming, labor intensive, and limited in flexibility. This method has been the mainstay for developing gene-specific markers in most animal genome projects. There is a need to develop more efficient methods. This is particularly important for animal genome projects, where scientific resources are more limited. One method that has excellent potential is the cross-species polymerase chain reaction (PCR). This method has been used successfully for the study of a number of individual genes but has not been applied on a genomewide basis for the purpose of map development. To study a single gene, the cost associated with the failure of a few primers sets to amplify the correct target is negligible and new primer sets can be easily redesigned and synthesized. However, when primer sets are being designed for many genes, the cost for

failed primers can become substantial, in terms of both time and other resources, so we have refined the design method to minimize this problem.

We describe here, in detail, 11 primer sets that can amplify gene-specific targets of dogs and other mammalian species. Seventy-five additional primer sets are listed in the Appendixes. Because markers based on PCR primers are called sequence-tagged sites [STSs (Olsen *et al.,* 1989)], we call these primer sets universal mammalian STSs (UM-STSs) because they should be useful for many mammalian genome projects.

MATERIALS AND METHODS

DNA Isolation

DNA from dog, human, pigtail macaque, horse, pig, rat, and mouse were isolated from various tissues by standard phenol–chloroform extraction methods (Sambrook *et al.,* 1989). Goat DNA was kindly supplied by Dr. Karen Friderici, Michigan State University. DNA was purified by standard methods from a canine liver cDNA library (Clontech) and from a canine genomic DNA library (Clontech) after growing 1×10^6 phage in *Escherichia coli* strain LE392 (Murray *et al.,* 1977) in liquid culture (Sambrook *et al.,* 1989).

Design of PCR Primers

Primers were designed to genes where the intron–exon structure was known in at least one species and where the nucleotide sequence was known in at least two species (the "index species") that are not closely related. Tandemly duplicated genes known to have undergone gene conversion in any species were avoided. Primers were generally designed so that the amplified product contained an intron. We have followed the human gene nomenclature system (ISGN, 1987) for naming the canine genes. The 11 loci described in detail in this paper, and their protein products, are as follows: *CFTR,* cystic fibrosis transmembrane regulator; *COL10A1,* type X collagen, α_1 chain; *CSFIR,* colony stimulating factor 1 receptor; *CYP1A1,* cytochrome P-450 1, α_1; *DCN1,* decorin; *FES,* c-*fes* (feline sarcoma) protooncogene; *GHR,* growth hormone receptor; *GLB1,* β-galactosidase; *PKLR,* pyruvate kinase—liver, RBC form; *PVALB,* parvalbumin; and *RB1,* retinoblastoma protein. The Genbank Accession numbers or reference for the sequence of the two index species for each locus are as follows: *CFTR,* M55129 and M60493; *COL10A1,* X65120 and X65121; *CSFIR,* X14720 and K01643; *CYP1A1* (Uchida *et al.,* 1990), X04300; *DCN1,* L01125 and Z12298; *FES,*

X06292 and J02088; *GHR*, Z11802 and J04811; *GLB1*, S59584 and M57734; *PVALB*, X63578 and M15452; *PKLR*, S59798 and M17088; and *RB1*, L11910 and M26391.

Primers were designed to highly conserved nucleotide sequences contained within coding regions. Additional considerations taken into account were degeneracy of underlying codons (Li and Grauer, 1987), overall amino acid mutability of the primer region (Collins and Jukes, 1994), placement of the 3′ end of the primer with respect to amino acid mutability, and good standard design practices such as avoidance of primer–dimers. Conservation of amino acids within multigene families was also taken into account, when possible. Where unavoidable nucleotide mismatches occurred between the two index species, the primer sequence was designed to match exactly one of the two, which we then call the "primary" index species. GC-rich genes were generally avoided due to the amplification difficulties that can occur, even with exactly matching primers. Primers were 20 bp in length on average. Each primer in a pair was adjusted to be of approximately the same annealing temperature (Breslauer *et al.*, 1986). All *sets* of primer pairs were designed to have approximately the same annealing temperature as well, in anticipation of performing multiplex amplifications. It was not always possible to follow every rule for every gene, given the actual circumstances; however, the majority of the rules was generally applicable. Primers were synthesized by either the Michigan State University Macromolecular Structure Facility or the University of Michigan DNA Synthesis Facility.

PCR Amplifications

Correct design and syntheses of the primers were examined by amplifying the DNA from the primary index species. Standard buffer, nucleotide, and primer concentrations were 50 m*M* Tris–HCl (pH 8.3 at room temperature), 50 m*M* KCl, 1.5 m*M* $MgCl_2$, 200 μ*M* dNTPs, 0.1 μg of each primer, and 0.5–1.0 μg of target DNA in a 25-μl reaction. Reactions were routinely boiled for 3 min prior to the addition of 2.0 U of Taq DNA polymerase. Optimal cycling conditions for the amplification of canine genomic DNA were usually found by testing one of several sets of conditions in general use in the laboratory. Occasionally it was necessary to use "hot-start" conditions (Bassam and Caetano-Anolles, 1993) in order to get stronger, cleaner amplifications. The presence of an amplification product was determined by electrophoresis of a portion of the reaction on a 1% agarose TBE gel (TBE = 90 m*M* Tris, pH 8.3, 90 m*M* sodium borate, 2.5 m*M* EDTA), followed by staining with ethidium bromide.

DNA Sequence Analysis

The identity of each amplified canine gene was confirmed by "single-pass" direct sequencing of PCR products using Sequenase or Taq cycle sequencing kits (United States Biochemical Corp., Cleveland, OH). The PCR products were gel purified with Qiaex (Qiagen Corp., Chatsworth, CA) or by elution from polyacrylamide gel slices (Bergenhem *et al.*, 1992) prior to their use in the sequencing reactions. The canine sequences were visually aligned with the sequences of the other species used to design the PCR primers in order to verify the degree of sequence identity.

RESULTS

The primer sets for the various UM-STSs reported here are given in Table I and efficient amplification conditions for the canine genes are given in Table II. It is probable that these conditions could be optimized further (e.g., reduction in the time in each cycle). However, the conditions reported here were found to work effectively while minimizing the number of conditions that had to be examined. A representative gel showing amplification of the canine target DNA along with the human target DNA is shown in Fig. 1. The human target serves as a positive control for the amplification system because these primers were designed to match the human sequence exactly. The ability to quickly screen genomic and cDNA libraries for the presence of sequences is also demonstrated in Fig. 1. The genomic clones for *GHR, COL10A1,* and *DCN1* [a very faint signal, stronger on other gels (data not shown)] are present in this particular canine genomic library. The presence of a decorin cDNA clone (encoded by the *DCN1* locus) in the canine liver cDNA library is shown by the presence of the 122-bp band; cDNA clones for *GHR* and *COL10A1* are not present. The *DCN1* PCR product from the cDNA library was sequenced and its identity confirmed (see Fig. 2). The human and canine genomic bands have different sizes for *GHR* and *DCN1* because of the intron size differences. The size for the *COL10A1* PCR product is the same between the species because an intron was not spanned, for this is the UM-STS. Although the PCR product bands in Fig. 1 are unique, a few UM-STS-species combinations sometimes contained one to several nonspecific amplification products. This is a minor problem with unique sequence primers, because it is almost always possible to deduce the correct band based upon staining intensity and on the similarity in size compared to the band of the primary index species.

The amplified products for all of the canine loci were sequenced to confirm their identity and the results are shown in Fig. 2. The degree of identity between the canine and the index species sequences for each locus is

Table I. Primer Sets for 11 Universal Mammalian Sequence-Tagged Sites

Locus	Index species[a]	Primer 1 (P1)[b]	Primer 2 (P2)	P1 name[c]	P2 name	P1 AA[d]	P2 AA	Human chromosome
CFTR	Human	CTAAGCCATGGCCACAAGCA	CATTGCTTCTATCCTGTGTTC	HCFTREX22D	HCFTREX23U	1346	1407	7q31–q32
	Mouse	..............T.....	...C.................					
COL10A1	Human	ATTCTCTCCAAAGCTTACCC	GCCACTAGGAATCCTGAGAA	HCOL10A1EX2D	HCOL10A1EX2U	505	693	6q21–q22
	Mouse	........T...........	....					
CSF1R	Human	TTCCAAAACACGGGGACCTA	CATGCCAGGGCGAGAAGGA	HCSF1REX3D	HCSF1REX4U	76	161	5q33–q35
	FeSV	....	.G.................					
CYP1A1	Dog	TTGGACCTCTTTGGAGCTGG	TGGTTGATCTGCCACTGGTT	DCYP1A1EX3D	DCYP1A1EX5U	319	417	7q31
	Human	....	....					
DCN1	Human	GTTGATGCAGCTAGCCTGAA	AAGTGAAGCTCCCTCAGATG	HDCN1EX6D	HDCN1EX7U	207	247	12q21–q23
	Rat	...........C........	.....G..............					
FES	Human	GGGGAACTTTGGCGAAGTGTT	TCCATGACGATGTAGATGGG	HFESEX14D	HFESEX15U	573	641	15q25–qter
	FeSV	............A........	....					
GHR	Human	CCAGTTCCAGTTCCAAAGAT	TGATTCTTCTGGTCAAGGCA	HGHREX9D	HGHREX10U	301	432	5p13.1–p12
	Rat	.....A..............	A...................					
GLB1	Human	GAATTCTATACTGGCTGGCT	CATTCCAATAGGCAAAATTGGT	HGLB1EX8D	HGLB1EX9U	268	319	3pter–p21
	Mouse	..G.................	....					
PKLR	Human	CGCCTCAAGGAGATGATCAA	ATGAGCCCGTCGTCAATGTA	HPKLREX4D	HPKLREX6U	72	193	1q21
	Rat	....	....					
PVALB	Human	ATGTGAAGAAGGTGTTTCACAT	TCTTTGTCTCCAGCAGCCAT	HPVALBEX3D	HPVALBEX4U	43	93	22q12–q13.1
	Rat	................C.....	..C.................					
RB1	Human	GTTCCAGAAAATAAATCAGATGGT	ACTCATTTCTGCCAGTTTCTG	HRB1EX25D	HRB1EX26U	844	905	13q14.2
	Mouse	...............C........	....					

[a]Primary index species listed first.
[b]Dots indicate identical nucleotides.
[c]First letter, primary index species; next letters, locus; EX + number, exon number; D, down; U, up.
[d]*Amino acid (AA) over which 5′ nucleotide of primer lies.*

Table II. Amplification Conditions for Canine UM-STSs

Locus	Temperatures (°C)	Times (min)	Size of PCR product (bp)	
			Human	Dog
CFTR	95, 57, 72	0.5, 1.5, 4	700	1000
COL10A1	94, 57, 72 (hs)[a]	1, 2, 3	384	384
CSF1R	94, 59, 72	1, 2, 3	730	730
CYP1A1	95, 57, 72	0.5, 1.5, 4	700	600
DCN1	94, 57, 72	1, 2, 3	1422	2000
FES	94, 57, 72	0.5, 1, 1.5	484	500
GHR	94, 57, 72	1, 2, 3	765	800
GLB1	94, 57, 72	1, 2, 3	238	240
PKLR	94, 59, 72	1, 2, 3	600	630
PVALB	94, 57, 72 (hs)	0.5, 1.5, 4	1400	1300
RB1	94, 59, 72	1, 2, 3	695	1300

[a]hs indicates "hot start" used.

within the range generally accepted (roughly 70 to 100%) as demonstrating homology between the genes of mammalian species (Li and Grauer, 1987). These results support the hypotheses that the canine PCR products are homologous to the respective index species' genes. The canine *COL10A1* sequence matched the human and mouse sequences to a similar extent (data not shown). The sequences for *PKLR* and *CYP1A1* exactly matched previously published canine coding sequences (Whitney *et al.*, 1994; Uchida *et al.*, 1990); the sequence for canine *FES* is given in Fig. 3. Although the majority of the canine sequence for *PVALB* is from an intron, we believe that the degree of sequence identity from this region is sufficient evidence to confirm that the PCR product is from the correct canine locus. As expected, the canine sequences tend to show greater identity with the human sequences than with the rodent sequences because of the faster evolutionary rate of the rodent genome (Gu and Li, 1993). A microsatellite repeat was found within the amplified product itself for *RB1.* Preliminary results show that the *RB1* repeat, $(GA)_{12(avg)}$, has moderate genetic variability within several canine breeds.

We hypothesized that each primer set should work for many mammals, given the evolutionary rate at which nucleotide substitutions occur (Li and Grauer, 1987) and the number of primer nucleotide mismatches that can be tolerated by PCR. We tested the "universal" utility of these primers on the DNAs from mammals representing several different orders. We used the same reaction conditions that were found to amplify the canine sequences. We have termed these reactions "Zoo PCRs." Figure 4 shows a representative experiment. The *FES* protooncogene was amplified from all of the

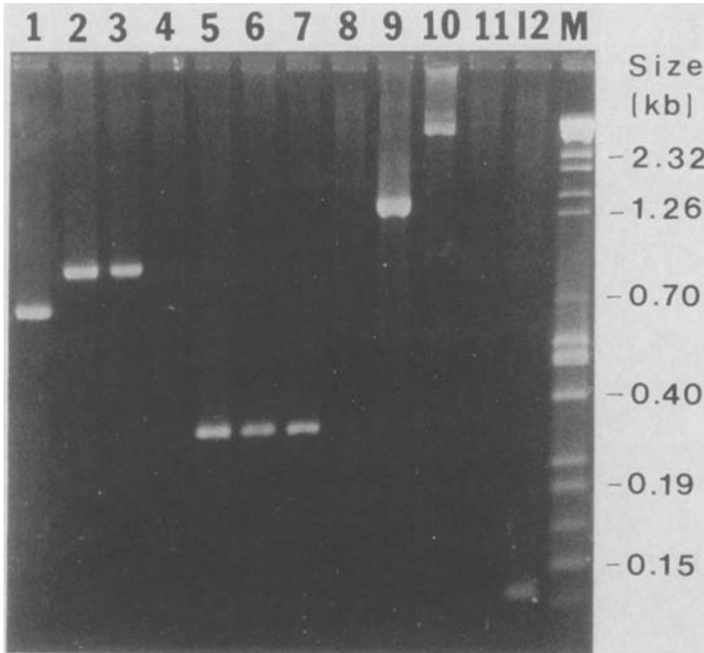

Fig. 1. Amplification of several canine gene segments using UM-STSs. The following lanes were amplified with the gene-specific primer sets (see Table I): lanes 1–4, *GHR;* lanes 5–8, *COL10A1;* and lanes 9–12, *DCN1.* Lane 13 contains a mixture of DNA size markers; λ bacteriophage DNA cut with the restriction endonuclease *Bst*EII and the plasmid pSK- (Stratagene) cut with *Msp*I. Lanes 1, 5, and 9 contain PCR products amplified from human genomic DNA. Lanes 2, 6, and 10 contain PCR products amplified from canine genomic DNA. Lanes 3, 7, and 11 contain PCR products amplified from DNA purified from a canine genomic library contained in a λ bacteriophage vector. Lanes 4, 8, and 12 contain PCR products amplified from a canine liver cDNA library.

DNAs examined. These PCR products were purified and sequenced directly without subcloning (see Methods and Materials). The sequences are tabulated in Fig. 3. The degree of sequence identity makes it highly likely that the canine PCR products are all homologous to the corresponding index species' genes. The pattern of nucleotide interchange is also what would be expected for homologous genes; members of the same mammalian order share more sequence similarity with one another than with those of other orders.

The data for the Zoo PCRs for the other UM-STS primer sets reported in this paper are given in Table III. More than 84% of the targets, excluding the index and canine species, amplified under the single condition used to

amplify the canine sequence. These species represent five mammalian orders: primates (human and macaque), carnivores (dog), artiodactyls (goat and pig), perissodactyls (horse), and rodents (mouse and rat). Limited experiments on other members of these orders (e.g., cat and ox) produced similar results (data not shown). Lack of amplification for *DCN1* for one of the artiodactyls (goat) would be predicted because there are four mismatches between the UM-STS primers and the sequence of the closely related bovine *DCN1* (Day *et al.,* 1987). We have found it difficult (although not impossible) to amplify DNA using primers that contain more than two mismatches with the target when using 20-mers (P.V., unpublished results). It is likely that the homologous gene from at least some of the nonamplifying species would appear using these primer sets if other PCR conditions were examined.

DISCUSSION

This study has shown the feasibility of generating a series of UM-STSs, useful for studies of many genomes, and addressed methodological considerations for their development. UM-STSs should serve as useful tools both for amplifying regions of interest from genomes and for isolation of clones from genomic and cDNA libraries and cross-species comparisons. The data reported in this paper indicate that approximately 85% of all carefully designed UM-STSs will be useful for any given mammalian species. We believe that this method is far more efficient, less costly, and considerably less labor intensive than traditional hybridization and Southern blotting-based methods. An additional important benefit is that the information for the necessary reagents (i.e., the primer sequences) is transmitted much more easily and quickly than the clones that are necessary for Southern blotting.

UM-STSs will also be useful for developing genetic markers within various genomes. We have found a microsatellite within 1 of the 11 loci reported here (*RB1*) and have found other microsatellite repeats associated with genomic clones isolated through the use of UM-STSs (unpublished results). Single-site variability should also be found directly in at least some of the amplified products by using one of a number of techniques developed for scanning for variability, such as the single-strand conformation polymorphism technique. For example, this method has been used to find two polymorphic sites in a study of the canine *ALAS2* gene in a PCR product of a size similar to those reported here (Boyer *et al.,* 1995). If the frequency of single site polymorphic variability for other mammals is as high as that estimated for humans (roughly 1 in 200–400 nucleotides), then a significant portion of UM-STSs will have these sites. We are currently screening for this

```
                                 CFTR
  A.A. 1346
         |                intron 22
Dog         -  -  -  -  -  -  -|  I  -  -  -  -
Mouse    -  -  -  -  -  -  -  -|  I  -  -  -  V  -  -  -  V  -  -  -  -  -  -
Human    E  P  S  A  H  L  D  P|  V  T  Y  Q  I  I  R  R  T  L  K  Q  A  F  A
Human   GAACCCAGTGCTCATTTGGATCC|AGTAACATACCAAATAATTAGAAGAACTCTAAAACAAGCATTTGCT
Mouse   ..G........C...C.A..C..|CA...........G.C...C..C..GT............C..C...
Dog       G....................|.A.............C
```

```
                                 COL10A1
   A.A. 569
         |
Dog      -  -  -  -  -  -  -  -  K  -  -  -  -  -  -  -  -  -  -  -  -  -  H
Mouse    -  I  Y  E  -  -  -  -  -  -  -  -  -  -  -  -  S  -  -  -  -  -  K
Human    P  F  D  K  I  L  Y  N  R  Q  Q  H  Y  D  P  R  T  G  I  F  T  C  Q
Human   CCATTTGATAAAATTTTGTATAACAGGCAACAGCATTATGACCCAAGGACTGGAATCTTTACTTGTCAG
Mouse   ..CA..T..G.G...C....C..T.....G........C........AT....T........C...A..
Dog     ...........G..C..........A.....................A...........C..C..C..C
```

```
                                 CSF1R
                A.A. 107
  intron 3          |
Dog     |-  -  V  -  -  -  Q  -  -  -  -  -  -  -  -  V  -  G  -  -  -  -  -
FeLV    |-  -  A  -  -  -  Q  -  -  -  -  -  -  T  -  L  -  G  -  -  -  -  -
Human   |D  P  A  R  P  W  N  V  L  A  Q  E  V  V  V  F  E  D  Q  D  A  L  L
Human   |ACCCTGCCCGGCCCTGGAACGTGCTAGCACAGGAGGTGGTCGTGTTCGAGGACCAGGACGCACTACTGC
FeLV    |.......T.....T.....G.....G..C.....A...ACG.....G..A.GT.....T..GT.G....
Dog     |......TT.....T.....G.....G..G........C......G......GG.....T..G..G....
```

```
                                 DCN1
   A.A. 231
         |                                 intron 6
Dog                           -  -  -  -  -  -  |-  -  -  -  -  -  -  -  -  -
Rat      -  -  -  -  -  -  -  -  N  S  -  -  S  |-  -  -  -  -  -  -  -  -  -
Human    V  D  A  A  S  L  K  G  L  N  N  L  A  |K  L  G  L  S  F  N  S  I  S
Human   GTTGATGCAGCTAGCCTGAAAGGACTGAATAATTTGGCTA|AGTTGGGATTGAGTTTCAACAGCATCTCT
Rat     ...........C............A..TC.......T...|..C....T.....C.....T......A.C
Dog                          ...................|........C.......T...........N
```

```
                                 GHR
   A.A. 333
          |
Dog       -  D  L  -  -  -  -  -  G  -  -  -  -  -  -  -  -  N  -  -  -  -  -
Rat       -  D  A  -  -  -  -  -  -  -  -  -  -  -  -  -  -  D  -  Q  -  -  -
Human     D  E  P  D  E  K  T  E  E  S  D  T  D  R  L  L  S  S  D  H  E  K  S
Human   TGATGAGCCAGATGAAAAGACTGAGGAATCAGACACAGACAGACTTCTAAGCAGTGACCATGAGAAATCA
Rat     ......TG.G.....G........A..G........C...............GA......G.........
Dog     ......C.T............C..A.G..........................AC...............
```

Fig. 2. Lineups of several canine gene sequences with homologous mammalian genes. The nucleotide and amino acid sequences are compared for each of several anchor loci between dog and two other species. The locations of PCR primers are underlined, although not all PCR primer sites are shown. Some of the lineups show intron sequence, whereas others simply identify the location of the introns. Genbank accession numbers for the canine sequences are as follows: *CFTR,* L77683 and L77689; *COL10A1,* L77672; *CSF1R,* L77670; *DCN1,* L77684; *GHR,* L77673; *GLB1,* L77671; *PVALB,* L77685 and L77686; and *RB1,* L77669.

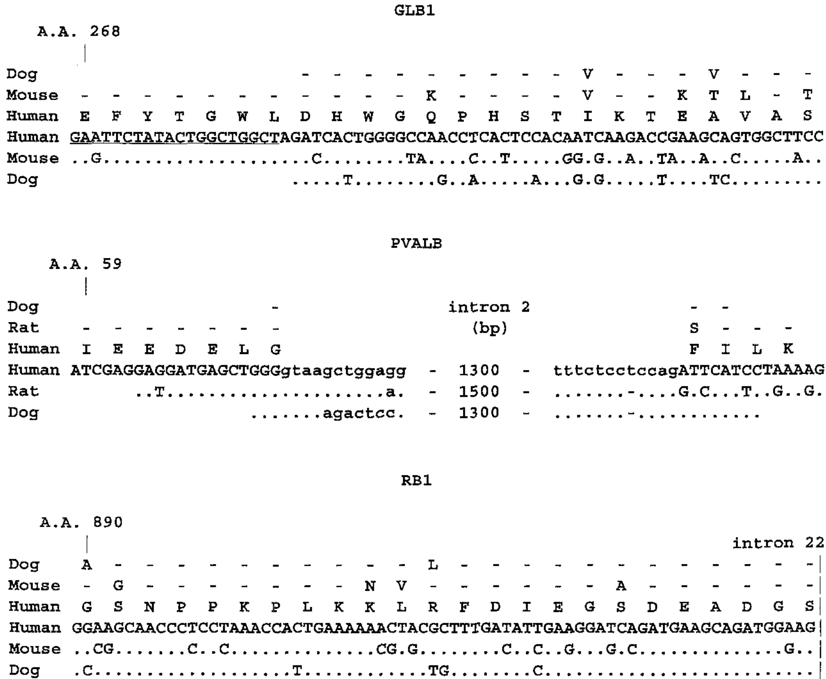

Fig. 2. (continued)

variability in the canine genome to estimate the frequency of such variation in the dog. It will be necessary to screen each species individually for genetic variability. However, the availability of previously designed UM-STS primer sets, such as those reported here, should make this work proceed more rapidly compared to the traditional method.

An example of the utility of cross-species comparisons is given by the case of Waardenburg syndrome. The clue to the location of one of the human Waardenburg syndrome genes—well-known for causing a syndromic hearing loss—was first gleaned from comparative mapping with the mouse (Asher and Friedman, 1990). The map locations in the mouse suggested possible locations of the human disease gene, one of which eventually was proven correct (e.g., Morell *et al.,* 1992). Because the identity of the gene in the mouse was not known at the time, this approach might more properly be called a "positional candidate" approach. UM-STSs will be useful for rapidly producing mammalian genetic maps so that the positional candidate approach can be applied to more species.

Very little is known about the location of genes within the canine genome. Indeed, except for genes located on the X chromosome (Meera-Khan, 1984; Deschenes *et al.,* 1994) and a few small unassigned linkage

```
MAC,CAT,FES                                                   I
HUM     A  D  N  T  L  V  A  V  K  S  C  R  E  T  L  P  P  D  L  K
HUM    GCCGACAACACCCTGGTGGCGGTGAAGTCTTGTAGAGAGACGCTCCCACCTGACCTCAAG
MAC    ............T.......A.................................A.....
CAT    ............T.......C.....A......C.C.....A............A.....
FES    ...........T........C.....A......C.C.....A.....A......A.....
DOG    ........T..T..............A.....CC.......C..................
COW    ..A.......................A......C.C.....A..................
GOA    ..A.......................A......C.......A..G..C............
PIG    ..A..T....................A.....CC....A.....................
HOR    ..T.......................A.....CC.............C..G.........
RAT    ..A..........C......T............C...................N......
MOU       .................T............C.....NNN..................


HUM     A  K  F  L  Q  E  A  R
HUM    GCCAAGTTTCTACAGGAAGCGAG GTGGGTGATAAACTAATGATCACCACGGGTCCCGCAT
MAC    ....................... .....................................
DOG    ....................... .......C....G...C...--..CA.A.CT..A...
CAT    ...........T........A.A ..A....AC...AG..C...--..CATAA.T.....C
FES    ...........T........A.A
COW    ...........C........... ..........G.AC.CCC...A.TGTA..C...CATA
GOA    ...........G........A.. ..........G.AC.CC...A.TGTA..C...T.C.C
PIG    ...........G........... ............AG..CC....TGTGATAAAAGA.CC
HOR    .................G..A.. ........C...A...CC....TGGTAT.CTAA.G..
RAT    ...........G..NNNN..... ......C.....A.GGGA.CAGT..A..T...TTGTG
MOU    ....................A.. ............A...AT
```

Fig. 3. Sequence of a portion of the *FES* protooncogene from several mammalian DNAs. Sequences are from exon 15 and intron 15. Notations for the sequence lineups are as follows: HUM, human; MAC, macaque; CAT, domestic cat; FES, feline sarcoma virus; DOG, dog; COW, ox; GOA, goat; HOR, horse; PIG, pig; RAT, rat; and MOU, mouse. The upper two lines for each block of text represent amino acid sequences and the lower lines represent nucleotide sequences. Dots indicate nucleotides in the various species that are identical to those of the human sequence. The human and cat sequences determined here exactly match the published sequences (Alcalay *et al.,* 1990; Roebroek *et al.,* 1987). The feline sarcoma virus sequence was not determined in this study but is included for comparative purposes. Only a single amino acid interchange was found among these sequences; isoleucine (I) for macaque, cat, and feline sarcoma virus and leucine (L) in all others. Sequence alignments for the intron were done visually and may not be optimal. Genbank accession numbers for these sequences are as follows: *MACFES,* L77678; *DOGFES,* L77674; *CATFES,* L77675; *COWFES,* L77677; *GOAFES,* L77681; *PIGFES,* L77679; *HORFES,* L77676; *RATFES,* L77680; and *MOUFES,* L77682.

groups (Meera-Khan, 1984), their locations (Holmes *et al.,* 1992; Ostrander *et al.,* 1993; Rothuizen *et al.,* 1994; Yuzbasiyan-Gurkan *et al.,* 1993) are based primarily on simple sequence repeats. The development of UM-STSs should help to identify rapidly the location of linkage groups on specific canine chromosomes. The identification of conserved syntenies will allow candidate linkages to be tested in the canine genome. The assignment of the proposed anchor loci (O'Brien *et al.,* 1993) as defined by UM-STSs to specific chromosomes can be accomplished by the somatic cell hybrid, flow sorted chromosome, and fluorescent *in situ* hybridization (FISH) methodologies. Other methods, such as assignment by use of linkage to previously mapped

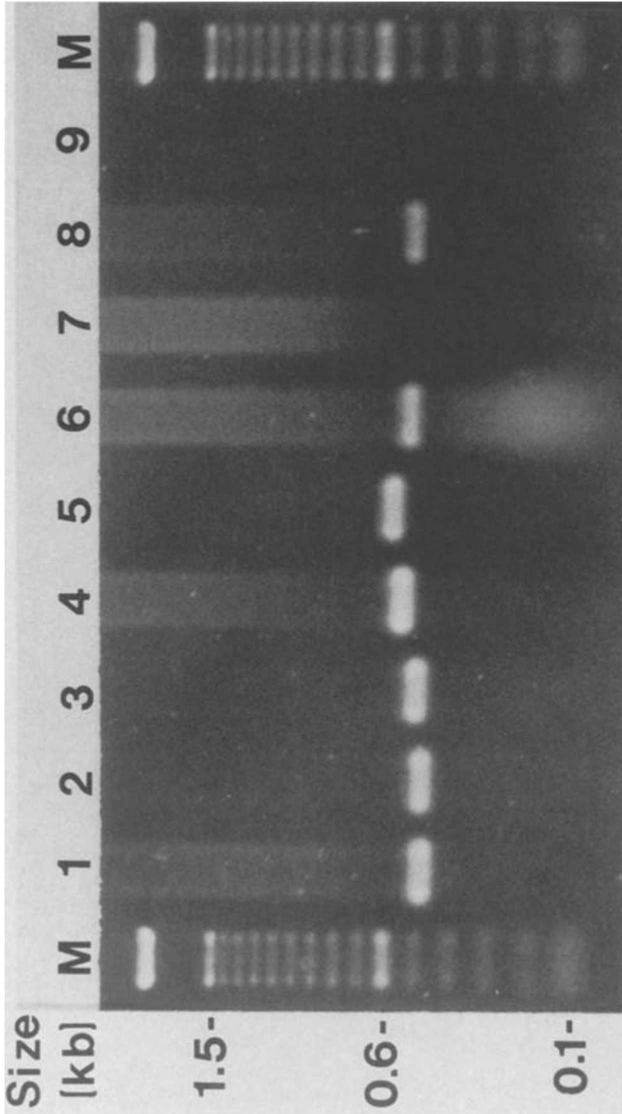

Fig. 4. Amplification of a portion of the *FES* protooncogene from several mammalian DNAs using UM-STS primers. Target DNAs for each lane are as follows: 1, human; 2, pigtailed macaque; 3, dog; 4, goat; 5, pig; 6, horse; 7, mouse; and 8, rat. The mouse DNA here was degraded; strong amplification was obtained with another lot (sequence shown in Fig. 3). The DNA marker lane (M) contains a 100-bp ladder.

Table III. Summary of Amplification Results for UM-STSs for Several Mammalian DNAs[a]

Locus	Human	Macaque	Dog	Goat	Pig	Horse	Mouse	Rat
CFTR	+[b]	+	+	−	+	+	+	+
COL10A1	+	+	+	−	+	+	+	+
CSF1R	+	+	+	+	−	−	+	+
CYP1A1	+	+	+	+	+	+	+	+
DCN1	+	+	+	−	+	+	+	+
FES	+	+	+	+	+	+	+	+
GHR	+	+	+	+	+	+	+	+
GLB1	+	+	+	+	+	+	+	+
PKLR	+	+	+	+	+	−	+	+
PVALB	+	+	+	+	+	+	−	−
RB1	+	+	+	−	+	+	+	+

[a]+, Amplification; −, no amplification.
[b]Boldface symbols indicate index species.

loci, are also possible. We have already assigned several genes by FISH to canine chromosomes using cosmids isolated with UM-STSs (Fujita *et al.*, 1996). Using the methods described here, we have developed a much greater number of UM-STSs that should cover, for linkage mapping purposes, a substantial portion of the canine and other mammalian genomes (see Appendixes 1 and 2).

APPENDIX 1

Table AI. Eighty-Six[a] Universal Mammalian Sequence-Tagged Sites—Human Chromosomal Locations and Names

Locus name	Gene product	Human chromosome	Primer 1 name	Primer 2 name
PND	Pronatriodilatin	1 p36	DPNDEX1D	DPNDEX2U
PKLR	Pyruvate kinase—RBC	1 q21	HPKLREX4D	HPKLREX6U
AT3	Antithrombin III	1 q23–q25	HAT3EX3D	HAT3EX4U
REN	Renin	1 q32	HRENEX8D	HRENEX9U
SFTP3	Pulmonary surfactant protein 3	2 p11.2	DSFTP3EX4D	DSFTP3EX5U
SPTBN1	β-Spectrin (non-RBC)	2 p21	HSPTBN1EX13D	HSPTBN1EX14U
APOB	Apolipoprotein B	2 p24–p23	HAPOBEX26D	HAPOBEX26U
IL1A	Interleukin 1α	2 q13	HIL1AEX2D	SIL1AEX3U
COL3A1	Collagen III α_1	2 q31–q32.3	HCOL3A1EX24D	HCOL3A1EX25U
ELN	Elastin	2 q31–qter	HELNEX32D	HELNEX33U
PAX3	Human paired domain 2	2 q34–q36	HHUP2EX2D	HHUP2EX3U
GCG	Glucagon	2 q36–q37	HGCGEX4D	HGCGEX5U
PIT1	Pituitary-specific transcription factor 1	3 p11	HPIT1EX4D	HPIT1EX5U

Table AI. (continued)

Locus name	Gene product	Human chromosome	Primer 1 name	Primer 2 name
GLB1	β-galactosidase	3 pter–p21	HGLB1EX8D	HGLB1EX9U
GPX1	Glutathione peroxidase 1	3 q11–q12	HGPX1EX1D	HGPX1EX2U
TF	Transferrin	3 q21	HTFEX7D	HTFEX8U
RHO1	Rhodopsin	3 q21–qter	HRHOEX3D	HRHOEX4U
GLUT2	Glucose transport-like 2	3 q26.1–q26.3	HGLUT2EX9D	HGLUT2EX10U
SST	Somatostatin	3 q28	HSSTEX1D	HSSTEX2U
HOX7	Homeobox 7	4 p16.1	HHOX7EX2D	HHOX7EX2U
PDEB	cGMP phosphodiesterase β	4 pter	HPDEBEX14D	HPDEBEX15U
ALB	Albumin	4 q11–q13	HALBEX4D	HALBEX5U
KIT	c-KIT protooncogene	4 q12–q13	HKITEX18D	HKITEX20U
FGG	Fibrinogen γ	4 q28	HFGGEX8D	HFGGEX9U
GHR	Growth hormone receptor	5 p13.1–p12	HGHREX9D	HGHREX10U
HEXB	β-Hexosaminidase	5 q13	HHEXBEX12D	HHEXBEX13U
IL4	Interleukin 4	5 q23–q31	HIL4EX1D	HIL4EX2U
ADRB2	Adrenergic receptor β_2	5 q31–q32	HADRB2EX1D	HADRB2EX1U
CSF1R	CSF-1 receptor	5 q33–q35	HCSF1REX3D	HCSF1EX4U
TNFA	Tumor necrosis factor α	6 p21.3	HTNFAEX1D	HTNFAEX4U
EDN1	Endothelin 1	6 p24–p23	HEDN1EX3D	HEDN1EX4U
COL9A1	Collagen IX αI	6 q12–q14	HCOL9A1EX3D	HCOL9A1EX4U
COL10A1	Collagen Type X αI	6 q21–q22	HCOL10A1EX2D	HCOL10A1EX2U
PLG	Plasminogen	6 q25–q27	HPLGEX18D	HPLGEX19U
EPO	Erythropoeitin	7 q21	HEPOEX2D	HEPOEX3U
CFTR	Cystic fibrosis trans. regulator	7 q31–q32	HCFTREX22D	HCFTREX23U
TCRB	T-cell receptor β	7 q35	DTCRBEX2D	DTCRBEX3U
SFTP2	Pulmonary suractant protein 2	8 p21	HSFTP2EX2D	HSFTP2EX4U
CA2	Carbonic anhydrase II	8 q22	CAUNIVEX3D	HCAIIEX4U
TG	Thyroglobulin	8 q24	HTGEX9D	HTGEX10U
ALDOB	Aldolase B	9 q21.3–q22.2	HALDOBEX7D	HALDOBEX8U
C5	Complement factor 5	9 q22–q34	HC5EX36D	HC5EX37U
ABL	ABL protooncogene	9 q34	HABLEX10D	HABLEX11U
RET	RET protooncogene	10 q11.2	HRETEX19D	HRETEX20U
TDT	Terminal transferase	10 q23–q24	HTDTEX9D	HTDTEX10U
OAT	Ornithine aminotransferase	10 q26	HOATEX7D	HOATEX8U
WT1	Wilms tumor 1	11 p13	HWT1EX8D	HWT1EX9U
LDHA	Lactate dehydrogenase A	11 p14–15.5	HLDHAEX3D	HLDHAEX4U
INS	Insulin	11 p15.5	DINSEX2D	DINSEX3U
CD20	CD20	11 q12–q13.1	HCD20EX6D	HCD20EX7U
ROM1	Rod outer segment protein-1	11 q13	HROM1EX1D	HROM1EX1U
APOC3	Apolipoprotein C3	11 q23–qter	DAPOC3EX2D	DAPOC3EX3U
VWF	von Willebrand's factor	12 p	HVWFEX46D	HVWFEX47U
LDHB	Lactate dehydrogenase B	12 p12.1–12.2	HLDHBEX3D	HLDHBEX4U
IL6	Interleukin 6	12 p12.2–p12	HIL6EX3D	DIL6EX4U
TPI	Triosphosphate isomerase	12 p13	HTPIEX2D	HTPIEX5U

Table AI. (continued)

Locus name	Gene product	Human chromosome	Primer 1 name	Primer 2 name
COL2A1	Collagen II α1	12 q14.3	HCOL2A1EX2D	HCOL2A1EX3U
DCN1	Decorin	12 q21–q23	HDCNEX6D	HDCNEX7U
IGF1	Insulin-like growth factor 1	12 q22	HIGF1EX3D	HIGF1EX4U
PLA2	Phospholipase A2	12 q23–qter	DPLA2EX2D	DPLA2EX3U
RB1	Retinoblastoma 1	13 q14.2	HRB1EX25D	HRB1EX26U
F7	Clotting factor VII	13 q34	HF7EX7D	HF7EX8U
CHY	Chymase (mast cell)	14 q11.2	DCHYEX4D	DCHYEX5U
CKBB	Creatine kinase brain	14 q32.3	DCKBEX6D	DCKBEX8U
TCRA	T-cell receptor α	14 q34	DTCRAEX3D	DTCRAEX4U
B2M	β-2-Microglobulin	15 q21–q22.2	HB2MEX2D	HB2MEX3U
CYP1A1	Cytochrome P-450 (AHH)	15 q22–q24	DCYP1A1EX3D	DCYP1A1EX5U
PKM	Pyruvate kinase—muscle	15 q22–qter	HPKMEX2D	HPKMEX3U
FES	FES protooncogene	15 q25–qter	HFESEX14D	HFESEX15U
HGBA	α-Hemoglobin	16 p13.3	HHGBAEX2D	HHGBAEX3U
GOT2	Glutamate oxaloacetate transaminase 2	16 q21–q22	HGOT2EX5D	HGOT2EX7U
CTRB	Chymotrypsinogen	16 q22.3–q23.2	DCTRBEX5D	DCTRBEX6U
APRT	Adenosine PR transferase	16 q24	HAPRTEX3D	HAPRTEX5U
TP53	Tumor protein 53	17 p13.1	HTP53EX5D	HTP53EX7U
NF1	Neurofibromatosis 1	17 q11.2	HNF1EX6D	HNF1EX7U
SCN4A	Skeletal muscle sodium channel	17 q23.1–q25.3	HSCN4AEX23D	HSCN4AEX24U
TS	Thymidylate synthetase	18 pter–q12	HTSEX5D	HTSEX6U
APOC2	Apolipoprotein C2	19 q13.2	DAPOC2EX3D	DAPOC2EX4U
CKMM	Creatine kinase muscle	19 q13.2–q13.3	DCKMEX2D	DCKMEX3U
PVALB	Parvalbumin	22 q12–q13.1	HPVALBEX3D	HPVALBEX4U
DYS	Dystrophin	X p21	DDYSEX7D	DDYSEX7U
MNK	Menkes protein	X q12–q13.3	HMNKEX4	HMNKEX4
HPRT	Hypoxanthine PR transferase	X q26	HHPRTEX7D	HHPRTEX8U
F9	Clotting factor IX	X q26.3–q27.1	DF9EX7D	DF9EX8U
F8	Clotting factor VIII	X q28	HF8EX24D	HF8EX25U
SRY	Sex determining region—Y	Y p11.3	HSRYEX1D	HSRYEX1U

[a]For convenience, the 11 loci described in detail are included in the Appendixes.

APPENDIX 2

Table AII. Eighty-Six Universal Mammalian Sequence-Tagged Sites—Sequences and Sizes

Locus name	Primer 1 sequence	Primer 2 sequence	PCR product size: Human genomic	PCR product size: Dog genomic[a]
PND	GCAGACCTGCTGGATTTCAAG	CAGTCCGCTCTGGGCTCCAAT	360	360
PKLR	CGCCTCAAGGAGATGATCAA	ATGAGCCCGTCGTCAATGTA	660	500
AT3	CTTCTTTGCCAAACTGAACTG	GGGCTGAACTTTGACTTCCA	658	660
REN	ACACTCCCCGACATCTCTTT	CGCCGATCAAACTCTGTGTA	137	137
SFTP3	GGAAGTTCCTGGAGCATGAG	CACAGGCCCAGGTGCTTACA	308	310
SPTBN1	TCTCAAGACTATGGCAAACA	CTGCCATCTCCCAGAAGAA	640	800
APOB	GTAAAAGCTCAGTATAAGAAAAAC	GTGCCCTCTAATTTGTACTG	460	460
IL1A	AGAAGTCAAGATGGCCAAAGT	TGATTCAGAGACAGATGGTC	1900	1900
COL3A1	GGACCAGGAAGTGATGGGAA	ACTTTCTCCTTGACTTCCCT	752	1400
ELN	GCTGCAGCCGCTAAAGCAG	AGGACACCTCCAAGGCCAG	600	1300
PAX3	GCCACAAGATCGTGGAGATG	GGTTCTCTCTTTTGTATTCCTC	1020	1170
GCG	TTCATTGCTTGGCTGGTGAA	GTGTTCATCTCATCAGAGAA	700	600
PIT1	TTCAGTCAAACAACAATCTG	GCTCCCACTTTTTCATTGTA	700	1000
GLB1	GAATTCTATACTGGCTGGCT	CATTCCAATAGGCAAAATTGGT	700	1000
GPX1	GACTACACCCAGATGAACGA	CAGGAACTTCTCAAAGTTCC	633	633
TF	GCTGACAGGGACCAGTATGA	AACAGCAGGTCCTTCCCATG	1700	585
RHO1	TACATGTTCGTGGTCCACTT	TGGTGGGTGAAGATGTAGAA	1479	553
GLUT2	TGGATGAGTTATGTGAGCAT	GACTTTCCTTTGGTTTCTGG	364	364
SST	GACTCCCGAGGCTTCCTCTTTG	ATACTGCAGGAGAGAGAAGAA	1200	1200
HOX7	AAGTTCCGCCAGAAGCAGTA	ATCTTCAGCTTCTCCAGCTC	400	400
PDEB	CTGAAGAGCTACTACACGGA	TGACACTTGTTCATCCACCA	300	300
ALB	GGCTGACTGCTGTGCAAAACA	AAGTAAGGATGTCTTCTGGC	730	730
KIT	CCTGTGAAGTGGATGGCACC	GCATCCCAGCAAGTCTTCAT	1000	1000
FGG	CAATATAAAGAAGGATTTGGACA	TGACACTTGTTCATCCACCA	1422	3000
GHR	CCAGTTCCAGTTCCAAAGAT	TGATTCTTCTGGTCAAGGCA	238	200
HEXB	TTCATTGGTGGAGAAGCTTG	ATCTTTGGAACTCCAGAGTC	1400	1000
IL4	CTATTAATGGGTCTCACCTCCCAACT	TCAACTCGGTGCACAGAGTCTTGG	469	450

Table AII. (continued)

Locus name	Primer 1 sequence	Primer 2 sequence	PCR product size: Human genomic	PCR product size: Dog genomic[a]
ADRB2	CCCATTCAGATGCACTGGTA	GCAGCCAGCAGAGGGTGAA	381	281
CSF1R	TTCCAAAACACGGGGACCTA	CATGCCAGGGCGAGAAGGA	1200	800
TNFA	CTCAGCCTCTTCTCCTTCCT	ATGGGCTCATACCAGGGCTT	1198	1200
EDN1	CCAAAAAGACAAGAAGTGCTG	TGGAACAGTCTTTTCCTTTCTT	1400	800
COL9A1	ATCAGGATTGGCCAAGATGA	GGAATCCTGAAGTCTACATT	484	500
COL10A1	ATTCTCTCCAAAGCTTACCC	GCCACTAGGAATCCTGAGAA	340	340
PLG	CAGCTCCCTGTGATTGAGAA	TAGACACCAGGCTTATTGGG	1100	1100
EPO	CTCCCTCTGGGCCTCCCAGT	CCATCCTCTTCCAGGCATAGAA	478	600
CFTR	CTAAGCCATGGCCACAAGCA	CATTGCTTCTATCCTGTGTT	765	800
TCRB	GACTGTGGCTTCACCTCGG	GATCTCATAGAGGATGGTGG	238	238
SFTP2	CAGAAACACACGGAGATGGT	GCCATCTTCATGATGTAGCA	500	600
CA2	CAGTTCCATTTTCAGTGGGG	GGCCAGTCCATCAGGTTGCT	350	1500
TG	TTCACCTCAGAGTGCTACTG	GCTTCTCTGTAGCTCATGATCTT	650	650
ALDOB	GTGACTGCTGGACATGCCTG	TTTGCAGCCTTGCCACCCC	463	450
C5	TGTGTACGATTCCGGATATTTGA	GCTCCTTCACAGACTTTCTG	300	300
ABL	TCAGACGAAGTGGAAAAGGA	AGAAGGCGCTCATCTTCATT	151	151
RET	CCCTTCCACATGGATTGAAA	CATCCAGTTAGCATATACAC	1800	1800
TDT	ACCTGGAAGGCCATCCGTGT	CGCCGGAGGTCTCTCTCAAA	1200	1250
OAT	CGTGCTCTTCAGGATCCAAA	GCCAGCCATCTACCAGTTCT	153	153
WT1	GAGAAACCATACCAGTGTGA	GTTTTACCTGTATGAGTCCT	820	850
LDHA	AACTCCAAGCTGGTCATTAT	GAATCCAGATTGCAACCACT	786	264
INS	GAGCGCGGCTTCTTCTACAC	GGTAGAGGGAGCAGATGCTGG	650	650
CD20	CTCTTTGCTGCCATTTCTGGAAT	TGGAAGAAGGCAAAGATCAGCAT	793	800
ROM1	CAGAGGACGGGCCACAGAA	GTTAAACACCACAGAGGCCTT	900	900
APOC 3	CAGGAACAGAGGTGCCATGC	TGCGCCACCTGGGACTCCTG	1900	1800
VWF	CCAGAGCGCATGGAGGCCTG	CTGCACTCCAGCTTGAATCC	700	850
LDHB	TTCCTCAGATCGTCAAGTACA	CTGCTGGGATGAATGCCAAG	650	650
IL6	GCACTGGCAGAAAACAACCT	ATCTGAAACTCCACAAGACC	172	172
TPI	TATATCGACTTCGCCCGGCA	ATGGCCCACACAGGCTCATA	300	500
COL2A1	CTCTGCGACGACATAATCTG	TCTCCAGGTTCTCCTTTCTG	600	1100

Table AII. (continued)

Locus name	Primer 1 sequence	Primer 2 sequence	PCR product size: Human genomic	PCR product size: Dog genomic[a]
DCN1	GTTGATGCAGCTAGCCTGAA	AAGTGAAGCTCCCTCAGATG	1200	1300
IGF1	GGCATCGTGGATGAGTGCTG	CTCCTTCTGTTCCCCTCCTG	950	900
PLA2	GACTACGGCTGCTACTGTG	TTACAGCTGGCCAGTTTCTT	420	300
RB1	GTTCCAGAAAATAAATCAGATGGT	ACTCATTTCTGCCAGTTTCTG	1600	1600
F7	AATGGAGCTCAGTTGTGTGG	CGATGTCGTGGTTGGTGGT	1600	700
CHY	GTCCCACCTGGGAGAATGTG	TGGGAGATTCGGGTGAAGAC	900	900
CKBB	TGGATCAACGAGGAAGACCA	TTCACACCATCCACCACCAT	200	950
TCRA	ACTGTCTGCCTGTTCACCGATTT	GTAACAACTTGGCATCACAGGAAT	900	900
B2M	TTCAGCAAGGACTGGTCTTT	CTGCTTACATGTCTCGATCT	1244	1100
CYP1A1	TTGGACCTCTTTGGAGCTGG	TGGTTGATCTGCCACTGGTT	800	860
PKM	GCCTTCATTCAGACCCAGCA	ATTCCAGACTTAATCATCTCCTT	1200	1200
FES	GGGGAACTTTGGCGAAGTGTT	TCCATGACGATGTAGATGGG	467	436
HGBA	CCCACCACCAAGACCTACTT	CGGTATTTGGAGGTCAGCAC	548	480
GOT2	TTTAAGTTCAGCCGAGATGT	CTTGGTAGGCCATGTCAAA	1400	1300
CTRB	AACGACATCACCCTGCTGTT	TGCAGGAGGAGACGCCACT	592	600
APRT	GACTCCCGAGGCTTCCTCTTTG	ATACTGCAGGAGAGAGAAGAA	695	1300
TP53	TACAAGCAGTCACAGCACAT	TCTTCCAGTGTGATGATGGT	1600	1600
NF1	ATTCACTCTCTGTGTACTTG	CAAAGCTTCTGTGACTGTTT	440	350
SCN4A	CTCAAGGTGGACATCCTGTACAA	AGCAGCGTCCGGATGCCCTT	1177	1100
TS	TGCCAGTTCTATGTGGTGAA	AGGTAAATATGTGCATCTCC	915	750
APOC2	GAATCACTCTACAGTTACTGG	AGCTGCTGTGCTTTTGCTGTA	201	201
CKMM	AAGAAGCTGCGGGACAAGGA	CAGCCCACGGTCATGATGAAA	1119	1100
PVALB	ATGTGAAGAAGGTGTTTCACAT	TCTTTGTCTCCAGCAGCCAT	420	830
DYS	GTTTCAGGCCAGACCTCTTT	TACCGACCTTCAGGATCAAG	500	500
MNK	GGCATGACTTGTAATTCCTG	CATCAAATCCCATGTCTTCTAT	669	750
HPRT	AGCTTGCTGGTGAAAAGGAC	TTATAGTCAAGGGCATATCC	766	650
F9	TGGGTGGTAACTGCAGCCCACT	CTACGCACACTCTTCACCCCA	650	650
F8	GATGCACAGATTACTGCTTC	GTAAGCAGAGATTTTACTCCCTG	779	800
SRY	AAGCGACCCATGAACGCATT	TTCGGGTATTTCTCTCTGTG	2500	2500

[a]All PCR products have been sequenced or, in a few cases, are derived from primers made to the published canine sequence.

ACKNOWLEDGMENTS

We thank Ya Shiou Yu and Murat Gurkan for their valuable technical assistance. We also thank Tracy Hammer, Neal Dittmer, Jessica Nadler, Elizabeth Tullett, Marc Crotteau, and Kristen Penner for their contributions to the development of a number of the primer sets. This work was supported by the American Kennel Club, the Orthopedic Foundation for Animals, and the Morris Animal Foundation. We also thank the Washington Regional Primate Facility for supplying the pigtail macaque tissues to Dr. Richard E. Tashian, with whom one of us (P.V.) originally isolated the DNA.

REFERENCES

Alcalay, M., Antolini, F., Van de Ven, W. J., Lanfrancone, L., Grignani, F., and Pelicci, P. G. (1990). Characterization of human and mouse c-fes cDNA clones and identification of the 5′ end of the gene. *Oncogene* **5**:267.

Asher, J. H. J., and Friedman, T. B. (1990). Mouse and hamster mutants as models for Waardenburg syndromes in humans. *J. Med. Genet.* **27**:618.

Barendse, W., Armitage, S. M., Kossarek, L. M., Shalom, A., Kirkpatrick, B. W., Ryan, A. M., Clayton, D., Li, L., Neibergs, H. L., Zhang, N., Grosse, W. M., Weiss, J., Creighton, P., McCarthy, F., Ron, M., Teale, A. J., Fries, R., McGraw, R. A., Moore, S. S., Georges, M., Soller, M., Womack, J. E., and Hetzel, D. J. S. (1994). A genetic linkage map of the bovine genome. *Nat. Genet.* **6**:227.

Bassam, B. J., and Caetano-Anolles, G. (1993). Automated "hot start" PCR using mineral oil and paraffin wax. *Biotechniques* **14**:30.

Bergenhem, N. C. H., Venta, P. J., Hopkins, P. J., and Tashian, R. E. (1992). Variation in coding exons of two electrophoretic alleles at the pigtail macaque carbonic anhydrase I locus as determined by direct, double-stranded sequencing of polymerase chain reaction (PCR) products. *Biochem. Genet.* **30**:279.

Boyer, G., Nonneman, D., Shibuya, H., Stoy, S. J., O'Brien, D., and Johnson, G. S. (1995). A PCR-RFLP marker for the erythroid aminolevulinate synthase gene (ALSA2) on canine chromosome X. *Anim. Genet.* **26**:206.

Breslauer, K., Frank, R., Blocker, H., and Markey, L. A. (1986). Predicting DNA duplex stability from the base sequence. *Proc. Natl. Acad. Sci. USA* **83**:3746.

Buchanan, F. C., Adams, L. J., Littlejohn, R. P., Maddox, J. F., and Crawford, A. M. (1994). Determination of evolutionary relationships among sheep breeds using microsatellites. *Genomics* **22**:397.

Collins, D. W., and Jukes, T. H. (1994). Rates of transition and transversion in coding sequences since the human-rodent divergence. *Genomics* **20**:386.

Day, A. A., McQuillan, C. I., Termine, J. D., and Young, M. R. (1987). Molecular cloning and sequence analysis of the cDNA for small proteoglycan II of bovine bone. *Biochem. J.* **248**:801.

Deschenes, S. M., Puck, J. M., Dutras, A. S., Somberg, R. L., Felsburg, P. J., and Henthorn, P. S. (1994). Comparative mapping of canine and human proximal Xq and genetic analysis of canine X-linked severe combined immunodeficiency. *Genomics* **23**:62.

Dietrich, W., Katz, H., Lincoln, S. E., Shin, H.-P., Friedman, J., Dracopoli, N. C., and Lander, E. S. (1992). A genetic map of the mouse suitable for typing intraspecific crosses. *Genetics* **131**:423.

Ellegren, H., Johansson, M., Sandberg, K., and Andersson, L. (1992). Cloning of highly polymorphic microsatellites in the horse. *Anim. Genet.* **23**:133.

Fujita, M., Loechel, R., McFarlin, K., Brewer, G. J., Venta, P., and Yuzbasiyan-Gurkan, V.

(1996). Assignment and localization of a set of genes to canine chromosomes using fluorescent in situ hybridization. *Mammal. Genome* **7**:268.

Gu, X., and Li, W.-H. (1993). Higher rates of amino acid substitution in rodents than in humans. *Mol. Phylogen. Evol.* **1**:211.

Holmes, N. G., Mellersh, C. S., Humphreys, S. J., Holliman, A., Curtis, R., and Sampson, J. (1992). Isolation and characterization of microsatellites from the canine genome. *Anim. Genet.* **24**:289.

ISGN (1987). Guidelines for human gene nomenclature: An international system for human gene nomenclature (ISGN, 1987). *Cytogenet. Cell Genet.* **46**:11.

Li, W.-H., and Grauer, D. (1987). *Fundamentals of Molecular Evolution,* Sinauer Associates, Sunderland, MA.

Meera Khan, P., Brahe, C., and Wijnen, L. M. M. (1984). Gene map of the dog: Six conserved and three disrupted syntenies. *Cytogenet. Cell Genet.* **37**:537.

Morell, R., Friedman, T. B., Moeljopawiro, S., Hartono, Soewito, and Asher, J. H., Jr. (1992). A frameshift mutation in the HuP2 paired domain of the probable human homolog of murine Pax-3 is responsible for Waardenburg syndrome type 1 in an Indonesian family. *Hum. Mol. Genet.* **1**:243.

Murray, N. E., Brammer, W. J., and Murray, K. (1977). Lambdoid phages that simplify the recovery of in vitro recombinants. *Mol. Gen. Genet.* **156**:53.

O'Brien, S. J. (1986). Molecular genetics in the domestic cat and its relatives. *Trends Genet.* **2**:137.

O'Brien, S. J., Womack, J. E., Lyons, L. A., Moore, K. J., Jenkins, N. A., and Copeland, N. G. (1993). Anchored reference loci for comparative genome mapping in mammals. *Nat. Genet.* **3**:103.

Olsen, M., Hood, L., Cantor, C., and Botstein, D. (1989). A common language for physical mapping of the human genome. *Science* **245**:1434.

Ostrander, E. A., Sprague, G. F., and Rine, J. (1993). Identification and characterization of dinucleotide repeat $(CA)_n$ markers for genetic mapping in dog. *Genomics* **16**:207.

Roebroek, A. J. M., Schalken, J. A., Onnekink, C., Bloemers, H. P. J., and Van de Ven, W. J. (1987). Structure of the feline c-fes/fps proto-oncogene: Genesis of a retroviral oncogene. *J. Virol.* **61**:2009.

Rothuizen, J., Wolfswinkel, J., Lenstra, J. A., and Frants, R. R. (1994). The incidence of mini- and micro-satellite repetitive DNA in the canine genome. *Theor. Appl. Genet.* **89**:403.

Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989). *Molecular Cloning. A Laboratory Manual* (2nd ed.), Cold Springs Harbor Laboratory Press, Cold Springs Harbor, NY.

Serikawa, T., Kuramoto, T., Hilbert, P., Mori, M., Yamada, J., Dubay, C. J., Lindpainter, K., Ganten, D., Guenet, J.-L., Lathrop, G. M., and Beckmann, J. S. (1992). Rat gene mapping using PCR-analyzed microsatellites. *Genetics* **131**:701.

Uchida, T., Komori, M., Kitada, K., and Kamataki, T. (1990). Isolation of cDNAs coding for three different forms of liver microsomal cytochrome P-450 form polychlorinated biphenyl-treated beagle dogs. *Mol. Pharmacol.* **38**:644.

Weissbach, J., Gyapay, G., Dib, C., Vignal, A., Morissette, J., Millasseau, P., Vaysseix, G., and Lathrop, M. (1992). A second-generation linkage map of the human genome. *Nature* **359**:794.

Whitney, K. M., Goodman, S. A., Bailey, E. M., and Lothrop, C. D., Jr. (1994). The molecular basis of canine pyruvate kinase deficiency. *Exp. Hematol.* **22**:866.

Winterø, A. K., Fredholm, M., and Thomsen, P. D. (1991). Variable $(dG\text{-}dT)_n\text{-}(dC\text{-}dA)_n$ sequences in the porcine genome. *Genomics* **12**:281.

Yuzbasiyan-Gurkan, V., Wagnitz, S., Blanton, S. H., and Brewer, G. J. (1993). Linkage studies of the esterase D and retinoblastoma genes to canine copper toxicosis: a model for Wilson disease. *Genomics* **15**:86.