A Second-Generation Genetic Linkage Map of the Domestic Dog, *Canis familiaris*

Mark W. Neff,* Karl W. Broman,† Cathryn S. Mellersh,‡ Kunal Ray,§ Gregory M. Acland,§ Gustavo D. Aguirre,§ Janet S. Ziegle, Elaine A. Ostrander‡ and Jasper Rine***

**Department of Molecular and Cell Biology, University of California, Berkeley, California 94720, †Marshfield Medical Research Foundation, Marshfield, Wisconsin 54449, ‡Clinical Research Division, Fred Hutchinson Cancer Research Center, Seattle, Washington 98109, §Center for Canine Genetics and Reproduction, James A. Baker Institute for Animal Health, Cornell University, Ithaca, New York 14853 and **PE Applied Biosystems, Foster City, California 94404*

Manuscript received August 3, 1998
Accepted for publication November 5, 1998

ABSTRACT

Purebred strains, pronounced phenotypic variation, and a high incidence of heritable disease make the domestic dog uniquely suited to complement genetic analyses in humans and mice. A comprehensive genetic linkage map would afford many opportunities in dogs, ranging from the positional cloning of disease genes to the dissection of quantitative differences in size, shape, and behavior. Here we report a canine linkage map with the number of mapped loci expanded to 276 and 10-cM coverage extended to 75–90% of the genome. Most of the 38 canine autosomes are likely represented in the collection of 39 autosomal linkage groups. Eight markers were sufficiently informative to detect linkage at distances of 10–13 cM, yet remained unlinked to any other marker. Taken together, the results suggested a genome size of about 27 M. As in other species, the genetic length varied between sexes, with the female autosomal distance being ~1.4-fold greater than that of male meioses. Fifteen markers anchored well-described genes on the map, thereby serving as landmarks for comparative mapping in dogs. We discuss the utility of the current map and outline steps necessary for future map improvement.

GENETIC maps took on a new significance with the recognition that sequence polymorphisms are common in all species and that comprehensive linkage maps could be rooted in simple molecular variation (Botstein *et al.* 1980). Genome maps are now routinely used to clone genes based on position, to dissect quantitative traits, and to pinpoint ancestral mutations using identity-by-descent methods. These approaches can be applied in any species for which a genetic map has been assembled and thus represent the most powerful means yet for understanding the basis of natural variation.

The technology used to assay molecular polymorphisms has progressed from restriction fragment length polymorphisms detected by DNA-blotting (Botstein *et al.* 1980) to single nucleotide polymorphisms detected by hybridization on microchips (Wang *et al.* 1998). For a wide range of applications, however, microsatellite markers have emerged as the best compromise of cost, ease, and power (Weber and May 1989). Microsatellite-based linkage maps spanning entire genomes are now available for a variety of species, particularly those of medical (Dib *et al.* 1996; Dietrich *et al.* 1996; Bihoreau *et al.* 1997) and commercial importance (Marklund *et al.* 1996; Kappes *et al.* 1997; Robic *et al.* 1997; de Gortari *et al.* 1998).

The domestic dog is uniquely suited among mammalian species to complement genetic analyses in humans and mice, but a comprehensive linkage map is not yet available. Over 400 breeds exist worldwide showing pronounced differences in size, shape, and behavior (Wilcox and Walkowicz 1993). This diversity can be exploited genetically to identify unconstrained developmental pathways that have likely played important roles in mammalian evolution. In addition, breeding primarily for morphological conformation has led to a high incidence of heritable defects in the dog. More than 400 genetic diseases are well documented and show simple patterns of inheritance (Nicholas *et al.* 1998). Any defect shared by dogs within a breed presumably stems from a single ancestral mutation inherited identical-by-descent. Popular sire effects and historical bottlenecks have further contributed to breed homogeneity. Thus genetic factors normally confounding linkage studies (*e.g.*, heterogeneity and modified expressivity) are less of a concern because much of the population has a simplified genetic architecture. Whether defining a direct model of human disease or identifying additional steps in a disease-related pathway, genetic analysis in dogs will help to elucidate the molecular mechanisms underlying human inborn errors. A comprehensive linkage map will serve as the principal tool for eliminating Mendelian canine disease, while also providing molecular access to morphological and behavioral breed differences.

An earlier linkage map was assembled on the basis of meiotic information from a canine reference resource, the Cornell Families (CF; Mellersh *et al.* 1997). Here

Corresponding author: Jasper Rine, Department of Molecular and Cell Biology, University of California, Berkeley, 401 Barker Hall #3204, Berkeley, CA 94720-3204. E-mail: jrine@uclink4.berkeley.edu

we report a second-generation linkage map based on the same mapping resource. The number of mapped markers was expanded from 150 to 276. The interval genetic distance (*i.e.*, the distance flanked by markers) increased from 900 to 1500 cM (almost 55% of the predicted genome size). An estimated 75–90% of the genome was within 10 cM of a molecular marker. A novel feature of the linkage groups presented here was the inclusion of standard errors for interlocus distances. The new version of the map was assembled with markers that were selected for compatibility with high-throughput genotyping, thereby increasing the feasibility of linkage studies. Fifteen polymorphic markers represented genes of known function and thus were also informative for comparative mapping. To facilitate future map improvement, we have provided primary genotype data in digital format at this journal's Internet site (http://www.genetics.org/supplemental/).

MATERIALS AND METHODS

Canine reference panel: The majority of animals typed for this study were from the CF shown in Figure 1. The CF consists of one large multi-generation pedigree (CF-I) composed of 16 sibships described previously (Mellersh *et al.* 1997) and a smaller three-generation pedigree (CF-II) composed of 2 sibships. CF-I and CF-II share common ancestors, but these animals were not studied. Together, the families consist of 218 animals of which 163 are F_2 progeny. The genetic background is an admixture of Miniature and Toy Poodles, Beagle, Norwegian Elkhound, Siberian Husky, and Irish Setter. In addition to the CF mapping resource, an intercross family having 20 F_2 animals was also typed and analyzed for map assembly. The two parental breeds were Border collie and Newfoundland (BC/Newf, Figure 1). The total number of animals analyzed was 247. All animals were housed under protocols approved by the Institutional Animal Care and Use Committees of the sponsoring institutions.

DNA for genotyping was prepared from blood or tissue using standard procedures (Bell *et al.* 1981). DNA concentration was estimated by spectrophotometry (A_{260}), and samples were diluted to 10 ng/μl. Each sample was initially assayed by PCR using primer pair CXX.2171 (5′ CTATGCCACTCTA GATTCTGGG 3′ and 5′ ACACATGCAGATAGGTCCTGG 3′) for 27 cycles. Concentrations of templates showing yields significantly different from that of a control template were adjusted further.

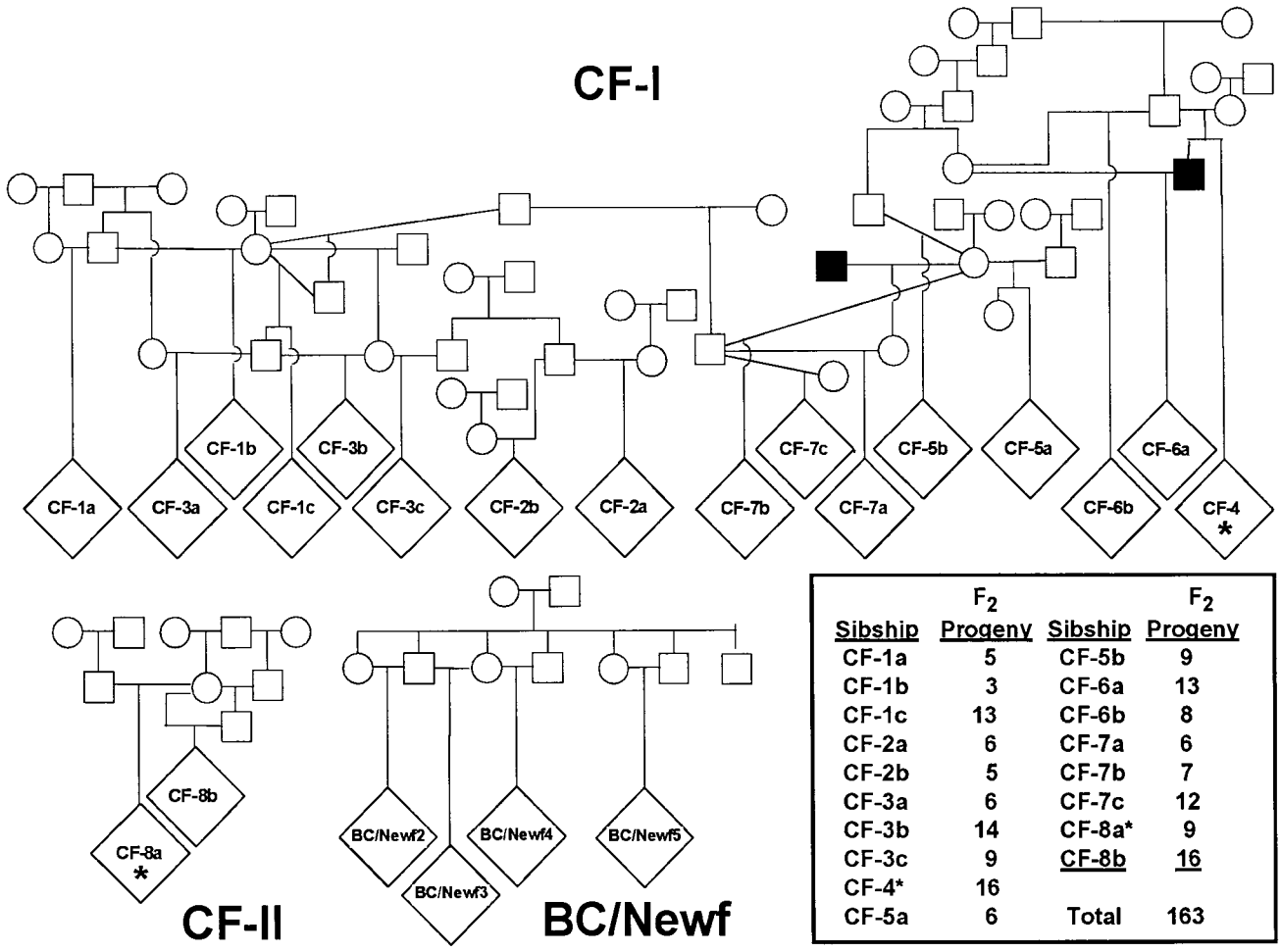
Genetic markers and genotyping: Markers used for genotyping are described in Table 1. To test a published multiplex strategy (Shuber *et al.* 1995), eight markers were modified by adding a universal sequence (5′ GCGGTCCCAAAAGGGT CAGT 3′) to the 5′ ends of both the forward and reverse primers (designated M-CPH in Table 1). Some dinucleotide marker names incorporate specific canine chromosome numbers as outlined in Ostrander *et al.* (1993). Primer pairs reported here for the first time are listed in Table 2.

Fifteen markers were typed at the Fred Hutchinson Cancer Research Center using autoradiographic methods described previously (Ostrander *et al.* 1995; Mellersh *et al.* 1997). Briefly, a ^{32}P-end-labeled primer was incorporated into PCR product during amplification, the product was size-fractionated by denaturing gel electrophoresis, and the resulting fragments were visualized by autoradiography. The remaining markers were typed using semiautomated fluorescent analysis. For each primer pair, the forward oligonucleotide was labeled at the 5′ end with one of three fluorescent dyes: 6-FAM, TET, or HEX (PE Applied Biosystems, Foster City, CA). PCRs were performed in microtiter format using 0.2-ml thin-walled tubes (USA-Scientific, Ocala, FL). DNA samples were added to each tube, the plate was loosely covered with parafilm, and the samples were left to desiccate at room temperature for 3 days. Samples were stable at room temperature for 1 month. Each 10 μl reaction contained 50 ng desiccated template, 0.01 mm tetramethylammonium chloride (Sigma, St. Louis), 1.5 mm $MgCl_2$, 1× buffer (Perkin-Elmer, Norwalk, CT), 200 μm each deoxynucleotide triphosphate (dNTP), and 0.25 units of TaqGold DNA Polymerase (Perkin-Elmer). Tetramethylammonium chloride at low concentrations can increase the specificity and yield of PCR (Hung *et al.* 1990). For amplification of single markers, primers were used at a final concentration of 0.5 μm. Some markers were amplified in multiplex PCR sets (M. Neff, unpublished results). Primers used in multiplex PCR were first purified by high-performance liquid chromatography. Final primer concentrations within multiplex sets ranged from 0.125–1.00 μm. Multiplix reaction conditions were the same as described above except for the concentration of TaqGold DNA Polymerase (1.0 unit/reaction). Four PCR programs were used, each having an initial denaturation step of 94° for 13 min (to activate the TaqGold DNA Polymerase) and a final extension step of 72° for 0.5–2 hr (to drive nontemplated single-base enzymatic additions to completion). All PCRs were amplified with PE 9600 instruments. Two programs were used for multiplex PCR. Program 1 included 35 cycles of 94° for 20 sec, 36 sec ramp time, 55° for 20 sec, 38 sec ramp time, and 72° for 1 min, 48 sec ramp time. Program 2 included 35 cycles of 94° for 20 sec, 70° for 15 sec, 59 sec ramp time, 57° for 20 sec, 30 sec ramp time, and 72° for 1.5 min. Program 3 was used for PCRs of dinucleotide markers assayed individually and included 35 cycles of 94° for 30 sec, 55° for 30 sec, and 72° for 1.5 min. Program 4 was used for tetranucleotide markers assayed individually and included 35 cycles of 94° for 30 sec, 58° for 30 sec, and 72° for 1.5 min. After PCR, fluorescent product was stored up to 2 months at −20°.

Products were electrophoretically separated and fluorescently detected with ABI 373 or ABI 377 instruments. Prior to gel loading, products for multiple markers were pooled, diluted, and combined with deionized formamide, loading dye, and an internal lane standard according to the manufacturer's instructions. After denaturation at 94° for 2–5 min, samples were stored on ice until loading. For ABI 373 runs, 8% LongRanger gels (AT Biochem, Malvern, PA) were used with 36-cm well-to-read plates. Electrophoresis conditions were 800 V, 40 mA, and 30 W for 6–8 hr. For ABI 377 runs, 4.5% LongRanger gels were used with 36-cm well-to-read plates. Electrophoresis conditions were 3000 V, 60 mA, and 200 W for 2 hr. Digitized fluorescent data were collected and stored with *GENESCAN* software. Gel data files were automatically sized and manually checked. Results were imported into *GENOTYPER* software where templates were constructed to semiautomate allele determination. Alleles were labeled as base-pair integers consistent among all families analyzed.

Data integrity: Three steps were taken to identify and correct errant typings prior to map construction. First, some markers from the first-generation map were retyped and the data compared to corresponding genotypes obtained previously (Mellersh *et al.* 1997). Second, a subset of individuals was chosen randomly for duplicate typing to allow for continuous quality assessment of newly acquired data. Third, genotype data were inspected for typings inconsistent with Mendelian inheritance. All discordant and non-Mendelian genotypes were either repeated or deleted from the primary data.

Map construction: Genetic linkage maps were constructed



Sibship	F_2 Progeny	Sibship	F_2 Progeny
CF-1a	5	CF-5b	9
CF-1b	3	CF-6a	13
CF-1c	13	CF-6b	8
CF-2a	6	CF-7a	6
CF-2b	5	CF-7b	7
CF-3a	6	CF-7c	12
CF-3b	14	CF-8a*	9
CF-3c	9	CF-8b	16
CF-4*	16		
CF-5a	6	Total	163

Figure 1.—Canine pedigrees and sibships. Sibships are represented as diamonds. The two symbols in black represent the same male animal, noted twice in CF-I for diagrammatic purposes. The F_2 generation in the BC/Newf family contains 20 progeny. This intercross family is not considered a reference resource for future map additions. A pair of presumed identical twins was found in each of the two CF sibships marked with an asterisk.

using the CRI-MAP program (Green *et al.* 1990). The *twopoint* option of CRI-MAP was used to obtain an estimated recombination fraction and LOD score for every pair of markers. Linkage groups were formed using a LOD threshold of 5. Twelve markers were typed at PE-AgGen; these typings did not include all animals and so were added only after initial linkage groups had been formed. Markers not placed in a linkage group at the initial stringency (LOD $\geq$5) were added subsequently if they showed linkage only to markers within a single group and with LOD scores $\geq$3.

For each group, a framework map was created consisting of a subset of markers whose order could be established using a LOD threshold of 2. For linkage groups containing eight or fewer markers, all possible marker orders were considered using the *all* option of CRI-MAP. The largest possible framework map was chosen. For linkage groups with more than eight markers, an initial map was formed using the *build* option of CRI-MAP, beginning with a pair of markers and positioning additional markers one at a time in order of decreasing informativeness. Additional analyses were performed using the *flips* option, and the largest subset of markers whose order could be established to within a LOD threshold of 2 was identified. Probable locations of markers not included in the framework maps were determined using the *all* option.

The *chrompic* option of CRI-MAP was used to identify tight multiple recombination events indicative of genotyping errors, mutations, or gene conversions. All typings suggesting multiple crossovers in short intervals ($\leq$5 cM) were retyped or removed from the data set.

Sex-averaged, male and female autosomal genetic distances, as well as the female-specific distance on the X chromosome, were estimated using CRI-MAP with the Kosambi map function. Standard errors for genetic distances were estimated using a bootstrap method (Manly 1997). The genotypes of the parents and grandparents were fixed, as were the allelic phases of both parents. Progeny data were simulated under the assumption that estimated marker orders and genetic distances were correct. The pattern of missing genotypes present in the data was retained in the simulation data. The *fixed* option of CRI-MAP was applied to estimate the genetic distances from the data set. The process was repeated 1000 times for each linkage group, and the standard deviations were reported as estimated standard errors for each interlocus distance.

Estimating genome length: The total sex-averaged genetic length of the canine genome was estimated using a modified version of the method of Chakravarti *et al.* (1991). This method is based on the distribution of recombination fractions among all pairs of autosomal loci. Chakravarti *et al.* considered a backcross experiment with codominant marker data and so could simply count recombinants. To convert the

TABLE 1

Information on canine molecular markers typed in this study

Marker name	Group	Repeat type	Allele range (bp)	No. alleles	No. inf. meioses	Parental HET	Ease of typing	PCR program
AHT103	L21	Di-	074–087	6	227	0.63	3	1
AHT106	Unlinked	Di-	082–088	3	51	0.14	2	4
AHT110	L8	Di-	115–126	5	93	0.26	3	4
AHT111	L2	Di-	074–093	9	237	0.67	4	4
AHT116	L8	Di-	227–244	6	92	0.26	2	3
AHT117	L1	Di-	082–088	3	143	0.44	3	4
AHT120	L35	Di-	071–078	2	162	0.42	4	1
AHT123	L30	Di-	076–090	6	87	0.32	3	4
AHT124	L23	Di-	130–136	2	125	0.29	1	3
AHT125	L22	Di-	088–111	11	331	0.88	3	4
AHT127	L26	Di-	170–187	5	67	0.16	2	4
AHT128	L35	Di-	078–092	7	191	0.49	2	3
AHT130	L14	Di-	105–124	9	263	0.73	4	1
AHT131	L12	Di-	104–117	7	251	0.88	3	4
AHT132	L2	Di-	170–184	6	280	0.70	3	3
AHT133	L29	Di-	145–162	6	204	0.47	2	4
AHT134	L14	Di-	122–139	9	169	0.42	2	4
AHT136	L10	Di-	080–109	11	277	0.70	3	3
AHT137	L10	Di-	131–156	11	268	0.70	2	4
AHT138	L1	Di-	101–115	7	222	0.57	2	3
AHT139	L20	Di-	146–160	7	156	0.44	1	4
AHT140	L15	Di-	100–118	8	301	0.83	1	4
AHT141	CFA5	Di-	104–112	4	139	0.35	2	4
AHT142	L16	Di-	181–202	7	286	0.72	1	3
AHTk18	L11	Di-	083–098	7	62	0.14	2	3
AHTk200	L28	Di-	063–071	3	165	0.49	4	3
AHTk211	L34	Di-	099–102	5	127	0.37	2	1
AR	*CFAX*	*Di-*	*178–189*	*4*	*231*	*0.25*	*1*	*3*
C0X.314	CFAX	Di-	144–165	3	213	0.14	1	1
C05.414	*CFA5*	*Di-*	*156–167*	*5*	*68*	*0.16*	*2*	*3*
C09.173	CFA9	Di-	101–115	9	279	0.72	1	3
C09.250	*CFA9*	*Di-*	*126–157*	*8*	*146*	*0.40*	*4*	*1*
C09.474	*CFA9*	*Di-*	*108–130*	*11*	*242*	*0.60*	*2*	*1*
C09.891	CFA9	Di-	097–104	3	163	0.38	3	4
C09.901	CFA9	Di-	092–099	3	144	0.40	2	3
C20.253	CFA20	Di-	101–114	7	200	0.55	1	3
C20.374	CFA20	Di-	187–204	7	82	0.22	4	1
C20.446	CFA20	Di-	183–202	8	242	0.68	3	1
C20.610	CFA20	Di-	210–216	3	97	0.28	1	3
C20.622	*CFA20*	*Di-*	*210–225*	*6*	*187*	*0.47*	*1*	*3*
CO4107	L8	Di-	≈165	ND	135	0.44	ND	4
CXX.2	*L19*	*Di-*	*207–220*	*9*	*205*	*0.70*	*2*	*1*
CXX.13	L5	Di-	088–103	5	129	0.39	3	4
CXX.16	L8	Di-	178–186	5	117	0.37	4	1
CXX.20	L10	Di-	112–135	8	152	0.85	2	1
CXX.30	*L2*	*Di-*	*149–161*	*9*	*205*	*0.58*	*3*	*3*
CXX.123	*L6*	*Di-*	*126–151*	*6*	*297*	*0.79*	*3*	*1*
CXX.140	L35	Di-	128–143	10	161	0.44	2	4
CXX.156	L14	Di-	131–142	4	252	0.58	3	3
CXX.164	L1	Di-	144–171	6	87	0.28	2	1
CXX.172	L29	Di-	156–164	4	83	0.21	1	3
CXX.176	L17	Di-	187–203	5	160	0.51	1	3
CXX.188	*L19*	*Di-*	*128–132*	*2*	*166*	*0.47*	*3*	*1*
CXX.204	L24	Di-	199–218	7	138	0.40	2	3
CXX.213	L15	Di-	135–167	6	205	0.69	3	1
CXX.236	L1	Di-	270–278	3	56	0.32	2	1
CXX.246	*L1*	*Di-*	*123–131*	*3*	*212*	*0.60*	*1*	*4*

(*continued*)

TABLE 1

(Continued)

Marker name	Group	Repeat type	Allele range (bp)	No. alleles	No. inf. meioses	Parental HET	Ease of typing	PCR program
CXX.251	L1	Di-	138–153	7	142	0.42	3	3
CXX.279	*L3*	*Di-*	*117–133*	*7*	*315*	*0.94*	*1*	*3*
CXX.342	L2	Di-	184–201	5	159	0.77	4	1
CXX.390.2	L25	Di-	304–321	6	160	0.78	4	1
CXX.402	L36	Di-	164–181	4	268	0.71	2	3
CXX.404	L8	Di-	152–168	8	269	0.65	2	3
CXX.406	L9	Di-	169–192	7	210	0.53	2	1
CXX.410	*L16*	*Di-*	*096–125*	*12*	*369*	*0.93*	*2*	*3*
CXX.420	L34	Di-	160–168	5	41	0.07	1	3
CXX.424	*L1*	*Di-*	*176–195*	*7*	*309*	*0.77*	*2*	*3*
CXX.434	L17	Di-	098–111	6	208	0.49	3	3
CXX.436	*L11*	*Di-*	*225–251*	*12*	*131*	*0.50*	*4*	*4*
CXX.438	L2	Di-	261–292	9	209	0.53	2	1
CXX.442	*L11*	*Di-*	*159–171*	*7*	*179*	*0.42*	*2*	*1*
CXX.452	L26	Di-	170–177	5	108	0.26	3	3
CXX.460	*L14*	*Di-*	*124–144*	*9*	*259*	*0.72*	*1*	*1*
CXX.466	*L2*	*Di-*	*149–166*	*7*	*200*	*0.51*	*2*	*1*
CXX.468	L3	Di-	186–203	6	226	0.65	4	4
CXX.502	*L11*	*Di-*	*161–174*	*7*	*151*	*0.37*	*2*	*4*
CXX.602	L8	Di-	168–192	8	324	0.84	3	3
CXX.606	L8	Di-	164–178	6	224	0.58	1	1
CXX.608	L20	Di-	132–149	8	192	0.42	2	1
CXX.618	*L16*	*Di-*	*188–206*	*9*	*276*	*0.70*	*1*	*1*
CXX.620	L5	Di-	189–199	10	230	0.53	2	1
CXX.636	L27	Di-	132–171	10	246	0.60	2	1
CXX.646	L13	Di-	176–192	6	204	0.67	3	1
CXX.672	*Unlinked*	*Di-*	*151–162*	*5*	*320*	*0.78*	*3*	*1*
CXX.852	L9	Di-	098–106	3	198	0.59	3	4
CXX.864.A	L2	Di-	244–255	6	134	0.65	4	1
CXX.864.B	L4	Di-	259–274	7	211	0.69	4	1
CXX.865	L8	Di-	127–156	6	121	0.34	2	4
CXX.866	L25	Di-	242–256	7	184	0.74	2	1
CXX.868	L10	Di-	209–233	7	262	0.65	2	3
CXX.873	L10	Di-	133–157	11	336	0.81	3	4
CXX.876	L12	Di-	098–120	8	254	0.70	2	3
CXX.877	L4	Di-	173–189	7	224	0.73	3	3
CXX.883	L19	Di-	164–189	7	276	0.79	2	1
CXX.889	L32	Di-	107–120	6	57	0.23	2	4
CXX.894	L2	Di-	141–165	10	288	0.72	2	3
CXX.895	L4	Di-	133–141	3	259	0.70	3	3
CXX.900	L7	Di-	121–132	4	121	0.28	3	3
FH2010	*L22*	*Tetra-*	*220–247*	*11*	*180*	*0.52*	*2*	*3*
FH2050	*L24*	*Tetra-*	*249–264*	*3*	*61*	*0.60*	*1*	*3*
FH2054	*L9*	*Tetra-*	*138–185*	*13*	*267*	*0.94*	*1*	*4*
FH2079	*L22*	*Tetra-*	*261–300*	*11*	*326*	*0.83*	*2*	*4*
FH2087L	L15	Tetra-	≈250	ND	148	0.53	ND	4
FH2087U	L2	Tetra-	105–145	ND	291	0.97	ND	4
FH2508	L6	Tetra-	≈190	ND	63	0.26	ND	4
FH2538	L3	Tetra-	≈250	ND	41	0.21	ND	4
GALK1	CFA9	Di-	≈185	ND	66	0.21	ND	4
GLUT4	CFA5	Di-	≈185	ND	230	0.76	ND	4
LEI001	L10	Di-	120–135	6	213	0.53	3	4
LEI002	L11	Di-	130–152	8	234	0.63	3	1
LEI004	Unlinked	Di-	086–112	6	241	0.67	3	1
LEI006	L17	Di-	075–079	2	183	0.56	1	3
M-CPH2	L31	Di-	136–150[a]	4	317	0.86	2	3
M-CPH4	L20	Di-	173–195[a]	7	145	0.47	2	1

(*continued*)

TABLE 1

(Continued)

Marker name	Group	Repeat type	Allele range (bp)	No. alleles	No. inf. meioses	Parental HET	Ease of typing	PCR program
M-CPH7	L2	Di-	202–218[a]	8	256	0.74	2	3
M-CPH8	L23	Di-	239–255[a]	7	244	0.65	2	3
M-CPH10	L33	Di-	194–216[a]	7	225	0.56	3	4
M-CPH12	L16	Di-	234–250[a]	5	201	0.50	3	3
M-CPH14	CFA5	Tri-	231–247[a]	6	265	0.64	2	1
M-CPH16	CFA20	Di-	190–220[a]	9	310	0.81	3	4
M-CPH18	CFA5	Di-	297–307[a]	5	159	0.45	1	3
M-CPH19	L4	Di-	195–209[a]	7	323	0.86	3	3
M-CPH20	L5	Di-	133–153[a]	6	162	0.45	3	3
PDEA	L21	Di-	≈135	ND	91	0.32	ND	4
PEZ1	L5	Di-	095–133	9	180	0.55	2	4
PEZ2	Unlinked	Tri-	120–134	4	117	0.79	1	4
PEZ3	L23	Tetra-	095–153	18	208	0.79	3	1
PEZ5	L9	Tetra-	096–121	7	228	0.67	3	1
PEZ6	L11	Tetra-	166–203	15	266	0.76	2	4
PEZ7	Unlinked	Tetra-	189–206	5	65	0.44	3	1
PEZ8	L33	Tetra-	213–260	16	253	0.73	1	4
PEZ9	L16	Tetra-	210–314	11	120	0.72	2	1
PEZ10	L25	Tetra-	282–302	4	139	0.86	3	1
PEZ11	CFA9	Tetra-	121–173	18	209	0.55	1	1
PEZ12	L4	Tetra-	266–313	12	281	0.79	2	4
PEZ13	L21	Tetra-	170–234	6	116	0.68	3	1
PEZ15	L11	Tetra-	200–245	10	124	0.81	1	4
PEZ16	L11	Tetra-	281–317	9	101	0.57	1	4
PEZ17	L21	Tetra-	199–227	10	110	0.70	1	4
PEZ18	L11	Tetra-	211–272	14	80	0.39	2	4
PEZ19	L3	Tetra-	186–208	6	97	0.74	2	4
PEZ22	L5	Tetra-	171–189	4	67	0.67	2	1
PGKAM[b]	L1	ND	≈150	ND	121	0.50	ND	4
RARA	CFA9	Di-	≈130	ND	100	0.38	ND	4
TAT	Unlinked	Di-	≈155	ND	25	0.09	ND	4
TBP	L9	ND	≈140	ND	51	0.18	ND	4
TETRA	L5	Tetra-	104–130	7	200	0.53	1	1
THRA1	CFA9	Di-	≈180	ND	68	0.24	ND	4
TK1	CFA9	Di-	≈125	ND	72	0.29	ND	4
TSHB	L33	Di-	133–137	2	104	0.24	1	4
UOR0421	L31	Tri-	333–348	4	26	0.08	4	1
UOR0442	L20	Tri-	224–240	4	100	0.28	1	4
UOR4101	L16	Tri-	156–168	4	53	0.12	2	4
UOR4107	L4	Tri-	220–232	4	294	0.70	2	4
vWF	*L11*	*Hexa-*	*150–177*	*7*	*299*	*0.88*	*2*	*4*
WILMS-TF	*L14*	*Tetra-*	*277–301*	*12*	*299*	*0.88*	*2*	*3*

An extended table of canine marker information (including primer sequences) for all loci presently on the linkage map is available at this journal's Internet site (http://www.genetics.org/supplemental/). Markers typed in duplicate between the first- and second-generation maps are shown in italics. Approximate product sizes are given in place of allele ranges for markers typed with autoradiographic methods. Ease-of-typing scores 1–4 are as described in Figure 2. PCR programs 1–4 are as described in materials and methods. Inf., informative; ND, not determined.

[a] Allele ranges are 40 bases longer than the reported sizes (Fredholm and Wintero 1995) owing to a 20-base-pair sequence added to each primer as described in materials and methods.

[b] Marker PGKAM is an autosomal locus that coamplifies with primers for the X-linked PGK marker. Primer pair sequences are from the following references: Holmes *et al.* (1993a,b, 1994, 1995, 1999), Ostrander *et al.* (1993, 1995), Deschenes *et al.* (1994), Shibuya *et al.* (1994), Fredholm and Wintero (1995), Fischer *et al.* (1996), Mellersh *et al.* (1997), Stanford *et al.* (1997), Thomas *et al.* (1997), Werner *et al.* (1997), and Yuzbasiyan-Gurkan *et al.* (1997).

present data into a comparable form, the equivalent number of informative meioses was calculated as

$$n = \mathrm{LOD}/[r \log_{10} r + (1 - r)\log_{10}(1 - r) + \log_{10} 2],$$

where r is the estimated recombination fraction for a pair of loci.

RESULTS

Marker screening: Most markers included in the earlier linkage map of the dog (Mellersh *et al.* 1997) were conducive to high-throughput typing, having been constructed to amplify under uniform PCR conditions. Tetranucleotide repeat loci were used to anchor the map owing to their high polymorphism content and their relative ease of typing (Francisco *et al.* 1996). Available markers not yet mapped were mostly dinucleotide microsatellites developed in many laboratories. These markers were originally optimized under a variety of PCR conditions. For example, magnesium concentrations ranged from 1.0–3.5 mm and annealing temperatures ranged from 50–63°. To identify a subset of markers amenable to efficient, large-scale typing, all available markers were assayed with four PCR cycling parameters. Of 224 markers screened, 155 performed satisfactorily for subsequent genotyping and linkage analysis. The variation in the electropherogram profiles of these markers is described in Figure 2. Further differences in ease of typing were also found, including inconsistent yield, nonallelic PCR products, and alleles differing in size by a single base pair. The ease-of-typing scores for markers listed in Table 1 reflect these differences.

Marker and reference family performance: Markers were typed on 218 animals from the CF mapping resource. Briefly, this panel consisted of two extended pedigrees composed of 18 sibships and 163 F_2 (Figure 1). On average, a sibship consisted of about 9 progeny. For all sibships except one, all parents and grandparents were included. The sibships were similar in marker informativeness as shown in Figure 3A. The average parental heterozygosity for all markers between sibships was 0.57 ± 0.04. In addition, animals from an F_2 intercross involving two distinct breeds were typed and analyzed. The mean parental heterozygosity for the 4 intercross sibships (BC/Newf) was 0.65 ± 0.02. A purebred family of Doberman pinschers (Stanford) typed previously for map assembly (Mellersh *et al.* 1997) had exhibited a low level of parental heterozygosity (0.33) and so was not typed for this study.

The distribution of polymorphism content among all markers mapped with the CF is shown in Figure 3B. When partitioned into types of repeat unit, the mean heterozygosity values were 0.54 ± 0.20 for dinucleotide loci and 0.62 ± 0.20 for tetranucleotide loci (significantly different, $P < 0.01$). Previous work has documented the higher polymorphism content of tetranucleotide repeat loci in dogs (Francisco *et al.* 1996), presumably owing to a higher rate of mutation. Trinucleotide repeat loci showed the lowest mean parental heterozygosity (0.44 ± 0.31), but only 6 markers of this type were tested. Three markers out of 155 were not sufficiently polymorphic among CF parents to warrant further genotyping.

Data integrity: Errors in genotype data can considerably inflate genetic distances (Buetow 1991; Lincoln and Lander 1992). The integrity of newly acquired genotype data was monitored continuously by duplicate typings. At the conclusion of the project, more than 10,500 typings had been done in duplicate. Among duplicate typings there was a 98.0% concordance. Almost 60% of discordant typings were attributable to just 6% of the markers (9/153). Many of the discrepancies were due to systematic errors, which were remedied once detected. For example, many discrepant typings involved alleles differing by a single base pair. In addition to genotypes duplicated within the present data set, 26 previously typed markers (Mellersh *et al.* 1997) were retyped. These duplicate data showed a 99.3% concordance (4139 out of 4168 genotypes).

Because the CF included two multi-generation families with many F_2 progeny, there was considerable power to detect non-Mendelian genotypes. To take full advantage of the pedigree structures, genotypes needed to be coded with absolute alleles consistent among all sibships. Previous genotype data for 150 markers were coded with relative alleles (Mellersh *et al.* 1997), which were inconsistent between sibships and thus precluded analysis of the CF as two whole families. Partitioning sibships into smaller three-generation families to accommodate relative allele assignment decreased the total number of meioses by $6 \pm 4\%$ (data not shown). The power to detect errant typings was also compromised. For these reasons, the new genotype data consisted of alleles labeled as base-pair integers consistent among all families and sibships.

Further inspection of the data revealed two pairs of littermates having nearly identical genotypes (100 and 99.3% identity, respectively). The genotype data were consistent with four animals being sets of monozygotic twins, although one pair was listed as being of opposite sexes. Twinning could not be distinguished from misprocurement of biological samples because all four animals were deceased. The requirement in linkage mapping of independent meiotic events was satisfied by dropping one set of genotype data from each pair of animals prior to map assembly.

Map construction: The genotype data were merged with earlier map data (Mellersh *et al.* 1997), and the complete set was analyzed with CRI-MAP (Green *et al.* 1990). A total of 268 markers (97%) were linked to at least one other marker at LOD $\geq$3; 8 markers were not linked to any other marker. Linkage groups were formed based on highly significant interlocus distances (LOD $\geq$5). Figure 4 shows the spacing and ordering of markers along 40 linkage groups. No markers mapped

TABLE 2

Previously unpublished primer sequences for mapped microsatellite markers

Name	Forward primer	Reverse primer	Source[a]
C09.891	CCAAGGGTTTCTGTTTTATC	CTTTGGATGTCATTCTGTC	C. Gaiser
C09.901	GCTATCCACTGTCACCTT	GTAGTTCTGAGAAGCAAATACATA	C. Gaiser
CXX.13	GAACCTGACCACTCACTC	ACCATCTCTGAGGACAAG	C. Gaiser
CXX.852	ACAGAAAGGGATAAAAGTCT	TTATTTGGCAGTTATTATTCA	C. Gaiser
CXX.864	TGACTGTAGATTTGATTTGAA	AGAAACTAGGAGAATAAAGGTT	C. Gaiser
CXX.865	AGTGTATGTAAGCCTGGAG	TAACTGATGTTATCACTCTCTGC	C. Gaiser
CXX.866	TGTCATAATAGTTGGAATGAC	TTAGAGCTTACTCATGATATCTG	C. Gaiser
CXX.868	AGTAGAGCACAGGGAGAC	AAGAGTATTTTCTTCCATATCC	C. Gaiser
CXX.873	CTGGCAGATTACAGGTAGC	GTTCTCCAAAGCACTCAT	C. Gaiser
CXX.876	CATGGATTCTGCATTTAC	GGTGGAGAACATACAAGAATATAC	C. Gaiser
CXX.877	ATAAGAGGAAAACATTCCG	TAAGGTTAGTTACCACATCTATC	C. Gaiser
CXX.883	ACAGGGAAAGGACAAATA	AATTAATTTTAGTTTGCCAGG	C. Gaiser
CXX.889	TGTATGCACACAGATAAAGT	CACACTTCCCTATACTTACACATA	C. Gaiser
CXX.894	TCAGCATCTAGAAAATTAGGT	ACTCATTTTCTCTTATTCTGCAG	C. Gaiser
CXX.895	CCCTACCTCTGTTCATGT	TATCTGCTTTGTAGGTAATCC	C. Gaiser
CXX.900	TTGGACTTCTAATTTTTCATT	CAACTGACTAAATCTCCTAATG	C. Gaiser
FH2508	GAACAACTGAGTGTCCACATG	TTCTTCCATGTTATCTTCCAGG	This article
FH2538	CAGAGACAAAGGCTTCCCTG	CCCCTCTACTCCTCCTGCTT	This article
PDEA	TAAAGACAGCAGAGCTGAGGG	GGGGCTGCTAATAACCCATT	Y13199
PEZ1	GGCTGTCACTTTTCCCTTTC	CACCACAATCTCTCTCATAAATAC	J. Halverson
PEZ2	TCCTCTCTAACTGCCTATGC	GCCCTTGAATATGAACAATGACACTGTATC	J. Halverson
PEZ3	CACTTCTCATACCCAGACTC	CAATATGTCAACTATACTTC	J. Halverson
PEZ5	GCTATCTTGTTTCCCACAGC	TCACTGTATACAACATTGTC	J. Halverson
PEZ6	ATGAGCACTGGGTGTTATAC	ACACAATTGCATTGTCAAAC	J. Halverson
PEZ7	ATCCTGGAGACCTGGGATTG	GATTGAGTCATCAATAGATG	J. Halverson
PEZ8	TATCGACTTTATCACTGTGG	ATGGAGCCTCATGTCTCATC	J. Halverson
PEZ9	ACAGTTATCCAACAATGAGG	ACGCCTGAACTTAATCCTGG	J. Halverson
PEZ10	CTTCATTGAAGTATCTATCC	CCTGCCTTTGTAAATGTAAG	J. Halverson
PEZ11	ATTCTCTGCCTCTCCCTTTG	TGTGGATAATCTCTTCTGTC	J. Halverson
PEZ12	GTAGATTAGATCTCAGGCAG	TAGGTCCTGGTAGGGTGTGG	J. Halverson
PEZ13	AGTCTGGTGATTTAATTCGG	GTCTAGTCCCCAGTCTAGTTCACTGCCC	J. Halverson
PEZ15	CTGGGGCTTAACTCCAAGTTC	CAGTACAGAGTCTGCTTATC	J. Halverson
PEZ16	GCTCTTTGTAAAATGACCTG	GTGGGAATCGTCCTAAAACCC	J. Halverson
PEZ17	CTAAGGGACTGAACTTCTCC	GTGGAACCTGCTTAAGATTC	J. Halverson
PEZ18	GAGAAGATAAAGCAATTCTC	AAGTCATTAATCTCTCCTCG	J. Halverson
PEZ19	GACTCATGATGTTGTGTATC	TTTGCTCAGTGCTAAGTCTC	J. Halverson
PEZ22	TGGGGAGATCTACAGACCAC	CTAATGTGTCTCTCAAGCCG	J. Halverson
TAT	AGCAATGTGGTGAAATGGTATG	GGCACTCAGAACTATAGGGCC	L47165
TBP	CCTTTCTCTTCTGGAGGAAC	CTGCTGGGATGTCGACTG	L47973
TSHB	TTTCCATGATCAAGGATAAAAGG	GTCTCATTGCCCAGTACCAATTC	A. McGraw
UOR0421	AAAGTTCCTTATTGTCAAGGC	TTAATGTGAGTACCTTGGAGGC	D. R. Meeks-Wagner
UOR0442	TCAGCTGGTTAATGATAGGTGG	ATATTCCTTTGCTCTGAGACGC	D. R. Meeks-Wagner
UOR4101	CCTACCATGGCAAGTGCC	TTCACGGTTGTGAGATGGAG	D. R. Meeks-Wagner
UOR4107	TGACCCTTCTACAACTCGGG	TGTGACCAGTCACTGCTTCC	D. R. Meeks-Wagner

[a] Source lists the GenBank accession number or the person who communicated the primer sequences.

to the pseudoautosomal region of the sex chromosomes, so X-chromosome distances were female specific.

Some linkage groups were assigned to specific canine autosomes based upon prior cytogenetic and linkage data. The canine karyotype consists of 38 mostly small, acrocentric autosomes, a minute metacentric Y chromosome, and a large metacentric X chromosome (Selden *et al.* 1975). Previous results with fluorescence *in situ* hybridization indicate that GLUT4 is on CFA5 (Werner *et al.* 1997) and GALK1, TK1, RARA, and THRA1 are on CFA9 (Werner *et al.* 1997). Moreover, linkage data from Lingaas *et al.* (1997) in combination with results of fluorescence *in situ* hybridization from Fischer *et al.* (1996) show that CPH16 is on CFA20 and CXX.213 (from L15) is on one of the smaller, indistinguishable autosomes (chromosomes 22–38). Chromosomes have been designated in accordance with the karyotype standard for dogs (Switonski *et al.* 1996).

Map statistics: The sex-averaged interval genetic distance of the assembled map was 15.1 ± 0.4 M. The

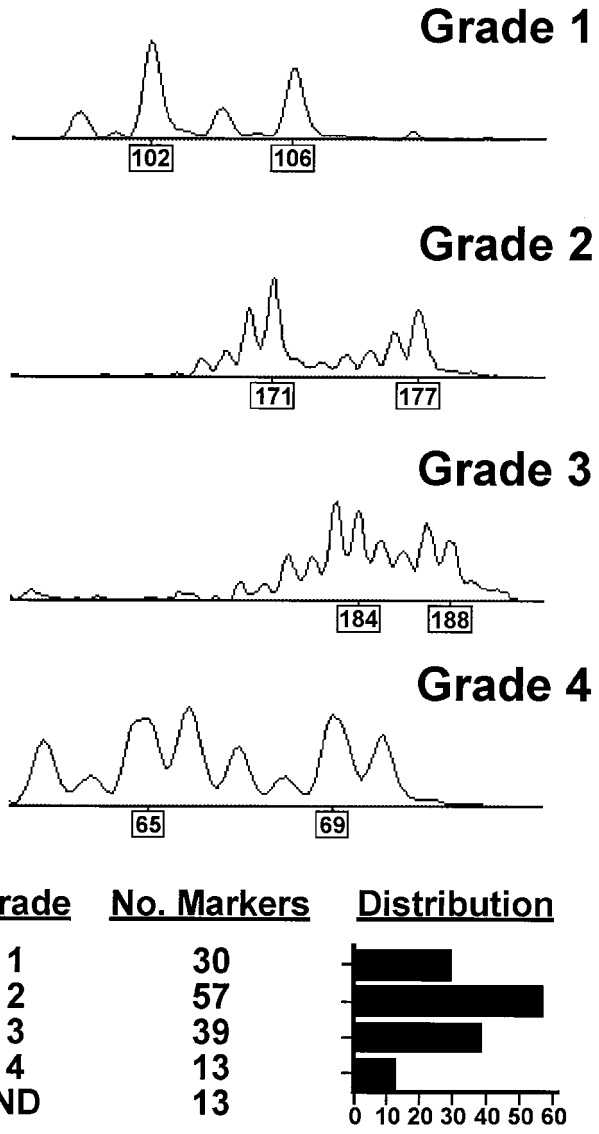


Figure 2.—Qualitative assessment of molecular markers. Four representative electropherograms illustrate variation in the quality and complexity of allele profiles. The grades are subjective assessments of reliability that reflect the user's experience in determining alleles for a given marker and are listed in order of ease of typing from most reliable (Grade 1) to most difficult (Grade 4). Relative fluorescence units are plotted on the *y*-axis and fragment size in base pairs is plotted on the *x*-axis. Each profile depicts a heterozygous genotype. The number and distribution of markers by their qualitative grades is charted on the bottom.

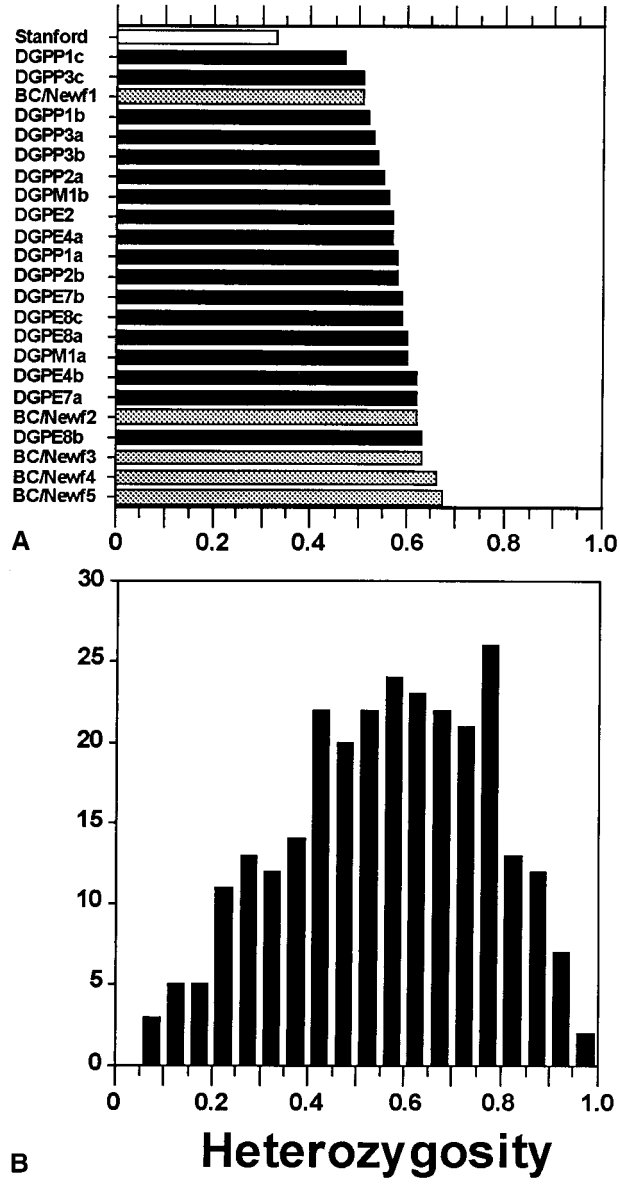


Figure 3.—Informativeness of sibships and markers. (A) The mean heterozygosity values for all markers on the map were calculated for the parents of each sibship. Results for CF sibships are depicted in black while nonreference sibships are depicted in gray (interbreed cross, BC/Newf) or white (purebred cross, Stanford). (B) The mean parental heterozygosity values were calculated for each marker on the map. Markers are grouped in 0.05 heterozygosity increments.

largest detected map interval was 33 cM, with the average being 9.3 ± 6.8 cM. Eight markers were sufficiently informative to detect linkage at 10–30 cM, yet remained unlinked.

As with other mammals, the genetic map of the dog showed pronounced variation in the rate of meiotic recombination between the sexes. The total female interval genetic length was 1.4-fold greater than the male interval genetic length (18.2 ± 0.8 M compared to 12.9 ± 0.5 M, respectively). As expected, the distribution of this variability was not uniform across the genome. Of 162 intervals on the map, 32 (20%) showed a greater female than male distance ($P < 0.05$); 12 (7%) showed a greater male than female distance ($P < 0.05$); and 118 (73%) showed no significant sex difference in recombination.

Assuming markers were randomly distributed, an estimate of the total genetic size of the canine genome was made from the genotype data using a maximum

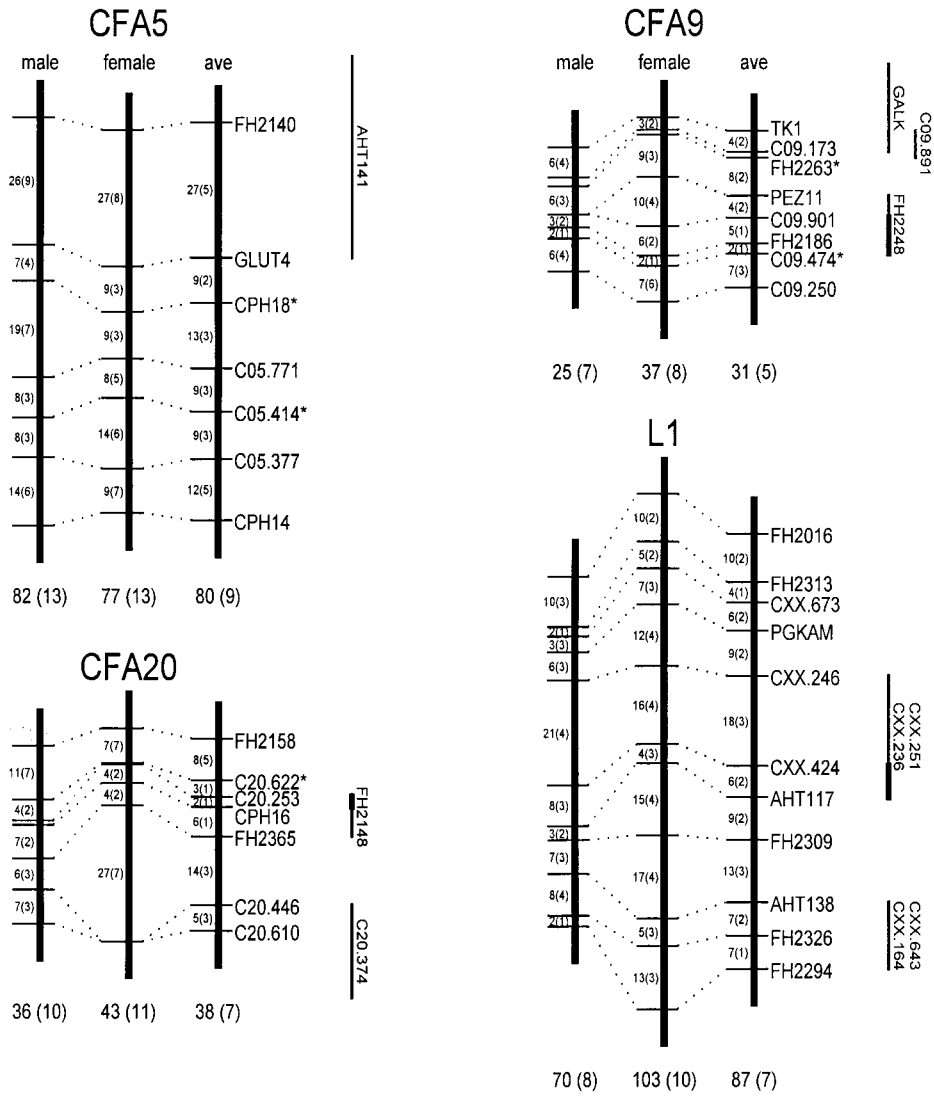


Figure 4.—Male, female, and sex-averaged genetic maps of the canine genome. Total estimated genetic lengths in centimorgans are given below each map. Estimated standard errors for genetic distances (in parentheses) were calculated as described in materials and methods. The order of the markers on the framework maps was favored with a likelihood ratio of at least 1000:1, except for markers indicated by an asterisk. These markers were ordered with a lower likelihood ratio, between 100:1 and 1000:1. The intervals to the right of the maps indicate the most likely placement of some markers, with the likelihood ratio compared to other positions greater than 100:1. Thickened lines indicate intervals for which the likelihood ratio compared to other positions is greater than 10:1. Groups assigned to specific canine autosomes are listed first, followed by unspecified groups numbered and arranged in order of decreasing maximal genetic size (male, female, or sex-averaged). The unlinked markers were LEI004, AHT106, CXX.672, PEZ2, PEZ7, FH2244, TAT, and FH2247.

likelihood computation (Chakravarti *et al.* 1991). The results suggested that the canine genome was 26.5 ± 1.1 M (24.3 to 28.7 M, with 95% confidence).

DISCUSSION

We have presented a linkage map of the domestic dog composed of 276 polymorphic microsatellite markers typed on almost 250 animals. The linkage map was based on meiotic information from the CF mapping resource, which had various breed contributions and a correspondingly high polymorphism content. These multigeneration pedigrees provided 325 phase-known meioses and allowed linkage distances of up to 33 cM to be detected. This mapping panel was especially well suited for early stages of map construction when marker density was low and interval distances were great. Markers mapped in this study were first screened for compatibility with high-throughput genotyping, such that all markers presently on the map were amplifiable with just a few PCR cycling conditions.

The linkage map: The first effort toward a canine linkage map was made by Lingaas *et al.* (1997) who reported 16 linkage groups, 3 of which had ordered markers. Their mapping panel, however, was limited and could not support further map additions. Mellersh *et al.* (1997) published a linkage map based on genotype data acquired from the CF reference panel. Thirty linkage groups were assembled, 14 of which had ordered markers.

The map reported here was built upon the earlier CF-based map and consisted of 268 markers assigned to 40 linkage groups. Twenty-nine groups had ordered markers. Three groups could be assigned to specific canine autosomes (Fischer *et al.* 1996; Lingaas *et al.* 1997; Werner *et al.* 1997), while 36 linkage groups remained anonymous. Five markers were located on the X chromosome, and no markers were Y-linked.

The total sex-averaged interval distance was ~1500 cM, up from 900 cM (Mellersh *et al.* 1997). Eight sufficiently informative markers (3%) remained unlinked. Using a statistical method for assessing genome

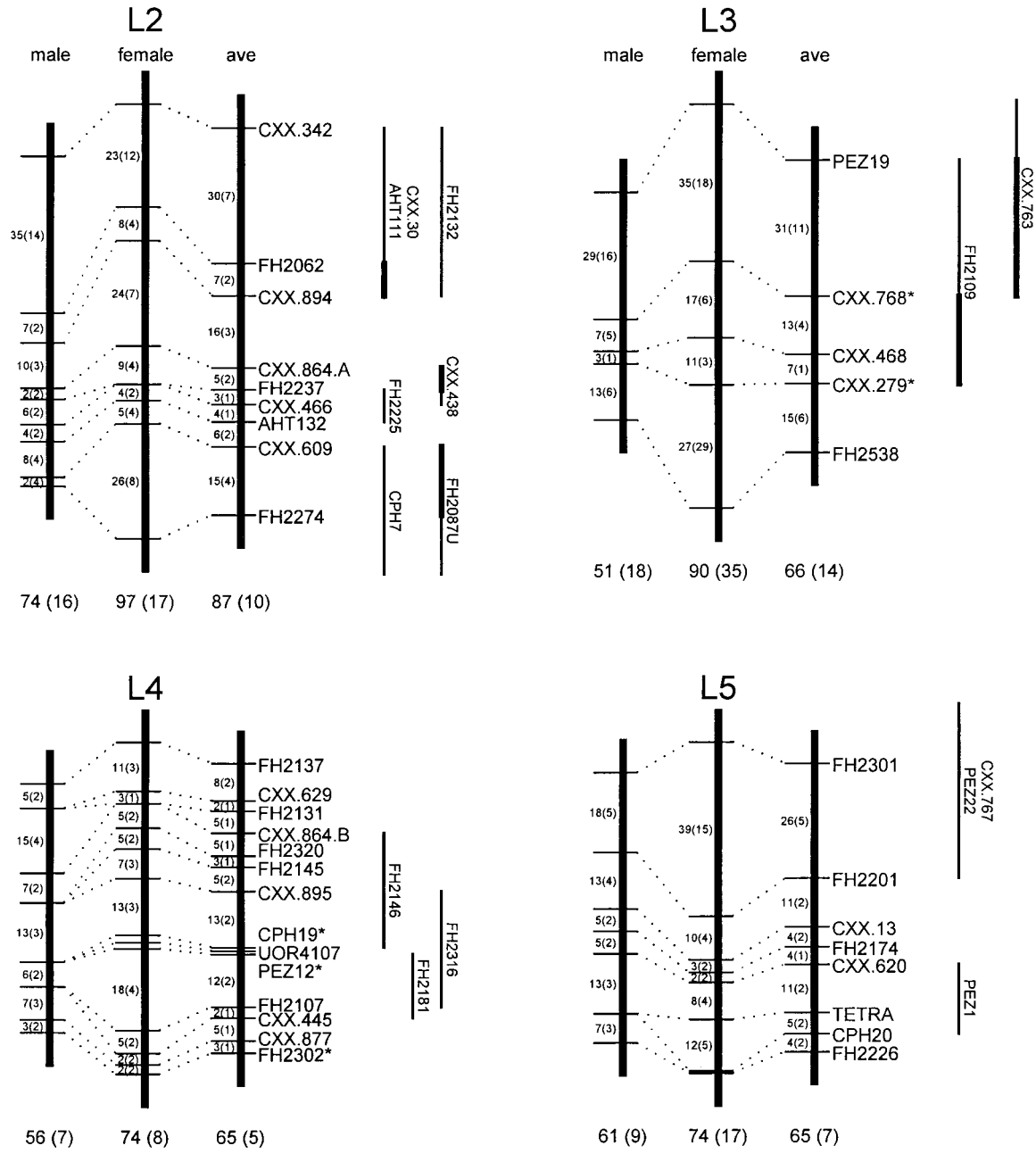


Figure 4.—*Continued.*

size from genotype data, we estimated the total genetic distance of the canine genome to be about 27 M. Given this estimate, more than 55% of the total genetic length was flanked by molecular markers, up from 30% (Mellersh *et al.* 1997).

Map coverage: Although assessing map coverage is difficult, a practical estimate can be made that reflects typical studies where pedigrees are ascertained to allow for detecting linkage at 10 cM. If 10 cM were added to the ends of each linkage group ($n = 40$) and to both sides of unlinked markers ($n = 8$), the genetic length covered by the present map would cover 24 M, or 90% of the estimated genome size. This estimate is likely to be an overestimate as some terminal markers may reside less than 10 cM from the ends of chromosomes.

Map coverage is also a function of marker informativeness. In mouse, the informativeness of markers is known *a priori* for interspecific crosses (Dietrich *et al.* 1996). In humans, highly polymorphic markers can be selected from the more than 8000 markers currently on the genetic map (Broman *et al.* 1998). In dogs, however, molecular diversity may vary considerably among breeds and even between lines of the same breed. Mixed-breed dogs show a relatively constant mean heterozygosity of about 0.55 (Ostrander *et al.* 1993; Zajc *et al.* 1997). In contrast, mean heterozygosity values for purebred

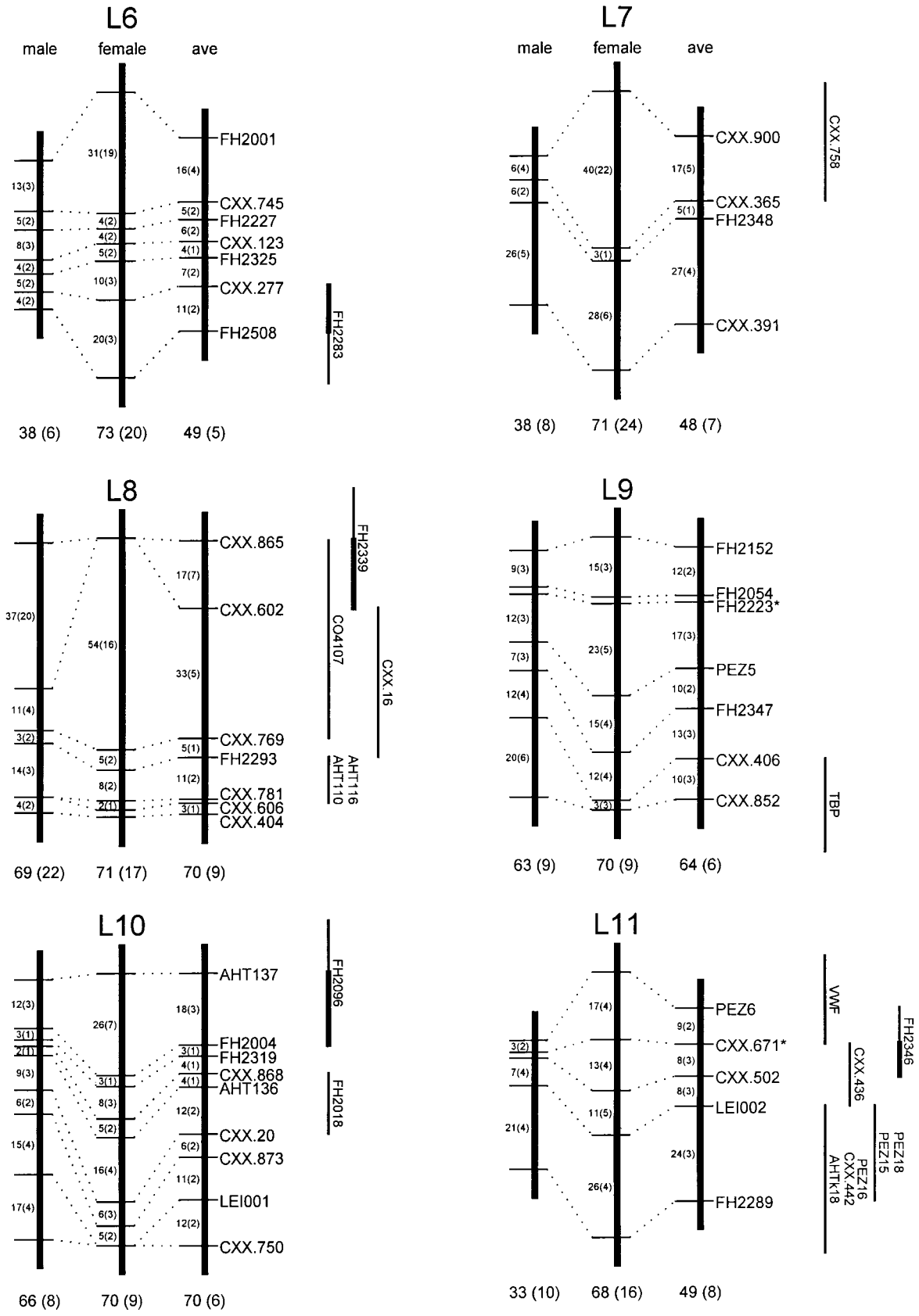


Figure 4.—*Continued.*

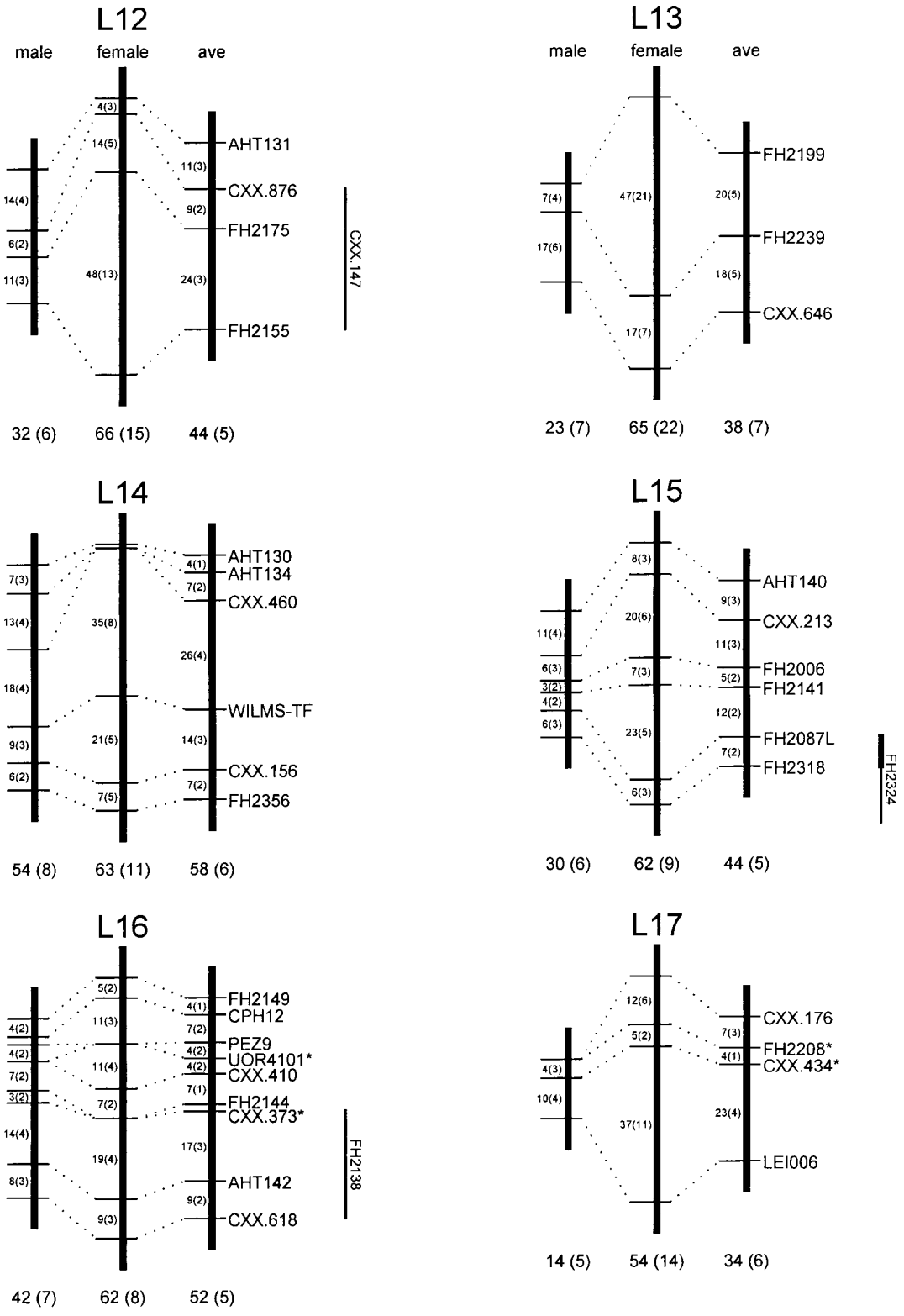


Figure 4.—*Continued.*

dogs range broadly—from 0.35 (Zajc *et al.* 1997) to 0.55 (Fredholm and Wintero 1995), depending on the breed being sampled and its geographical location. The values of heterozygosity calculated from our genotype data are in general agreement with the reported trends. The F_1 generation from the Border collie and Newfoundland intercross showed a mean heterozygosity of 0.65, the mixed-breed parents from the CF showed a mean heterozygosity of 0.55, and the parents of a large purebred pedigree typed in an earlier study (Mellersh

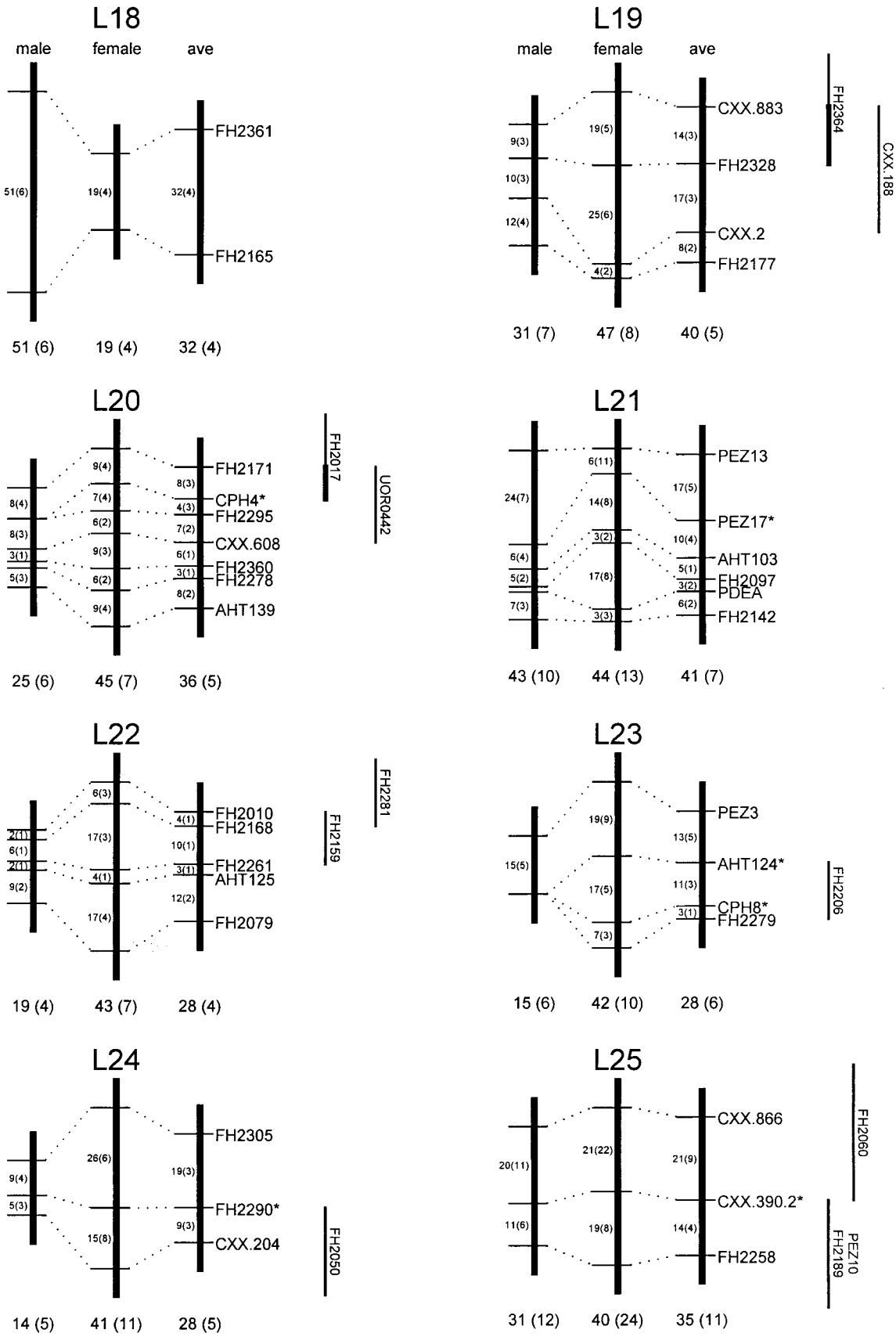


Figure 4.—*Continued.*

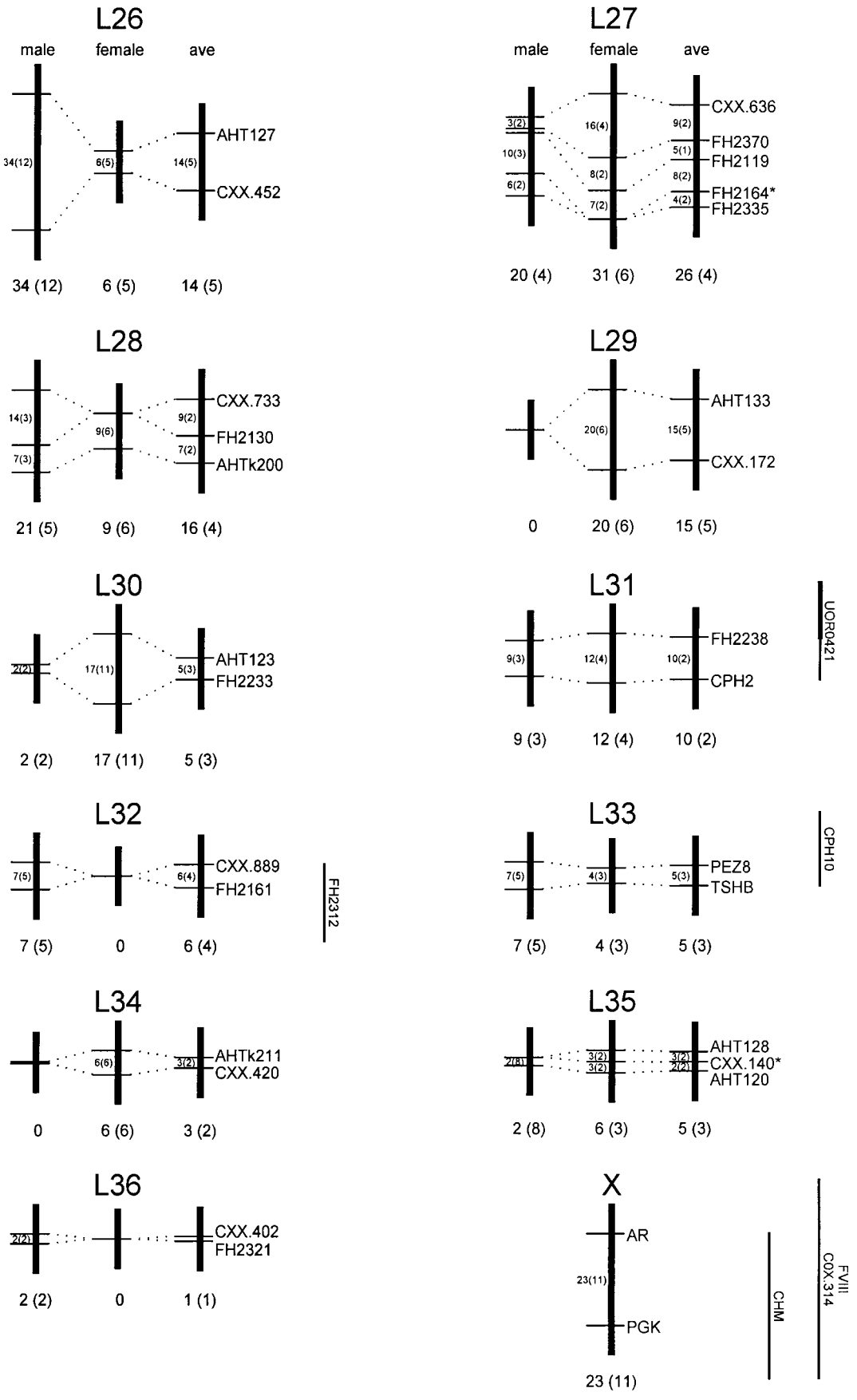


Figure 4.—*Continued.*

TABLE 3

Summary of 15 loci informative for canine comparative mapping

Locus	Gene	Dog position	Mouse chromosome (cM)	Human chromosome
GLUT4	Solute carrier family 2, member 4	CFA5	11 (40)	17p13
THRA1	Thyroid hormone receptor, alpha	CFA9	11 (57)	17q11-2
RARA	Retinoic acid receptor	CFA9	11 (57)	17q12
TK1	Thymidine kinase	CFA9	11 (78)	17q23-5
GALK1	Galactokinase	CFA9	11 (78)	17q24
TBP	TATA box binding protein, TFIID	L9	17 (08)	6q27
VWF[a]	von Willebrand factor	L11	6 (59)	12p13
WILMS-TF[a]	Wilms' tumor factor	L14	2 (58)	11p13
PDEA	Phosphodiesterase alpha	L21	18 (31)	5q31-4
TSHB	Thyroid stimulating hormone, beta	L33	3 (48)	1p13
TAT	Tyrosine aminotransferase	UL	8 (40)	16q22
F VIII[a]	Coagulation factor 8	X	X (30)	Xq28
AR[a]	Androgen receptor	X	X (36)	Xq11-2
PGK[a]	Phosphoglycerate kinase	X	X (45)	Xq13
CHM[a]	Choroidermia	X	X (96)	Xq21

All mouse and human assignments are taken from the Mouse Genome Data Base (http://www.informatics.jax.org/) and the Human Genome Data Base (http://gdbwww.gdb.org/). UL, unlinked.

[a] Loci assigned in the first-generation linkage map (Mellersh *et al.* 1997).

et al. 1997) showed a mean heterozygosity of 0.33. The context-dependence of marker informativeness in dogs indicates that additional characterizations of molecular variation within breeds, where Mendelian disease is most common, would aid greatly in the design of canine linkage studies.

Anchor sites for comparative mapping: Comparative maps use conserved gene sequences to identify larger chromosomal segments that have remained intact during evolution (O'Brien *et al.* 1988; Nadeau 1995). A comparative map provides the opportunity to draw upon map information from one species and apply it to another. Thus determinants of a phenotype in one species can be directly tested for their involvement in a similar phenotype in another species. A robust comparative map facilitates directly testing candidate genes for linkage. Nine newly mapped markers were tied to genes of known function. Table 3 lists all 15 genes presently on the canine linkage map and summarizes their position in the mouse and human genomes. Many of these loci have already been successfully applied to comparative mapping of canine disease genes (Werner *et al.* 1997; Acland *et al.* 1998). Increasing the density of gene markers on the linkage map will further establish evolutionarily conserved gene clusters and help to predict the location of candidate genes in the dog genome. In the future, some disease genes will be mapped and cloned first in the dog, thereby reversing the flow of information and providing candidates for human linkage studies.

Toward map closure: Previous mapping studies revealed linkage among molecular markers (Lingaas *et al.* 1997; Werner *et al.* 1997), some of which overlap with those now on the CF-based map. Table 4 lists 21 markers whose location on the CF linkage map can be inferred from previous associations. Most of these loci are positioned in regions of high marker density and therefore do not add to map coverage or aid in map closure. Filling vacant map regions with markers chosen randomly becomes less efficient as coverage increases. Simulation studies show that as many as 800 randomly chosen markers may be necessary to achieve map closure (*i.e.*, 90% of the simulations produced maps where the largest remaining interval distance was $\leq$32 cM, a linkage distance detectable with the CF; data not shown). Some regions may be more problematic than described by the simulation data. For example, the paucity of markers on the X and the Y chromosomes in many mammals reflects a decreased frequency of microsatellites in these regions (Dietrich *et al.* 1996). In dogs, only one random marker was assigned to the X chromosome (COX.314), even though it is the largest chromosome in the canine karyotype (Selden *et al.* 1975).

A directed approach could considerably accelerate map closure. In human and mouse, new microsatellite markers are now mapped with whole-genome radiation hybrid (WGRH) panels and placed on the linkage maps only if they fall in marker-deficient regions (Gyapay *et al.* 1996; McCarthy *et al.* 1997; Slonim *et al.* 1997; Stewart *et al.* 1997). Radiation hybrid mapping uses radiation-induced breakage and coretention of markers as analogs for meiotic recombination and cosegregation of coupled alleles, respectively. The resolution of radia-

TABLE 4

Linkage group associations inferred from previous mapping results

Linkage group	Locus	Primary anchor	Reported distance (cM)	LOD	Secondary anchor
CFA9	MYL4[a]	TK1	1.5	17.9	THRA1
CFA9	GH1[a]	TK1	1.5	18.5	THRA1
CFA9	NF1[a]	RARA	31.2	3.4	THRA1
CFA9	CRYBA1[a]	RARA	33.6	2.7	THRA1
CFA20	C20.630	CPH16	12.2	4.1	
L3	LEI005	CXX.279	5.6	23.3	
L5	VIAS-D10	CPH20	8.3	3.6	
L6	453	CXX.123	1.8	9.6	
L6	TF	CXX.123	3.0	13.3	
L6	CPH6	CXX.123	10.8	7.5	
L15	LEI032	CXX.213	21.8	3.5	
L15	AHTk120	CXX.213	8.3	8.1	
L16	P11	CXX.618	3.9	5.4	
L19	CPH9	CXX.2	6.5	6.9	CXX.188
L22	AHT118	AHT125	2.1	10.5	FH2010
L22	CXX.363	AHT125	7.1	9.3	FH2010
L22	CXX.130	AHT125	15.0	9.0	FH2010
L23	LEI024	CPH8	11.1	5.9	
L23	LEI025	CPH8	2.1	10.8	
L26	CPH13	AHT127	6.9	10.1	
L29	CXX.359	AHT133	7.0	8.9	

[a] Associations from Werner *et al.* (1997). All other associations from Lingaas *et al.* (1997).

tion hybrid mapping can be customized to accommodate marker density by varying the radiation dose. Polymorphism is not required to assay marker retention in hybrid lines, so each locus has comparable mapping potential. A WGRH panel for the dog that could aid in closure of the canine linkage map has recently become commercially available (Research Genetics, Birmingham, AL). Once framework genetic markers have been typed, the WGRH panel could be used to screen for markers anchoring previously unmapped regions. Subsequent typing of these markers on the CF panel would drive closure of the linkage map.

Continued map improvement: A salient feature of maps is that they improve with use. The CF resource has been valuable for assembling a genetic linkage map of the dog genome and will continue to support map refinement. DNA from the CF mapping resource will be made available to investigators seeking to place additional public domain markers on the linkage map, through arrangements to be announced. All CF genotype data should in turn be made publicly accessible. Accordingly, we deposit the merged primary genotype data for the present linkage map at this jounal's Internet site (http://www.genetics.org/supplemental/). It is hoped that the reference families and genotype data reported here will facilitate future improvements of the canine linkage map.

The authors thank N. Holmes, M. Binns, A. McGraw, D. R. Meeks-Wagner, J. Halverson, and C. Gaiser for providing primer sequences prior to publication. We thank M. Bozzini, L. Issel-Tarver, N. Wiegand, P. Lauer, and T. Speed for technical advice and helpful discussions. We thank other members of our laboratories, especially S. Okamura, D. Metallinos, and A. Dillin, for providing critical comments about the manuscript. CF pedigrees were developed as part of a project sponsored by the National Institutes of Health (EY-06855), the Foundation Fighting Blindness, and the Progressive Retinal Atrophy Research Fund. This work was funded in part by the Canine Health Foundation of the American Kennel Club (grant 1291 to J.R. and grant 1268 to E.O.), by the American Cancer Society (grant VM-165 to E.O.), by the Gatsby Charitable Foundation, and by the private support of Martin and Enid Gleich. Additional support was provided by the following breed clubs: Chief Solano Kennel Club, Golden Gate Labrador Retriever Club, Two Cities Kennel Club, Northern California Siberian Husky Club, and the San Joaquin Kennel Club. C.S.M. was supported by a Wellcome Prize Travelling Research Fellowship. M.W.N. was supported by a U.S. Department of Energy Human Genome Distinguished Postdoctoral Fellowship administered by the Oak Ridge Institute for Science and Education (IF32GM16262-01), and by a California Division-American Cancer Society Senior Fellowship (1-30-97B).

LITERATURE CITED

Acland, G. M., K. Ray, C. S. Mellersh, W. Gu, A. A. Langston *et al.*, 1998 Linkage analysis and comparative mapping of canine progressive rod-cone degeneration (*prcd*) establishes potential locus homology with retinitis pigmentosa (*RP17*) in humans. Proc. Natl. Acad. Sci. USA **95:** 3048–3053.

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

Bihoreau, M. T., D. Gauguier, N. Kato, G. Hyne, K. Lindpaintner *et al.*, 1997 A linkage map of the rat genome derived from three F_2 crosses. Genome Res. **7:** 434–440.

Botstein, D., R. L. White, M. Skolnick and R. W. Davis, 1980 Construction of a genetic linkage map in man using restriction fragment length polymorphisms. Am. J. Hum. Genet. **32:** 314–331.
Broman, K. W., J. C. Murray, V. C. Sheffield, R. L. White and J. L. Weber, 1998 Comprehensive human genetic maps: individual and sex-specific variation in recombination. Am. J. Hum. Genet. **63:** 861–869.
Buetow, K. H., 1991 Influence of aberrant observations on high-resolution linkage analysis outcomes. Am. J. Hum. Genet. **49:** 985–994.
Chakravarti, A., L. K. Lasher and J. E. Reefer, 1991 A maximum likelihood method for estimating genome length using genetic linkage data. Genetics **128:** 175–182.
de Gortari, M. J., B. A. Freking, R. P. Cuthbertson, S. M. Kappes, J. W. Keele *et al.*, 1998 A second-generation linkage map of the sheep genome. Mamm. Genome **9:** 204–209.
Deschenes, S. M., J. M. Puck, A. S. Dutra, R. L. Somberg, P. J. Felsburg *et al.*, 1994 Comparative mapping of canine and human proximal Xq and genetic analysis of canine X-linked severe combined immunodeficiency. Genomics **23:** 62–68.
Dib, C., S. Faure, C. Fizames, D. Samson, N. Drouot *et al.*, 1996 A comprehensive genetic map of the human genome based on 5,264 microsatellites. Nature **380:** 152–154.
Dietrich, W. F., J. Miller, R. Steen, M. A. Merchant, D. Damron-Boles *et al.*, 1996 A comprehensive genetic map of the mouse genome. Nature **380:** 149–152.
Fischer, P. E., N. G. Holmes, H. F. Dickens, R. Thomas, M. M. Binns *et al.*, 1996 The application of FISH techniques for physical mapping in the dog (*Canis familiaris*). Mamm. Genome **7:** 37–41.
Francisco, L. V., A. A. Langston, C. S. Mellersh, C. L. Neal and E. A. Ostrander, 1996 A class of highly polymorphic tetranucleotide repeats for canine genetic mapping. Mamm. Genome **7:** 359–362.
Fredholm, M., and A. K. Wintero, 1995 Variation of short tandem repeats within and between species belonging to the Canidae family. Mamm. Genome **6:** 11–18.
Green, P., K. Fall and S. Crooks, 1990 Documentation for CRI-MAP, version 2.4.
Gyapay, G., K. Schmitt, C. Fizames, H. Jones, N. Vega-Czarny *et al.*, 1996 A radiation hybrid map of the human genome. Hum. Mol. Genet. **5:** 339–346.
Holmes, N. G., S. J. Humphreys, M. M. Binns, R. Curtis, A. Holliman *et al.*, 1993a Characterization of canine microsatellites. Exper. Suppl. **67:** 415–420.
Holmes, N. G., C. S. Mellersh, S. J. Humphreys, M. M. Binns, A. Holliman *et al.*, 1993b Isolation and characterization of microsatellites from the canine genome. Anim. Genet. **24:** 289–292.
Holmes, N. G., N. J. Strange, M. M. Binns, C. S. Mellersh and J. Sampson, 1994 Three polymorphic canine microsatellites. Anim. Genet. **25:** 200.
Holmes, N. G., H. F. Dickens, H. L. Parker, M. M. Binns, C. S. Mellersh *et al.*, 1995 Eighteen canine microsatellites. Anim. Genet. **26:** 132–133.
Holmes, N. G., H. F. Dickens, M. W. Neff, J. M. Mee, J. Sampson *et al.*, 1999 Nine canine microsatellites. Anim. Genet. (in press).
Human Genome Database (GDB), The Jackson Laboratory, Bar Harbor, ME. World Wide Web (URL: http://gdbwww.gdb.org/).
Hung, T., K. Mak and K. Fong, 1990 A specificity enhancer for polymerase chain reaction. Nucleic Acids Res. **18:** 4953.
Kappes, S. M., J. W. Keele, R. T. Stone, R. A. McGraw, T. S. Sonstegard *et al.*, 1997 A second-generation linkage map of the bovine genome. Genome Res. **7:** 235–249.
Lincoln, S. E., and E. S. Lander, 1992 Systematic detection of errors in genetic linkage data. Genomics **14:** 604–610.
Lingaas, F., A. Sorensen, R. K. Juneja, S. Johansson, M. Fredholm *et al.*, 1997 Towards construction of a canine linkage map: establishment of 16 linkage groups. Mamm. Genome **8:** 218–221.
Manly, B. F. J., 1997 *Randomization, Bootstrap, and Monte Carlo Methods in Biology.* Chapman and Hall, London.
Marklund, L., M. Johansson Moller, B. Hoyheim, W. Davies, M. Fredholm *et al.*, 1996 A comprehensive linkage map of the pig based on a wild pig-Large White intercross. Anim. Genet. **27:** 255–269.
McCarthy, L. C., J. Terrett, M. E. Davis, C. J. Knights, A. L. Smith *et al.*, 1997 A first-generation whole genome-radiation hybrid map spanning the mouse genome. Genome Res. **7:** 1153–1161.
Mellersh, C. S., A. A. Langston, G. M. Acland, M. A. Fleming, K. Ray *et al.*, 1997 A linkage map of the canine genome. Genomics **46:** 326–336.
Mouse Genome Database (MGD), Mouse Genome Informatics, The Jackson Laboratory, Bar Harbor, ME. World Wide Web (URL: http://www.informatics.jax.org/).
Nadeau, J. H., 1995 A Rosetta stone of mammalian genetics. Nature **373:** 363–365.
Nicholas, F. W., S. C. Brown and P. R. Le Tissier, 1998 Online Mendelian Inheritance in Animals (URL: http://www.angis.su.oz.au/Databases/BIRX/omia/omia_form.html).
O'Brien, S. J., H. N. Seuanez and J. E. Womack, 1988 Mammalian genome organization: an evolutionary view, pp. 323–351 in *Annual Review of Genetics*, edited by A. Campbell. Annual Reviews Inc., Palo Alto, CA.
Ostrander, E. A., G. F. Sprague, Jr. and J. Rine, 1993 Identification and characterization of dinucleotide repeat $(CA)_n$ markers for genetic mapping in dog. Genomics **16:** 207–213.
Ostrander, E. A., F. A. Mapa, M. Yee and J. Rine, 1995 One hundred and one new simple sequence repeat-based markers for the canine genome. Mamm. Genome **6:** 192–195.
Robic, A., D. Milan, N. Woloszyn, J. Riquet, M. Yerle *et al.*, 1997 Contribution to the physically anchored linkage map of the pig. Anim. Genet. **28:** 94–102.
Selden, J. R., P. S. Moorhead, M. L. Oehlert and D. F. Patterson, 1975 The Giemsa banding pattern of the canine karyotype. Cytogenet. Cell Genet. **15:** 380–387.
Shibuya, H., B. K. Collins, T. H. Huang and G. S. Johnson, 1994 A polymorphic $(AGGAAT)_n$ tandem repeat in an intron of the canine von Willebrand factor gene. Anim. Genet. **25:** 122.
Shuber, A. P., V. J. Grondin and K. W. Klinger, 1995 A simplified procedure for developing multiplex PCRs. Genome Res. **5:** 488–493.
Slonim, D., L. Kruglyak, L. Stein and E. Lander, 1997 Building human genome maps with radiation hybrids. J. Comput. Biol. **4:** 487–504.
Stanford, J. L., J. J. Just, M. Gibbs, K. G. Wicklund, C. L. Neal *et al.*, 1997 Polymorphic repeats in the androgen receptor gene: molecular markers of prostate cancer risk. Cancer Res. **57:** 1194–1198.
Stewart, E. A., K. B. McKusick, A. Aggarwal, E. Bajorek, S. Brady *et al.*, 1997 An STS-based radiation hybrid map of the human genome. Genome Res. **7:** 422–433.
Switonski, M., N. Reimann, A. A. Bosma, S. Long, S. Bartnitzke *et al.*, 1996 Report on the progress of standardization of the G-banded canine (*Canis familiaris*) karyotype. Chromosome Res. **4:** 306–309.
Thomas, R., N. G. Holmes, P. E. Fischer, H. F. Dickens, M. Breen *et al.*, 1997 Eight canine microsatellites. Anim. Genet. **28:** 153–154.
Wang, D. G., J. B. Fan, C. J. Siao, A. Berno, P. Young *et al.*, 1998 Large-scale identification, mapping, and genotyping of single-nucleotide polymorphisms in the human genome. Science **280:** 1077–1082.
Weber, J. L., and P. E. May, 1989 Abundant class of human DNA polymorphisms which can be typed using the polymerase chain reaction. Am. J. Hum. Genet. **44:** 388–396.
Werner, P., M. G. Raducha, U. Prociuk, P. S. Henthorn and D. F. Patterson, 1997 Physical and linkage mapping of human chromosome 17 loci to dog chromosomes 9 and 5. Genomics **42:** 74–82.
Wilcox, B., and C. Walkowicz, 1993 *The Atlas of Dog Breeds of the World.* T.F.H. Publications, Neptune City, NJ.
Yuzbasiyan-Gurkan, V., S. H. Blanton, Y. Cao, P. Ferguson, J. Li *et al.*, 1997 Linkage of a microsatellite marker to the canine copper toxicosis locus in Bedlington terriers. Am. J. Vet. Res. **58:** 23–27.
Zajc, I., C. S. Mellersh and J. Sampson, 1997 Variability of canine microsatellites within and between different dog breeds. Mamm. Genome **8:** 182–185.

Communicating editor: D. Botstein