A Comprehensive Linkage Map of the Dog Genome

Aaron K. Wong,*,[1] Alison L. Ruhe,* Beth L. Dumont,[†] Kathryn R. Robertson,* Giovanna Guerrero,*,[2] Sheila M. Shull,* Janet S. Ziegle,[‡] Lee V. Millon,* Karl W. Broman,[§] Bret A. Payseur[†] and Mark W. Neff*,[3,4]

*Veterinary Genetics Laboratory, School of Veterinary Medicine, University of California, Davis, California 95616, [‡]Applied Biosystems, Foster City, California 94404, [§]Department of Biostatistics and Medical Informatics and [†]Laboratory of Genetics, University of Wisconsin, Madison, Wisconsin 53706

ABSTRACT

We have leveraged the reference sequence of a boxer to construct the first complete linkage map for the domestic dog. The new map improves access to the dog's unique biology, from human disease counterparts to fascinating evolutionary adaptations. The map was constructed with ~3000 microsatellite markers developed from the reference sequence. Familial resources afforded 450 mostly phase-known meioses for map assembly. The genotype data supported a framework map with ~1500 loci. An additional ~1500 markers served as map validators, contributing modestly to estimates of recombination rate but supporting the framework content. Data from ~22,000 SNPs informing on a subset of meioses supported map integrity. The sex-averaged map extended 21 M and revealed marked region- and sex-specific differences in recombination rate. The map will enable empiric coverage estimates and multipoint linkage analysis. Knowledge of the variation in recombination rate will also inform on genomewide patterns of linkage disequilibrium (LD), and thus benefit association, selective sweep, and phylogenetic mapping approaches. The computational and wet-bench strategies can be applied to the reference genome of any nonmodel organism to assemble a *de novo* linkage map.

G ENOMICS has broadened the exploration of natural variation, helping to bring the naturalist perspective back into the fold of modern biology. The dog is an example of a nonmodel organism with much to offer in the way of natural phenotypic diversity. Since Darwin (1883), the domesticated dog has been recognized as a model of mammalian evolution, with striking variation in form and function. Much of the phenotypic diversity in the dog appears to have evolved rapidly, accelerated by the intense pressure of artificial selection and the force of genetic drift through population bottlenecks. Selective breeding over centuries has served as the preliminary experiment in canine genetics: the mutants have been screened and the strains have been established. The architectures of domestic traits are likely enriched for genes of large effect, given that artificial selection can act only on discernible phenotypic differ-

ences (ANDERSSON 2001; NEFF and RINE 2006). It follows that the diversity in size, shape, and behavior in the dog is genetically tractable. Moreover, this "adaptive" variation can be couched in the developmental context of an extant progenitor, the wolf (VILÀ *et al.* 1997), to illuminate the morphological and behavioral antecedents from which breed-defining traits have been derived. In addition to propagating purposefully bred traits, managed breeding has had unintended consequences. The dog suffers many of the same diseases as man, from Mendelian defects (*e.g.*, deafness) to complex diseases (*e.g.*, cancer susceptibilities) (OSTRANDER *et al.* 1993). These health issues follow breed predilection, implying that ancestral mutations have been trapped within the now-closed gene pools, presumably from founder effects at the inception of breed registries.

Variation in the dog, both adaptive and maladaptive, can be understood genetically through several phenotype-driven approaches (reviewed in KARLSSON and LINDBLAD-TOH 2008). Linkage analysis of Mendelian traits and monogenic diseases is particularly powerful owing to large sibships and pedigree ascertainment through breeders. Genomewide association mapping leverages the strengths of population structure and is driven by a breed-based linkage disequilibrium (LD) that is 20- to 100-fold more extensive than that observed in human populations (SUTTER *et al.* 2004; LINDBLAD-TOH *et al.* 2005). Less conventional methods for detecting

genotype–phenotype correlations, such as *in silico* association (GRUPE *et al.* 2001; JONES *et al.* 2008) and selective sweep (POLLINGER *et al.* 2005) mapping, also hold promise in the dog, especially for understanding hallmark traits that have been bred to fixation, thereby presenting a "segregation problem" to geneticists.

Each of these approaches to genetic mapping is more powerfully applied with a thorough understanding of recombination rate. This includes a relevance to LD mapping that may be particularly important in canine genetics. In man, LD tends to extend over a kilobase scale, for which genetic and physical distances are reasonably correlated. In contrast, LD in dog breeds extends to a megabase scale (KARLSSON *et al.* 2007); understanding local variation in recombination rate (cM/Mb) could facilitate interpretations of LD.

Given the strengths of canine genetics, a principal unmet need is a comprehensive description of genome-wide variation in recombination rate. The current linkage map is built upon the success of previous efforts (LINGAAS *et al.* 1997; MELLERSH *et al.* 1997; DOLF 1999; LANGSTON *et al.* 1999; NEFF *et al.* 1999); however, those maps were constructed with <600 markers and lack the coverage, precision, and accuracy needed to optimize linkage analysis and subserve LD mapping (DUKES-MCEWAN and JACKSON 2002). A comprehensive map would enhance genetic analysis in several ways. It would allow assessment of empiric coverage during linkage scans, provide essential information for multipoint and fine-resolution mapping, and serve as a foundation for interpreting LD across the genome. The map can also serve as a comparative resource for continuing to improve our understanding of the basic processes governing meiotic recombination and faithful chromosome segregation.

Toward this end, we have leveraged the high-quality draft sequence (7.6×) of the dog genome (LINDBLAD-TOH *et al.* 2005) to assemble a *de novo* linkage map. The reference sequence afforded an abundance of molecular polymorphisms, which were computationally "cloned" for even spacing. The physical scaffold aided map assembly by providing a first approximation of marker order that was ultimately tested with genetic data. Several measures of map integrity indicated that our estimates of recombination rate were both accurate and reasonably precise. Thus a comprehensive map will join a list of important canine resources, from high-resolution radiation hybrid maps (VIGNAUX *et al.* 1999; BREEN *et al.* 2001; GUYON *et al.* 2003) to the high-quality draft sequence, annotated with 2.4 million SNPs. We discuss the biological implications of this new canine map and provide online resources to facilitate its application.

MATERIALS AND METHODS

Samples and markers: The familial resources used in this study are summarized in Table 1. The Cornell families (CF) were a genetically admixed research colony inclusive of six breed backgrounds (NEFF *et al.* 1999). The F_2 family was developed previously as an experimental intercross of a border collie (sire) and a Newfoundland (dam) (NEFF *et al.* 1999). The extended Silken Windhound family was a privately managed pedigree provided for this study. A panel of 36 purebred dogs was assembled and surveyed to test marker performance across breed populations (supporting information, Table S1). DNA was prepared from blood, tissue, or buccal swab samples using previously described methods (BELL *et al.* 1981; OBERBAUER *et al.* 2003).

Microsatellite loci were computationally mined from *canFam1* and *canFam2* (LINDBLAD-TOH *et al.* 2005) by exploiting the RepeatMasker track of the University of California Santa Cruz (UCSC) Genome Browser (www.genome.ucsc. edu). Perfect $(CA)_n$ microsatellites ($n > 12$) were chosen for even spacing across the genome (∼750-kb steps). PCR primers were designed with Primer3 using standardized parameters (Table S2). Candidate oligo sequences were screened against the genome sequence using BLAST to ensure unique primer binding sites. Markers with a primer sequence having ≥95% identity to a secondary annealing site were replaced by a suitable adjacent locus. Table S3 lists the autosomal molecular markers developed in this study. An electronic file with the full marker set is also available (File S2).

Microsatellite genotyping: Microsatellite markers were multiplexed in sets of 3–9 loci on the basis of differential dye labels and expected differences in PCR product size. Marker loci were amplified according to a published method that uses dye-specific M13 oligonucleotides to incorporate fluorescent label during the PCR (OETTING *et al.* 1995). In addition, a leader sequence (GTTTCTT) was appended to the 5′ end of a subset of reverse primers to catalyze the nontemplated enzymatic addition of a nucleotide to the 3′ end of the labeled product strand (BROWNSTEIN *et al.* 1996). Final PCR concentrations for forward (40 mers) and reverse primers (27 mers) were 0.18 and 1.8 μM, respectively. M13 primers for labeling product were included in the PCR at a final concentration of 0.36 μM.

PCRs were performed in 17-μl reaction volumes with 50 ng genomic DNA template or 2 μl buccal swab extract, and final reaction conditions of 2.5 mM $MgCl_2$, 1× buffer, 200 μM each dNTP (GenScript), and 0.15 units of Taq DNA polymerase (ABGene). Reactions were covered with inert Chill-Out (BioRad) to permit hot-start PCR. Thermocycling consisted of an initial denaturation step of 93° for 3 min and a final extension step of 72° for 20 min. Marker amplifications were for 7 cycles of 93° for 20 sec, 65° for 30 sec, 72° for 2 min; 5 cycles of 93° for 20 sec, 58° for 30 sec, 72° for 2 min; and 25 cycles of 93° for 20 sec, 55° for 30 sec, 72° for 2 min.

An aliquot of PCR product (0.4 μl) was combined with 0.5 μl of GeneScan 500 LIZ (Applied Biosystems) and 10 μl Hi-Di formamide and denatured 3 min at 95°. An aliquot (1 μl) was separated by electrophoresis and detected by fluorescence using ABI 3730 capillary instruments. The command line option in GeneMapper 4.0 was used to automatically preprocess ABI 3730 files remotely. Genotypes were scored with allele assignments that were consistent across families. All genotypes were manually curated.

SNP genotyping: Pedigree members of the F_2 intercross family were genotyped with a commercially available Infinium CanineSNP20 BeadChip according to the manufacturer's instructions. The canine BeadChip assays 22,362 unique SNPs distributed across the genome with ∼110 kb spacing. Data were collected using an Illumina BeadStation scanner and dedicated data collection software. Genotypes were generated with BeadStudio.

Fluorescence *in situ* hybridization: BAC clones corresponding to genomic regions were obtained from CHORI. BAC-derived DNA was labeled with Spectrum Green or Spectrum

Orange using nick translation. Paired sets of chromosome-specific probes (300 ng each) were differentially labeled and cohybridized to canine metaphase spreads prepared from cultured canine lymphocytes. Fluorescence *in situ* hybridization (FISH) preparations were washed (0.4× SSC, 0.3% NP-40) and counterstained with DAPI. Slides were analyzed and images were captured using a Genus Cytogenetics digital microscopy station.

Data integrity: Microsatellite and SNP data were inspected for Mendelian inheritance using PedCheck (O'CONNELL and WEEKS 1998). When a single non-Mendelian inheritance was detected for a marker within a family, all potentially errant genotypes among the siblings or parent (*i.e.*, PedCheck output) were removed from the data set. When multiple genotypes for a locus were inconsistent with inheritance in a family, all of the genotypes for the marker for that family were removed. If the number of errant genotypes detected by inheritance exceeded 10% for a given locus across families, the locus was discarded.

Map assembly algorithm: Markers were positioned according to their respective sequence coordinates in the *canFam2* reference genome (www.ncbi.nim.nih.gov). Correct chromosome assignment was assessed by testing all pairwise marker combinations for linkage using the *twopoint* option of CRIMAP (GREEN *et al.* 1990). Markers showing stronger linkage to at least two loci on a different chromosome were reassigned to that chromosome. The *all* option of CRIMAP was used to scrutinize the order of markers along each chromosome. Every marker was iteratively removed and reassigned solely on the basis of genetic data. Markers that could be placed elsewhere on the chromosome with equivalent statistical support (within one LOD) were deemed either error prone, poorly informative, or incorrectly positioned on the sequence build. These markers were not included in the framework map. The *flips* option of CRIMAP was used to further test local marker order. Maximum likelihoods were calculated for each permuted order in a three-marker sliding window along the chromosome. The order inferred from physical coordinates was maintained only if it was within one LOD unit of the highest likelihood obtained.

The effects of positive interference render tight double recombinants biologically unlikely. Double crossover events (DCO), which usually involve a single marker out of phase with the surrounding markers, are most likely indicative of genotyping errors, mutations, gene conversions, sequence misassemblies, or segregating chromosome polymorphisms. The *chrompic* option of CRIMAP was used to identify multiple crossovers over short genetic distances. The output was graphed with a custom tool that visually emphasized interruptions in the parental origin of phased chromosomes (*kodachrompic*; A. K. WONG, unpublished data). A double crossover event isolated to a single offspring was interpreted as a genotyping error of that individual. DCOs shared among siblings suggested a genotype error in a parent. The distribution of alleles in the family was scrutinized to identify the most probable errant genotype, which was removed. Terminal markers, which could not be evaluated in the same way, were scored twice by independent readers to improve the reliability of the genotype data.

Map expansion caused by undetected errors was estimated by iteratively (1) removing a marker, (2) rebuilding the local map, (3) recalculating intermarker recombination fractions, and (4) comparing the original and new recombination fractions. Markers that led to map inflation (>2 cM) were excluded. Sex-averaged, male-specific, and female-specific autosomal genetic distances, as well as the female-specific distances of the *X* chromosome, were calculated with the *fixed* option of CRIMAP using the Kosambi map function.

Quantifying interindividual variation: The number of recombination events in parental meioses was inferred from offspring genotypes using the chrompic option in CRIMAP. Recombination events localized to chromosome termini (first and last three markers) were discounted; distal exchanges are challenging to distinguish from genotyping errors. The mean recombination rate in each mother (or father) was estimated as the average number of recombination events in the haploid gametes transmitted to her (or his) progeny. One-way ANOVA was applied to test for significant differences among mothers and fathers. The recombination count in a given meiotic product was used as the response variable, and parental identity was treated as the factor. The significance of resultant *F*-statistics was empirically determined using the quantile position of the realized statistic along the distribution of 10^4 *F*-values derived by permutation.

Sequence correlates: Markers on the framework map were binned into nonoverlapping 5- and 10-Mb windows across the genome. For each window harboring a sufficient number of markers ($n \geq 2$), the regional recombination rate was estimated by the slope of a simple linear regression of genetic map position in centimorgan units on physical genome position in megabases. Imposing a harsher criterion of $n \geq 5$ makers per window had no significant impact on our findings, but did vastly diminish the number of regions that could be included in the analysis, particularly at the 5-Mb scale. Results reported in the main text derive from the more lax criterion.

The number of repetitive elements (LINES, SINES, LTRs, poly-A repeats, poly-C repeats, poly-G repeats, poly-T repeats, simple repeats, AT-rich, CA-rich, satellites, DNA repeats, and low-complexity repeats), base composition measures (GC%, number of CpG dinucleotides, number of CpG islands), incidence of the CCNCCNTNNCCNC putatively recombinogenic motif (MYERS *et al.* 2008), the number of genes (from RefSeq and the N-SCAN gene prediction track implemented in the UCSC Genome Browser), and the distance to the telomere as a fraction of total chromosome length were computed for each window using the most recent genome assembly (*canFam2*). Spearman rank correlation tests between recombination rate and individual sequence variables were performed in the *R* environment for statistical computing.

RESULTS

Assembling markers and meioses: Perfect CA-repeat microsatellites were computationally selected from the reference sequence (LINDBLAD-TOH *et al.* 2005) to achieve submegabase and possibly subcentimorgan spacing on a linkage map. Because markers were newly developed and untested, the collection ($n = 3349$ loci) was preliminarily screened against a panel of 36 purebred dogs (Table S1). Results indicated that the markers were amenable to high-throughput genotyping and were sufficiently informative for map assembly (HET = 0.46 ± 0.16; Table S3 and File S1).

The markers were typed on familial resources comprising 450 mostly phase-known meioses (Table 1). The families represented an efficient meiotic mapping resource, with 281 dogs from three pedigrees totaling 24 sibships and an average of eight offspring per cross. Genotyping was performed in six phases, with each phase including markers to span the genome in ~4-Mb increments. This approach allowed a more frequent quality control of genotyping, as well as a new map assembly

TABLE 1

Meiotic mapping resources used in this study

Familial resource	No. of dogs	Max. meioses	Avg. parental HET	Avg. informative meioses
CF reference[a]	213	333	0.54 ± 0.08	149 ± 75
F$_2$ intercross[b]	28	52	0.62 ± 0.01	26 ± 18
Silken Windhound	40	65	0.50 ± 0.07	27 ± 19
Total resources	281	450	0.61 ± 0.09	201 ± 91

[a] CF, Cornell families, described in Neff *et al.* (1999).
[b] Parental types were Border Collie (paternal) and Newfoundland (maternal).

following each phase. Intermediate map builds were posted online during the course of the study (http://www.vgl.ucdavis.edu/dogmap/). Completion of the phases yielded a million genotypes for assembly of a final map. The meiotic contributions of families are summarized in Table 1. Summary statistics are given in Table S3.

Clarifying marker order: For map assembly, markers were ordered along chromosomes according to their physical coordinates in *canFam2*. Chromosome assignments and relative order were tested against the genetic data. Several loci were discrepant ($n = 17$), showing significantly stronger linkage to two or more loci on a different chromosome than the one to which they had been assigned in *canFam2*. Genetic data were given preference and discrepant markers were reassigned (Table S4). The vast majority of marker positions were concordant for both types of positional information (99.4%), indicating the *canFam2* build was of high quality. This implied there was considerable value in using the physical scaffold to guide and clarify marker order. It should be noted that the power to detect macroassembly errors was presumably greater than the power to discern microassembly errors. Fine-scale errors in the sequence build may have been undetected in our analyses.

Estimating genetic distances: Given an established marker order, genetic distances for intermarker intervals were estimated for male-specific, female-specific, and sex-averaged maps. Sex-averaged distances were tested for map inflation, which can result from cryptic errors producing artifactual crossovers. Each locus was iteratively removed from the map, and relevant intermarker distances were recalculated. Approximately 1% of markers ($n = 46$) were suspected of having undetected genotyping errors, as evidenced by a significant reduction in map length (>2 cM) upon their iterative removal. These potentially errant loci were excluded from further assembly.

Autosomal maps: Of autosomal markers genotyped ($n = 3549$), 1469 qualified as anchor loci on the basis of an ability to "sample" a large number of meioses (*i.e.*, >100 informative meioses) with concomitant statistical support for linkage (*i.e.*, pairwise LOD > 11). These markers formed the basis of a framework map. An additional 1606

markers were considered "map validators" (*i.e.*, a pairwise LOD > 3); these loci contributed modestly to estimates of recombination rate, but supported the mapping content of anchor loci (*e.g.*, by substantiating phase transitions of observed crossovers). Table 2 lists summary statistics for the framework and comprehensive maps.

A SNP-based map: In addition to microsatellite markers, ~22,000 SNPs distributed across the genome (every 104 kb ± 25 kb) were genotyped on the F$_2$ intercross family (Table 1). The quality of the SNP data was high, as reflected by call rate (99.8%), inheritance (0.09% non-Mendelian errors), and informativeness (15,375 polymorphic loci; average MAF ≥ 0.20). The resulting SNP-based map detected most of the same crossovers observed in the microsatellite-based map (718/836). Undetected crossovers were mostly attributable to uncertain phase of diallelic SNPs. The SNP-based map exhibited 64 crossovers at the ends of autosomes that had not been captured by microsatellite markers. This suggested the sex-averaged length of the canine autosomal genome might be somewhat greater than 21 M. Overall, the two maps were in strong agreement, consistent with the microsatellite-based map being of high integrity.

An X map and the pseudoautosomal region: Markers were analyzed separately for the X chromosome to accommodate female-specific recombination. Of 140 markers analyzed, 82 served as anchor loci and 55 served as validating loci. Markers were ordered according to physical coordinates, and genetic distances were com-

TABLE 2

Summary statistics of linkage maps

	Framework	Combined
No. of markers	1496	3075
Genetic length	1937 cM	2085 cM
Genetic spacing	1.3 ± 1.7 cM	0.7 ± 0.9 cM
Physical spacing	1.4 ± 1.0 Mb	0.7 ± 0.4 Mb
Gaps < 5 cM	94%	98%
Recombination rate	0.92 cM/Mb	0.97 cM/Mb
Genetic female/male	1.28	1.19
Avg. informative meioses	266 ± 57	201 ± 91

Not inclusive of the X chromosome.

TABLE 3

Physical and genetic lengths of dog chromosomes

CFA	Physical (Mb)[a]	Sex avg. (cM)	Male (cM)	Female (cM)	Avg. cM/Mb	No. of markers
1	120.9	82.5	69.3	96.5	0.68	174
2	85.4	74.9	70.3	77.6	0.88	120
3	89.3	61.7	50.2	74.0	0.69	121
4	88.2	65.8	59.5	72.1	0.75	129
5	86.8	73.8	69.8	76.6	0.85	123
6	75.0	53.2	42.0	65.2	0.71	100
7	80.4	60.7	51.6	69.8	0.75	108
8	72.7	60.4	53.0	68.9	0.83	101
9	60.8	67.3	69.1	65.5	1.11	85
10	67.5	52.5	45.9	60.1	0.78	99
11	73.3	61.5	56.5	65.9	0.84	94
12	72.2	63.2	60.5	66.1	0.87	107
13	62.9	56.6	47.3	67.8	0.90	87
14	60.1	53.3	49.3	57.2	0.89	83
15	63.4	50.5	53.0	47.8	0.80	92
16	55.4	49.9	43.2	59.4	0.90	80
17	64.0	57.0	62.7	55.0	0.89	87
18	52.5	46.1	34.5	58.0	0.88	80
19	53.7	50.0	49.5	52.1	0.93	76
20	57.2	53.4	50.5	57.6	0.93	83
21	49.5	50.7	45.8	54.4	1.02	74
22	59.4	49.7	47.3	53.5	0.84	91
23	49.5	50.7	47.9	54.1	1.02	71
24	47.4	53.6	45.2	63.8	1.13	80
25	50.5	53.2	45.9	60.6	1.05	70
26	37.7	46.2	48.9	43.2	1.22	57
27	44.3	52.1	42.8	62.5	1.18	62
28	39.1	51.3	45.1	57.3	1.31	58
29	41.3	47.5	49.4	46.1	1.15	58
30	37.7	51.3	38.6	61.8	1.36	53
31	35.1	45.4	38.1	51.9	1.29	51
32	37.8	51.9	54.3	49.8	1.37	52
33	31.4	48.0	49.5	46.8	1.53	47
34	37.9	44.6	33.8	54.2	1.17	53
35	26.3	56.7	49.5	63.5	2.15	34
36	30.0	46.1	49.2	43.0	1.54	51
37	30.6	42.4	42.6	44.3	1.38	42
38	22.3	49.4	48.4	52.6	2.22	42
X	124.9	NA	NA	111.1	0.89	137
Total	2274.4	2085.1	1910.0	2387.7	NA	3212

NA, not applicable.

[a] From the first and last marker coordinates in *canFam2*.

puted from informative maternal meioses. The results suggested a genetic length for the *X* chromosome of 111 cM (Table 3 and Figure S1).

The vast majority of *X*-linked loci showed a single allele in males, as expected for hemizygosity. However, five markers exhibited significant heterozygosity in males [average (avg.) HET = 0.80]. None of these markers exhibited fixed alleles, arguing against duplicated regions. The five loci clustered physically (Figure S1), and their allelic variation segregated with autosomal inheritance, consistent with their belonging to the canine pseudoautosomal region (PAR). The genotype

data for 565 SNPs along the *X* chromosome were checked for a similar pattern. Fifty SNPs showed heterozygosity in males (avg. HET = 0.56), and these were from the same physical region as the five microsatellites (Figure S1). This suggested the PAR on the metacentric *X* localized to the telomeric end (Figure S1). This corresponded to a female-specific length of 7 cM and a male-specific length of 28 cM. Reference sequence data were available from female dogs only (KIRKNESS *et al.* 2003; LINDBLAD-TOH *et al.* 2005), precluding a characterization of the physical arrangement of the PAR with *Y* synteny.
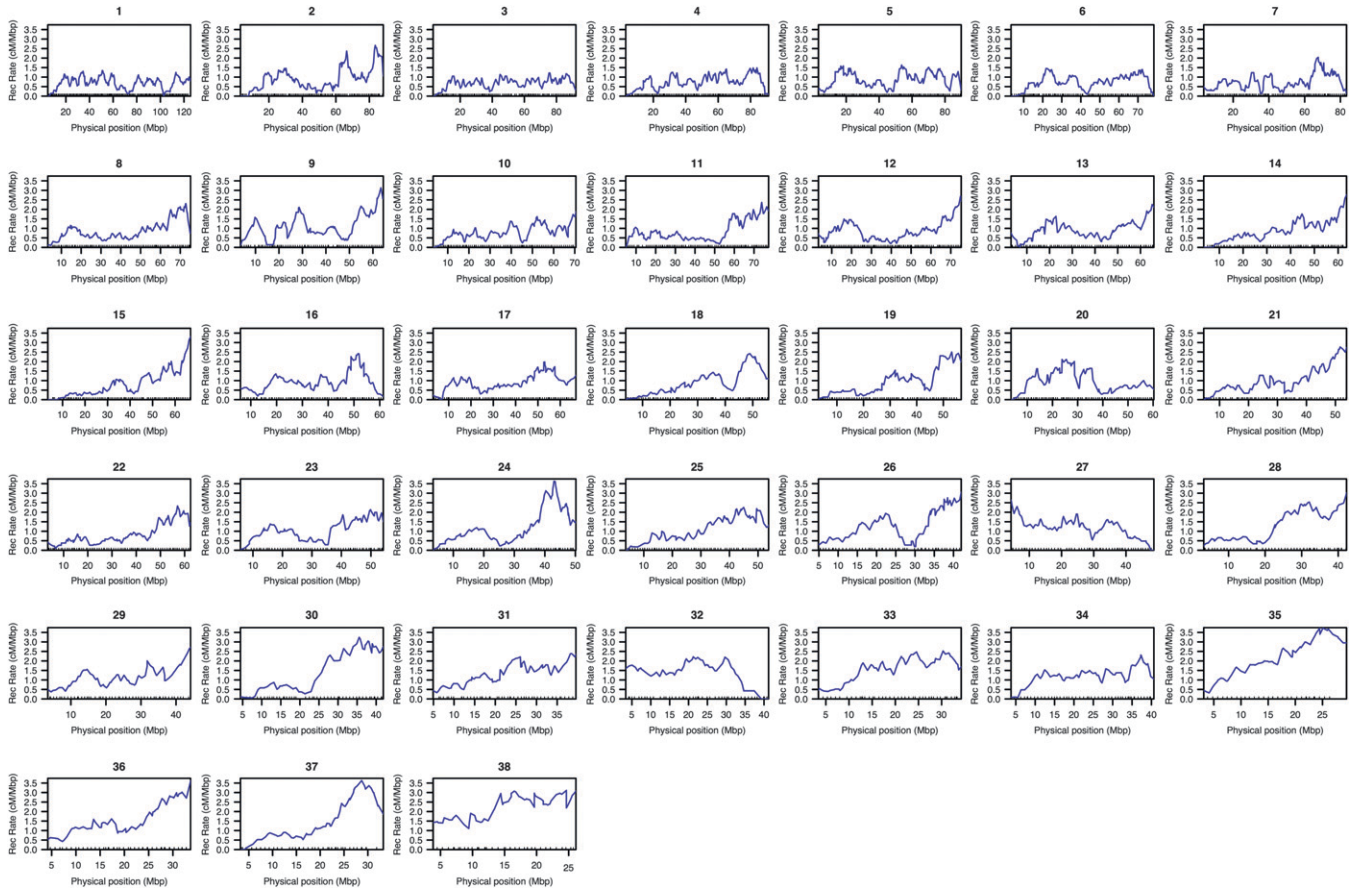
FIGURE 1.—Variation in recombination rate across the autosomal genome. Recombination rate along each of the 38 autosomes is shown from centromeric to telomeric end (left to right). *x*-axis, physical length, scaled to accommodate different chromosome sizes. *y*-axis, sex-averaged recombination rate. Sliding windows of 5 Mb were used to calculate cM/Mb along the chromosome. The pattern of recombination rate for two autosomes (*CFA27* and *CFA32*) suggested their orientation may be incorrectly specified in the *canFam2* build.

Observed variation in genomewide recombination: Although the map showed pronounced regional differences in recombination rates, several systematic patterns were evident. Nearly every autosome was characterized by low recombination rate at the centromeric end and high rate at the telomeric end (Figure 1). Two chromosomes (*CFA27* and *CFA32*) exhibited a reversed pattern (*i.e.*, high rates at the centromere and low rates at the telomere; Figure 1). Given the consistency of the recombinational profiles across the other 36 autosomes and the generality of this pattern in mammals (KONG *et al.* 2002; SHIFMAN *et al.* 2006), the simplest interpretation was that the orientation of the chromosomes was misspecified in *canFam2*.

We tested this prediction directly using FISH analysis. Differentially labeled probes were generated from BAC clones, which were anchored to the ends of each autosomal linkage map by markers. The probes were localized by FISH with canine metaphase spreads. The results strongly supported the chromosome orientation inferred from the linkage map rather than the one listed in *canFam2* (Figure S3; Table S5).

Recombination rate was inversely associated with the physical length of chromosomes (Figure 2). The small-

est chromosome (*CFA38*) exhibited a threefold greater rate than the largest chromosome (*CFA1*) (Table 3). These results were generally consistent with the expectation of at least one crossover per chromosome, a nearly universal requirement of meiosis from yeast to man (KABACK *et al.* 1989; KONG *et al.* 2002).

The rate of recombination also systematically varied by sex. Female meioses exhibited a 1.2-fold greater average rate than male meioses (Table 2). The influence of sex was not uniformly distributed across the genome (Figure S2), with differences in sex ratio being most striking at autosome ends. Female meioses exhibited a 4-fold greater rate near centromeres, whereas male meioses exhibited a 4-fold greater rate near telomeres. Similar observations of sex differences have been made in human (BROMAN *et al.* 1998) and mouse (PAIGEN *et al.* 2008).

Interindividual variation in recombination rate was also observed (Figure S4). The female with the highest recombination rate exhibited ∼20% more crossovers per meiosis than the female with the lowest rate, although these differences could be chance variation (one-way ANOVA: $F = 1.35$; $P = 0.16$). There was substantially
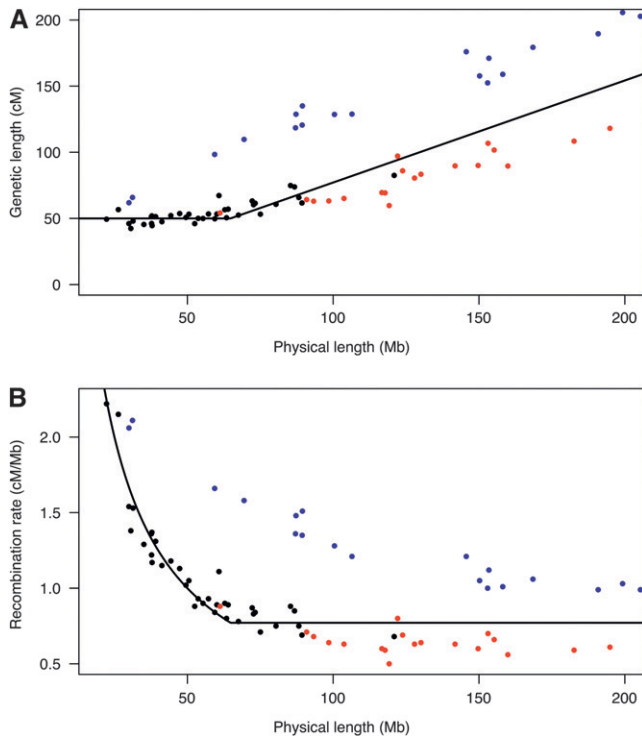
FIGURE 2.—Correlation of recombination with physical chromosome size. (A) Genetic length of autosomes plotted against physical length. (B) Mean recombination rate across autosomes plotted against physical length. Results are shown for canine (black), human (blue), and mouse (red) autosomes. Data describing physical sizes and sex-averaged recombination rates were obtained from KONG *et al.* (2002) and SHIFMAN *et al.* (2006) for human and mouse, respectively. Physical chromosome sizes in the dog were calculated from the *canFam2* coordinates of the first and last markers for each autosome. The average recombination rate for canine chromosomes was calculated from the autosomal linkage maps.

more variation in genomewide rates (nearly twofold) among male dogs, and these differences were statistically significant ($F = 1.77$; $P = 0.042$). This finding complements observations of individual differences in recombination in a variety of species, including human (BROMAN *et al.* 1998; KONG *et al.* 2004), Drosophila (CHINNICI 1971; KIDWELL 1972; BROOKS and MARKS 1986), Tribolium (DEWEES 1975), and laboratory strains of mice (REEVES *et al.* 1990; KOEHLER *et al.* 2002; PAIGEN *et al.* 2008). Interindividual variation may therefore be a pervasive characteristic of meiotic recombination in sexually reproducing species.

Sequence correlates of recombination rate: A sequence-explicit framework motivated investigation of the relationship between sequence variables and regional heterogeneity in recombination rate. Twenty sequence features were tested for rank correlation with recombination rate (Table 4). Nonoverlapping 5- and 10-Mb windows were tested across the genome. The strongest predictor of sex-averaged recombination rates for both window sizes was proximity to the telomere. In addition, several classes of repetitive elements, the

number of CpG dinucleotide sites, and the incidence of a recombinogenic sequence motif (MYERS *et al.* 2005, 2008) were also predictive of local recombination rates, similar to what has been observed in other mammals (KONG *et al.* 2002; JENSEN-SEAMAN *et al.* 2004; MYERS *et al.* 2005; SHIFMAN *et al.* 2006).

We examined these sequence correlates in the context of sex-specific influences. This analysis revealed distinct differences in males and females. Several GC-related measures (GC-rich repeats, CpG dinucleotide sites, and CpG islands) were positively correlated with male recombination rates, but weakly or negatively correlated with female recombination rates (Table 4). These findings were intriguing in light of the fact that males disproportionately drive the evolution of GC content in the human genome (DRESZER *et al.* 2007; DURET and ARNDT 2008), which may stem from a male-intensified, biased gene conversion toward G/C alleles during recombinational repair (MARAIS 2003). The putative recombinogenic motif (CCNCCNTNNCCNC) was also positively correlated with male recombination rates, but uncorrelated with female rates. This sex specificity in the dog was similar to that observed in the house mouse (SHIFMAN *et al.* 2006), but different from the motif's recombinogenic properties in both human sexes (MYERS *et al.* 2008).

Sex-dependent correlations were also found for several classes of repetitive elements. AT-rich repeats, satellites, low-complexity repeats, poly(T) repeats, poly(A) repeats, and long terminal repeats were positively correlated with female recombination rates, but noncorrelated with male rates. Together, these findings suggest that regional recombination is mediated differently in the two sexes by features of the sequence landscape.

Interpolated maps for SNPs and scanning set loci: The resolution and integrity of the canine map afforded an opportunity to place additional markers on the map through linear interpolation (KONG *et al.* 2002). Positions for 2.4 million publicly available SNPs were interpolated against the sex-averaged map. Results have been made available here (File S2) and are also available upon request as a custom track for a publicly available browser (www.genome.ucsc.edu). Microsatellite markers from the latest minimal scanning set (MMS3, 507 markers; SARGAN *et al.* 2007) were similarly positioned by interpolation to afford multipoint analyses (File S3). An electronic file with inferred intermarker genetic distance recombination fractions has also been made available (File S4).

DISCUSSION

We have addressed a principal unmet need in our field by assembling a linkage map that enables navigating the dog genome in a structured and systematic way. The map we have created (i) informs on the basic biology of meiosis and the evolution of recombination rate, (ii) provides important clues on the genomic de-

TABLE 4

Sequence correlates of genomewide variation in recombination rate

Sequence feature	5 Mb			10 Mb		
	Female	Male	Sex avg.	Female	Male	Sex avg.
GC%	−0.20**	0.10	−0.04	−0.23*	−0.01	−0.14
No. CpG sites	−0.05	0.32***	0.16*	−0.08	0.30**	0.14
No. CpG islands	−0.19**	0.25***	0.09	−0.22*	0.24*	0.07
Distance to telomere	−0.23***	−0.46***	−0.47***	−0.17*	−0.60***	−0.56***
Motif count	−0.02	0.29***	0.15*	−0.01	0.30**	0.16
No. RefSeq genes	−0.04	0.16*	0.05	0.11	0.17*	0.10
No. NScan genes	−0.06	0.22**	0.08	0.05	0.12	0.02
SINES	0.19**	0.14*	0.12*	0.33***	0.08	0.09
LINES	−0.08	−0.05	−0.08	0.31***	−0.04	−0.01
LTRs	0.28***	0.01	0.11	0.48***	−0.01	0.12
Poly(A)	0.30***	0.15*	0.23***	0.49***	0.12	0.24*
Poly(C)	−0.02	0.15*	0.06	0.09	0.06	0.02
Poly(G)	−0.04	0.17*	0.09	0.07	0.07	0.02
Poly(T)	0.28***	0.20**	0.24***	0.47***	0.12	0.23*
Simple repeats	0.32***	0.06	0.18*	0.54***	0.06	0.20*
AT-rich repeats	0.24***	−0.07	0.04	0.41***	−0.11	0.05
GC-rich repeats	−0.08	0.25***	0.12	−0.04	0.16	0.04
Satellites	−0.16*	−0.02	−0.09	−0.17*	−0.08	−0.15
Low complexity repeats	0.24***	−0.04	0.06	0.45***	−0.07	0.09
DNA repeats	0.30***	0.20**	0.23***	0.43***	0.09	0.16

Rank correlation coefficients were calculated in 5-Mb and 10-Mb sliding windows. *$P < 0.01$; **$P < 10^{-4}$; ***$P < 10^{-6}$.

terminants of regional heterogeneity and sex specificity of recombination rates, (iii) facilitates unbiased genetic access to natural variation and phenotypic diversity in canids, and (iv) lays out an integrated wet bench and bioinformatics approach for developing a *de novo* map for any species for which a draft sequence becomes available.

Basic biology of meiosis: Although classical genetics in model organisms has dissected the mechanisms governing meiosis, observations from linkage maps have uniquely informed on natural variation in underlying processes (DUMONT and PAYSEUR 2008; reviewed in COOP and PRZEWORSKI 2007). The canine genome represents a natural experiment in meiosis, with 38 pairs of unusually short, acrocentric autosomes. Despite the unique challenges such a karyotype might present to the meiotic apparatus, patterns of recombination in the genetic map suggest this karyotype has been accommodated by conventional means. An upregulation of recombination rate among physically smaller chromosomes, for instance, is a conserved feature of meiosis from yeast to man (KABACK *et al.* 1989; KONG *et al.* 2002). This nonrandom distribution ensures the obligatory crossover per chromosome arm that helps homologs attain bipolar orientation during meiosis I.

Despite the unique karyotypic features of the dog, the sex-averaged rate of recombination across the genome (0.97 cM/Mb) is within the range of other characterized mammals (0.5–1.1 cM/Mb). Genetic map-based estimates of recombination rate are available for two additional carnivore species, house cat (~1.1 cM/Mb;

MENOTTI-RAYMOND *et al.* 1999) and silver fox (~0.6 cM/Mb; KUKEKOVA *et al.* 2007). While large differences in map quality and coverage preclude detailed comparisons, the rate of recombination in dog is clearly not an outlier among carnivores. Although fine-scale recombination rates evolve on short evolutionary timescales (PTAK *et al.* 2005; WINCKLER *et al.* 2005), the dog genetic map adds evidence for more rigid evolutionary constraints on broader scale recombination rates (MYERS *et al.* 2005; DUMONT and PAYSEUR 2008).

The elevated recombination rate in female meioses, a salient feature of eutherian maps, was also evident in the dog, although the sex ratio was not as great as it is in man (1.2-fold in dog *vs.* 1.7-fold in humans) (BROMAN *et al.* 1998; KONG *et al.* 2002). In part, this may be attributable to karyotype—mouse and rat also have mostly acrocentric autosomes and similarly exhibit a muted sex difference relative to human (SHIFMAN *et al.* 2006). Karyotypic organization accounts for only some of the differences between sexes (HUNT and HASSOLD 2002). For reasons not yet clear, sex differences are modest in cow (IHARA *et al.* 2004) and sheep (MADDOX *et al.* 2001), for instance, and are reversed in the two metatherian mammals (marsupials) studied to date (ZENGER *et al.* 2002; SAMOLLOW *et al.* 2004).

As in other placental mammals, sex differences in the dog were not uniformly distributed across the genome. Female meioses exhibited a greater recombination rate near centromeres whereas male meioses showed a greater rate near telomeres. In general, it appears female meioses are at greater risk for nondisjunction in

man (ANTONARAKIS 1991), and this risk is exacerbated among chromosomes with distal crossover events (LAMB *et al.* 1996). Distal crossovers decrease fidelity in yeast as well, suggesting crossover placement fundamentally influences chromosome orientation on the meiotic spindle. Interestingly, segregation in yeast can be rescued with an experimental tether, but only if the tether is located at the centromere (LACEFIELD and MURRAY 2007). This implies that physical ties near the centromere (*e.g.*, crossovers) intrinsically promote bipolar orientation. If so, the greater recombination rate near the centromere in females could be compensatory—an adaptation that offsets the greater sensitivity to nondisjunction in oogenesis.

Evolution of patterned recombination: The canine linkage map may offer an unprecedented opportunity to address the evolution of recombination rate. By generating favorable allelic combinations and breaking down negative linkage disequilibrium, recombination can facilitate the response to selection (FISHER 1930; FELSENSTEIN 1965). This has led to the hypothesis that species under frequent and intense selection evolve toward increased recombination rates (BURT and BELL 1987). Recent work has shown that the purebred domesticated chicken exhibits higher recombination rates than Red Jungle fowl, its wild ancestral counterpart (GROENEN *et al.* 2009). In our case, describing the effects of domestication requires a genetic map assembled for the wolf, *Canis lupus.* There are logistical constraints to this, but a broad assessment of recombination patterns could be inferred from an analysis of LD in wolf populations, as has been done in man (MCVEAN *et al.* 2004).

The dog also presents an opportunity to address microevolutionary changes in recombination rate. Mouse strains show systematic differences in recombination (SHIROISHI *et al.* 1991; PAIGEN *et al.* 2008), and interindividual variation, from Drosophila to man, has been tied to causative genes (CHINNICI 1971; KIDWELL 1972; BROOKS and MARKS 1986; SHIROISHI *et al.* 1991; KOEHLER *et al.* 2002; COOP *et al.* 2008; KONG *et al.* 2008). Consistent with these observations, we documented significant variation among male dogs. If this variation is heritable, we might also expect differences in recombination rate among breeds. However, our use of admixed pedigrees for map assembly complicated the detection of such differences. The relatively short timescale separating dog breeds and the artificial selection that has shaped breed phenotypes should motivate the construction of breed-specific maps to understand the effects of recent isolation and strong selection on the evolution of recombination rate.

Genomic control of recombination rate: Sequence correlates provide mechanistic insights into recombination rate variation. Several sequence correlates found in human (YU *et al.* 2001; KONG *et al.* 2002), mouse (JENSEN-SEAMAN *et al.* 2004), and rat (JENSEN-SEAMAN *et al.* 2004) were also found in the dog. The number of CpG dinucleotides, for instance, is positively correlated with recombination rate in each species (KONG *et al.* 2002; JENSEN-SEAMAN *et al.* 2004), including the dog. Cross-species comparisons have also revealed differences. Although recombination is correlated with LINES in human, mouse, and rat (KONG *et al.* 2002; JENSEN-SEAMAN *et al.* 2004), no such correlation was observed in the dog. Distance from the centromere is the best predictor of recombination in man and dog, whereas sequence features remain the best predictors in mouse and rat (JENSEN-SEAMAN *et al.* 2004).

Many sequence-based correlations were significant, but most were relatively weak. In this respect, the dog joins a growing list of mammals for which the majority of genomic variation cannot be explained by sequence characteristics (KONG *et al.* 2002; JENSEN-SEAMAN *et al.* 2004; MYERS *et al.* 2005; SHIFMAN *et al.* 2006). Alternative sources of variation, for which there is mounting evidence, are epigenetics and chromatin state (WINCKLER *et al.* 2005; NEUMANN and JEFFREYS 2006; SANDOVICI *et al.* 2006; SIGURDSSON *et al.* 2009). CpG islands, the principal targets of methylation, are correlated with recombination rate in human, mouse, rat, and dog. Interestingly, CpG association in the dog differed in sign between the sexes, suggesting opposing processes in male and female meioses.

The canine genetic map provides important clues to the origins of sex differences in recombination. This dimorphism has long been a puzzle (MORGAN 1912; DUNN 1920; HALDANE 1922). The most striking result to date is a sequence variant in humans that increases recombination rates in males, but decreases recombination rates in females (KONG *et al.* 2008). Thus, causative genes and *cis*-sequence motifs may combine to account for sex-specific differences in recombination rate, possibly mediated by chromatin state (PETKOV *et al.* 2007).

Application to mapping phenotypes: The *de novo* map makes known the full genetic landscape of the dog genome and thus allows assessing empirical coverage provided by linkage-scanning sets. The genetic positions of markers in commonly used sets were interpolated so that researchers could derive these benefits. Known positions and intermarker distances are also essential for multipoint analyses to maximize the power and resolution of fine mapping and quantitative trait loci (QTL) mapping. The genetic map may also inform on the pattern and extent of LD in the purebred dog. LD is 20- to 100-fold more extensive in breed isolates than in human populations (SUTTER *et al.* 2004; LINDBLAD-TOH *et al.* 2005), a consequence of introgression and historical bottlenecks. Recombination is one of several forces that shape LD; knowledge of local recombination rates may therefore be useful for discerning the contributions of other factors (*e.g.*, selection).

Map-building extensibility: An available reference sequence presented us with an uncommon opportunity to assemble a map efficiently. The genome sequence

afforded an abundance of putative markers that could be computationally mined from a physical framework and selected for even spacing and complete coverage. These markers performed well, and their standardized design facilitated wet bench genotyping. Systematic marker ascertainment affords other opportunities, such as studying mutation rates in the context of the "slippery genome" hypothesis of dog domestication (FONDON and GARNER 2004). Positional information for markers was valuable in guiding and clarifying locus order. This resulted in a map that was more inclusive of typed markers. These strategies for building a map *de novo* are applicable to any natural species with a draft sequence. This exemplifies how genomics is continuing to enhance access to natural variation in a broader array of nonmodel organisms.

Online maps and ancillary data: We have provided map figures (Figure S5 and Figure S6) and electronic files with marker and map content. We have also posted detailed map builds online with hyperlinks to ancillary data. The genetic maps and associated tabular data are available at: http://www.vgl.ucdavis.edu/dogmap/.

LITERATURE CITED

ANDERSSON, L., 2001 Genetic dissection of phenotypic diversity in farm animals. Nat. Rev. Genet. **2:** 130–138.

ANTONARAKIS, S. E., 1991 Parental origin of the extra chromosome in trisomy 21 as indicated by analysis of DNA polymorphisms. Down Syndrome Collaborative Group. N. Engl. J. Med. **324:** 872–876.

BELL, G. I., 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.

BREEN, M., S. JOUQUAND, C. RENIER, C. S. MELLERSH, C. HITTE *et al.*, 2001 Chromosome-specific single-locus FISH probes allow anchorage of an 1800-marker integrated radiation-hybrid/linkage map of the domestic dog genome to all chromosomes. Genome Res. **11:** 1784–1795.

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.

BROOKS, L. D., and R. W. MARKS, 1986 The organization of genetic variation for recombination in *Drosophila melanogaster*. Genetics **114:** 525–547.

BROWNSTEIN, M. J., J. D. CARPTEN and J. R. SMITH, 1996 Modulation of non-templated nucleotide addition by Taq DNA polymerase: primer modifications that facilitate genotyping. Biotechniques **20:** 1004–1006, 1008–1010.

BURT, A., and G. BELL, 1987 Mammalian chiasma frequencies as a test of two theories of recombination. Nature **326:** 803–805.

CHINNICI, J. P., 1971 Modification of recombination frequency in Drosophila. I. Selection for increased and decreased crossing over. Genetics **69:** 71–83.

COOP, G., and M. PRZEWORSKI, 2007 An evolutionary view of human recombination. Nat. Rev. Genet. **8:** 23–34.

COOP, G., X. WEN, C. OBER, J. K. PRITCHARD and M. PRZEWORSKI, 2008 High-resolution mapping of crossovers reveals extensive variation in fine-scale recombination patterns among humans. Science **319:** 1395–1398.

DEWEES, A. A., 1975 Genetic modification of recombination rate in *Tribolium castaneum*. Genetics **81:** 537–552.

DOLF, G., 1999 DogMap: an international collaboration toward a low-resolution canine genetic marker map. J. Hered. **90:** 3.

DRESZER, T. R., G. D. WALL, D. HAUSSLER and K. S. POLLARD, 2007 Biased clustered substitutions in the human genome: the footprints of male-driven biased gene conversion. Genome Res. **17:** 1420–1430.

DUKES-MCEWAN, J., and I. J. JACKSON, 2002 The promises and problems of linkage analysis by using the current canine genome map. Mamm. Genome **13:** 667–672.

DUMONT, B. L., and B. A. PAYSEUR, 2008 Evolution of the genomic rate of recombination in mammals. Evolution **62:** 276–294.

DUNN, L. C., 1920 Linkage in mice and rats. Genetics **5:** 325–343.

DURET, L., and P. F. ARNDT, 2008 The impact of recombination on nucleotide substitutions in the human genome. PLoS Genet. **4:** e1000071.

FELSENSTEIN, J., 1965 The effect of linkage on directional selection. Genetics **52:** 349–363.

FISHER, R. A., 1930 *The Genetical Theory of Natural Selection*. Oxford University Press, Oxford.

FONDON, J. W., and H. R. GARNER, 2004 Molecular origins of rapid and continuous morphological evolution. Proc. Natl. Acad. Sci. USA **101:** 18058–18063.

GREEN, P., K. FALLS and S. CROOKS, 1990 Documentation for CRI-MAP, version 2.4. Washington University School of Medicine, St. Louis.

GROENEN, M. A., P. WAHLBERG, M. FOGLIO, H. H. CHENG, H. J. MEGENS *et al.*, 2009 A high-density SNP-based linkage map of the chicken genome reveals sequence features correlated with recombination rate. Genome Res. **19:** 510–519.

GRUPE, A., S. GERMER, J. USUKA, D. AUD, J. K. BELKNAP *et al.*, 2001 In silico mapping of complex disease-related traits in mice. Science **292:** 1915–1918.

GUYON, R., T. D. LORENTZEN, C. HITTE, L. KIM, E. CADIEU *et al.*, 2003 A 1-Mb resolution radiation hybrid map of the canine genome. Proc. Natl. Acad. Sci. USA **100:** 5296–5301.

HALDANE, J., 1922 Sex ratio and unisexual sterility in hybrid animals. J. Genet. **12:** 101–109.

HUNT, P. A., and T. J. HASSOLD, 2002 Sex matters in meiosis. Science **296:** 2181–2183.

IHARA, N., A. TAKASUGA, K. MIZOSHITA, H. TAKEDA, M. SUGIMOTO *et al.*, 2004 A comprehensive genetic map of the cattle genome based on 3802 microsatellites. Genome Res. **14:** 1987–1998.

JENSEN-SEAMAN, M. I., T. S. FUREY, B. A. PAYSEUR, Y. LU, K. M. ROSKIN *et al.*, 2004 Comparative recombination rates in the rat, mouse, and human genomes. Genome Res. **14:** 528–538.

JONES, P., K. CHASE, A. MARTIN, P. DAVERN, E. A. OSTRANDER *et al.*, 2008 Single-nucleotide-polymorphism-based association mapping of dog stereotypes. Genetics **179:** 1033–1044.

KABACK, D. B., H. Y. STEENSMA and P. DE JONGE, 1989 Enhanced meiotic recombination on the smallest chromosome of Saccharomyces cerevisiae. Proc. Natl. Acad. Sci. USA **86:** 3694–3698.

KARLSSON, E. K., and K. LINDBLAD-TOH, 2008 Leader of the pack: gene mapping in dogs and other model organisms. Nat. Rev. Genet. **9:** 713–725.

KARLSSON, E. K., I. BARANOWSKA, C. M. WADE, N. H. SALMON HILLBERTZ, M. C. ZODY *et al.*, 2007 Efficient mapping of mendelian traits in dogs through genome-wide association. Nat. Genet. **39:** 1321–1328.

KIDWELL, M. G., 1972 Genetic change of recombination value in Drosophila melanogaster. I. Artificial selection for high and low recombination and some properties of recombination-modifying genes. Genetics **70:** 419–432.

KIRKNESS, E. F., V. BAFNA, A. L. HALPERN, S. LEVY, K. REMINGTON et al., 2003 The dog genome: survey sequencing and comparative analysis. Science **301:** 1898–1903.

KOEHLER, K. E., J. P. CHERRY, A. LYNN, P. A. HUNT and T. J. HASSOLD, 2002 Genetic control of mammalian meiotic recombination. I. Variation in exchange frequencies among males from inbred mouse strains. Genetics **162:** 297–306.

KONG, A., D. F. GUDBJARTSSON, J. SAINZ, G. M. JONSDOTTIR, S. A. GUDJONSSON et al., 2002 A high-resolution recombination map of the human genome. Nat. Genet. **31:** 241–247.

KONG, X., K. MURPHY, T. RAJ, C. HE, P. S. WHITE et al., 2004 A combined linkage-physical map of the human genome. Am. J. Hum. Genet. **75:** 1143–1148.

KONG, A., G. THORLEIFSSON, H. STEFANSSON, G. MASSON, A. HELGASON et al., 2008 Sequence variants in the RNF212 gene associate with genome-wide recombination rate. Science **319:** 1398–1401.

KUKEKOVA, A. V., L. N. TRUT, I. N. OSKINA, J. L. JOHNSON, S. V. TEMNYKH et al., 2007 A meiotic linkage map of the silver fox, aligned and compared to the canine genome. Genome Res. **17:** 387–399.

LACEFIELD, S., and A. W. MURRAY, 2007 The spindle checkpoint rescues the meiotic segregation of chromosomes whose crossovers are far from the centromere. Nat. Genet. **39:** 1273–1277.

LAMB, N. E., S. B. FREEMAN, A. SAVAGE-AUSTIN, D. PETTAY, L. TAFT et al., 1996 Susceptible chiasmate configurations of chromosome 21 predispose to non-disjunction in both maternal meiosis I and meiosis II. Nat. Genet. **14:** 400–405.

LANGSTON, A. A., C. S. MELLERSH, N. A. WIEGAND, G. M. ACLAND, K. RAY et al., 1999 Toward a framework linkage map of the canine genome. J. Hered. **90:** 7–14.

LINDBLAD-TOH, K., C. M. WADE, T. S. MIKKELSEN, E. K. KARLSSON, D. B. JAFFE et al., 2005 Genome sequence, comparative analysis and haplotype structure of the domestic dog. Nature **438:** 803–819.

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.

MADDOX, J. F., K. P. DAVIES, A. M. CRAWFORD, D. J. HULME, D. VAIMAN et al., 2001 An enhanced linkage map of the sheep genome comprising more than 1000 loci. Genome Res. **11:** 1275–1289.

MARAIS, G., 2003 Biased gene conversion: implications for genome and sex evolution. Trends Genet. **19:** 330–338.

MCVEAN, G. A., S. R. MYERS, S. HUNT, P. DELOUKAS, D. R. BENTLEY et al., 2004 The fine-scale structure of recombination rate variation in the human genome. Science **304:** 581–584.

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.

MENOTTI-RAYMOND, M., V. A. DAVID, L. A. LYONS, A. A. SCHÄFFER, J. F. TOMLIN et al., 1999 A genetic linkage map of microsatellites in the domestic cat (Felis catus). Genomics **57:** 9–23.

MORGAN, T. H., 1912 Complete linkage in the second chromosome of the male of Drosophila. Science **36:** 718–720.

MYERS, S., L. BOTTOLO, C. FREEMAN, G. MCVEAN and P. DONNELLY, 2005 A fine-scale map of recombination rates and hotspots across the human genome. Science **310:** 321–324.

MYERS, S., C. FREEMAN, A. AUTON, P. DONNELLY and G. MCVEAN, 2008 A common sequence motif associated with recombination hot spots and genome instability in humans. Nat. Genet. **40:** 1124–1129.

NEFF, M. W., and J. RINE, 2006 A fetching model organism. Cell **124:** 229–231.

NEFF, M. W., K. W. BROMAN, C. S. MELLERSH, K. RAY, G. M. ACLAND et al., 1999 A second-generation genetic linkage map of the domestic dog, Canis familiaris. Genetics **151:** 803–820.

NEUMANN, R., and A. J. JEFFREYS, 2006 Polymorphism in the activity of human crossover hotspots independent of local DNA sequence variation. Hum. Mol. Genet. **15:** 1401–1411.

O'CONNELL, J. R., and D. E. WEEKS, 1998 PedCheck: a program for identification of genotype incompatibilities in linkage analysis. Am. J. Hum. Genet. **63:** 259–266.

OBERBAUER, A. M., D. I. GROSSMAN, M. L. EGGLESTON, D. N. IRION, A. L. SCHAFFER et al., 2003 Alternatives to blood as a source of DNA for large-scale scanning studies of canine genome linkages. Vet. Res. Commun. **27:** 27–38.

OETTING, W. S., H. K. LEE, D. J. FLANDERS, G. L. WIESNER, T. A. SELLERS et al., 1995 Linkage analysis with multiplexed short tandem repeat polymorphisms using infrared fluorescence and M13 tailed primers. Genomics **30:** 450–458.

OSTRANDER, E. A., J. RINE, G. H. SACK and L. C. CORK, 1993 What is the role of molecular genetics in modern veterinary practice? J. Am. Vet. Med. Assoc. **203:** 1259–1262.

PAIGEN, K., J. P. SZATKIEWICZ, K. SAWYER, N. LEAHY, E. D. PARVANOV et al., 2008 The recombinational anatomy of a mouse chromosome. PLoS Genet. **4:** e1000119.

PETKOV, P. M., K. W. BROMAN, J. P. SZATKIEWICZ and K. PAIGEN, 2007 Crossover interference underlies sex differences in recombination rates. Trends Genet. **23:** 539–542.

POLLINGER, J. P., C. D. BUSTAMANTE, A. FLEDEL-ALON, S. SCHMUTZ, M. M. GRAY et al., 2005 Selective sweep mapping of genes with large phenotypic effects. Genome Res. **15:** 1809–1819.

PTAK, S. E., D. A. HINDS, K. KOEHLER, B. NICKEL, N. PATIL et al., 2005 Fine-scale recombination patterns differ between chimpanzees and humans. Nat. Genet. **37:** 429–434.

REEVES, R. H., M. R. CROWLEY, B. F. O'HARA and J. D. GEARHART, 1990 Sex, strain, and species differences affect recombination across an evolutionarily conserved segment of mouse chromosome 16. Genomics **8:** 141–148.

SAMOLLOW, P. B., C. M. KAMMERER, S. M. MAHANEY, J. L. SCHNEIDER, SCOTT J. WESTENBERGER et al., 2004 First-generation linkage map of the gray, short-tailed opossum, Monodelphis domestica, reveals genomewide reduction in female recombination rates. Genetics **166:** 307–329.

SANDOVICI, I., S. KASSOVSKA-BRATINOVA, J. E. VAUGHAN, R. STEWART, M. LEPPERT et al., 2006 Human imprinted chromosomal regions are historical hot-spots of recombination. PLoS Genet. **2:** e101.

SARGAN, D. R., J. AGUIRRE-HERNANDEZ, F. GALIBERT and E. A. OSTRANDER, 2007 An extended microsatellite set for linkage mapping in the domestic dog. J. Hered. **98:** 221–231.

SHIFMAN, S., J. T. BELL, R. R. COPLEY, M. S. TAYLOR, R. W. WILLIAMS et al., 2006 A high-resolution single nucleotide polymorphism genetic map of the mouse genome. PLoS Biol. **4:** e395.

SHIROISHI, T., T. SAGAI, N. HANZAWA, H. GOTOH and K. MORIWAKI, 1991 Genetic control of sex-dependent meiotic recombination in the major histocompatibility complex of the mouse. EMBO J. **10:** 681–686.

SIGURDSSON, M. I., A. V. SMITH, H. T. BJORNSSON and J. J. JONSSON, 2009 HapMap methylation-associated SNPs, markers of germline DNA methylation, positively correlate with regional levels of human meiotic recombination. Genome Res. **19:** 581–589.

SUTTER, N. B., M. A. EBERLE, H. G. PARKER, B. J. PULLAR, E. F. KIRKNESS et al., 2004 Extensive and breed-specific linkage disequilibrium in Canis familiaris. Genome Res. **14:** 2388–2396.

VIGNAUX, F., C. HITTE, C. PRIAT, J. C. CHUAT, C. ANDRE et al., 1999 Construction and optimization of a dog whole-genome radiation hybrid panel. Mamm. Genome **10:** 888–894.

VILÀ, C., P. SAVOLAINEN, J. E. MALDONADO, I. R. AMORIM, J. E. RICE et al., 1997 Multiple and ancient origins of the domestic dog. Science **276:** 1687–1689.

WINCKLER, W., S. R. MYERS, D. J. RICHTER, R. C. ONOFRIO, G. J. MCDONALD et al., 2005 Comparison of fine-scale recombination rates in humans and chimpanzees. Science **308:** 107–111.

YU, A., C. ZHAO, Y. FAN, W. JANG, A. J. MUNGALL et al., 2001 Comparison of human genetic and sequence-based physical maps. Nature **409:** 951–953.

ZENGER, K. R., L. M. MCKENZIE and D. W. COOPER, 2002 The first comprehensive genetic linkage map of a marsupial: the tammar wallaby (Macropus eugenii). Genetics **162:** 321–330.

Communicating editor: I. HOESCHELE

GENETICS

A Comprehensive Linkage Map of the Dog Genome

Aaron K. Wong, Alison L. Ruhe, Beth L. Dumont, Kathryn R. Robertson, Giovanna
Guerrero, Sheila M. Shull, Janet S. Ziegle, Lee V. Millon, Karl W. Broman,
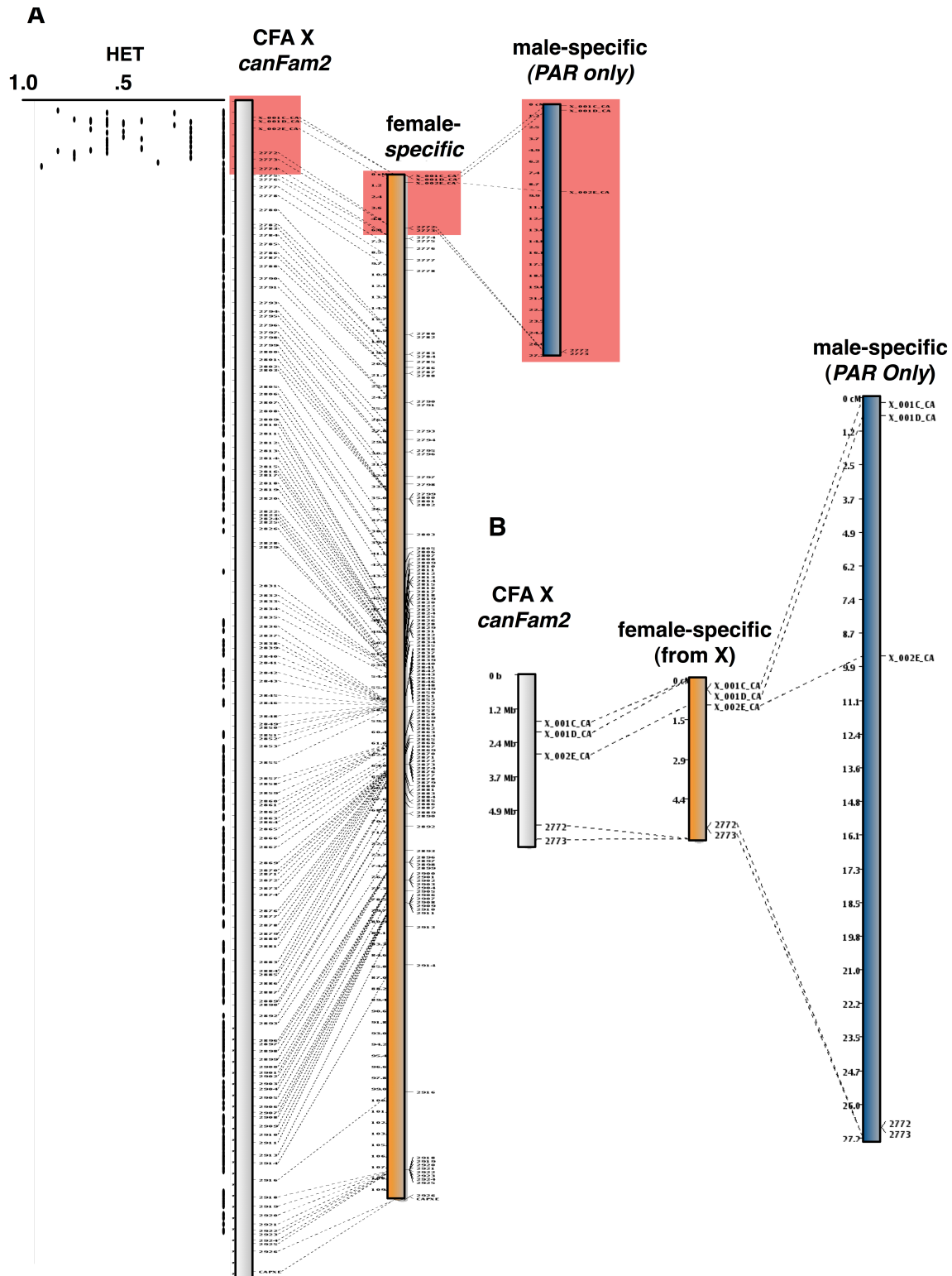Bret A. Payseur and Mark W. Neff

FIGURE S1.—An *X* Chromosome linkage map with denoted PAR. (A) Graph showing the physical clustering of SNPs that exhibited heterozygosity in males and ideogram of the *X* showing physical (Mb, left), female-specific (cM, middle), and male-specific (cM, right) genetic distances. Presumptive PAR intervals are depicted with red-shaded insets. (B) Expanded view of presumptive PAR
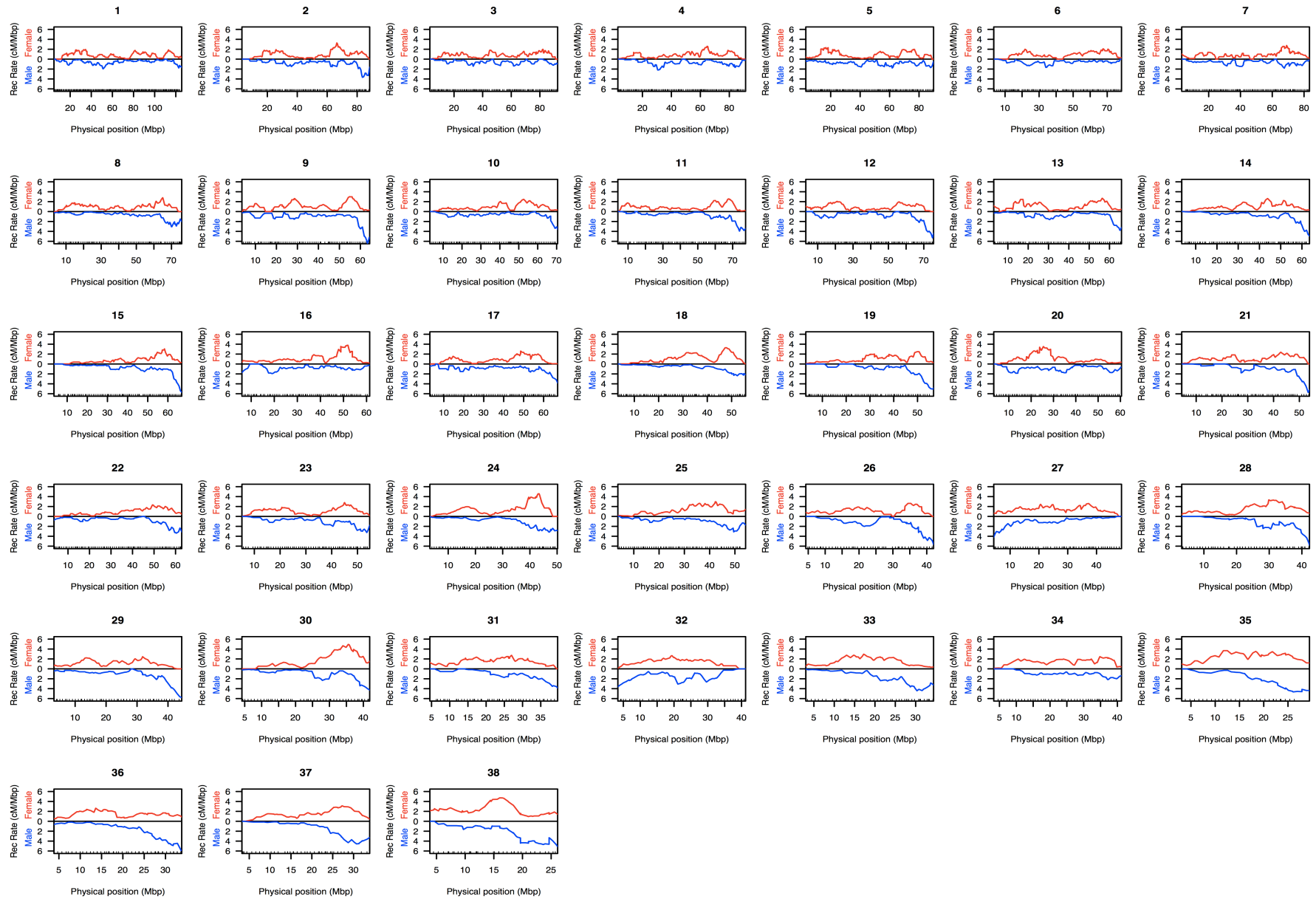
FIGURE S2.— Ratio of sex-specific recombination rates across autosomes. Female- and male-specific rates are shown from the centromeric to telomeric ends (left to right). X-axes are scaled to accommodate different physical sizes of autosomes. Sliding windows of 6 Mb were used to calculate cM/Mb.

A. K. Wong *et al.*
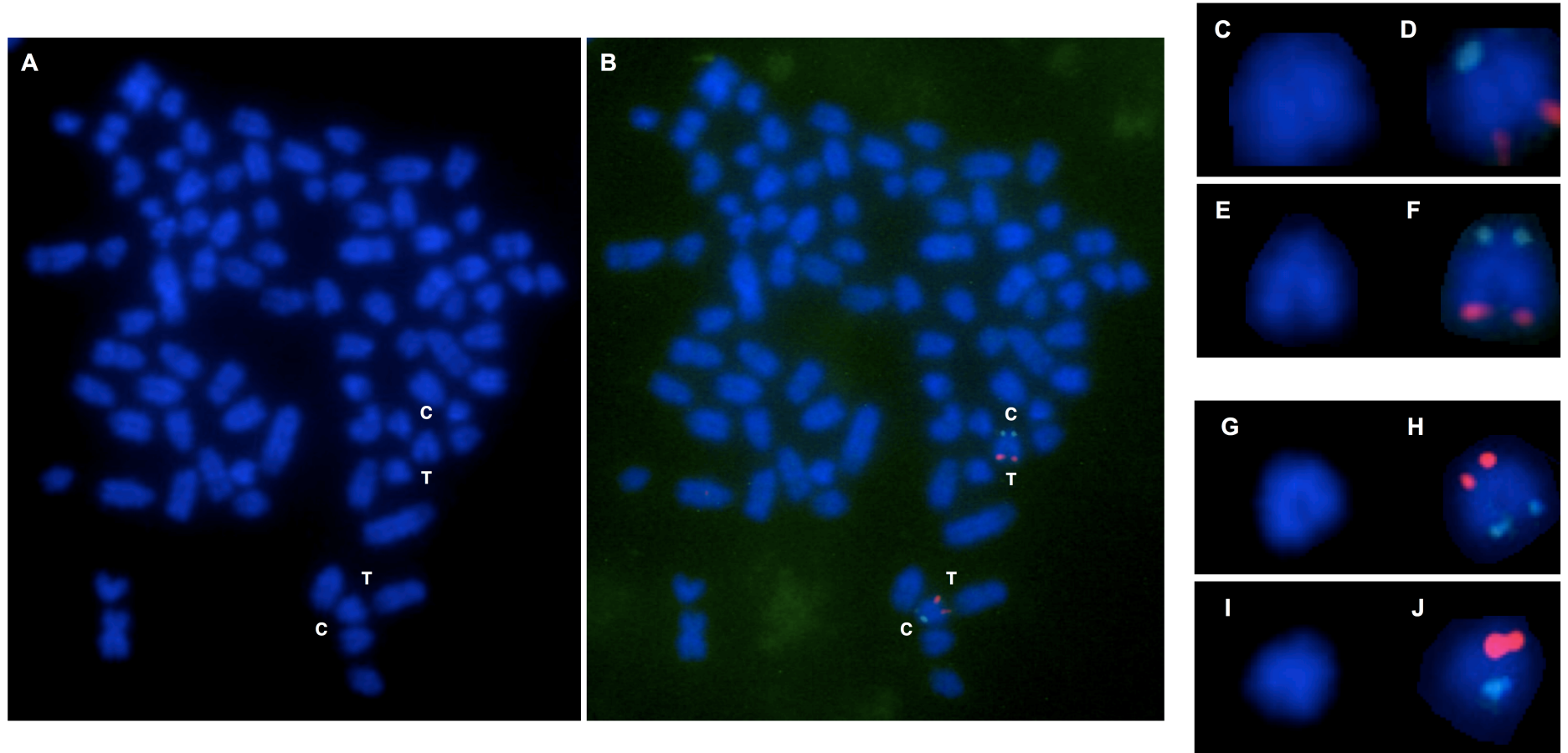
FIGURE S3.—FISH analysis of suspected mis-orientation in *canFam2* of *CFA 27* and *CFA 32* . Cytogenetic analysis by (A) DAPI-staining and (B) FISH with a representative metaphase spread (c, centromeric; t, telomeric). (C-F) Analysis of *CFA27*. (C and E) DAPI-staining of *CFA27* homologs from the same metaphase spread. (D and F) FISH with probes to compare the chromosome orientation specified in *canFam2* to the one suggested by the linkage map. (G-J) Analysis of *CFA32*. Arrangement same as above, except fluorescent labels for distal and proximal probes were reversed. Results from examining multiple spreads are summarized in Table S5.
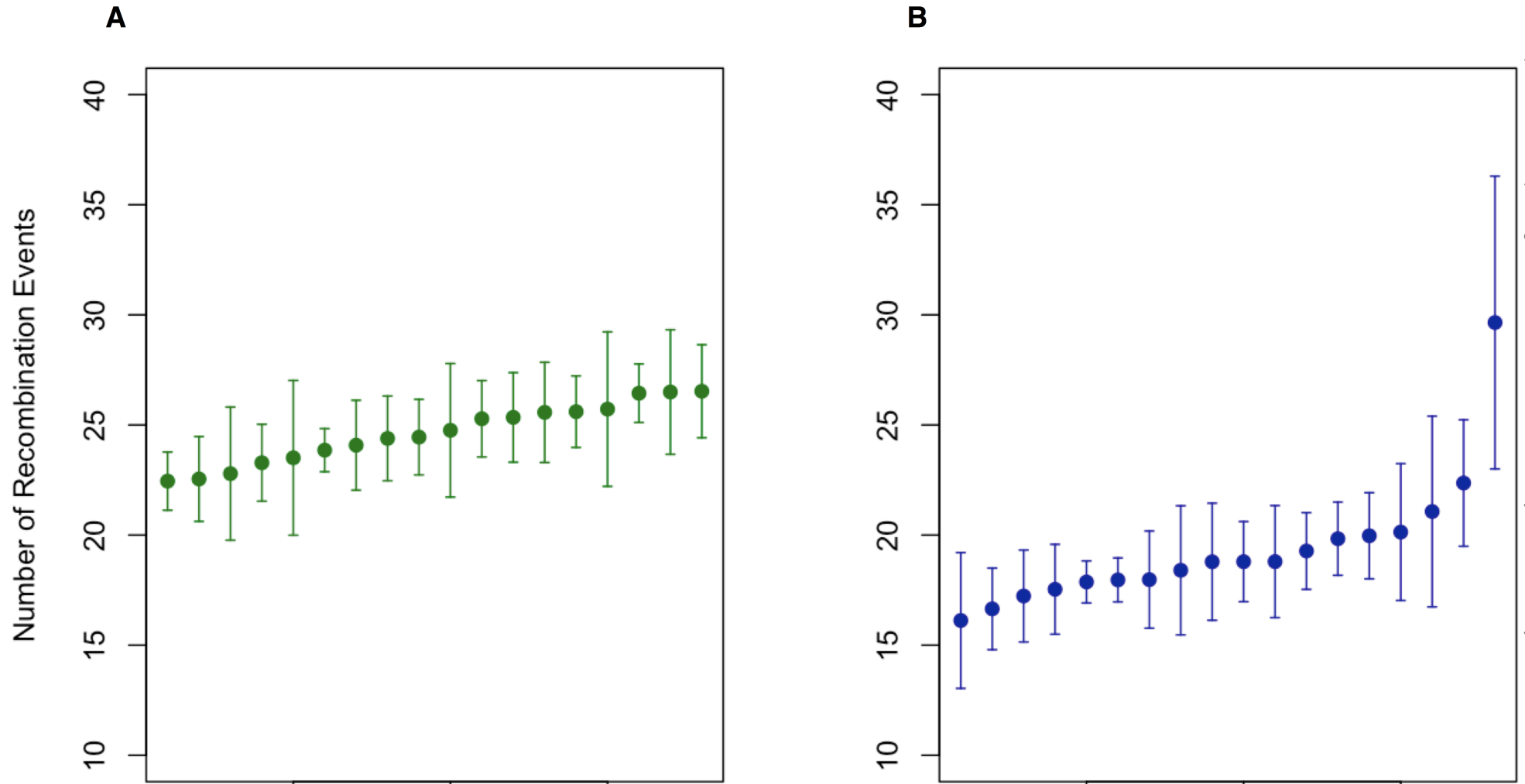
FIGURE S4.—Inter-individual variation in global recombination frequency during meiosis. (A) Female inter-individual differences. (B) Male inter-individual differences. Individuals arranged from lowest to highest mean frequency.
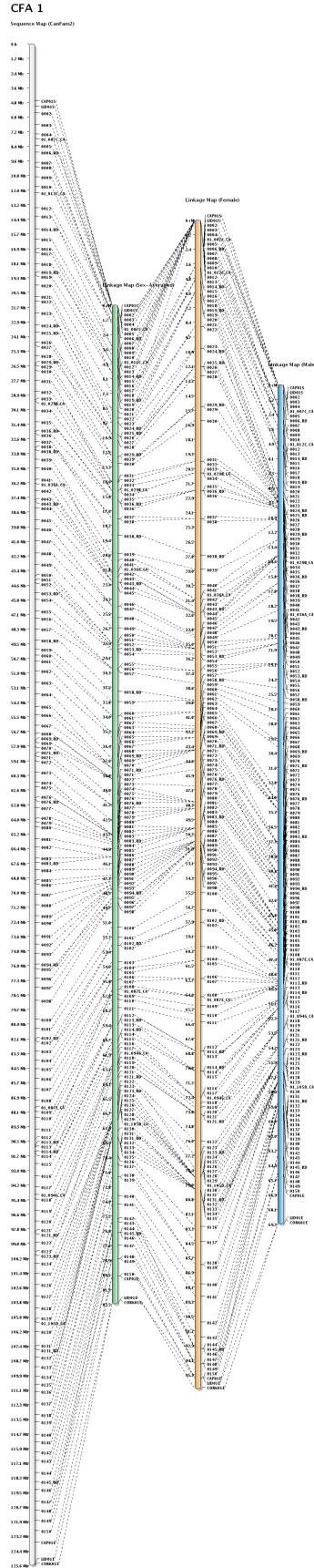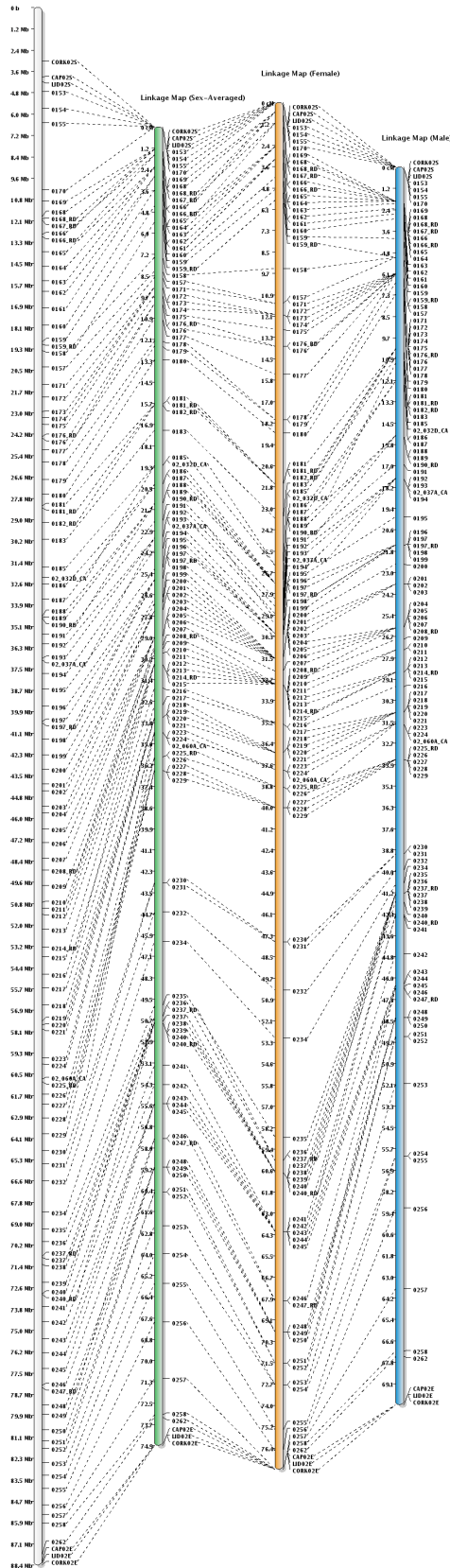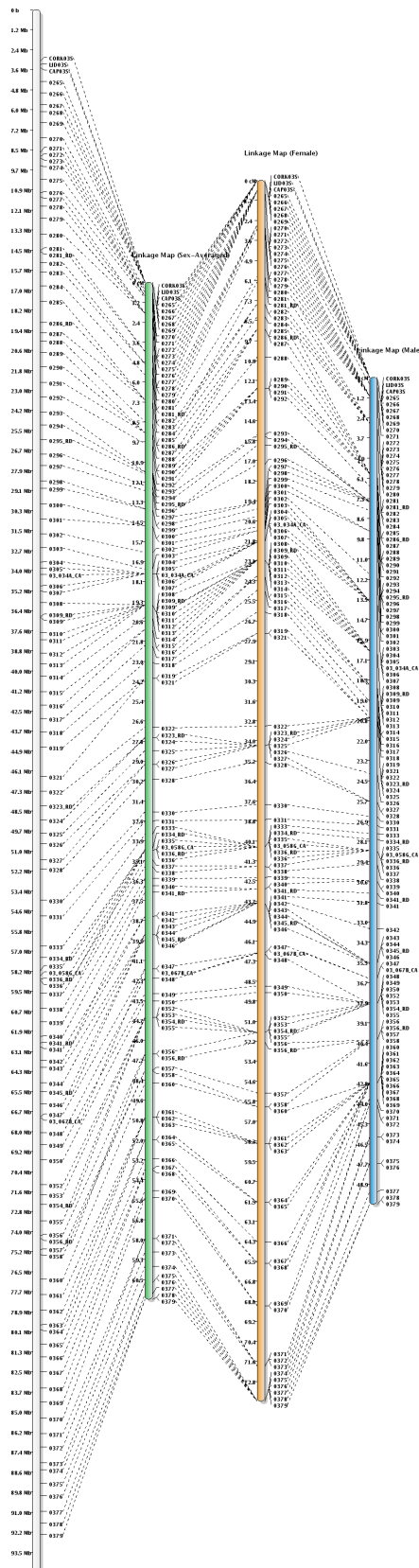
A. K. Wong *et al.*

CFA 1

A. K. Wong *et al.*

CFA 3

CFA 4

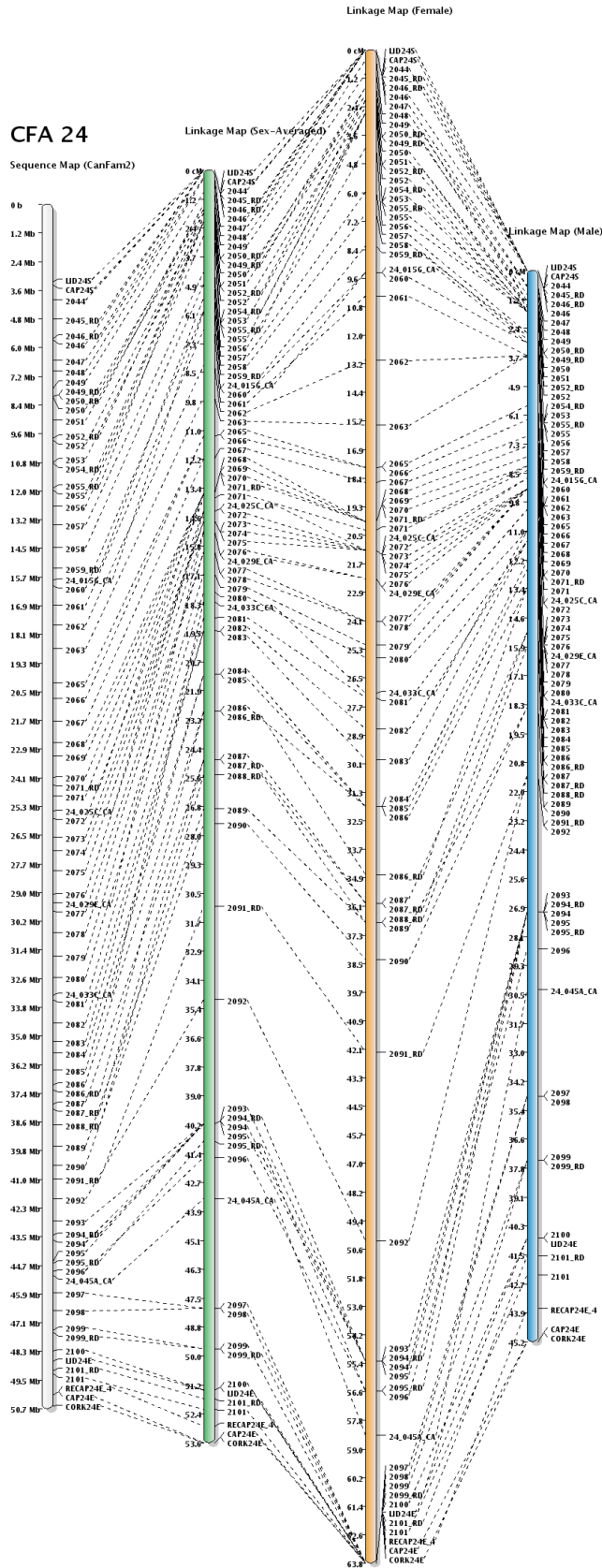A. K. Wong *et al.*

CFA 5

CFA 6

A. K. Wong *et al.*

CFA 7

CFA 8

Sequence Map (CanFam2)

Linkage Map (Female)

Linkage Map (Sex–Averaged)

Linkage Map (Male)

A. K. Wong *et al.*

CFA 9
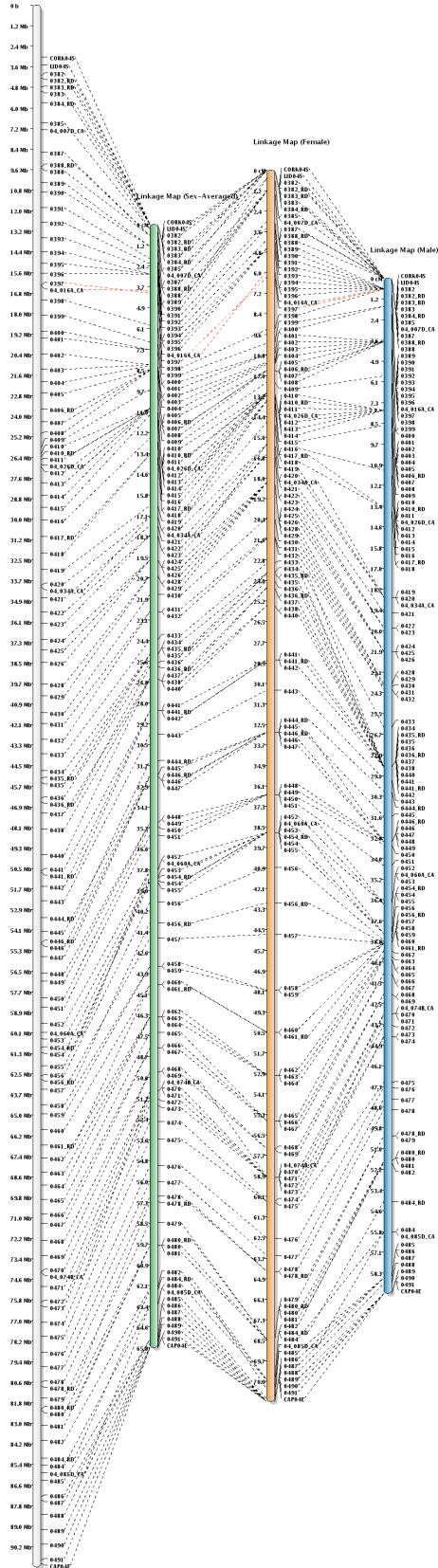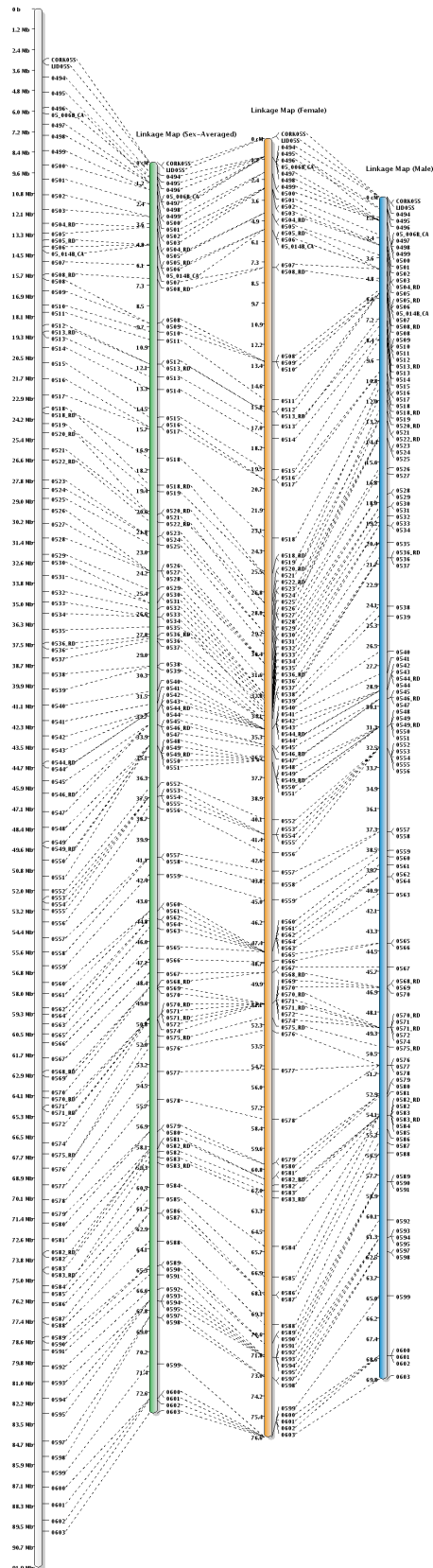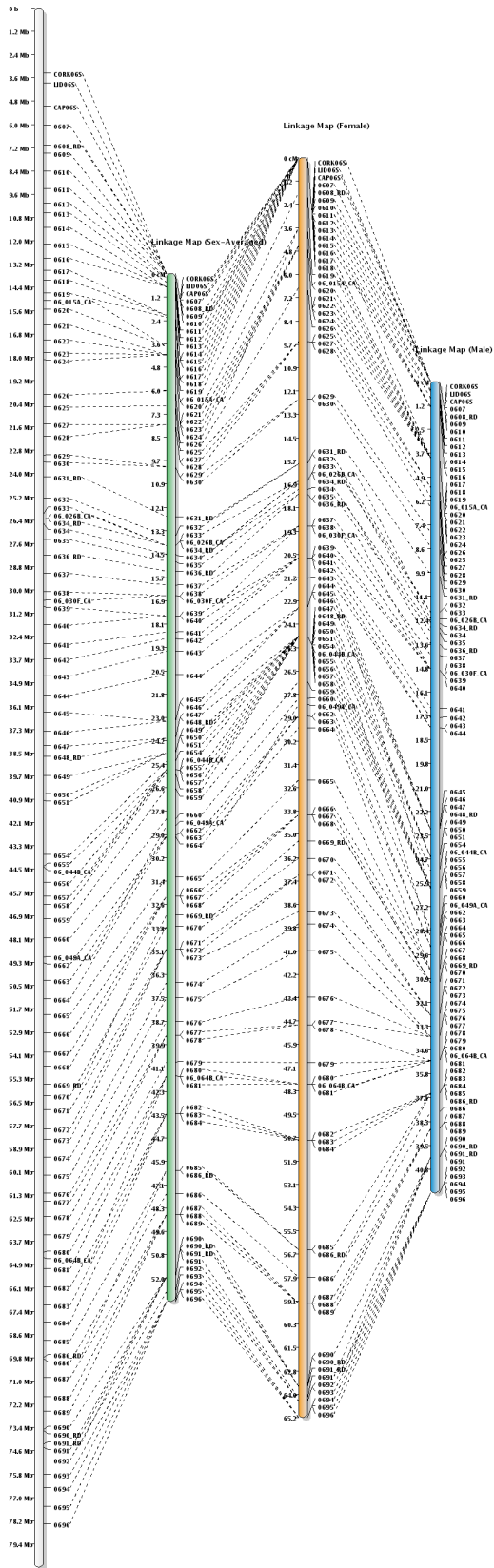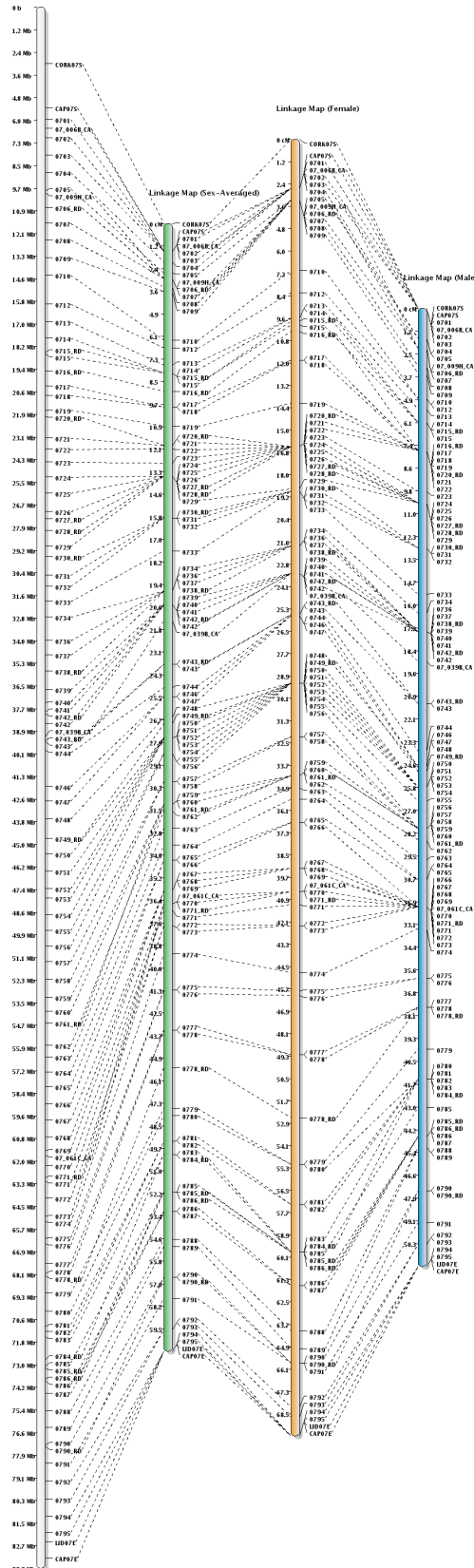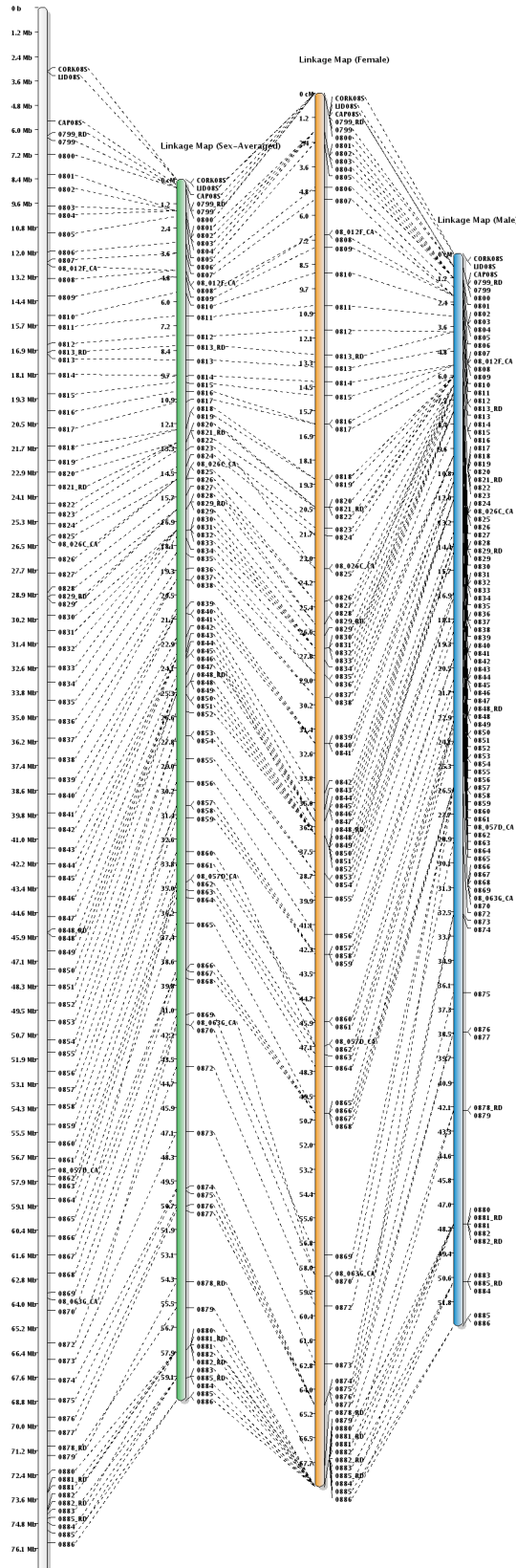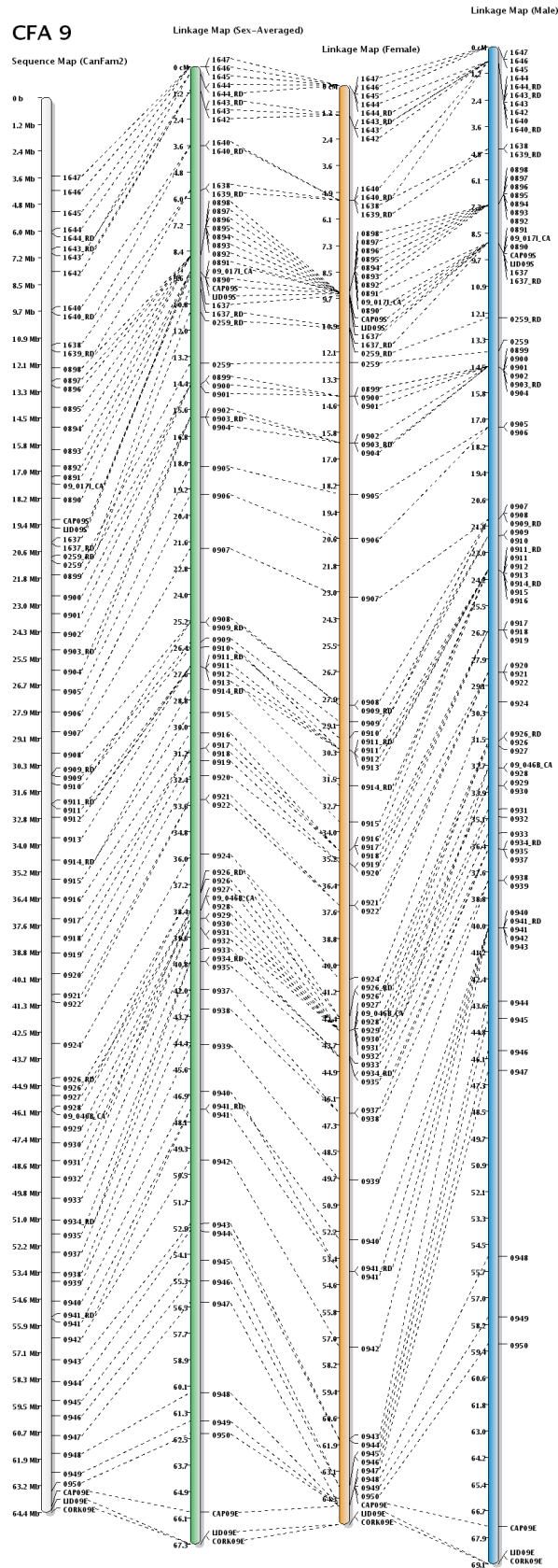
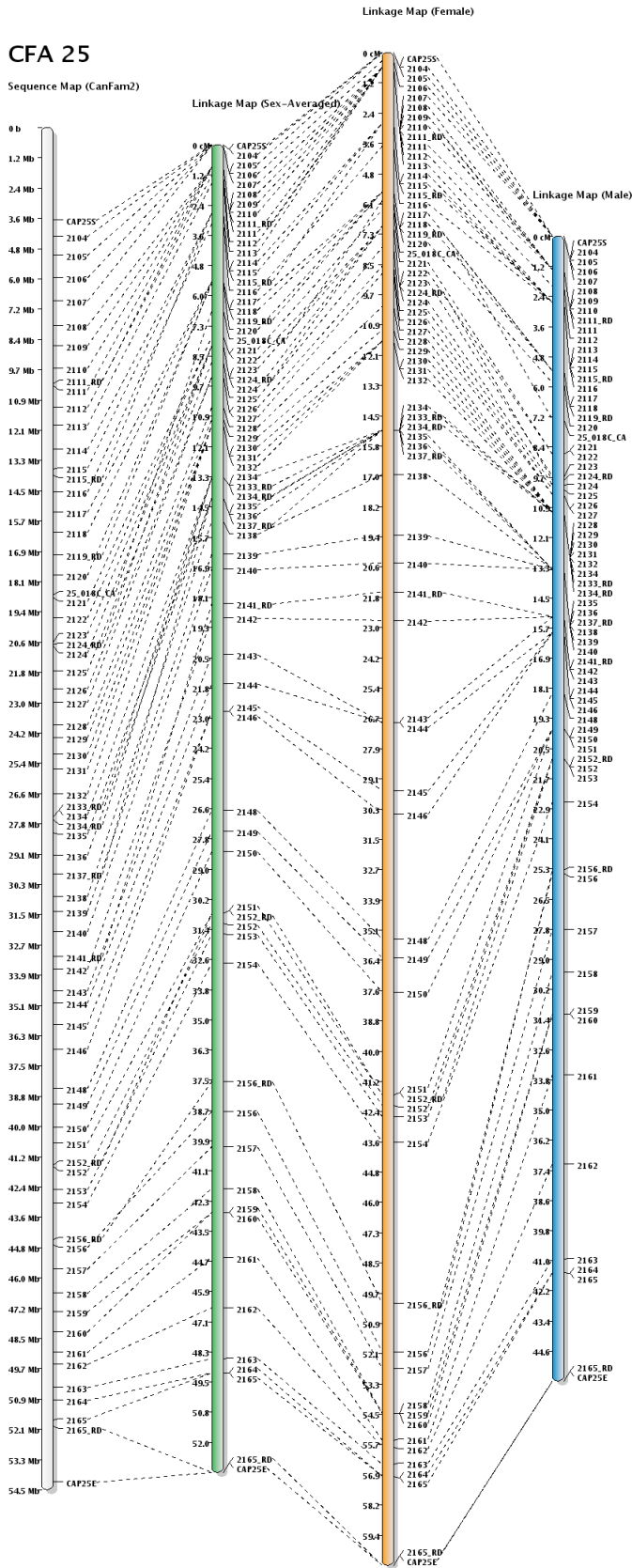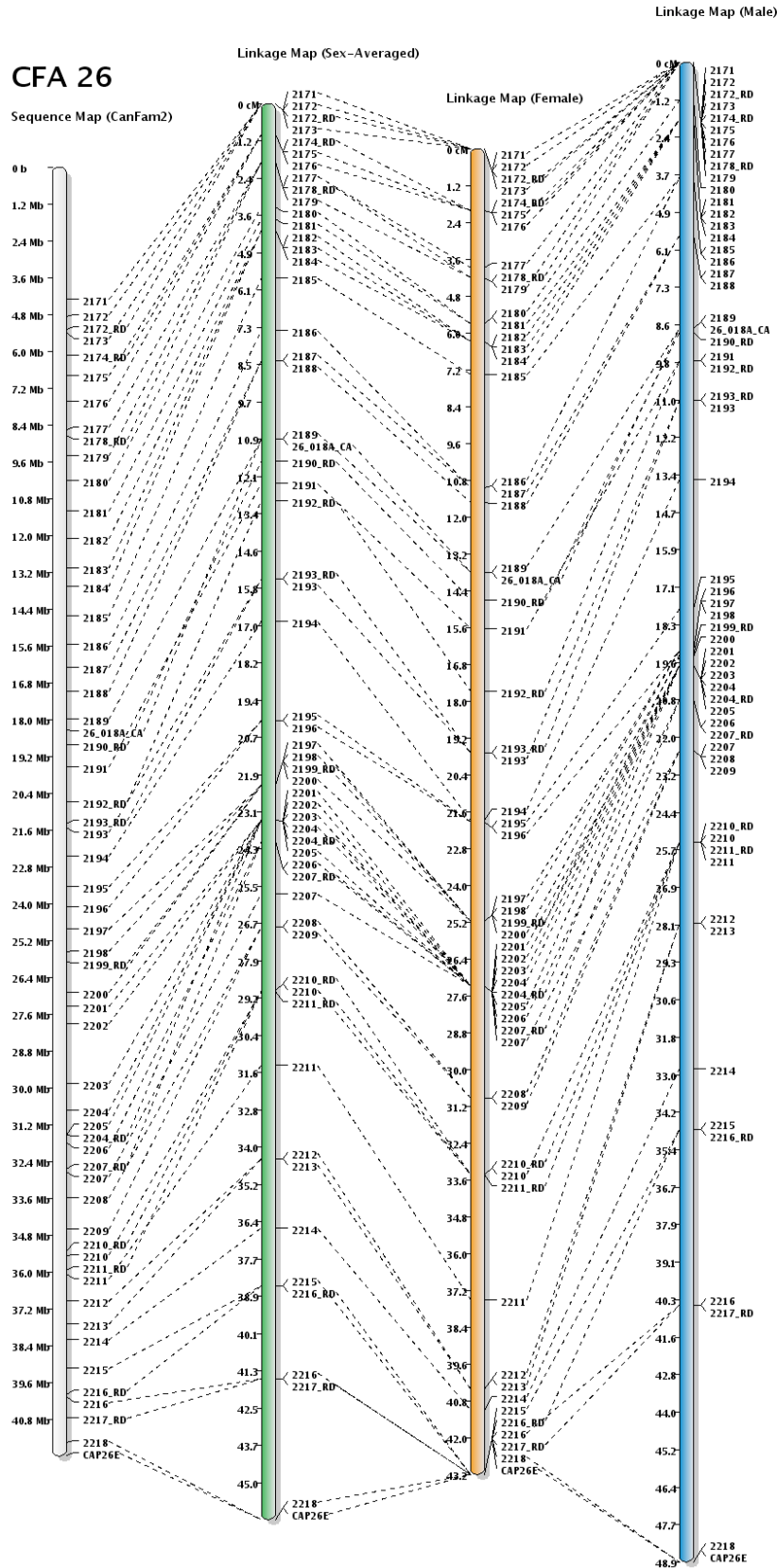# CFA 10

Sequence Map (CanFam2)

A. K. Wong *et al.*
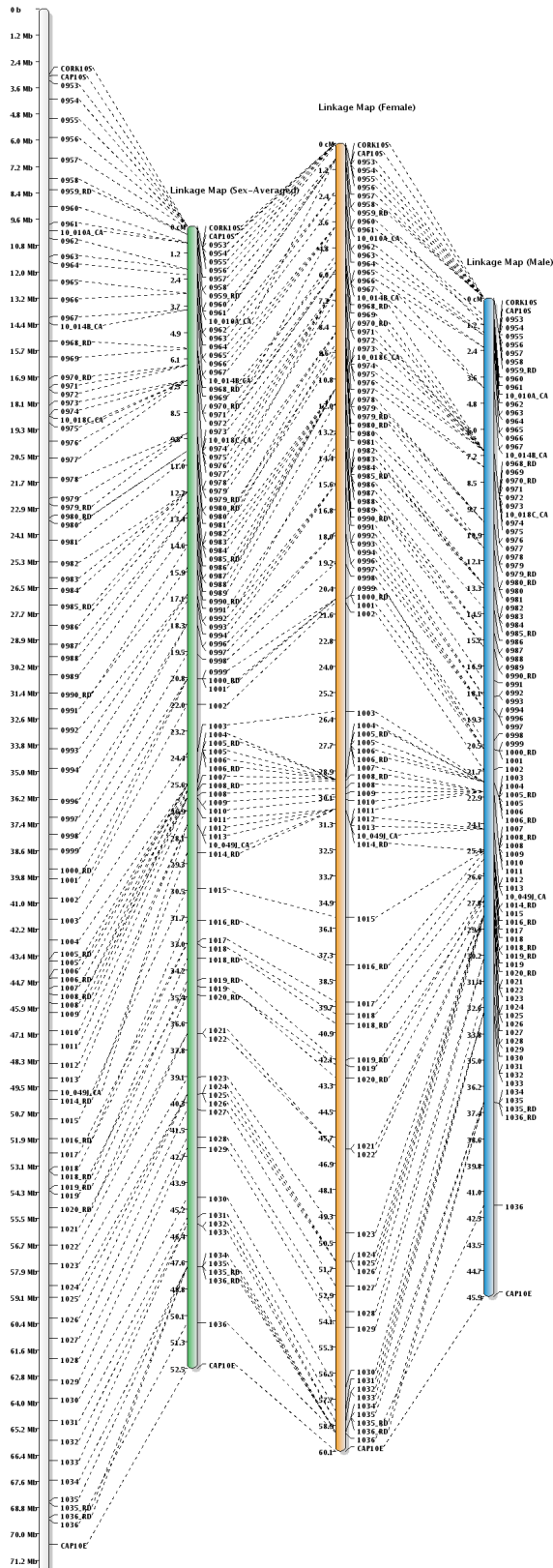
CFA 11
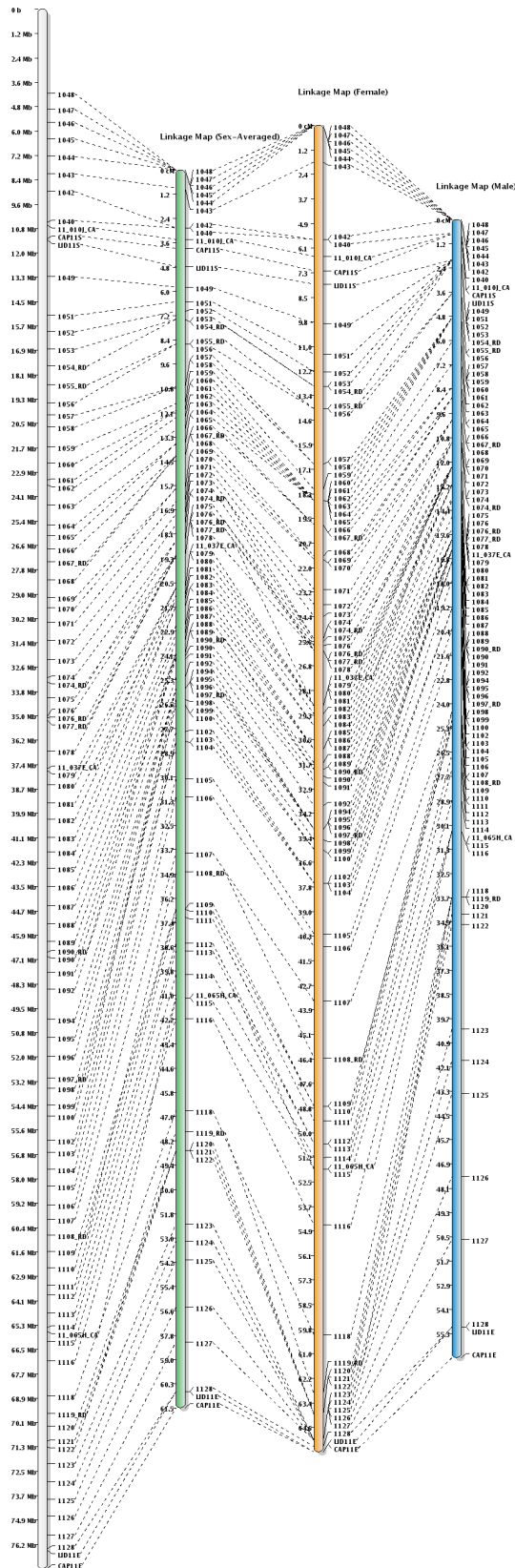
CFA 12

Sequence Map (CanFam2)
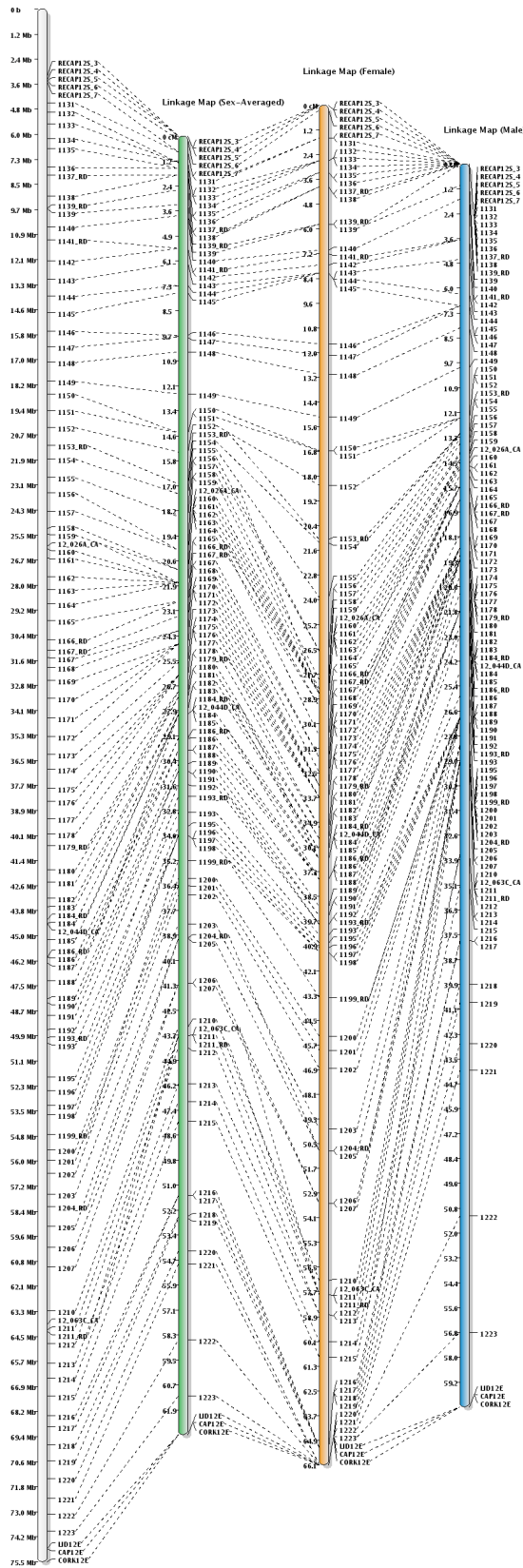
Linkage Map (Female)

Linkage Map (Sex–Averaged)

Linkage Map (Male)

A. K. Wong *et al.*

CFA 13

CFA 14

A. K. Wong *et al.*
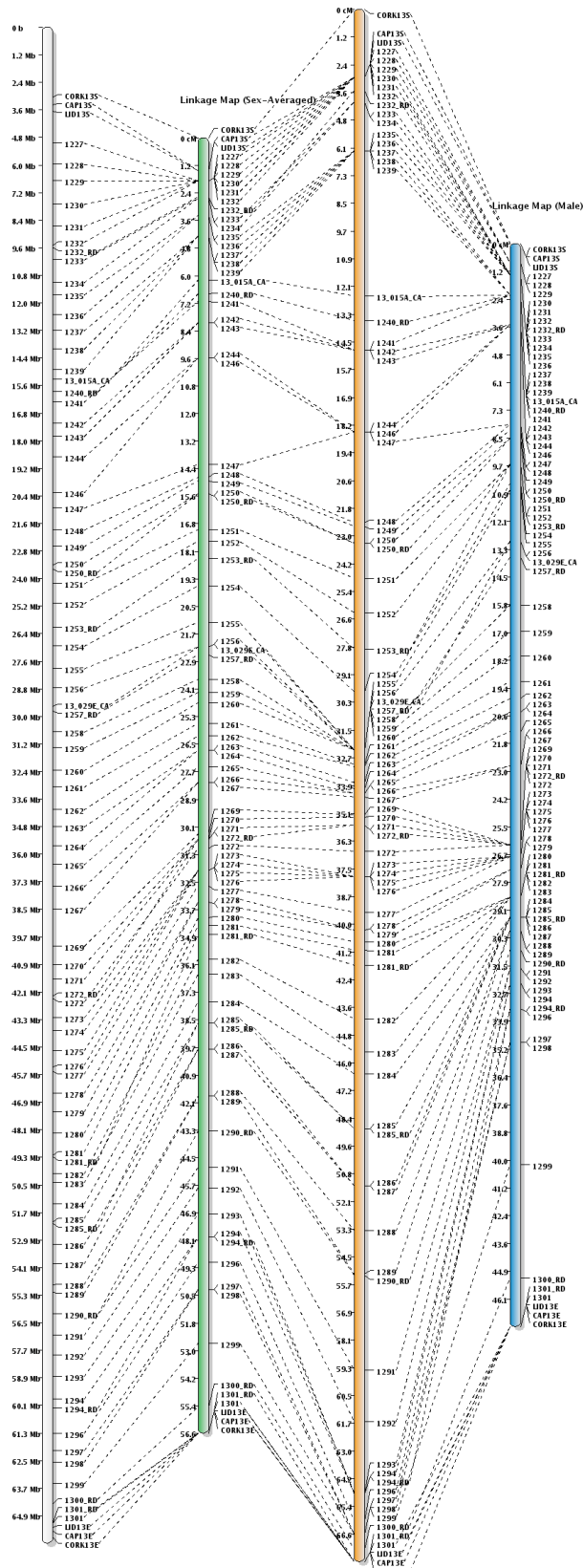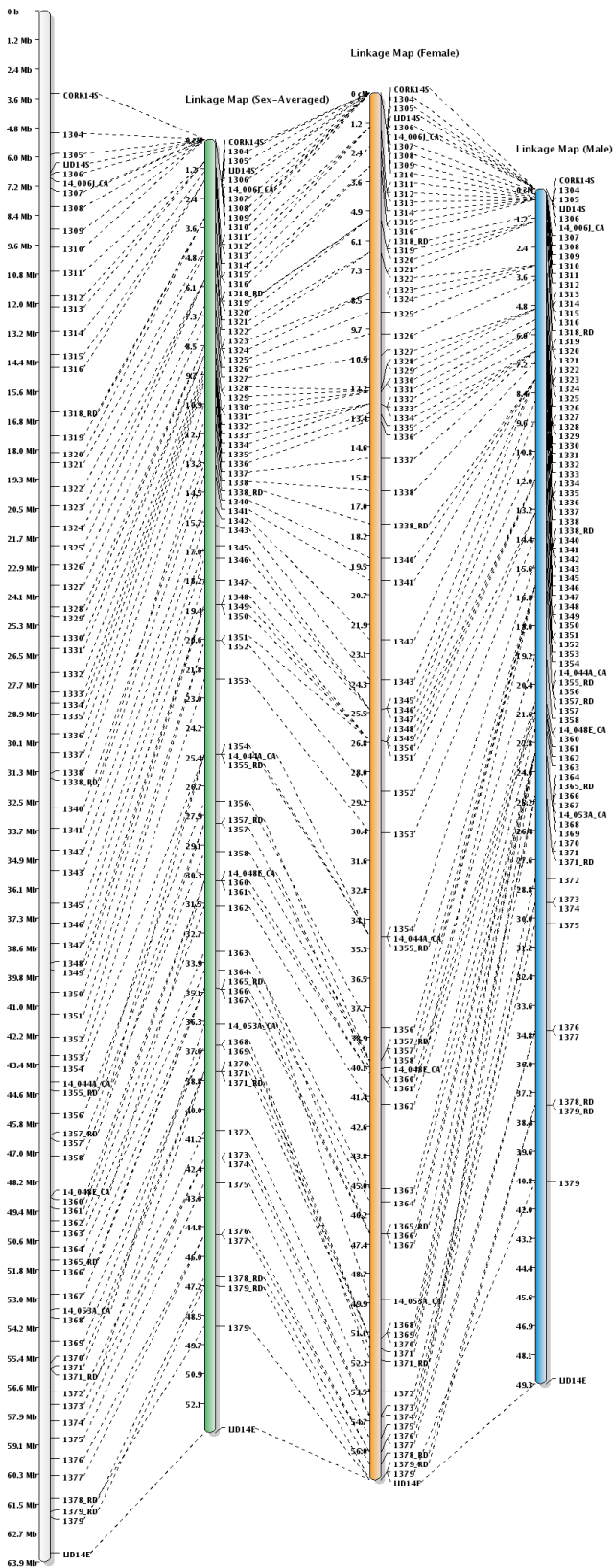
CFA 15
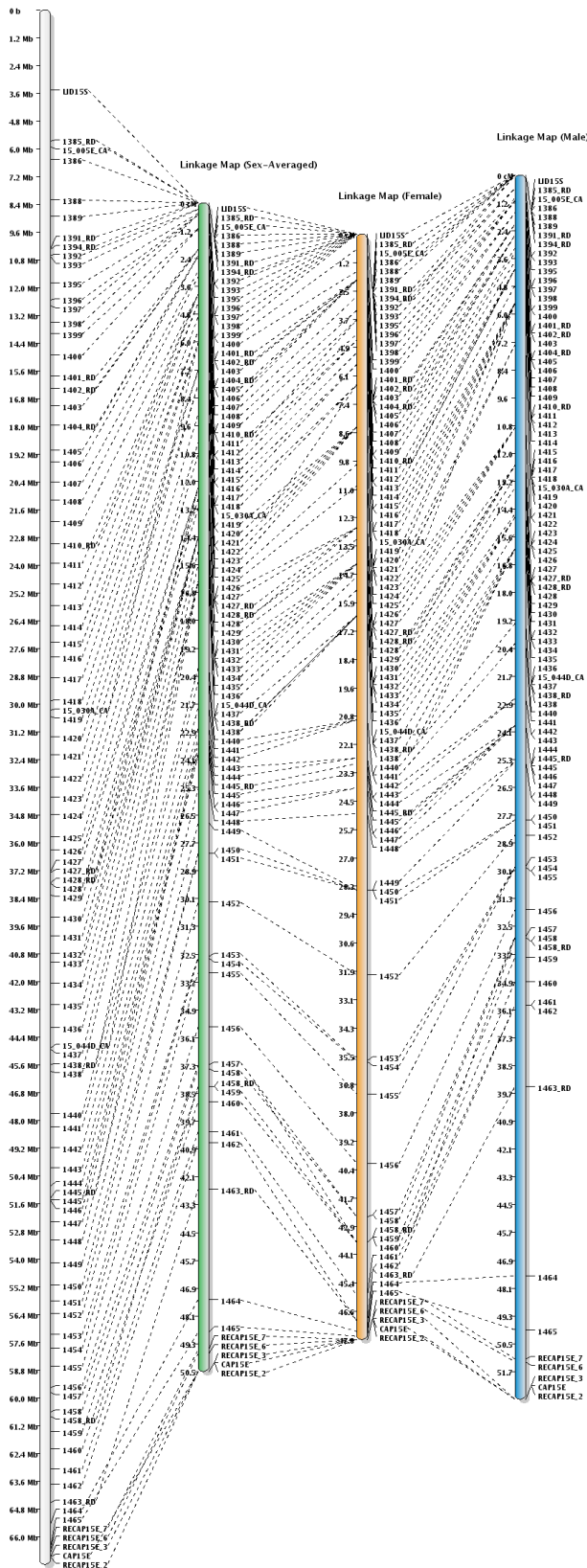
CFA 16

Sequence Map (CanFam2)
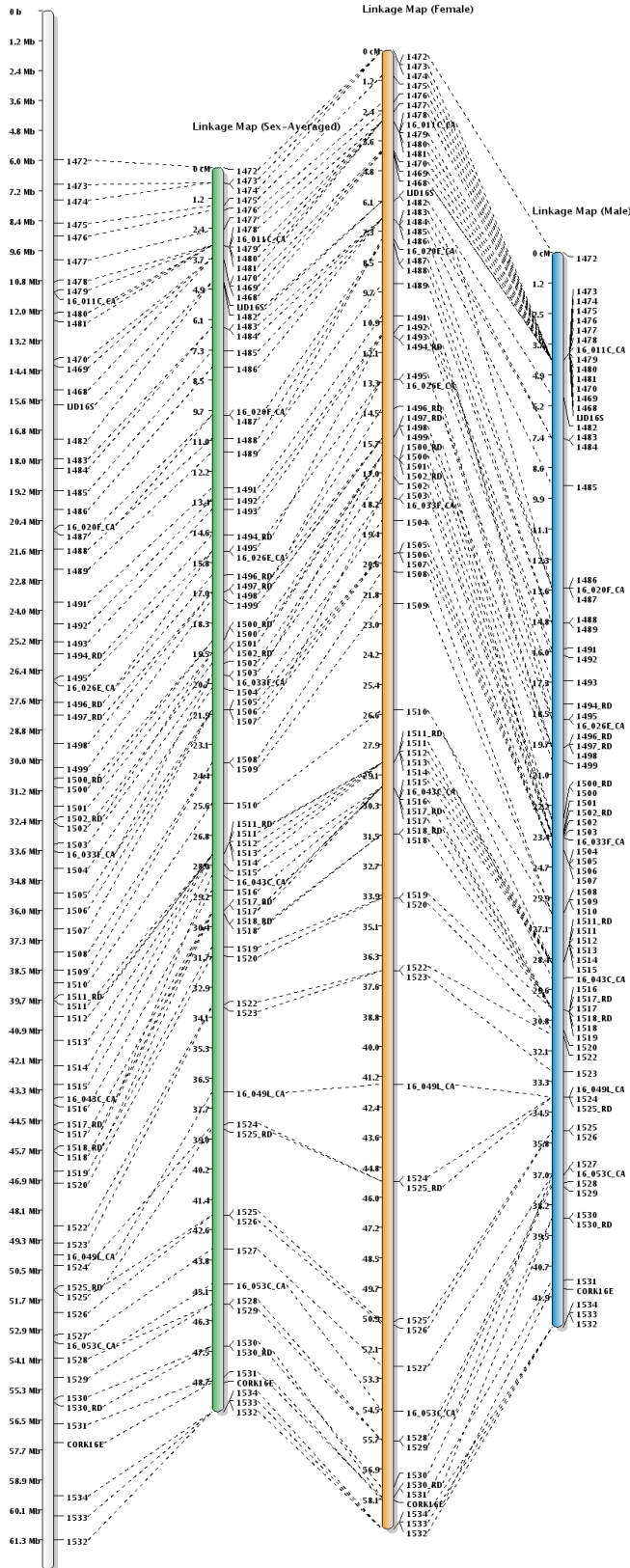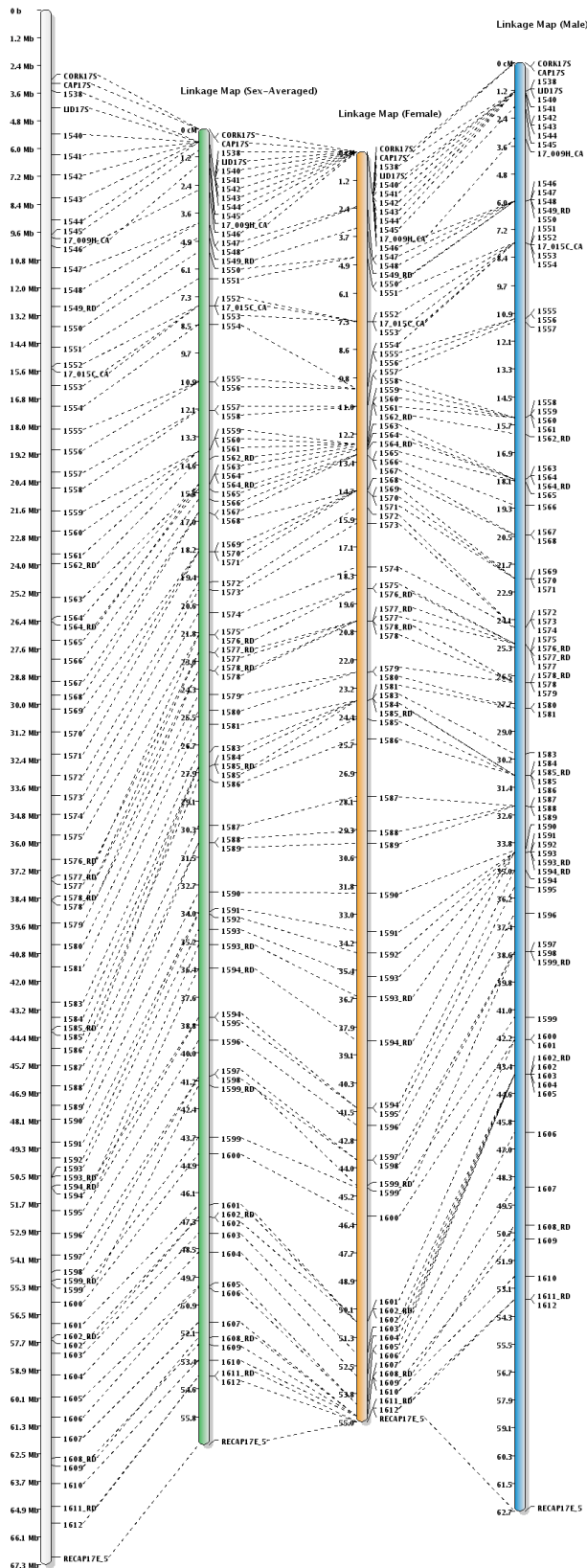
Linkage Map (Sex–Averaged)

Linkage Map (Female)

Linkage Map (Male)

A. K. Wong *et al.*

CFA 17

CFA 18

Sequence Map (CanFam2)

Linkage Map (Female)

Linkage Map (Sex–Averaged)

Linkage Map (Male)

A. K. Wong *et al.*
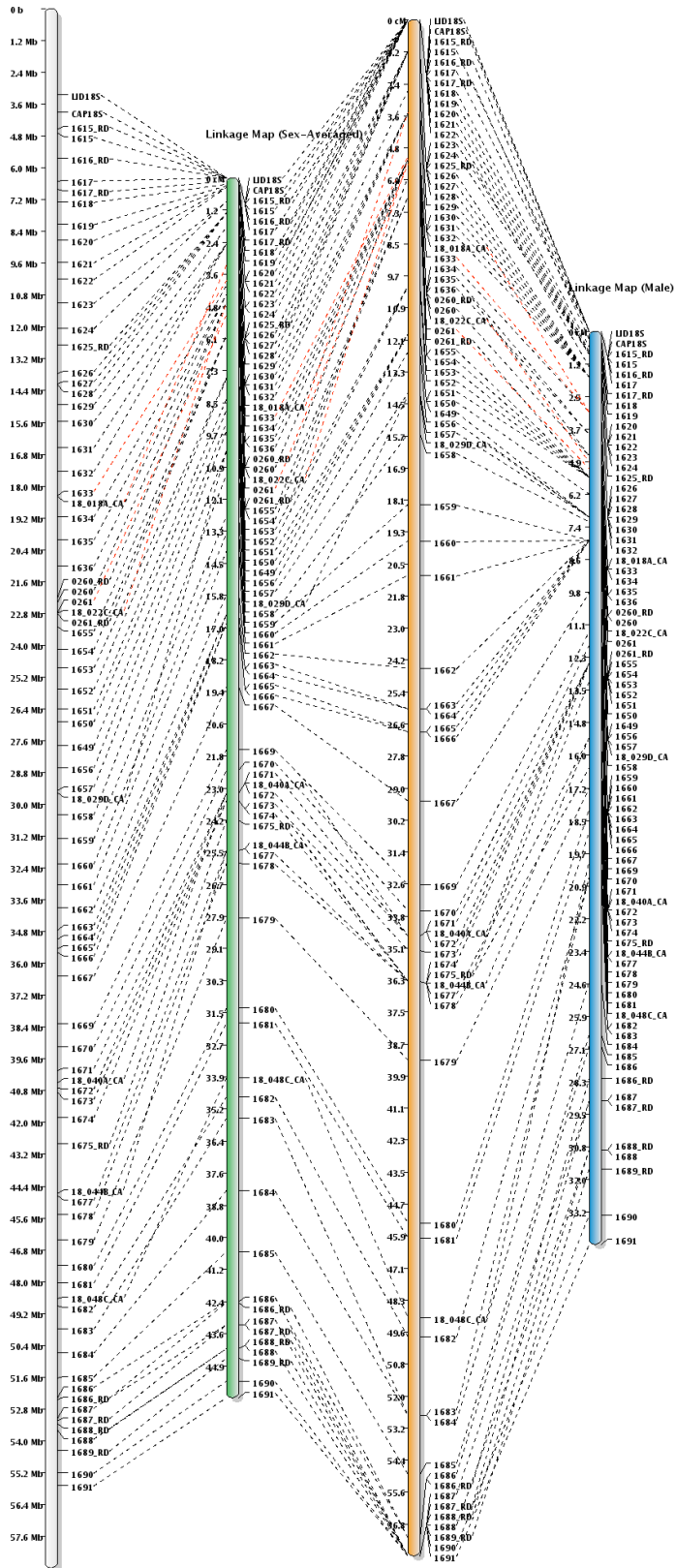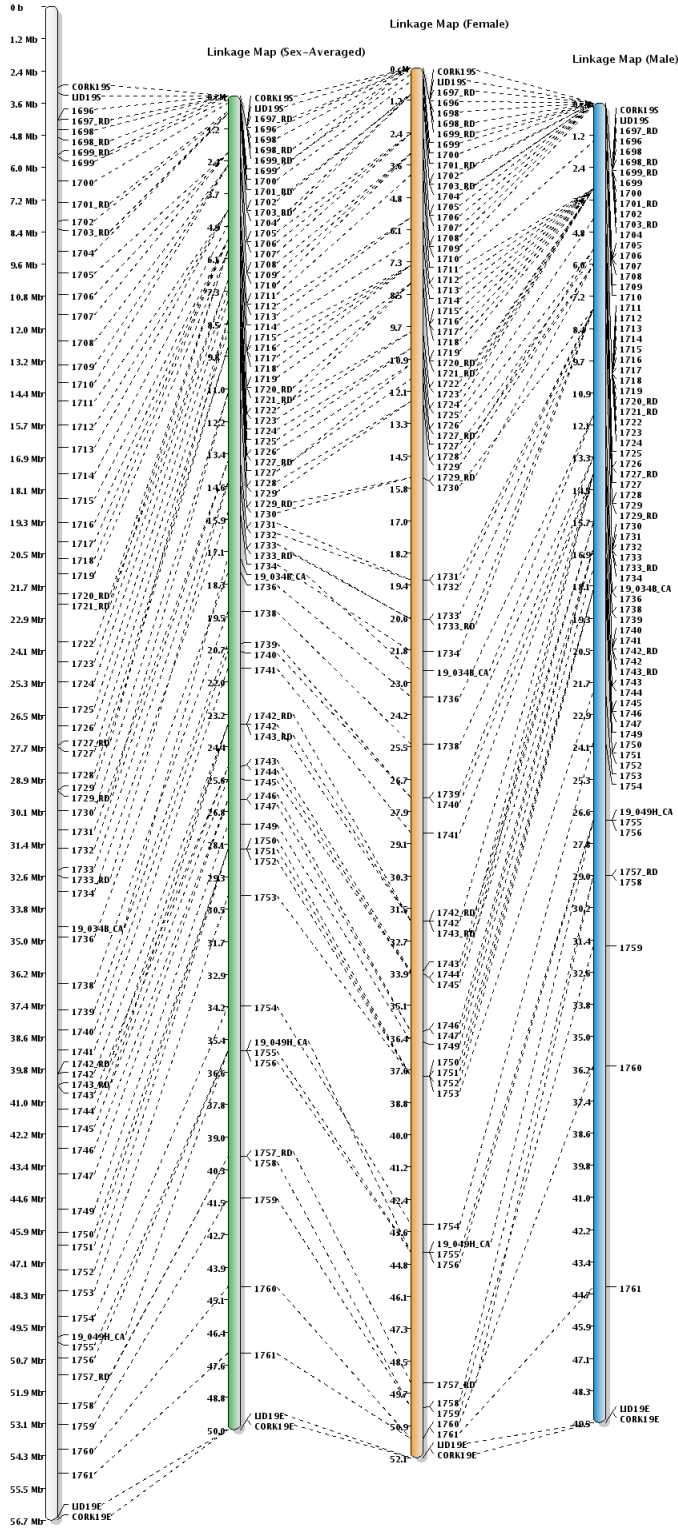
FIGURE S5.—Chromosome -specific linkage maps (CFA1-19)

CFA 20

Sequence Map (CanFam2)
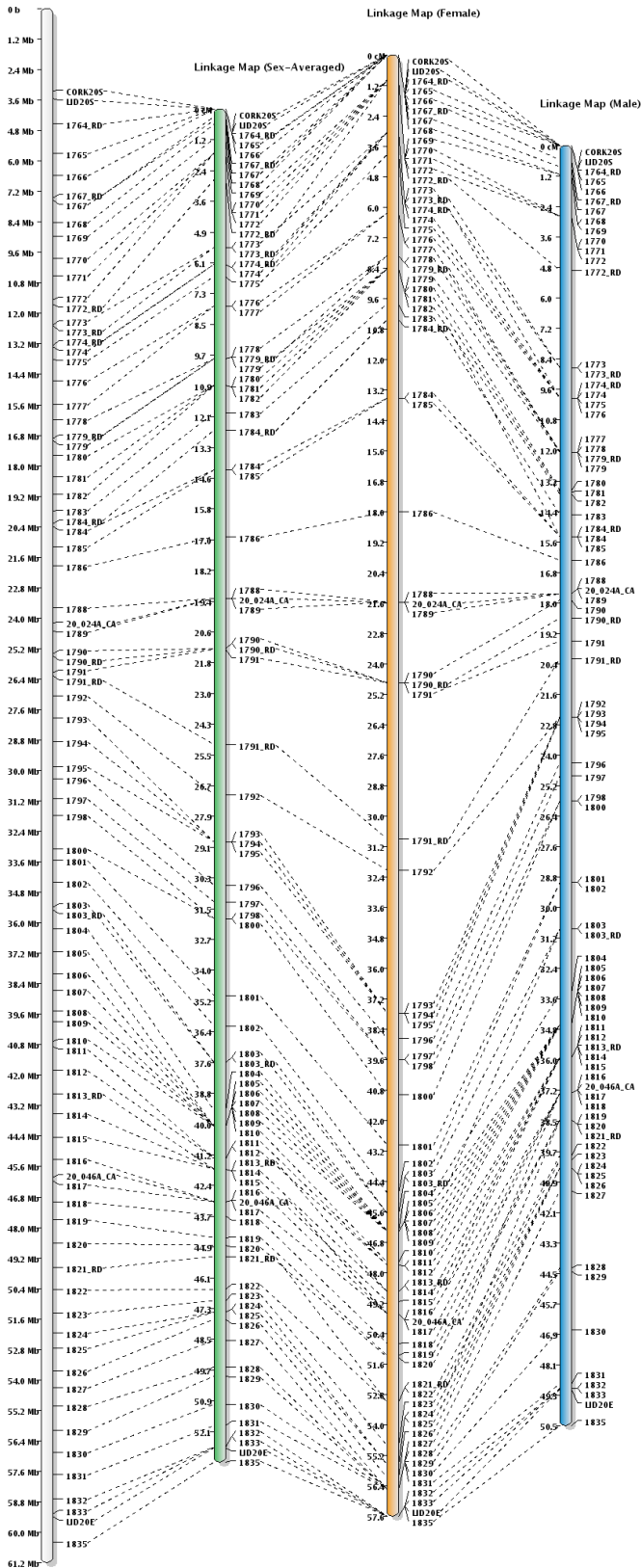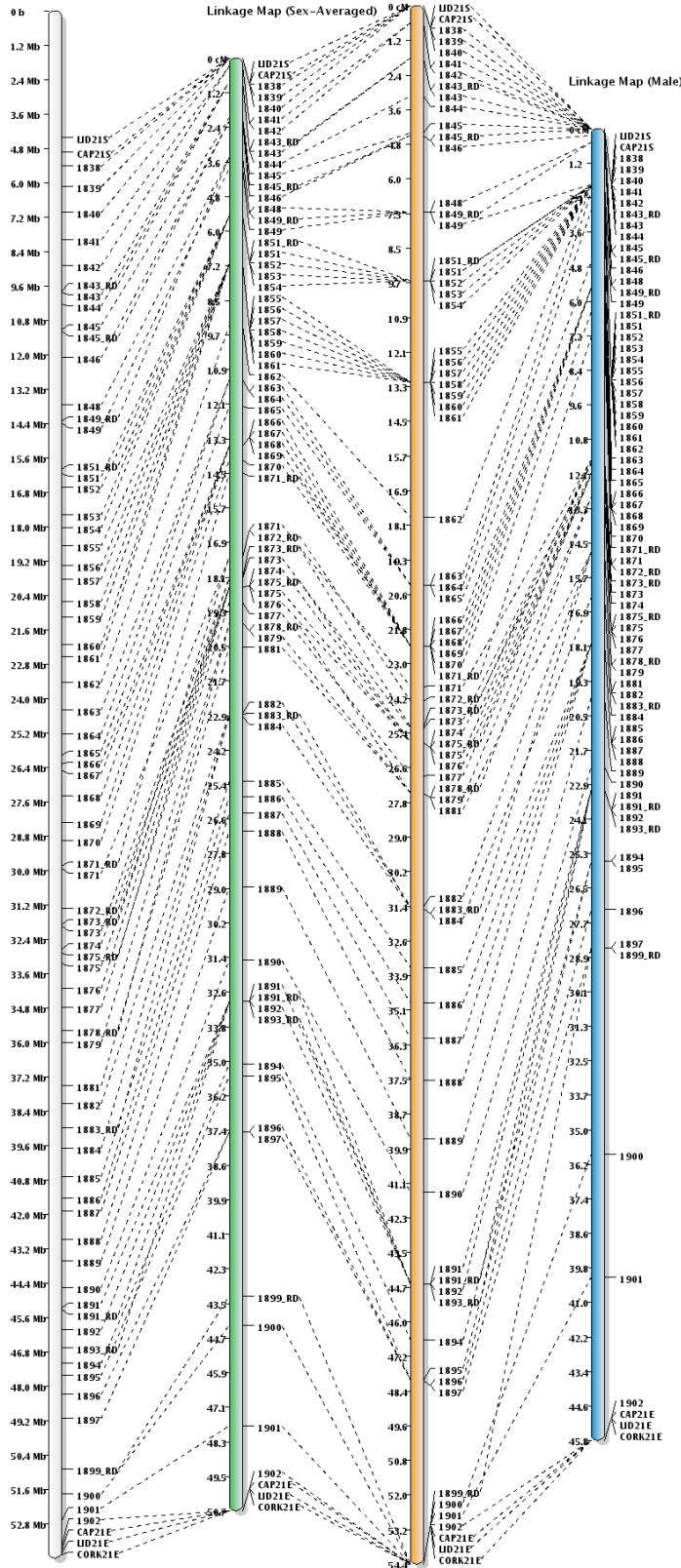
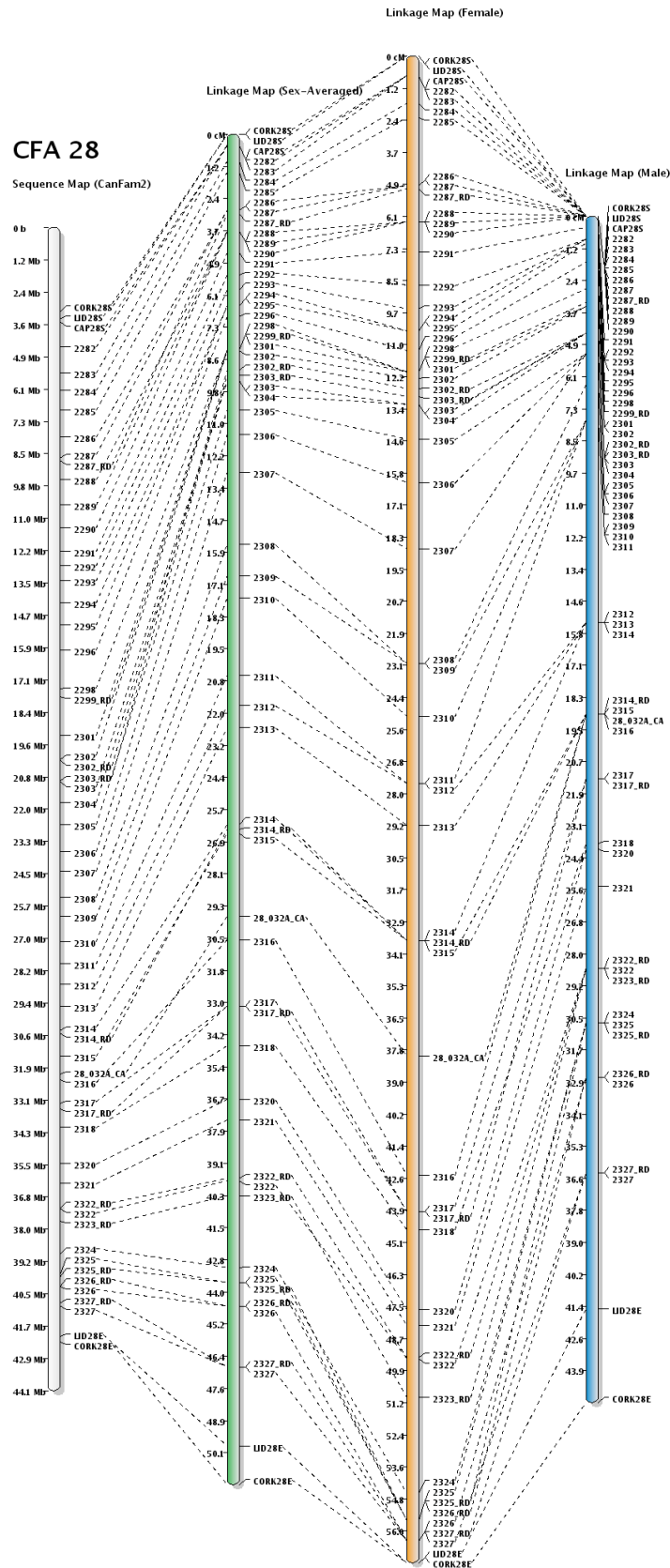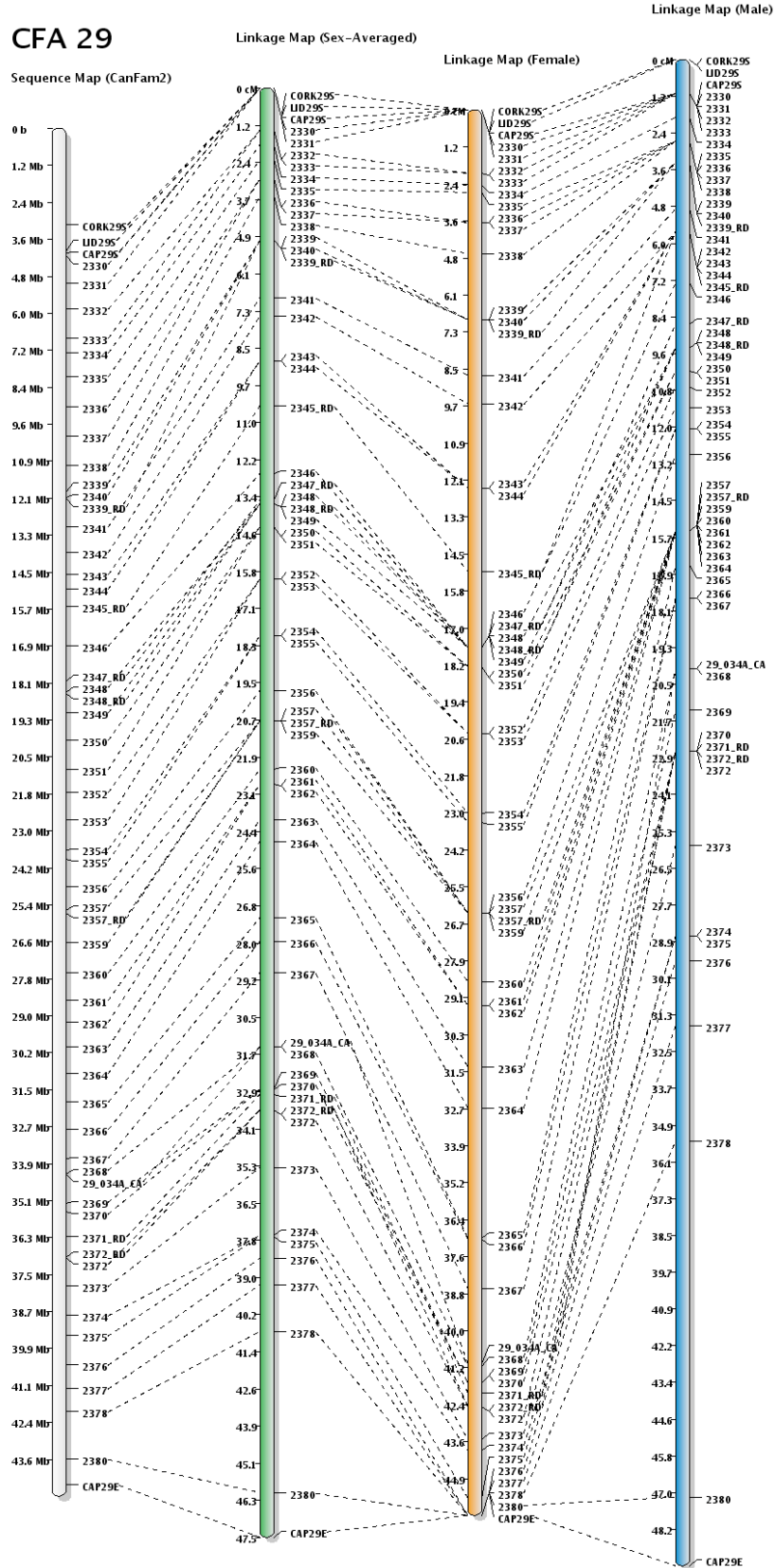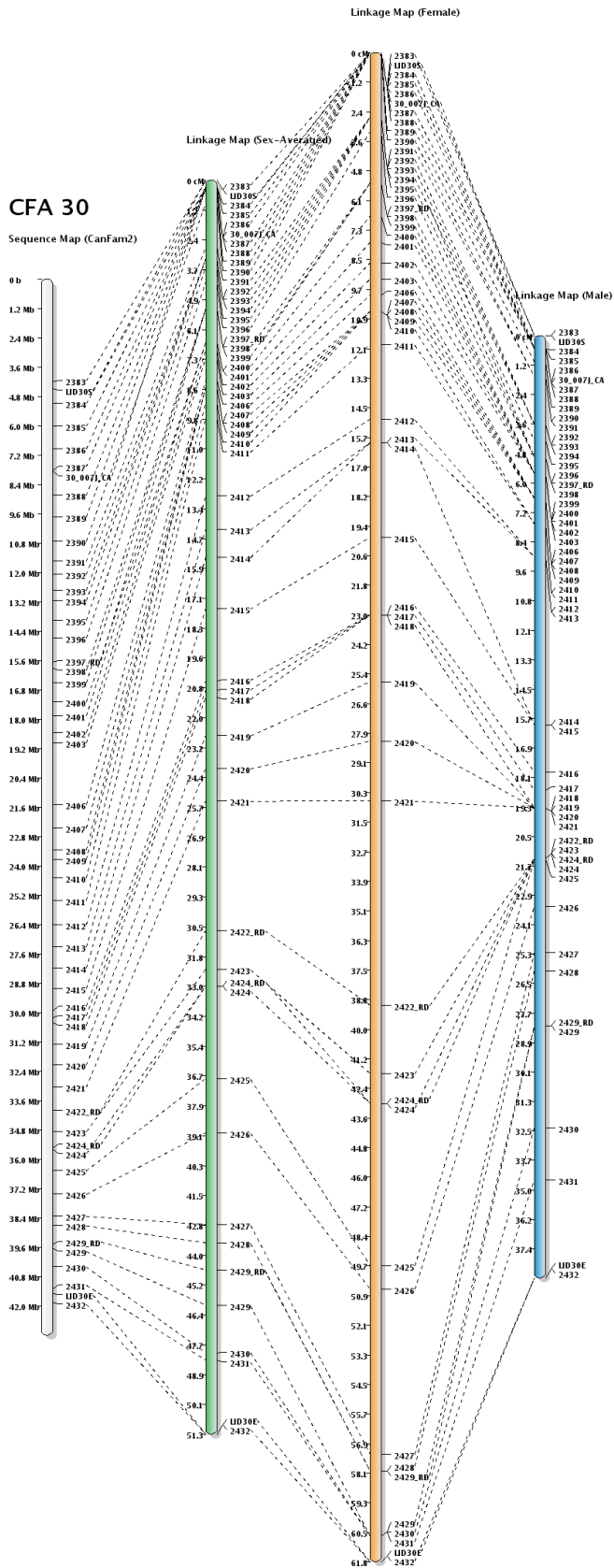A. K. Wong *et al.*

CFA 21

CFA 22

A. K. Wong *et al.*

CFA 23

CFA 24

Linkage Map (Female)

Linkage Map (Sex–Averaged)

Sequence Map (CanFam2)

Linkage Map (Male)

A. K. Wong *et al.*

CFA 25

CFA 26

A. K. Wong *et al.*

CFA 27

CFA 28

A. K. Wong *et al.*

A. K. Wong *et al.*

CFA 31

A. K. Wong *et al.*

CFA 33

CFA 34

Sequence Map (CanFam2)

Linkage Map (Sex–Averaged)

Linkage Map (Female)

Linkage Map (Male)

CFA 35

CFA 36

CFA 37
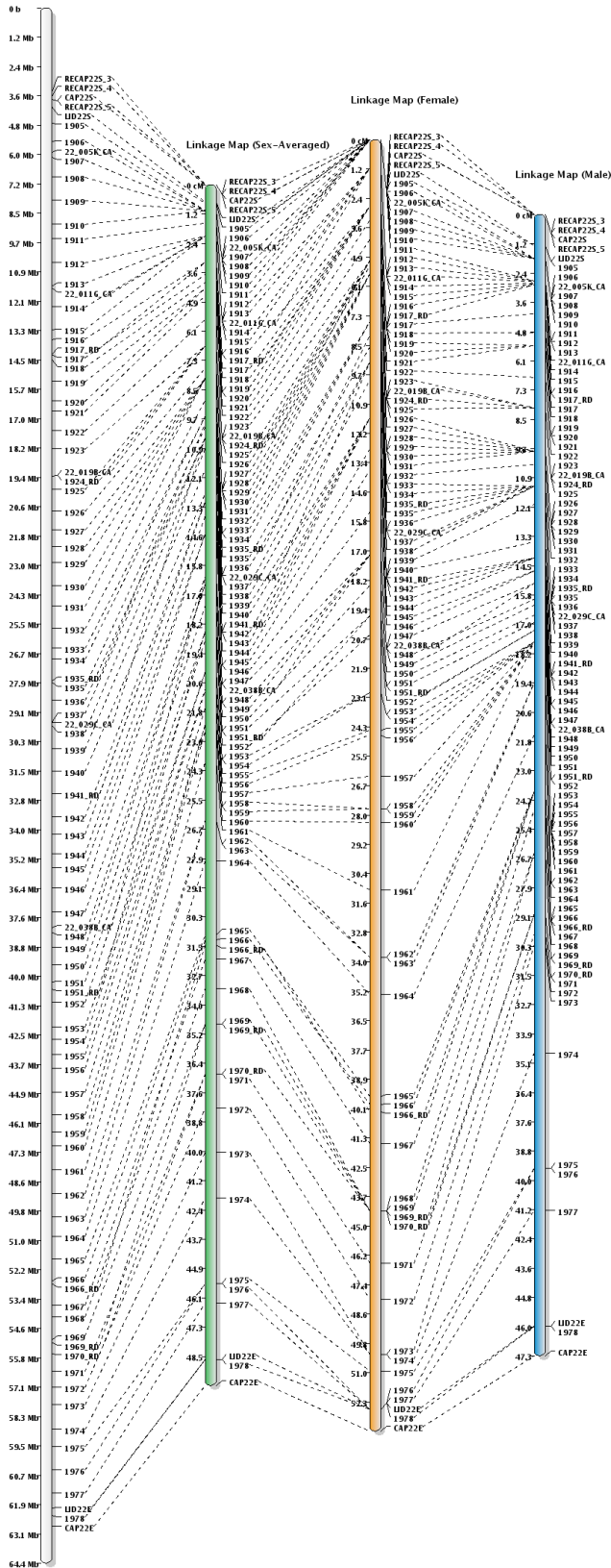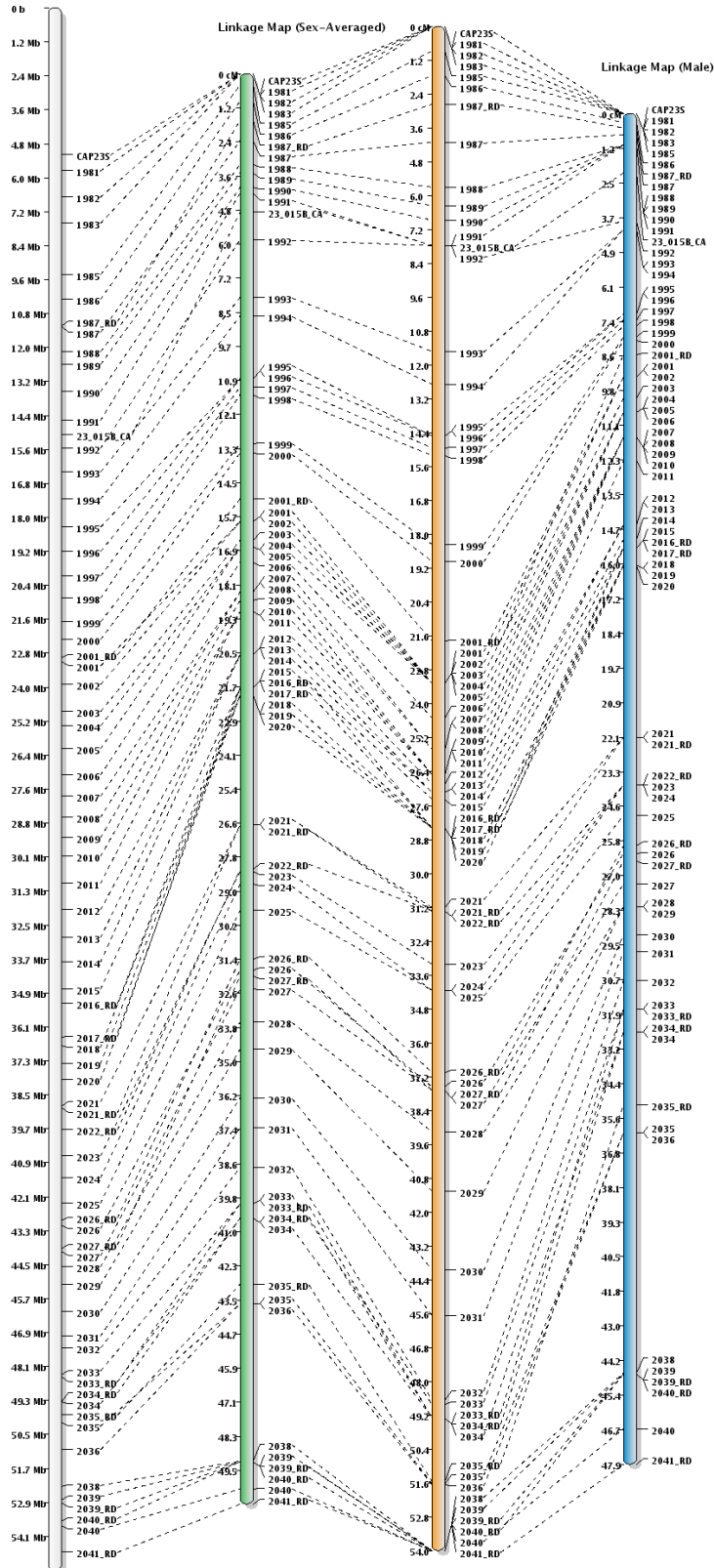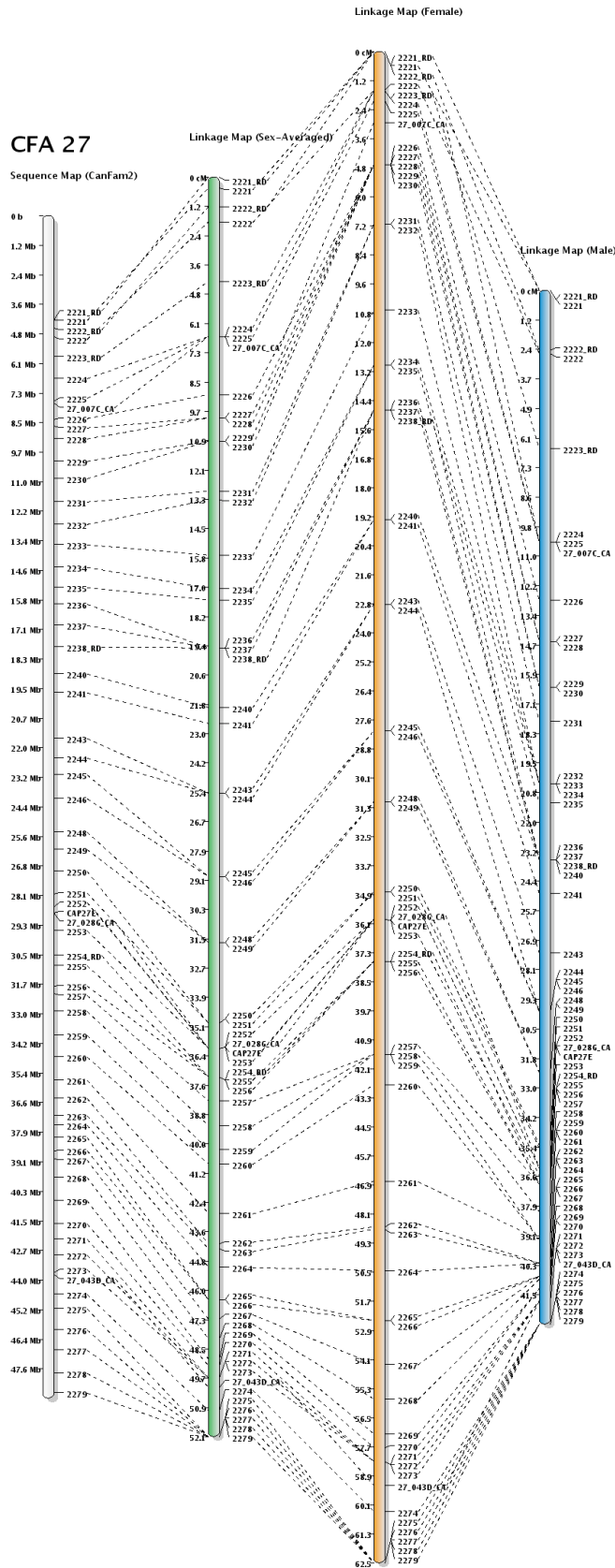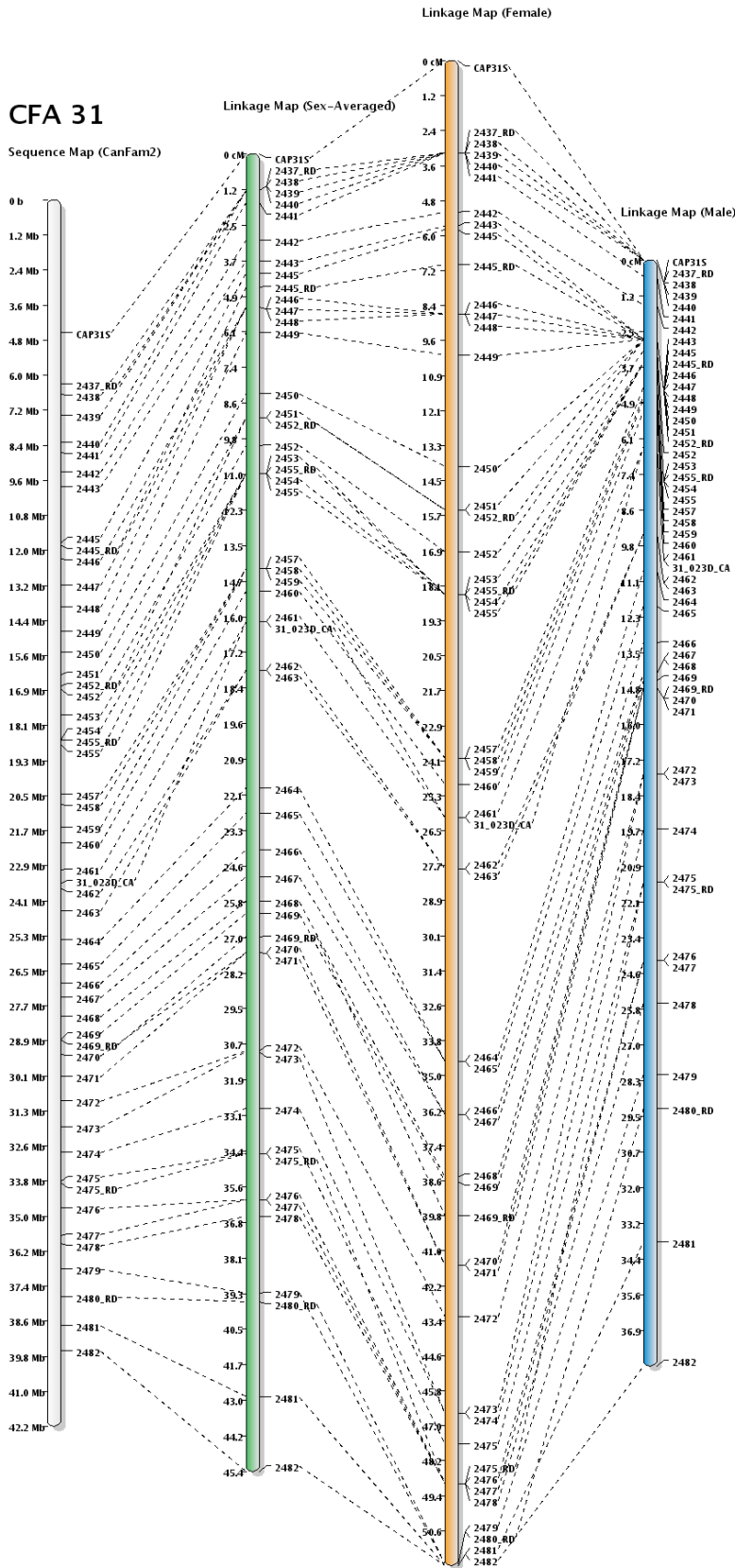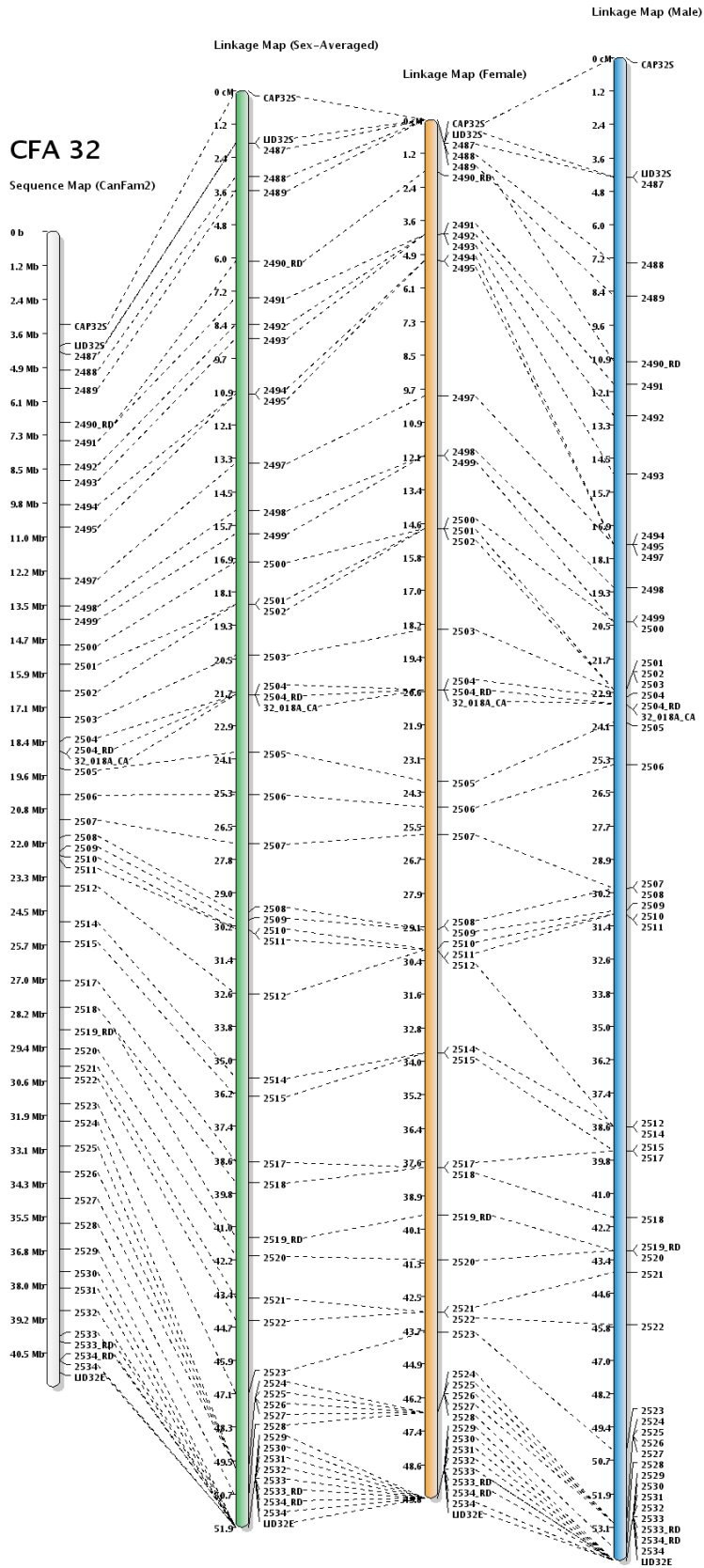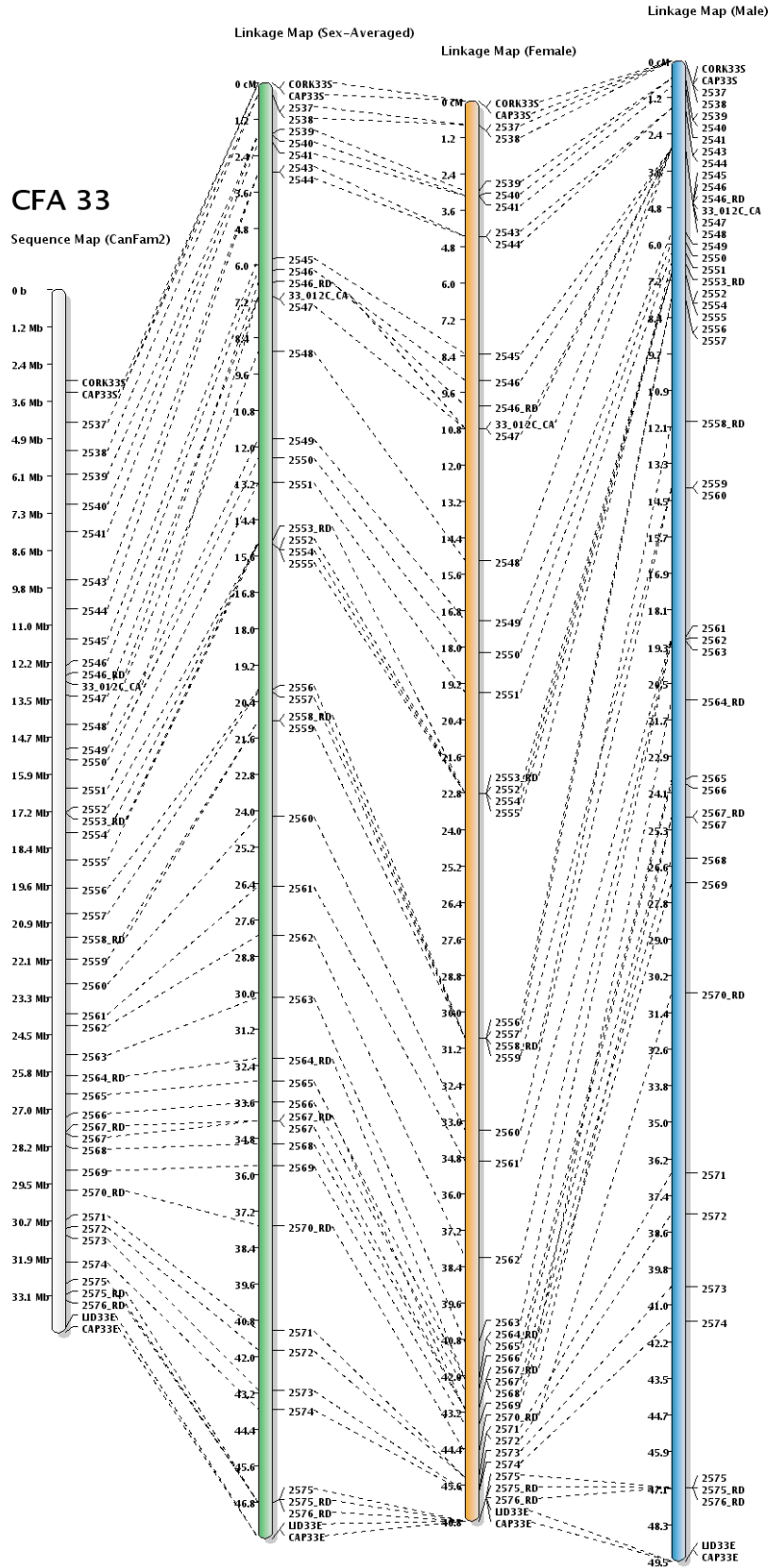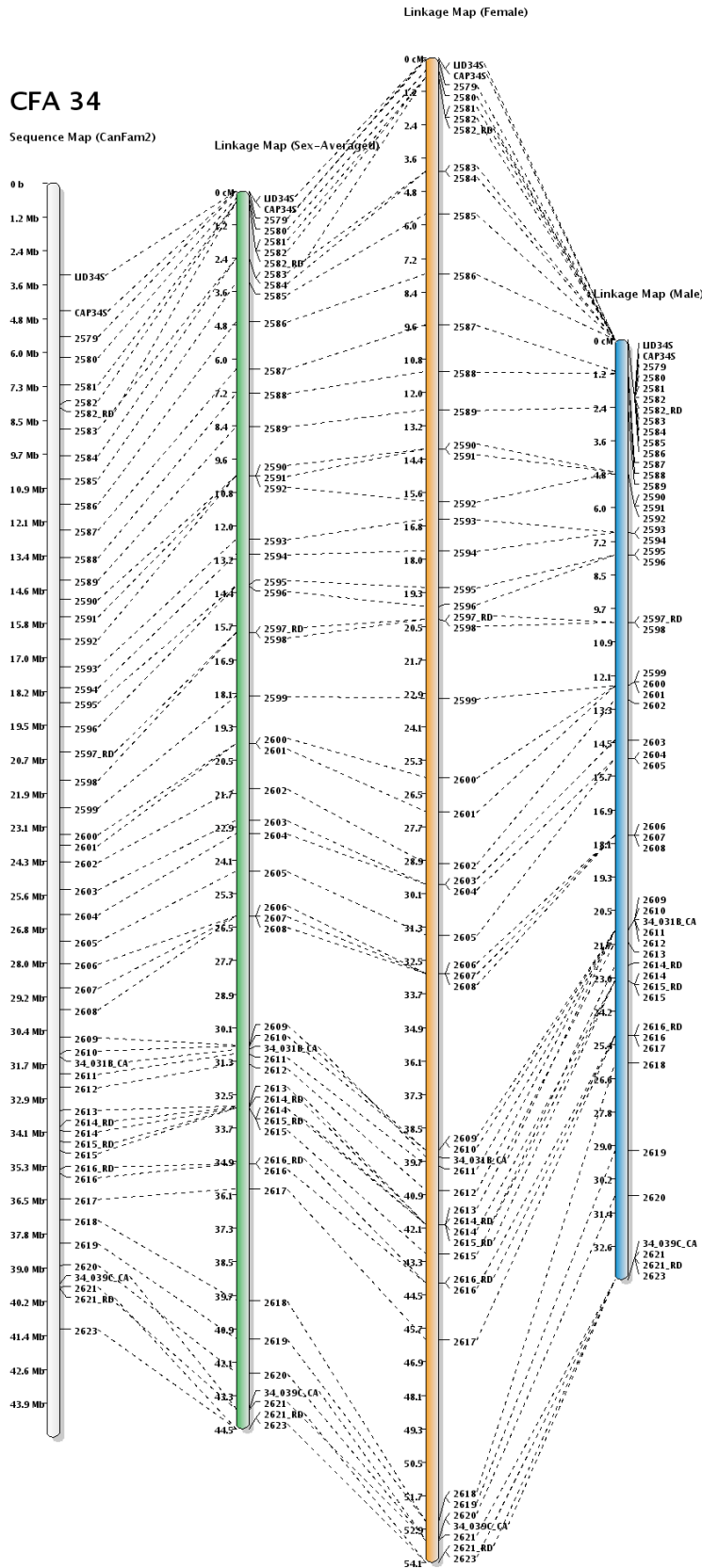
FIGURE S6.—Chromosome -specific linkage maps (CFA20-38)

TABLE S1

Panel of 36 purebred dogs for assessing marker performance

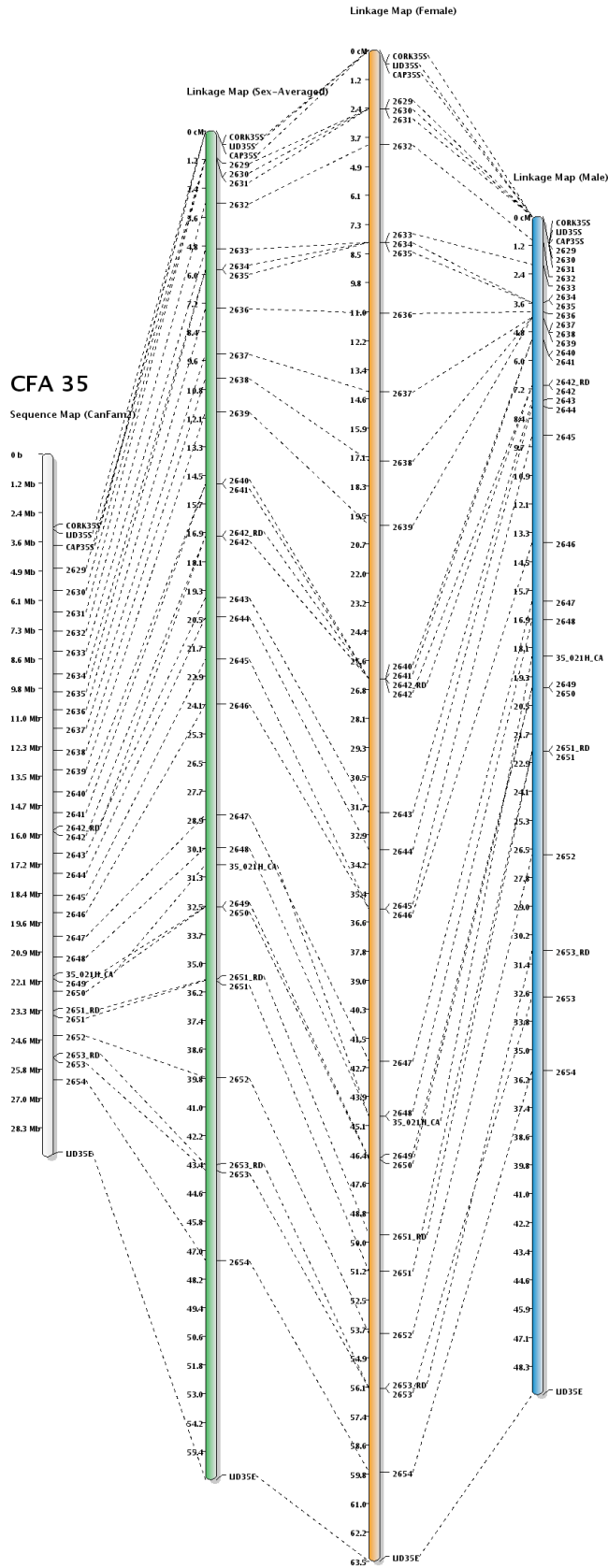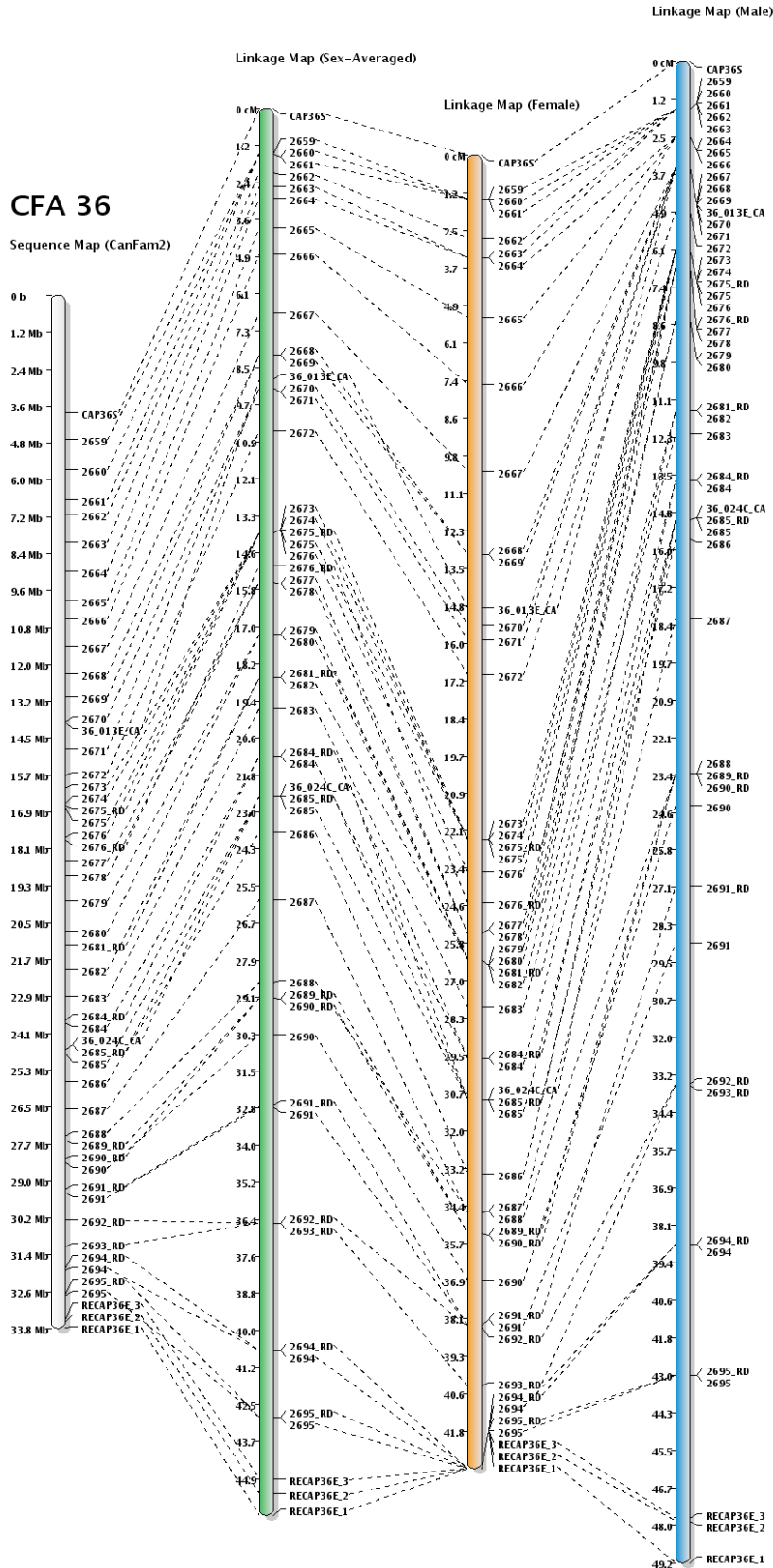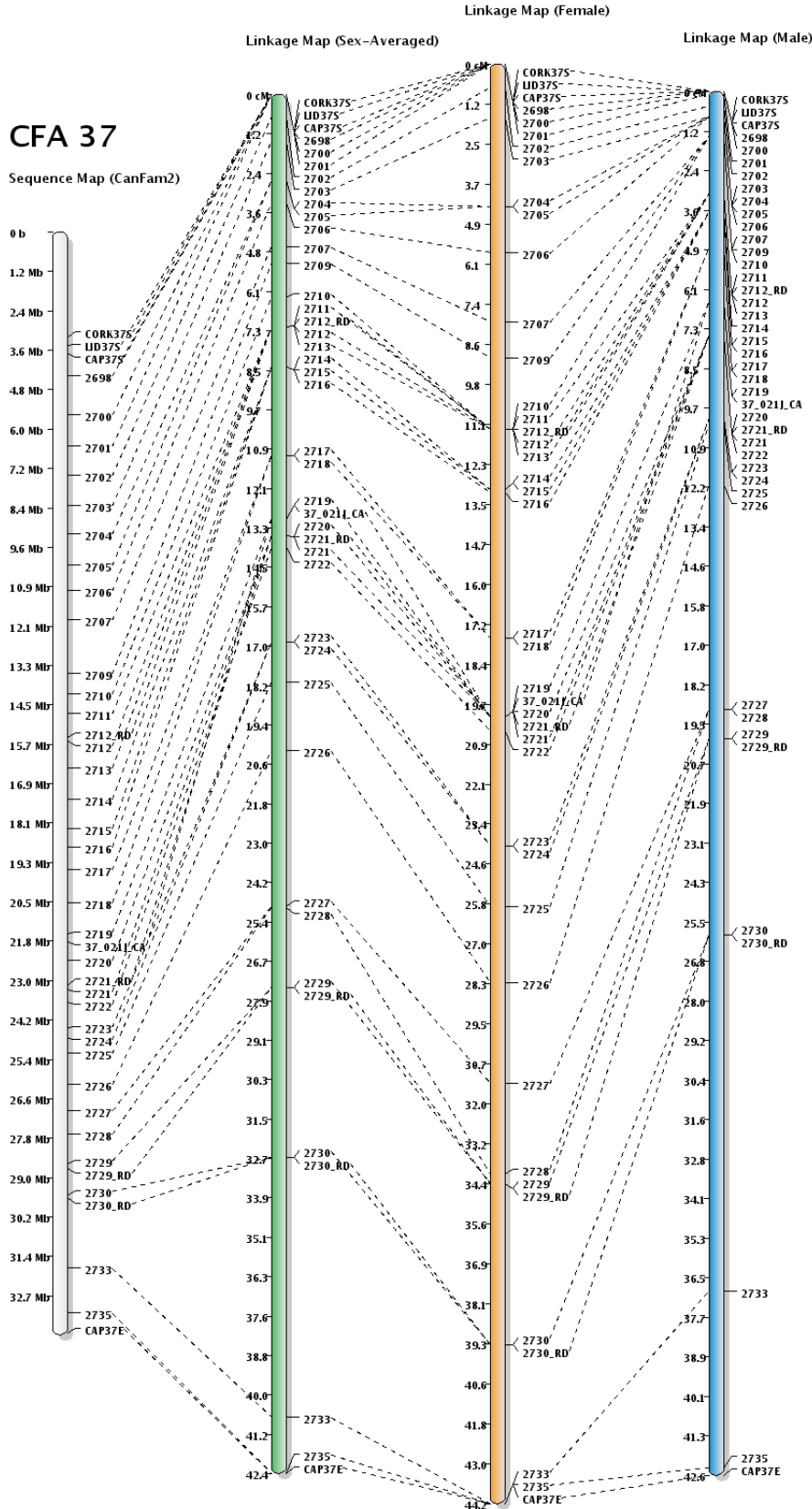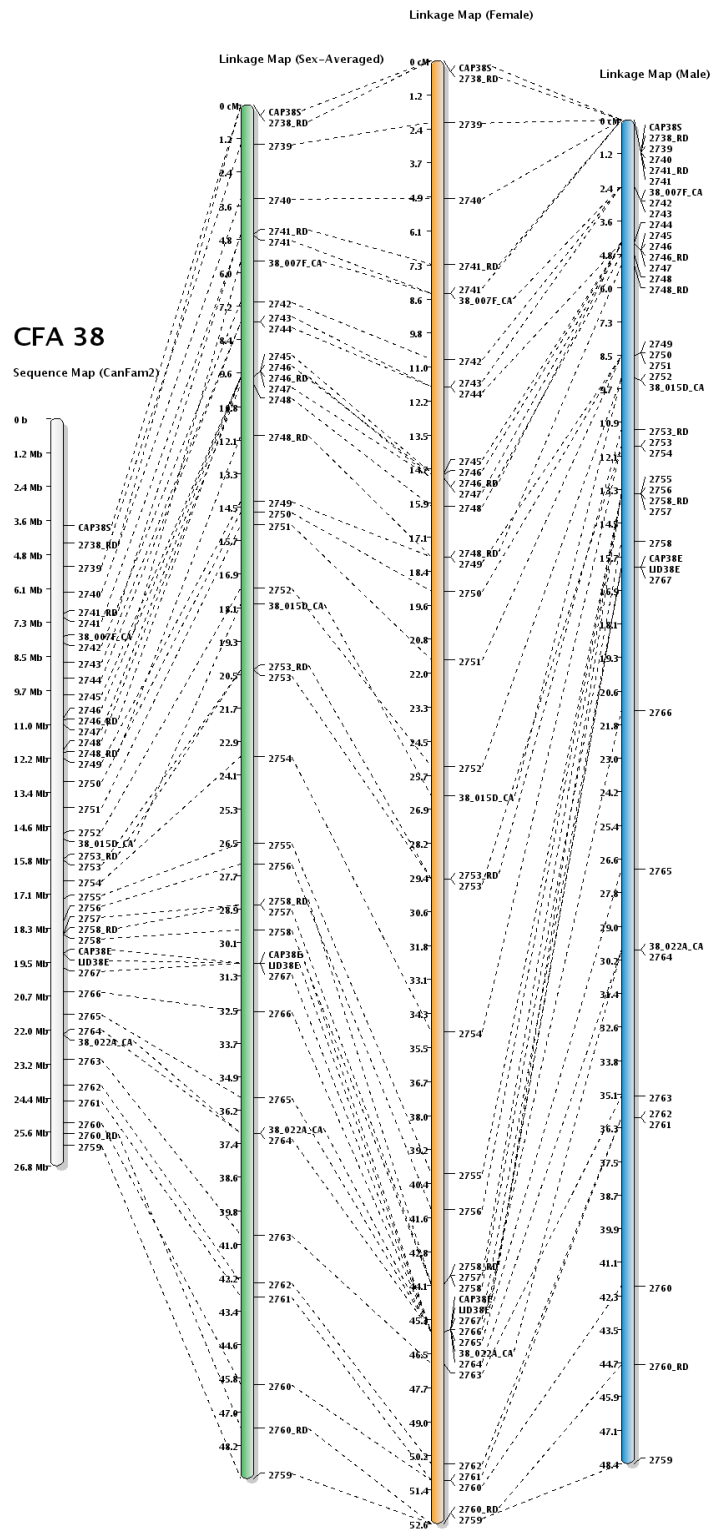| Breed	Avg. HET[a]
Afghan Hound	0.41
American Foxhound	0.47
American Staffordshire Terrier	0.49
Australian Shepherd	0.45
Bedlington Terrier	0.25
Belgian Malinois	0.49
Bichon Frise	0.30
Border Collie	0.42
Border Terrier	0.37
Boston Terrier	0.44
Boxer	0.26
Brittany	0.54
Cocker Spaniel, American	0.37
Collie	0.31
Dachshund, Miniature, Smooth	0.53
Dalmatian	0.39
Doberman Pinscher	0.40
Field Spaniel	0.33
Giant Schnauzer	0.29
Golden Retriever	0.46
Jack Russell Terrier	0.55
Kerry Blue Terrier	0.46
Labrador Retriever	0.45
Maltese	0.47
Newfoundland	0.43
Norwich Terrier	0.38
Poodle, Standard	0.39
Portuguese Water Dog	0.43
Pug	0.27
Rottweiler	0.37
Schipperke	0.38
Shetland Sheepdog	0.38
Shih Tzu	0.41
Silken Windhound	0.41
Staffordshire Bull Terrier	0.45
Whippet	0.41

[a]HET, heterozygosity

TABLE S2

Design parameters for markers in this study

Parameter	Criteria
No. of $[CA]_n$ repeats[a]	$n > 12$
Amplicon Size	100-400 bp
Primer Length	19-22 bases
T_m	59-64°C

[a]CA-repeat loci were from RepeatMasker (www.genome.ucsc.edu)

A. K. Wong *et al.*

TABLE S3

Information on markers in this study

Table S3 is available for download as an Excel file at http://www.genetics.org/cgi/content/full/genetics.109.106831/DC1

TABLE S4

Loci tagging discrepancies between map and *canFam2*

Marker	Original *CFA*[a]	Coordinates (bp)[a]	Map-based *CFA*[b]	Regional Placement (bp)[c]	
0184_RD	2	30,900,273	24	3,020,165	7,458,106
0295	3	25,980,927	25	20,208,952	27,716,347
0329	3	53,000,822	25	38,469,614	39,097,474
0652_RD	6	41,820,223	5	61,900,396	62,939,992
0698_RD	6	80,083,719	9	7,920,421	28,720,464
LID07S	7	3,522,026	8	3,141,423	8,895,661
0923_RD	9	42,018,622	12	43,243,877	53,776,997
0925_RD	9	43,986,116	23	47,077,772	50,137,361
11_005D_CA	11	5,714,494	19	33,186,887	37,614,077
CAP12S	12	3,731,042	37	32,158,941	33,876,070
1365	14	51,530,920	15	4,372,086	31,159,141
CPH4	15	42,118,984	35	18,542,875	20,254,794
1592_RD	17	49,748,339	10	37,922,525	47,598,445
1834_RD	20	59,720,787	7	22,020,315	31,919,589
LEI005	22	50,370,265	27	25,472,414	29,564,494
2054	24	11,083,049	23	47,077,772	48,456,919
2168_RD	25	52,818,756	32	3,370,596	8,996,523

[a]Physical assignment in *canFam2* assembly

[b]Genetic assignment in linkage map

[c]Physical interval of *canFam2* with greatest support for new marker assignment

A.K. Wong *et al.*

TABLE S5

Testing autosome orientation against observed chromosomal patterns of recombination rate variation

Ch.	BAC[1]	BAC Start (bp)	BAC End (bp)	Expected *canFam2*	Expected Meiotic Map	Observed (FISH)	STRP Locus[2]
27	334N08	4,595,727	4,805,510	CEN	TEL	TEL (centromere-distal)	2222_RD
27	179L15	47,755,449	47,950,156	TEL	CEN	CEN (centromere-proximal	2278
32	501F3	4,032,914	4,211,467	CEN	TEL	TEL (centromere-distal)	LID32S
32	123F17	41,075,715	41,289,726	TEL	CEN	CEN (centromere-proximal	LID32E

[1] BAC clones were obtained from the Children's Hospital Oakland Research Institute (CH82 Library).

[2] STRPs anchored individual BAC clones to specific linkage map positions. BAC clones were confirmed by STRP genotyping.

FILES S1-S4

Files S1-S4 are available for donwload as Excel files at http://www.genetics.org/cgi/content/full/genetics.109.106831/DC1.

File S1: Information on markers in this study

File S2: Interpolated map positions for SNP loci on Illumina CanineBeadSNP20

File S3: Interpolated Map Positions for previously published Minimal Screening Sets

File S4: Inter-marker map distances for multipoint analysis