Am. J. Hum. Genet. 63:861–869, 1998

Comprehensive Human Genetic Maps: Individual and Sex-Specific Variation in Recombination

Karl W. Broman,[1] Jeffrey C. Murray,[2,3] Val C. Sheffield,[2,4] Raymond L. White,[5] and James L. Weber[1]

[1]Marshfield Medical Research Foundation, Marshfield, WI; Departments of [2]Pediatrics and [3]Biology, University of Iowa, and [4]Howard Hughes Medical Institute, Iowa City; and [4]Eccles Institute for Human Genetics, University of Utah, Salt Lake City

Summary

Comprehensive human genetic maps were constructed on the basis of nearly 1 million genotypes from eight CEPH families; they incorporated >8,000 short tandem-repeat polymorphisms (STRPs), primarily from Généthon, the Cooperative Human Linkage Center, the Utah Marker Development Group, and the Marshfield Medical Research Foundation. As part of the map building process, 0.08% of the genotypes that resulted in tight double recombinants and that largely, if not entirely, represent genotyping errors, mutations, or gene-conversion events were removed. The total female, male, and sex-averaged lengths of the final maps were 44, 27, and 35 morgans, respectively. Numerous (267) sets of STRPs were identified that represented the exact same loci yet were developed independently and had different primer pairs. The distributions of the total number of recombination events per gamete, among the eight mothers of the CEPH families, were significantly different, and this variation was not due to maternal age. The female:male ratio of genetic distance varied across individual chromosomes in a remarkably consistent fashion, with peaks at the centromeres of all metacentric chromosomes. The new linkage maps plus much additional information, including a query system for use in the construction of reliably ordered maps for selected subsets of markers, are available from the Marshfield Website.

Received February 26, 1998; accepted for publication July 7, 1998; electronically published August 7, 1998.

Address for correspondence and reprints: Dr. Karl W. Broman, Center for Medical Genetics, Marshfield Medical Research Foundation, 1000 North Oak Avenue, Marshfield, WI 54449. E-mail: BromanK@cmg.mfldclin.edu

0002-9297/98/6303-0028$02.00

Introduction

Polymorphic DNA markers and their corresponding maps are an essential resource for localization of genes via linkage analysis, for characterization of meiosis, and for providing a foundation for the construction of physical maps. Although physical maps, including genome sequences, can provide the order of tightly linked polymorphisms, the physical maps do not provide genetic distances or other recombination data.

The era of human genome-scale genetic-map construction was heralded by the landmark paper by Botstein et al. (1980), in which both the use of DNA polymorphisms, as opposed to protein polymorphisms or other measurable phenotypes, in linkage mapping and an effective strategy for human genetic map construction were introduced. Development of the CEPH reference families (Dausset et al. 1990), which allowed data from many labs to be combined into a single comprehensive map, was another critical step in genetic-map construction. Although the first published genomewide human linkage map was based primarily on low-informativeness diallelic polymorphisms (Donis-Keller et al. 1987), it was the more informative and easier to type multiallelic short tandem-repeat (microsatellite) polymorphisms (STRPs) (Weber and May 1989) that permitted cost-effective construction of dense human genetic maps. Many human genetic maps covering individual chromosomes or segments of chromosomes have been described (e.g., see Tomfohrde et al. 1992; Weber et al. 1993; O'Connell et al. 1994; Shen et al. 1994; Zahn and Kwiatkowski 1995; Fain et al. 1996). Several human genomewide genetic maps also have been published, some of which include markers developed in only one laboratory (Buetow et al. 1994; Gyapay et al. 1994; Utah Marker Development Group 1995; Dib et al. 1996) and some of which include markers from many labs (NIH/CEPH Collaborative Mapping Group 1992; Matise et al. 1994; Murray et al. 1994). The most recent comprehensive maps (Matise et al. 1994; Murray et al. 1994) are now, however, outdated.

Genotyping errors dramatically inflate genetic-map

distances and usually can be identified as highly improbable tight double-recombination events that involve only a single allele in phase opposite to that of alleles from adjacent markers (Buetow 1991). Previous single-chromosome map–construction efforts have shown that the great majority of tight double-recombination events are due to genotyping errors (Tomfohrde et al. 1992; Zahn and Kwiatkowski 1995). Mutation events in which one parental allele mutates into the other also can masquerade as tight double-recombination events. Short tandem repeats are known to have relatively high germline mutation rates, on the order of 5×10^{-3}–10^{-4}/locus/gamete/generation, with changes in single repeat units greatly predominant over changes of multiple repeat units (Weber and Wong 1993; Banchs et al. 1994; Gyapay et al. 1994; Zahn and Kwiatkowski 1995; Heyer et al. 1997). The lymphoblastoid cell lines from which the CEPH family DNA was prepared also are known to have higher mutation rates than those of lymphoid cells in vivo (Weber and Wong 1993; Banchs et al. 1994). Gene-conversion events, although not yet known to occur at short tandem repeats, may be another source of tight double-recombination events.

In this study, new comprehensive human genetic maps based on >8,000 STRPs are described. Evidence is presented for significant variation in total number of recombination events per meiosis in females, but not males. Peaks in the female:male recombination ratio were found at the centromeres of all metacentric chromosomes.

Material and Methods

Genetic Markers and Genotype Data

The genetic maps incorporate a total of 8,325 STRPs, including 5,283 from Généthon (Dib et al. 1996), 1,543 from CHLC (Sheffield et al. 1995; Sunden et al. 1996), 940 from the Utah Marker Development Group (1995), 316 from Marshfield, 20 telomeric markers (Rosenberg et al. 1997), and 223 miscellaneous markers. Those from Généthon and Marshfield are almost entirely dinucleotide-repeat markers; the CHLC and Utah markers are largely trinucleotide- and tetranucleotide-repeat markers. All genotyping except that with the 20 telomeric markers was completed either at the University of Utah, Généthon, or Marshfield. Genotypes for eight CEPH families (1331, 1332, 1347, 1362, 1413, 1416, 884, and 102) were used, although the Utah markers were typed on only four of these eight families (1331, 1332, 1362, and 884). Two individuals (1332-09 and 1416-10) were typed with only a very small number of the markers.

In some cases, markers were typed more than once. These duplicate genotypes were combined, to give a single set of genotypes for each marker. In the case of a discrepancy between duplicate genotypes, the typing that did not lead to a double recombinant was retained in the data.

Map Construction

The genetic maps were constructed by use of the CRI-MAP program (Lander and Green 1987; P. Green, K. Falls, and S. Crooks, documentation for CRI-MAP, version 2.4). An initial map was formed by use of the *build* option of CRI-MAP. Additional markers were added one at a time, in decreasing order of informativeness, and were placed at the location with the highest relative likelihood. The *flips* option of CRI-MAP, using overlapping segments of ~50 markers, was used to further improve the marker orders. Large intermarker distances on the maps could indicate problems in marker order. When such gaps were observed, segments of markers were removed and then were replaced one at a time, in order of decreasing informativeness. During the process of identification of tight double recombinants (described below), we paid special attention to the appearance both of three recombination events within a short distance on a gamete and of double recombinants on several gametes in the same region; in such situations, we again attempted to improve marker order by removing a segment of markers and replacing them one at a time. Sex-averaged and sex-specific genetic distances were estimated by use of CRI-MAP with the Kosambi map function.

Cryptic Duplicate Markers

Groups of markers that we call "cryptic duplicates" were identified; these markers are at the same genetic locus and correspond to the same polymorphism, but have (in all—or nearly all—cases) different PCR primers. Pairs of markers whose members were located <20 cM from each other were compared by three methods: (i) searching for overlap between the DNA sequences at the corresponding loci; (ii) comparing allele sizes for the individuals in the eight CEPH families, looking for a simple shift in size; and (iii) looking for a concordance of individuals who were homozygous and heterozygous at the markers. DNA sequences were obtained for ~90% of the markers from Genbank (Benson et al. 1998), and sequence comparisons were performed by use of the *overlap* program in the Wisconsin Package, version 9.0 (Genetics Computer Group). Allele-size information was available for the markers from Généthon, CHLC, and Marshfield but not for those from Utah. For each group of cryptic duplicate markers, only one was retained on the maps, generally that with the largest number of completed genotypes. The others were removed, after the genotype information was combined. In the case of a discrepancy between the genotypes for two cryptic du-

plicate markers, the typing that did not result in a double recombinant was retained in the data.

Tight Double Recombinants

The data were carefully screened for the presence of tight double-recombination events, which can have a substantial impact on both the length of the genetic maps and the inferred marker orders. The data on the Généthon markers had been screened, in a previous study (Dib et al. 1996), for such double recombinants. Genotypes indicated to be errors, mutations, or gene conversions were removed from the data set. Several cycles of identification of tight double recombinants and of reordering of the markers were performed.

Tight double recombinants were found by inspection of the output from the *chrompic* option of CRI-MAP, in a search for two recombination events separated by small genetic distances (in general, <5 cM), especially when a single informative marker was present between the two recombinations. In some cases, a third recombination event was observed near the double recombinant, which could indicate an error in marker order rather than in the genotype data, and so, in such situations, a reordering of the markers was considered, prior to the removal of any genotype data.

We paid special attention to phase errors, in which all (or many) progeny in a family show double recombinants at the same location and on the same side (i.e., either all paternal or all maternal) (Tomfohrde et al. 1992). The segregation of alleles in a family was often an important clue; for example, if the two parents in a family had genotypes 1/2 and 3/4, and if the typed progeny were all either 1/3 or 2/3, with several of them showing tight double-recombination events around the marker, one could conclude that the parent observed as 3/4 probably had genotype 3/3.

Another important case was that in which an individual had double recombinants, on both the maternal chromosome and the paternal chromosome, at a particular location. This was often seen to occur in a child with heterozygous parents sharing one allele; for example, if the parents had genotypes 1/2 and 1/3 and the child had genotype 1/2, then an error causing the child's genotype to be observed as 1/3 would lead to an apparent double recombinant on both sides.

Characterization of Human Meiosis

For study of the variation in the female:male ratio for genetic distance across the chromosomes, this ratio was computed, on each chromosome, within windows of markers spanning $\geqslant 5$ cM in sex-averaged genetic distance and $\geqslant 1\frac{1}{4}$ cM in both female genetic distance and male genetic distance. These widths were chosen so that true variation would be observed and the effect of sampling error in the distance estimates would be attenuated; in particular, if the male distance within the window is allowed to be too small, the ratio may vary widely. These data were then smoothed by use of splines (Venables and Ripley 1994). The locations of the centromeres on each chromosome were estimated by use of the MIT radiation-hybrid maps (Hudson et al. 1995).

The relationship between each of female or male genetic distance and physical distance along chromosome 7 was examined further. Estimated physical locations for 263 Généthon markers on chromosome 7, described by Bouffard et al. (1997), were obtained from the NCBI Website. Because slight marker-order differences between the genetic and physical maps were observed, a subset of 238 markers for which the orders on the two maps were consistent were examined. The genetic distance:physical distance ratio, with female and male genetic distances considered separately, were then computed by use of windows spanning $\geqslant 5$ cM in sex-averaged genetic distance and $\geqslant 1\frac{1}{4}$ cM in both female genetic distance and male genetic distance, and splines were used for smoothing of the data.

For study of individual variation in recombination, the total numbers of recombination events on the 22 autosomes in each meiosis were calculated by use of the

Table 1

No. of Markers, Genetic Lengths, and Female:Male Length Ratio, for Each Chromosome

Chromosome	No. of Markers	Length (cM)			Female:Male Ratio
		Sex Averaged	Female	Male	
1	670	290	365	216	1.7
2	643	269	331	209	1.6
3	529	228	270	190	1.4
4	457	212	264	160	1.6
5	480	198	245	151	1.6
6	465	193	254	131	1.9
7	424	182	231	134	1.7
8	400	168	224	113	2.0
9	304	169	193	143	1.3
10	428	173	211	138	1.5
11	402	148	180	115	1.6
12	395	171	214	128	1.7
13	248	115	130	95	1.4
14	261	138	155	117	1.3
15	213	122	136	110	1.2
16	262	134	169	101	1.7
17	304	126	149	105	1.4
18	218	126	156	97	1.6
19	231	105	114	96	1.2
20	221	101	121	81	1.5
21	111	58	65	51	1.3
22	101	62	74	49	1.5
X	264		184		
Overall	8,031	3,488	4,435	2,730	1.6

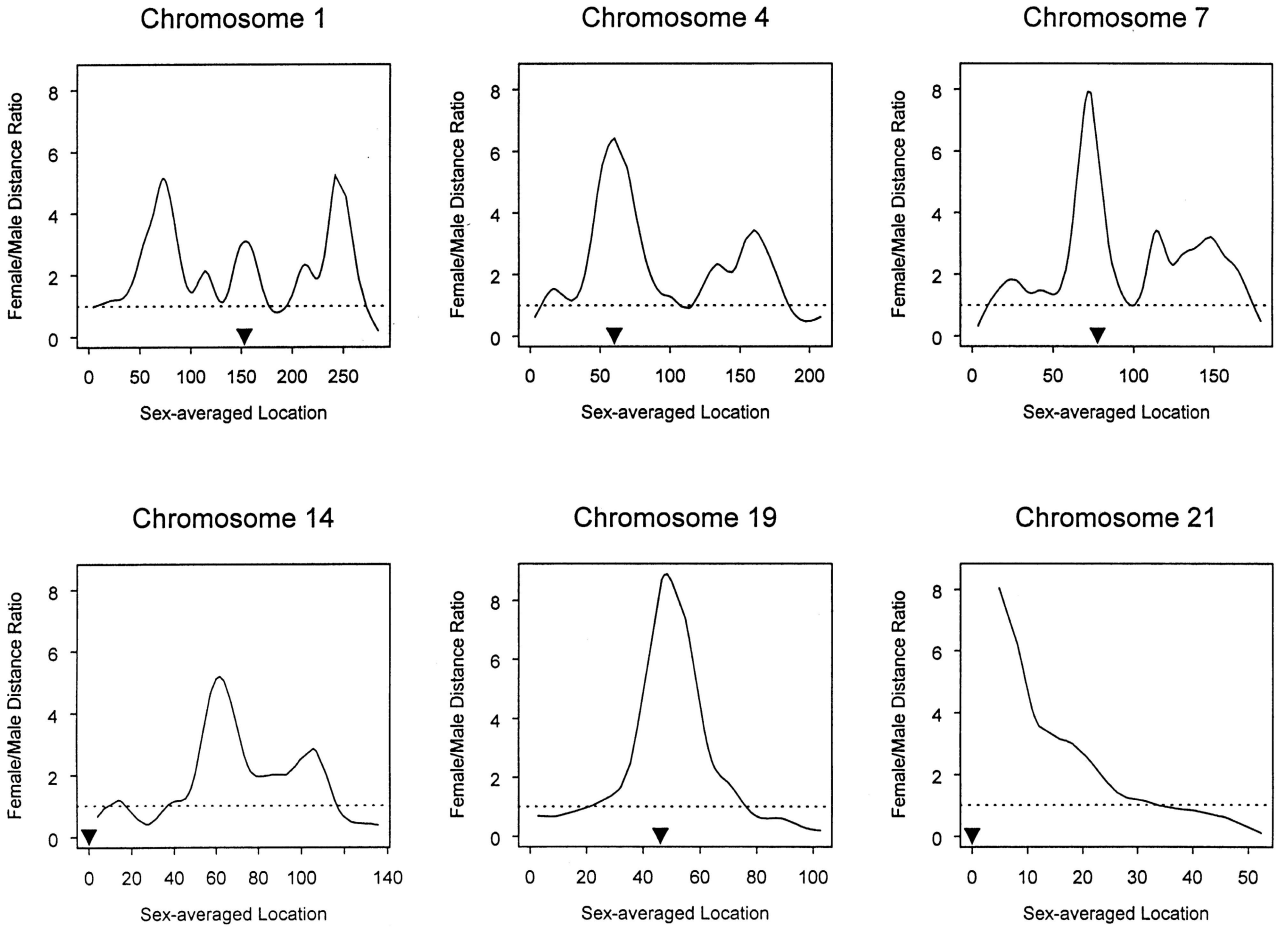

Figure 1 Plots of the female:male genetic-distance ratio against sex-averaged genetic location (in cM) along six selected chromosomes. Approximate locations of the centromeres are indicated by the triangles. The dashed lines correspond to equal female and male distances.

output from the *chrompic* option of CRI-MAP. In the counting of recombinations, events occurring within the five most telomeric markers on each end of each chromosome were ignored, since it is often difficult to determine whether these are the result of true recombination events or, rather, are genotyping errors, mutations, or gene conversions. To test the hypotheses of no recombinational variation among the mothers and among the fathers, permutation tests were performed (Manly 1997), with use of the *F* statistic from an analysis of variance. On the basis of data on the dates of birth of the parents and children in these eight CEPH families (generously provided by Mark Leppert, University of Utah), the age of each parent at the birth of each of her or his children was calculated. To test for a relationship between the age of a parent at the birth of a child and the total number of observed recombination events in the corresponding meiosis, permutation tests again were performed, this time with use of the correlation coefficient between the ages and the numbers of observed recombination events.

Results

The detailed sex-specific and sex-averaged genetic maps that we constructed are available from the Marshfield Website. Table 1 contains the number of markers; the estimated sex-averaged, male, and female genetic lengths; and the female:male length ratio for each chromosome. The female length of the entire map is 44 M, whereas the male length is 28 M. The female:male genetic-length ratio over the entire genome is 1.6. The ratios vary from 1.2 (for chromosome 15 and 19) to 2.0 (for chromosome 8). Because male recombination generally is highest at the telomeres and because terminal markers are not perfectly informative, these ratios are likely to be somewhat inflated.

Markers (or unresolved clusters of markers that were

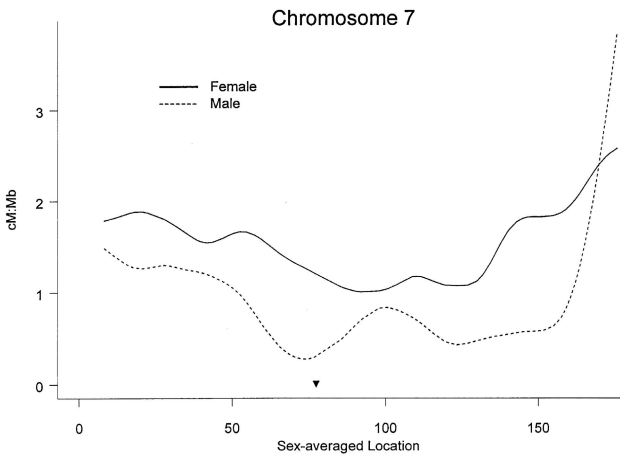

Figure 2 Plot of sex-specific genetic distance:physical distance ratio (in cM/Mb) against sex-averaged genetic location. The unbroken line was obtained by use of female genetic distance; the dashed line was obtained by use of male genetic distance. The triangle indicates the approximate location of the centromere.

not separated by recombination events) were spaced at an average ± SD of 1.3 ± 1.0, 1.8 ± 1.5, and 1.5 ± 1.5 cM in sex-averaged, female, and male genetic distance, respectively. The maximum intermarker spacing was 8.7, 12.3, and 12.9 cM in sex-averaged, female, and male genetic distance, respectively. The genetic distances are generally accurate but are subject to a large sampling error due to the relatively small number of meioses examined (188 for the autosomes and 94 for the X chromosome).

We uncovered 267 sets of cryptic duplicate markers, which corresponded to the exact same genetic locus, although with different PCR primers; these sets included 243 pairs, 21 triples, and 3 quadruples of markers. One marker from each group was retained on the map; the other 294 markers were removed. Tables displaying the observed cryptic duplicate markers are available at the Marshfield Website.

A total of 764 of 969,425 genotypes (~0.08%) were indicated to be either genotyping errors, mutations, or gene conversions and were removed from the data. As a result of removal of these genotypes and subsequent revision of the marker order, the total female and male genetic lengths changed from 58 and 39 M to 44 and 28 M, respectively, a decrease of ~25%. Of the 764 genotypes removed, 51 were removed from parents or grandparents, to eliminate phase errors; the other 713 genotypes were removed from progeny. In all but a handful of cases, a single marker, rather than two or more markers, separated the two recombination events. A subset of markers accounts for a large fraction of the genotypes removed. Eleven markers had >5% of their genotypes removed. Fifty-three markers had >2.5% of their genotypes removed, so that <1% of the markers account for >20% of the removed genotypes.

In figure 1 the female:male genetic-distance ratios across a set of six selected chromosomes are displayed. Similar plots for the remaining 16 autosomes may be obtained at the Marshfield Website. Estimates of the sampling error in these curves (data not shown) suggest that the major peaks and valleys in these curves are real. Marked variation in female:male genetic distance is observed along the chromosomes. At the telomeres of nearly all chromosomes, the female:male genetic-distance ratio approaches and often dips below 1, so that males exhibit equal or greater recombination rates in the telomeric regions. Except for the acrocentric chromosomes, distinct peaks in the female:male genetic-distance ratio are observed at all centromeres.

In figure 2 the genetic distance:physical distance ratio (in cM:Mb) is displayed for females and males, separately, along chromosome 7. Although both female and male genetic distances show large variation in physical distance across the chromosome, a distinct dip in male genetic distance, observed near the centromere and corresponding to nearly 20 Mb without male recombinations, is largely responsible for the peak in the female:male distance ratio at the centromere of chromosome 7. This region, from marker AFM288vb5 (D7S670) to marker AFM220ya3 (D7S639), spans ~20 cM in female genetic distance.

In figure 3*A* and *B,* the total number of recombination events on the 22 autosomes in the female and male meioses, respectively, by family, are displayed. The within-family means ± SDs of these totals are exhibited in table 2. Although very little variation is observed among the fathers (a permutation test gives $P \approx .30$), marked variation is observed among the mothers ($P < .001$). In figure 3*C* and *D,* the relationship between the number of observed recombinations and the parent's age at the birth of the corresponding child is shown. There is no evidence for an age-dependent effect ($P > .2$).

The following is a summary of the information available at the Marshfield Website: the detailed sex-averaged and sex-specific genetic maps, tables listing the markers' probe and locus names, tables listing the markers that were found to be cryptic duplicates, a table displaying the estimated locations of the centromeres on the genetic maps, and figures displaying the female:male genetic-distance ratio along each of the 22 autosomes (analogous to fig. 1).

In addition, the Marshfield Website provides a query system that, when given a list of markers, provides the locations of those markers on our comprehensive maps, a table giving the number of observed recombinations between the members of each pair of markers, and a framework map for the queried markers, which gives the largest subset of those markers which may be ordered

reliably. A portion of the output from this program, for a sample of 10 markers from chromosome 9p, is displayed in table 3; the numbers of informative meioses for each marker are on the diagonal, the numbers of meioses in which both members of a pair of markers were informative are above the diagonal, and the numbers of observed recombinations between the members of a pair of markers are below the diagonal. Marker GGAA20C01 (i.e., marker 5), for example, was informative in 72 meioses. Markers GGAA20C01 and AFMa312zh1 (i.e., marker 10) were jointly informative in 59 meioses, and, in those 59 meioses, exactly one recombination event was observed between them.

The type of information shown in table 3, as well as its use in the formation of reliable maps, was first described by Fain et al. (1995). Pairs of markers between whose members zero recombinations occur cannot, of course, be ordered, whereas pairs of markers between whose members one or, better, two or more recombinations occur generally can be ordered. We have developed an algorithm (to be described elsewhere) that uses the pairwise recombination information to identify the longest possible framework map that can be produced from a given set of markers, where adjacent markers on the framework map are always separated by at least a single recombination event. These maps are a further output of our Website query system.

Table 2

Observed Recombination Events on the 22 Autosomes, for Male and Female Meioses, within Each Family

	Mean ± SD No. of Recombination Events in	
Family	Mother	Father
1416	44 ± 6	22 ± 4
1413	44 ± 7	24 ± 3
1362	37 ± 4	24 ± 4
1347	47 ± 7	24 ± 3
1332	33 ± 4	22 ± 3
1331	38 ± 7	21 ± 4
884	40 ± 7	22 ± 4
102	39 ± 8	23 ± 4
Overall	40 ± 8	23 ± 4

Discussion

The chromosomal lengths on our maps (table 1) are generally consistent with those of previously published maps. Notable exceptions, however, are the female and male genetic lengths for chromosome 7, compared with those reported by Dib et al. (1996). The explanation appears to be a simple clerical error in switching the male and female intermarker distances, on the Généthon map, over the last third of the chromosome, starting at approximately marker AFM240vh4 (D7S647). The

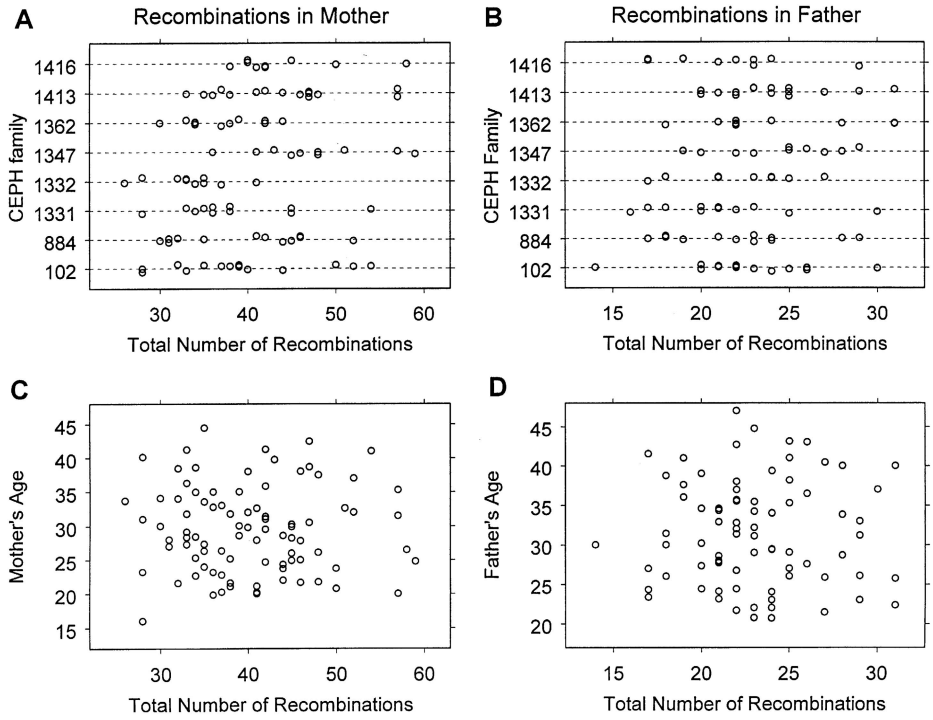

Figure 3 Total no. of observed recombination events in the 22 autosomes in each male and female meiosis, plotted by family (*A* and *B*) and against the age of the parent at the birth of the corresponding child (*C* and *D*).

Table 3

Nos. of Informative Meioses, Meioses in Which a Pair of Markers Was Informative, and Recombinations, for a Set of 10 Markers on Chromosome 9p

Marker Number (Name)	1	2	3	4	5	6	7	8	9	10
1 (ATA6D06)	102	40	67	57	62	71	94	65	59	79
2 (AFM274xe1)	0	69	40	54	22	34	57	49	29	58
3 (AFM026tg9)	0	0	106	80	41	83	84	56	74	79
4 (Mfd141)	3	4	7	135	38	69	121	83	49	111
5 (GGAA20C01)	4	2	2	0	72	53	72	52	42	59
6 (AFM345ta9)	3	2	5	1	0	108	97	72	54	74
7 (AFM242xh6)	7	4	7	4	1	1	160	109	64	125
8 (AFM362td1)	3	3	3	2	0	0	0	111	34	88
9 (AFMa184wh5)	5	3	5	0	1	0	0	0	86	59
10 (AFMa312zh1)	7	6	8	2	1	0	0	0	0	147

NOTE.—The nos. along the diagonal (underlined) are the no. of informative meioses for each marker, the nos. above the diagonal are the no. of meioses in which both members of a pair of markers were informative, and the nos. below the diagonal are the no. of observed recombinations between the members of a pair of markers.

1993–94 version of the Généthon genetic map of chromosome 7 (Gyapay et al. 1994) conforms closely to our map.

The cryptic duplicate markers are troublesome because investigators can waste substantial resources in analysis of polymorphisms that have different names and that appear to be at different loci yet, in reality, differ only in the sequences of the PCR primers. The list of cryptic duplicates described here greatly expands the list previously published by Dib et al. (1996). Although we cannot be certain that we have identified every single cryptic duplicate within the 8,325 STRPs incorporated into the maps, careful use of three different approaches to identify the duplicates (see Material and Methods) gives us confidence that the great majority have been detected.

A total of 764 genotypes (0.08%) that led to highly improbably tight double-recombination events were removed from the genotyping data set during map construction. From previous reports (e.g., see Tomfohrde et al. 1992) and on the basis of the bias, with respect to the tight double recombinants, toward a few problematic markers (see Results), we know that many of the tight double recombinants are simple lab errors. We also know that many other tight double recombinants are due to in vitro or in vivo mutation events (Weber et al. 1993). Meiotic gene-conversion events between homologues have been reported in humans (Jeffreys et al. 1994) and may be responsible for a small fraction of the tight double recombinants but probably are too rare to have a major effect on the maps. Finally, if the tight double recombinants are due to true meiotic crossing-over, then it would be expected that, because the maps are so dense, many events would involve at least a few markers. However, nearly every single genotype that was removed led to a *single* allele of the opposite phase. In conclusion, it is certainly possible that a few true recombination events were deleted through our "data-cleaning" efforts, but we strongly assert that the maps that we have produced are a much more accurate reflection of true meiotic recombination than are maps produced through the inclusion of the tight double recombinants. Such data cleaning has been a routine part of the development of many human genetic maps (e.g., see Dib et al. 1996).

The female:male genetic-distance ratio shows great variation between chromosomes, as well as along the chromosomes. Except for the acrocentric chromosomes, the telomeric regions show a female:male genetic-distance ratio $\leqslant 1$, whereas the regions around the centromeres generally show a very high female:male genetic-distance ratio (fig. 1). Increased male telomeric recombination is well known (e.g., see Weber et al. 1993; Shen et al. 1994; Zahn and Kwiatkowski 1995); however, the great excess of female recombination at the centromeres is relatively new. Mohrenweiser et al. (1998) recently described this behavior for chromosome 19; we show that it holds more generally in the genome. By comparison of sex-specific genetic distance to physical distance, on chromosome 7 (fig. 2), it was determined that the peak in the female:male genetic-distance ratio, at the centromere, was largely due to a nearly 20-Mb region in which there was no male recombination. It may be that this effect on the other chromosomes has a similar cause. Also, it is conceivable that other, noncentromeric peaks in the female:male recombination ratio may be due to the presence of suppressed latent centromeres (Choo 1998; du Sart et al. 1998).

We observed striking individual variation in the overall extent of recombination on the 22 autosomes among females, a result that did not appear in the males. This variation was clearly not a result of an age-dependent effect. This approach to the study of individual recombinational variation differs markedly from that of Yu et

al. (1996), in which, for each of five men, many sperm were genotyped at two closely spaced markers on chromosome 6. The types of variation that will be observed by the two different approaches are qualitatively different and are likely the result of quite different mechanisms (Robinson 1996).

The dramatic recombinational variation between the two sexes, across and along chromosomes, as well as the more subtle but still observable whole-genome individual recombinational variation among females, may have an impact on the statistical analyses in gene mapping. The large differences between the female and male genetic maps raises the question of whether it would be best to use sex-specific genetic maps in the search for genes, rather than the sex-averaged maps that generally are used at present. The number of meioses available for each of the sex-specific maps is half that for the sex-averaged map, however, and so the decreased precision in the sex-specific maps may counteract the benefit that their improved accuracy might provide.

Our new comprehensive maps will be useful in a wide variety of research and clinical applications. Because the maps were constructed from relatively few meioses, many of the markers cannot be ordered. Marker orders near telomeres, where genotyping errors are difficult to distinguish from true recombinations, are particularly tenuous. Nevertheless, the lists of markers for specific chromosomal regions from the maps are a valuable starting point for higher-resolution mapping. In addition, our Website query system allows investigators automatically to obtain completely ordered maps for subsets of markers.

Acknowledgments

Klaven Embertson and Jeremy Heil assisted in the collection of the CEPH-family genotyping data. Brian Hoerneman and Matt Stevenson produced preliminary results for chromosome 12. Mark Neff generously provided comments for improvement of the manuscript. This work was supported in part by National Heart, Lung and Blood Institute contract N01-HV-48141 for the Mammalian Genotyping Service.

Electronic-Database Information

URLs for data in this article are as follows:

Marshfield, http://www.marshmed.org/genetics
MIT, http://www-genome.wi.mit.edu
NCBI, http://www2.ncbi.nlm.nih.gov

References

Banchs I, Bosch A, Guimerà J, Lázaro C, Puig A, Estivill X (1994) New alleles at microsatellite loci in CEPH families mainly arise from somatic mutations in the lymphoblastoid cell lines. Hum Mutat 3:365–372

Benson DA, Boguski MS, Lipman DJ, Ostell J, Ouellette BFF (1998) Genbank. Nucleic Acids Res 26:1–7

Botstein D, White RL, Skolnick M, Davis RW (1980) Construction of a genetic linkage map in man using restriction fragment length polymorphisms. Am J Hum Genet 32: 314–331

Bouffard GG, Idol JR, Braden VV, Iyer LM, Cunningham AF, Weintraub LA, Touchman JW, et al (1997) A physical map of human chromosome 7: an integrated YAC contig map with average STS spacing of 79 kb. Genome Res 7:673–692

Buetow KH (1991) Influence of aberrant observations on high-resolution linkage analysis outcomes. Am J Hum Genet 49: 985–994

Buetow KH, Weber JL, Judwigsen S, Scherpbier-Heddema T, Duyk GM, Sheffield VC, Murray JC (1994) Integrated genome-wide maps constructed using the CEPH reference panel. Nat Genet 6:391–393

Choo KHA (1998) Turning on the centromere. Nat Genet 18: 3–4

Dausset J, Cann H, Cohen D, Lathrop M, Lalouel JM, White R (1990) Centre d'Etude du Polymorphisme Humain (CEPH): collaborative genetic mapping of the human genome. Genomics 6:575–577

Dib C, Fauré S, Fizames C, Samson D, Drouot N, Vignal A, Millasseau P, et al (1996) A comprehensive genetic map of the human genome based on 5,264 microsatellites. Nature 380:152–154

Donis-Keller H, Green P, Helms C, Cartinhour S, Weiffenbach B, Stephens K, Keith TP, et al (1987) A genetic linkage map of the human genome. Cell 51:319–337

du Sart D, Cancilla MR, Earle E, Mao JI, Saffery R, Tainton KM, Kalitsis P, et al (1998) A functional neo-centromere formed through activation of a latent human centromere and consisting of non-alpha-satellite DNA. Nat Genet 16: 144–153

Fain PR, Kort EN, Chance PF, Nguyen K, Redd DF, Econs MJ, Barker DF (1995) A 2D crossover-based map of the human X chromosome as a model for map integration. Nat Genet 9:261–266

Fain PR, Kort EN, Yousry C, James MR, Litt M (1996) A high resolution CEPH crossover mapping panel and integrated map of chromosome 11. Hum Mol Genet 5: 1631–1636

Gyapay G, Morissette J, Vignal A, Dib C, Fizames C, Millasseau P, March S, et al (1994) The 1993–1994 Généthon human genetic linkage map. Nat Genet 7:246–339

Heyer E, Puymirat J, Dieltjes P, Bakker E, de Knijff P (1997) Estimating Y chromosome specific microsatellite mutation frequencies using deep rooting pedigrees. Hum Mol Genet 6:799–803

Hudson TJ, Stein LD, Gerety S, Ma J, Castle AB, Silva J, Slonim DK, et al (1995) An STS-based map of the human genome. Science 270:1945–1954

Jeffreys AJ, Tamaki K, MacLeod A, Monckton G, Neil DL, Armour JAL (1994) Complex gene conversion events in germline mutation at human minisatellites. Nat Genet 6: 136–145

Lander ES, Green P (1987) Construction of multilocus genetic linkage maps in humans. Proc Natl Acad Sci USA 84: 2363–2367

Manly BFJ (1997) Randomization, bootstrap and Monte Carlo methods in biology. Chapman & Hall, London

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

Mohrenweiser HW, Tsujimoto S, Gordon L, Olsen AS (1998) Regions of sex-specific hypo- and hyper-recombination identified through integration of 180 genetic markers into the metric physical map of human chromosome 19. Genomics 47:153–162

Murray JC, Buetow KH, Weber JL, Ludwigsen S, Scherpbier-Heddema T, Manion F, Quillen J, et al (1994) A comprehensive human linkage map with centimorgan density. Science 265:2049–2054

NIH/CEPH Collaborative Mapping Group (1992) A comprehensive genetic linkage map of the human genome. Science 258:67–86

O'Connell P, Leach RJ, Rains D, Taylor T, Gargia D, Ballard L, Holik P, et al (1994) A PCR-based genetic map for human chromosome 3. Genomics 24:557–567

Robinson WP (1996) The extent, mechanism, and consequences of genetic variation, for recombination rate. Am J Hum Genet 59:1175–1183

Rosenberg M, Hui L, Ma J, Nusbaum HC, Clark K, Robertson L, Dziadzio L, et al (1997) Characterization of short tandem repeats from thirty-one human telomeres. Genome Res 7: 917–923

Sheffield VC, Weber JL, Buetow KH, Murray JC, Even DA, Wiles K, Gastier JM, et al (1995) A collection of tri- and tetranucleotide repeat markers used to generate high quality, high resolution human genome-wide linkage maps. Hum Mol Genet 4:1837–1844

Shen Y, Kozman HM, Thompson A, Phillips HA, Holman K, Nancarrow J, Lane S, et al (1994) A PCR based genetic linkage map of human chromosome 16. Genomics 22:68–76

Sunden SL, Businga T, Beck J, McClain A, Gastier JM, Pulido JC, Yandava CN, et al (1996) Chromosomal assignment of 2900 tri- and tetranucleotide repeat markers using NIGMS somatic cell hybrid panel 2. Genomics 32:15–20

Tomfohrde J, Wood S, Schertzer M, Wagner MJ, Wells DE, Parrish J, Sadler LA, et al (1992) Human chromosome 8 linkage map based on short tandem repeat polymorphisms: effect of genotyping errors. Genomics 14:144–152

Utah Marker Development Group (1995) A collection of ordered tetranucleotide-repeat markers from the human genome. Am J Hum Genet 57:619–628

Venables WN, Ripley BD (1994) Modern applied statistics with S-Plus. Springer-Verlag, New York

Weber JL, May PE (1989) Abundant class of human DNA polymorphisms which can be typed using the polymerase chain reaction. Am J Hum Genet 44:388–396

Weber JL, Wang Z, Hansen K, Stephenson M, Kappel C, Salzman S, Wilkie PJ, et al (1993) Evidence for human meiotic recombination interference obtained through construction of a short tandem repeat–polymorphism linkage map of chromosome 19. Am J Hum Genet 53:1079–1095

Weber JL, Wong C (1993) Mutation in human short tandem repeat polymorphisms. Hum Mol Genet 2:1123–1128

Yu J, Lazzeroni L, Qin J, Huang M-M, Navidi W, Erlich H, Arnheim N (1996) Individual variation in recombination among human males. Am J Hum Genet 59:1186–1192

Zahn LM, Kwiatkowski DJ (1995) A 37-marker PCR-based genetic linkage map of human chromosome 9: observations on mutations and positive interference. Genomics 28: 140–146