Am. J. Hum. Genet. 66:1911–1926, 2000

Characterization of Human Crossover Interference

Karl W. Broman and James L. Weber

Marshfield Medical Research Foundation, Marshfield, WI

We present an analysis of crossover interference over the entire human genome, on the basis of genotype data from more than 8,000 polymorphisms in eight CEPH families. Overwhelming evidence was found for strong positive crossover interference, with average strength lying between the levels of interference implied by the Kosambi and Carter-Falconer map functions. Five mathematical models of interference were evaluated: the gamma model and four versions of the count-location model. The gamma model fit the data far better than did any of the other four models. Analysis of intercrossover distances was greatly superior to the analysis of crossover counts, in both demonstrating interference and distinguishing between the five models. In contrast to earlier suggestions, interference was found to continue uninterrupted across the centromeres. No convincing differences in the levels of interference were found between the sexes or among chromosomes; however, we did detect possible individual variation in interference among the eight mothers. Finally, we present an equation that provides the probability of the occurrence of a double crossover between two nonrecombinant, informative polymorphisms.

Introduction

Crossover interference may be defined as the nonrandom placement of crossovers along chromosomes in meiosis. Interference was identified soon after the development of the first working models for the recombination process (Sturtevant 1915; Muller 1916). Strong evidence for positive crossover interference (with crossovers more evenly spaced than would be expected with random placement) has been obtained in many species (Zhao et al. 1995*b*). Investigations of interference in experimental organisms have generally involved observed frequencies of rare multiple-recombination events in sets of adjacent intervals (Zhao et al. 1995*b*). Such an approach requires many thousands of meioses, each of which must be informative for the same set of markers. Weinstein (1936), for example, studied seven loci in 28,239 *Drosophila melanogaster* offspring.

Meiotic recombination occurs after the chromosomes have duplicated. Homologous chromosome pairs line up together, forming tight bundles of four chromatids. Nonsister chromatids then synapse and exchange material; the locations at which this occurs are called chiasmata. The chiasmata are resolved as crossovers in the four products of meiosis.

Interference is generally split into two aspects: chromatid interference and crossover interference. Chromatid interference is a dependence in the choice of strands involved in adjacent chiasmata. There is little consistent evidence for the presence of chromatid interference in experimental organisms (Zhao et al. 1995*a*), and any inference with regard to chromatid interference generally requires that data be available for all four products of meiosis (so-called "tetrad data"); such data are not available in humans. Thus, throughout the present study, we assume there is no chromatid interference. Crossover interference (also known as "chiasma interference") is defined as the nonrandom placement of chiasmata on individual chromatids. Under positive crossover interference, chiasmata are more evenly spaced, whereas, under negative crossover interference, they are more clustered. Meiosis in experimental organisms generally shows positive crossover interference (Zhao et al. 1995*b*), although exceptions do exist (Munz 1994). Interference is also under genetic control (Sym and Roeder 1994).

Interference is important in meiosis, in that, given a limited number of chiasmata per meiosis genomewide, interference results in chiasmata being more evenly distributed across chromosomes. Thus, interference ensures that the smallest chromosomes will have at least one chiasma, which is necessary for the proper segregation of chromosomes (reviewed by Egel [1995] and Roeder [1997]). Yeast mutants for which interference is absent show a greater rate of nondisjunction (Sym and Roeder 1994; Chua and Roeder 1997).

Good evidence for positive interference in humans exists, although a detailed characterization has not yet been achieved. Cytogenetic evidence for interference in males has been obtained by Hultén and colleagues (Hultén 1974; Laurie and Hultén 1985*a*, 1985*b*; Hultén et

Received February 7, 2000; accepted for publication March 24, 2000; electronically published May 11, 2000.

Address for correspondence and reprints: Dr. Karl W. Broman, Department of Biostatistics, Johns Hopkins University, 615 N. Wolfe Street, Baltimore, MD 21205. E-mail: kbroman@jhsph.edu

0002-9297/2000/6606-0020$02.00

al. 1990; Povey et al. 1992), through analysis of chiasma locations in spermatocytes. Several groups (Haines et al. 1993; Kwiatkowski et al. 1993; McInnis et al. 1993; Weber et al. 1993; Zahn and Kwiatkowski 1995) have shown that the distribution of the number of crossovers per chromosome differs significantly from that which is expected under the assumption of no interference. Similar observations have been made in mice (see, for example, Blank et al. 1988; Ceci et al. 1989; Kingsley et al. 1989). In copies of human chromosome 19 with exactly two recombination events, the distribution of the genetic distance between crossovers showed spacing that was more distant than that which would be expected in the absence of interference (Weber et al. 1993). Additional evidence has come from estimation of the Rao map function parameter p (Rao et al. 1977), which is obtained by combination of pairwise recombination fractions (Shields et al. 1991; Collins et al. 1992; Lawrence et al. 1993; Attwood et al. 1994; Collins et al. 1995; Cox et al. 1995; Forabosco et al. 1995; Collins et al. 1996). Finally, Arnheim and colleagues found evidence for positive interference by typing, in single human sperm, multiple markers in a defined pseudoautosomal interval (Schmitt et al. 1994).

A further feature of the recombination process is worthy of note: the centromere has often been assumed to play a special role in recombination, in that it has been thought to serve as the origin of chiasma formation and, thus, as a barrier to interference (Mather 1938). However, there appears to be little evidence for this assumption, and Colombo and Jones (1997) have provided evidence that, in the grasshopper, interference does act across the centromere and that, in fact, the level of interference around the centromere appears to be no different than that seen for any other segment of the chromosome. The results of cytogenetic studies in the human male (Laurie and Hultén 1985*b*) have similarly shown that the number of chiasmata on the two arms of chromosome 18 are not independent.

Traditional approaches to studying interference involve the use of either the coincidence measure for two disjoint intervals or map functions. The coincidence, c, is defined as the ratio of the probability that recombination events occur in both of two disjoint intervals to that which would be expected under no interference. Analysis generally proceeds by testing for $c = 1$ with the use of data on three adjacent markers—an approach that suffers from very low power to detect interference (Weeks et al. 1994). More commonly, map functions are used as a measure of interference. A map function relates the recombination probability for an interval to the average number of crossovers, per meiotic product, in that interval (i.e., its length in genetic distance). Examples of map functions include those of Haldane (1919), Kosambi (1944), Carter and Falconer (1951), Sturt (1976), and Rao et al. (1977). The meaning and use of map functions are reviewed elsewhere (Speed 1996; Zhao and Speed 1996).

Numerous mathematical models for recombination that incorporate interference have been developed (for reviews, see Karlin and Liberman [1994] and McPeek and Speed [1995]). In the present study, we consider two types of models: the count-location (CL) model and the gamma model. Karlin and Liberman (1978, 1979) and Risch and Lange (1979) independently developed the CL model, in which the number of chiasmata on the four-strand bundle follows some distribution $\mathbf{p} = (p_0, p_1, p_2, \ldots)$ and in which the locations of the chiasmata, given their number, are obtained by tossing them down uniformly at random. In other words, the chiasma locations are independent and identically distributed Uniform$(0,L)$, where L is the length of the chromosome in genetic distance. In considering a possible mechanism underlying the CL model, one might imagine that the number of chiasmata on the four-strand bundle is under biological control, but the locations of the chiasmata are allowed to vary freely. Several special cases of this model are important to mention. First, one may require that $p_0 = 0$ (i.e., there must be at least one chiasma on the four-strand bundle), which we will call the "obligate-chiasma CL model." Second, in the case that the distribution $\mathbf{p}$ is a Poisson distribution, where $p_x = e^{-\lambda}\lambda^x/x!$, one obtains the no-interference model. Finally, one may take $p_0 = 0$ and $p_x = e^{-\lambda}\lambda^x/[x!(1 - e^{-\lambda})]$ for $x > 0$, for a model in which there is an obligate chiasma but in which there otherwise is no interference. We will call this the "truncated Poisson model." This model is similar to that which was proposed by Sturt (1976).

Much recent interest has been focused on gamma models, in which the locations of the chiasmata on the four-strand bundle are determined according to a stationary renewal process, with increments gamma distributed with shape and rate parameters ν and 2ν, respectively, for $\nu > 0$. In other words, the distances between the chiasmata are independent and follow a gamma distribution with a mean and SD of 1/2 and $1/(2\sqrt{\nu})$ Morgans, respectively. (For a detailed discussion of renewal processes, see Cox [1962].) Under the assumption of no chromatid interference, the locations of crossovers on a random meiotic product are obtained by "thinning" the chiasma process: chiasmata on the four-strand bundle are retained as crossovers independently, with a probability of 1/2. The shape and rate parameters of the gamma model satisfy the constraint that the average interchiasma distance is 0.5 Morgans, and, therefore, the average intercrossover distance is 1 Morgan. The parameter ν is a unitless measure of the strength of interference: the case $\nu = 1$ corresponds to no interference, and $\nu > 1$ (< 1) corresponds to positive (negative) crossover interference. Such models have a

long history (see McPeek and Speed 1995), having first been proposed by Fisher et al. (1947). Foss et al. (1993) and Foss and Stahl (1995) revived interest in these models, after describing a mechanism for recombination that gives rise to such models. In their biological model, chiasmata must be separated by a fixed number, m, of intermediate gene-conversion events. If the locations of the chiasmata and the intermediate events are at random (i.e., according to a Poisson process), then the locations of the chiasmata are determined according to the χ^2 model, which is a special case of the gamma model, with $\nu = m + 1$ for a nonnegative integer m, and which is thus denoted because the gamma distribution with shape and rate parameters $m + 1$ and $2(m + 1)$, respectively, is a scaled version of a χ^2 distribution with $2(m + 1)$ df. The χ^2 model was also considered by other groups (Zhao et al. 1995*b*; Lin and Speed 1996).

We recently constructed comprehensive human genetic maps that were based on nearly 1 million genotypes for eight of the reference families from CEPH and that incorporated more than 8,000 short-tandem-repeat polymorphisms (STRPs) from several laboratories (Broman et al. 1998). These data, which consist of nearly complete genotypes on highly polymorphic markers at a density of approximately one marker per 0.5 cM, allow for a relatively precise localization of all recombination events in the corresponding meioses. The data were fit to five mathematical models for interference in recombination: the gamma model and four versions of the CL model (including the no-interference model). The results of the present study provide strong, genomewide evidence for positive human crossover interference. A proper understanding of interference in human meiosis is valuable, in that one may then (*a*) use realistic models in computer simulations of recombination, in studies of the properties of statistical methods to map disease genes, (*b*) develop new statistical methods for finding genes—methods that take proper account of crossover interference, and (*c*) better gauge the chance of a double crossover in an interval when one follows the transmission of a disease allele through a pedigree with the use of flanking genetic markers.

Material and Methods

Genetic Markers and Genotype Data

We used the data described by Broman et al. (1998)—data that we summarize again here. We considered eight of the CEPH families (families 102, 884, 1331, 1332, 1347, 1362, 1413, and 1416), excluding individuals 1332–09 and 1416–10, for whom few genotype data were available. These families include a total of 134 individuals, including 92 progeny, 16 parents, and 26 grandparents. We considered genotypes for 8,010 STRPs, including 5,279 from Généthon (Dib et al. 1996), 1,376 from the Cooperative Human Linkage Center (Sheffield et al. 1995; Sunden et al. 1996), 854 from the Utah Marker Development Group (1995), 285 from the Center for Medical Genetics, Marshfield Medical Research Foundation, 35 telomeric markers (Rosenberg et al. 1997), and 181 miscellaneous markers. The genotypes for families 884, 1331, 1332, and 1362 were 96% complete. Families 102, 1347, 1413, and 1416 were not typed for the Utah Marker Development Group markers, but their genotypes were 94% complete for the other 7,156 markers. All of these data are available from the Center for Medical Genetics, Marshfield Medical Research Foundation.

The genetic maps and marker order were the same as those determined by Broman et al. (1998). The CRI-MAP program (Lander and Green 1987; P. Green, K. Falls, and S. Crooks, documentation for CRI-MAP, version 2.4) was used to estimate the recombination fractions between adjacent markers, and the Kosambi map function (Kosambi 1944) was used to convert the recombination fractions to genetic distances. Because of the unusually high density of markers in these data, the choice of map function has little influence on the estimated genetic distances. The average distance between markers was ~0.5 cM. The location of the centromere on each chromosome was estimated by use of the radiation-hybrid maps from the Whitehead Institute for Biomedical Research/MIT Center for Genome Research (Hudson et al. 1995). Sex-specific genetic distances were used in all analyses, because of the large difference between female and male genetic distance.

Genotypes causing tight apparent double-recombination events, which are indicative of genotyping errors, mutations, or, possibly, gene conversions, were removed. (For further details, see the Results section and Broman et al. [1998].)

The *chrompic* option of CRI-MAP was used to infer phase in each of the 92 progeny, for each chromosome. These data were converted to crossover locations. Although, in reality, each crossover can only be localized to a position within the interval between the two informative markers flanking the corresponding recombination event (i.e., the locations of the crossovers are interval censored), the intervals into which the crossovers could be placed were generally quite small. For example, figure 1 shows the grandparental origins of DNA in maternal chromosomes 2 from the progeny in family 1331. The blackened bars, which represent censoring of the crossover locations, are quite small, especially in comparison to the distances between crossovers. Each crossover was assumed to have occurred at the midpoint of the interval between its two flanking informative markers, and interval censoring was ignored. We further assumed that all cross-

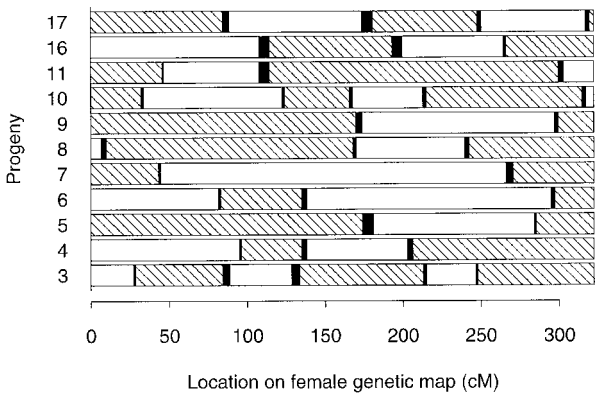

Figure 1 Grandparental phase in the maternal chromosomes 2, for the 11 progeny from family 1331. Hatched segments denote grandpaternal origin; unblackened segments, grandmaternal origin. The smaller blackened segments are the noninformative segments in which a recombination occurred.

overs were observed (i.e., that no double crossovers between informative markers occurred), and, therefore, we will frequently use the term *"crossover,"* rather than the term *"recombination event,"* throughout the remainder of the present article.

The assumption that all crossovers were observed relies on the high density of markers typed in these families. Because the telomeres exhibit a lower marker density and because it is difficult to clean the telomeric data appropriately (at the telomeres, it is difficult to distinguish true double-recombination events from false double-recombination events), we considered only the recombination data that were five or more markers away from the telomeres. In addition, we dropped the chromosomes where the maximum distance between adjacent informative markers was ⩾14 cM. In 12 cases, the data for all of the progeny in a family were dropped for a particular chromosome: for maternal chromosomes 4–6, 8, 12, and 18 and paternal chromosome 6 in family 884, for paternal chromosome 9 in family 1331, for maternal chromosome 15 and paternal chromosome 3 in family 1332, for paternal chromosome 2 in family 1413, and for maternal chromosome 1 in family 1416. In 17 other cases, data for just one of the progeny in a family were dropped for a particular chromosome. The numerous gaps in the chromosomes of family 884 were the result of long homozygous segments, which were likely the result of autozygosity, in the parents of that family (Broman and Weber 1999).

Fit of the CL Models

CL models were first described by Karlin and Liberman (1978, 1979) and Risch and Lange (1979). Under these models, the number of chiasmata on the four-strand bundle, n, follows some distribution $\mathbf{p} = (p_0,p_1,p_2,\ldots)$, and, given n, the locations of the chiasmata are independent and identically distributed Uniform(0,L), where L is the length of the chromosome in genetic distance. Under the assumption of no chromatid interference, the number of crossovers, m, on a random meiotic product has the distribution

$$\Pr(m = i) = \sum_{n=i}^{\infty} p_n \binom{n}{i} \left(\frac{1}{2}\right)^n .$$

We considered four CL models: the no-interference model, for which the distribution $\mathbf{p}$ is Poisson(λ), with $p_n = e^{-\lambda}\lambda^n/n!$; the truncated Poisson model (with no interference but with an obligate chiasma on the four-strand bundle), for which $p_n = e^{-\lambda}\lambda^n/[n!(1 - e^{-\lambda})]$ for $n \geqslant 1$; a model with the p_n varying freely (hereafter called *"the* CL model"); and a model with $p_0 = 0$ and with the other p_n varying freely (hereafter called the "obligate-chiasma CL model").

To fit the CL models, we may ignore the locations of the crossovers and may simply consider the number of crossovers, m_i, on each of the N meiotic products. The Poisson model is fit by taking simply $\hat{\lambda} = 2\sum_i m_i/N$. The other three models are fit by use of a version of the EM algorithm (Dempster et al. 1977), as described by Ott (1996). Let n_i be the (unobserved) number of chiasmata on the ith four-strand bundle. We begin with initial values for the distribution of the number of chiasmata on the four-strand bundle, $\hat{\mathbf{p}}^{(0)}$. At the kth iteration, we calculate, for each i and j,

$$w_{ij}^{(k)} = \Pr[n_i = j | m_i, \hat{\mathbf{p}}^{(k-1)}] = \frac{\hat{p}_j^{(k-1)} \binom{j}{m_i} \left(\frac{1}{2}\right)^j}{\sum_{s \geqslant m_i} \hat{p}_s^{(k-1)} \binom{s}{m_i} \left(\frac{1}{2}\right)^s} .$$

For the truncated Poisson model, we then obtain $\hat{\lambda}^{(k)} = \Sigma_{i=1}^{N}\Sigma_{j=1}^{J} j w_{ij}^{(k)}/N$. For the other two models, we take $\hat{p}_j^{(k)} = \Sigma_{i=1}^{N}\Sigma_{j=0}^{J} w_{ij}^{(k)}/N$. In both of the aforementioned cases, the sum over j is taken up to some maximum value J for which the probability has become quite small—for example, $J = 2\max\{m_i\} + 5$. It can be shown that, in the CL and obligate-chiasma CL models, one may take $J = 2\max\{m_i\} - 1$ when $\max\{m_i\} \geqslant 1$ (Yu and Feingold 1998). The truncated Poisson model, however, often requires a somewhat larger J.

Note that, in the no-interference and truncated Poisson models, the parameter λ is equivalent to the length of the chromosome in genetic distance L. Rather than use the value L estimated from the genetic data, we chose to reestimate λ from the data on crossover counts,

thereby enabling the best fit of the models to the crossover-count data. The resulting estimates $\hat{\lambda}$ correspond very closely to what would be expected, given the estimates L.

Fit of the Gamma Model

Let $x_0, x_1, x_2, \ldots$ be the genetic distances between chiasmata on the four-strand bundle, with x_0 denoting the distance from pter to the first chiasma. Under the gamma model, the locations of the chiasmata on the four-strand bundle are according to a stationary gamma renewal process, and, therefore, the x_i are independent and $x_1, x_2, \ldots$ follow a gamma distribution with shape and rate parameters ν and 2ν, respectively. Thus, the x_i have the density $f(x;\nu) = e^{-2\nu x}(2\nu)^{\nu}x^{\nu-1}/\Gamma(\nu)$. The density of x_0 is that which is required to give stationarity: $g(x;\nu) = 2[1 - F(x;\nu)]$, where $F(x;\nu)$ is the cumulative distribution function (cdf) of $f(x;\nu)$. Note that $\nu > 0$ and that $\nu = 1$ corresponds to no interference, whereas $\nu > 1$ (< 1) corresponds to positive (negative) crossover interference.

Let $y_0, y_1, y_2, \ldots$ be the genetic distances between crossovers on a random meiotic product. Under no chromatid interference, the locations of the crossovers are obtained by means of thinning the chiasma locations independently, with a probability of 1/2. Thus, the y_i are independent, and $y_1, y_2, \ldots$ have the density $f^*(y;\nu) = \Sigma_{k=1}^{\infty}(1/2)^k f_k(y;\nu)$, where $f_k(y;\nu)$ is the density of a gamma distribution with shape and rate parameters $k\nu$ and 2ν—the convolution of $f(y;\nu)$ with itself k times. The density of y_0 is

$$g^*(y;\nu) = (1/2)g(y;\nu) + \Sigma_{k=1}^{\infty}(1/2)^{k+1}(g * f_k)(y;\nu)$$
$$= 1 - F^*(y;\nu) \ ,$$

where $F^*(y;\nu)$ is the cdf of $f^*(y;\nu)$. Let $G^*(y;\nu)$ denote the cdf of $g^*(y;\nu)$.

Now let $y_{i0}, y_{i1}, \ldots, y_{im_i}$ be the genetic distances between crossovers on the maternal (or paternal) chromosome in individual i, with y_{i0} denoting the genetic distance from pter to the first crossover, with y_{im_i} denoting the genetic distance from the final crossover to qter, and with m_i denoting the number of crossovers on the chromosome. If $m_i = 0$, then $y_{i0} = L$, the genetic length of the chromosome. The ith individual's contribution to the log likelihood for ν is

$$l(\nu;\mathbf{y}_i) = \begin{cases} \log[1 - G^*(L;\nu)] & \text{if } m_i = 0 \\ \log\ g^*(y_{i0};\nu) + \log\ g^*(y_{i1};\nu) & \text{if } m_i = 1 \\ \log\ g^*(y_{i0};\nu) + \sum_{j=1}^{m_i-1} \log\ f^*(y_{ij};\nu) + \log g^*(y_{im_i};\nu) & \text{otherwise} \end{cases} .$$

Note that y_{im_i} has been treated as a right-censored observation, with density g^* if $m_i = 0$ and with density f^* if $m_i > 0$. The log likelihood for the complete data may be obtained as the sum over the individual contributions.

For calculation of the log-likelihood function at any value ν, $f^*(y;\nu)$ was calculated by use of its explicit formula, summing over k from 0 to 25. Calculation of $g^*(y;\nu)$ required a numerical approximation to the incomplete gamma function, which was obtained by an algorithm from Press et al. (1992, pp 216–219), and the convolution of $g(y;\nu)$ with $f^*(y;\nu)$, which was obtained by numerical integration using Simpson's rule (Press et al. 1992, pp 136–139). Calculation of $G^*(y;\nu)$ required the numerical integration of $g^*(y;\nu)$, which was also performed using Simpson's rule. The maximum-likelihood estimate $\hat{\nu}$ was obtained by maximization of the log likelihood by Brent's method (Brent 1973; Press et al. 1992, pp 402–405).

The variance of $\hat{\nu}$ was estimated by the reciprocal of the observed Fisher information, $-1/l''(\hat{\nu})$. The second derivative of the log-likelihood function was approximated numerically as $l''(\hat{\nu}) = [l(\hat{\nu} + h) - 2l(\hat{\nu}) + l(\hat{\nu} - h)]/h^2$, for sufficiently small h. (The value of h was gradually decreased until the estimated value of $l''(\hat{\nu})$ was stabilized.)

Likelihood-ratio tests were used to test hypotheses such as $\nu = 1$, $\nu = 2.6$, and $\nu_{\text{female}} = \nu_{\text{male}}$. For example, $2[l(\hat{\nu};\mathbf{y}) - l(1;\mathbf{y})]$ has approximately a χ^2 distribution with 1 df, under the hypothesis that $\nu = 1$.

Assessment of Model Fit

To assess how well the data follow each of the five fitted models, one must determine how well point-process data, such as the data displayed in figure 1, conform to what might be expected under these models. To do so, we considered two summary characteristics of the data: the number of recombination events per chromosome and the distances between recombination events.

First, we considered the distribution of the observed numbers of crossovers. This was compared to the fitted distributions under each of the five models. For the four CL models, the fitted distributions of the observed numbers of crossovers are simple functions of the estimated parameters. For the gamma model, the fitted distribution was obtained by computer simulation. (We assumed a model with the estimate of the parameter ν and the estimated genetic length of the chromosome, and we simulated the locations of crossovers on 10,000 meiotic products. The distribution of the observed numbers of crossovers among these 10,000 chromosomes was used as the fitted distribution for the gamma model.) To compare the observed distribution with the five fitted distributions,

χ^2 goodness-of-fit statistics were calculated; in this process, bins with <5 expected meiotic products were combined. For the Poisson, truncated Poisson, and gamma models, we provide approximate P values, which were calculated with df = number of bins − number of parameters − 1, where the number of parameters was one for the Poisson and truncated Poisson models (the average number of chiasmata λ) and was two for the gamma model (the interference parameter ν and the chromosome length L). P values were not calculated for the other two CL models, since they involve the estimation of a very large number of parameters and therefore generally give nearly perfect fits to this summary of the data.

Second, we considered the distribution of the distances between crossovers. Because the distribution of the distance between two adjacent crossovers depends on the number of observed crossovers on the corresponding chromosomes (when a chromosome has many crossovers, its crossovers will be closely spaced), one ideally should look at the distribution of the intercrossover distances, conditional on the number of crossovers observed. However, since the current data consist of, at most, 92 meiotic products, there is not sufficient material for such a procedure to be informative. Thus, we considered instead the marginal distribution of the intercrossover and crossover-to-terminal-marker distances. For both a particular chromosome number and a particular (parental) sex, we combined all of the intercrossover distances together, including the pter-to-crossover and crossover-to-qter distances and, in the case of transmission of whole chromosomes, the length of the chromosome. We then created a histogram of the resulting distances. To obtain the corresponding fitted distributions, we simulated 10,000 chromosomes from each model, given the parameter estimates and estimated chromosome length; calculated the intercrossover and crossover-to-terminal-marker distances for each simulated chromosome; and obtained a kernel-density estimate (Silverman 1996) of the distribution of distances, with a Gaussian kernel and a bandwidth of 20 cM.

Interference across the Centromere

To evaluate the possibility of interference across the centromere, we performed two further analyses. First, for all but the telocentric chromosomes (chromosomes 13–15, 21, and 22), we fit the gamma model and estimated its parameter ν, under the assumption of no interference across the centromere but of constant interference in each arm. This was done by splitting the chromosomes into their two arms and by treating the processes on the arms as independent.

Second, we performed a modified version of the analysis described by Colombo and Jones (1997). Consider the meiotic products showing crossovers on both arms; let x_p and x_q denote the distances from the centromere to the first observed crossovers on the p and q arms, respectively, and let L_p and L_q denote the lengths of the chromosome arms. We estimated the correlation function $\rho(d) = \text{corr}(x_p, x_q | \max\{x_p, x_q\} \leq d, \min\{L_p, L_q\} \geq d)$ for d = 20–100 cM. In doing so, we pooled data across sex and chromosome number, considering all chromosomes showing at least one crossover on each arm. Under the hypothesis that the presence of a crossover near the centromere on one arm inhibits the presence of a crossover near the centromere on the other arm, $\rho(d) < 0$. A nonparametric bootstrap (Manly 1997) was used to obtain pointwise confidence limits for $\rho(d)$; 1,000 bootstrap replicates were performed.

Note that Colombo and Jones (1997) considered instead the function $\rho^*(d) = \text{corr}(x_p, x_q | x_p + x_q \leq d)$. This approach is flawed, however, in that it can be shown that $\rho^*(d) < 0$, for $d > 0$, even in the case of no interference. The variation in the lengths of chromosome arms may similarly lead to a spurious correlation; therefore, we restricted ourselves to those meiotic products for which $\max\{x_p, x_q\} \leq d$ and $\min\{L_p, L_q\} \geq d$, in which case, under no interference, x_p and x_q would be independent. Thus, the modified version $\rho(d)$ has the property that $\rho(d) = 0$, for all d, in the case of no interference across the centromere.

Results

Genotype Data

In the present study, analysis of interference was based on the genotyping of 8,010 STRPs in eight large, mostly three-generation CEPH kindreds (for details, see the Material and Methods section and Broman et al. [1998]). The *chrompic* option of the CRI-MAP program was used to identify tight apparent double-recombination events indicative of genotyping errors, mutations, or, possibly, gene conversions. Approximately 0.08% of the genotypes were likely to be in error and were removed. It is important to justify this procedure, since, if these apparent double-recombination events were, in fact, real, then the removal of these genotypes would bias the results of our analysis in favor of strong positive interference. In figure 2, the distributions of the number of informative markers separating pairs of apparent recombination events, in both the raw and clean data, are displayed. The striking difference between the raw and clean data is that, in the raw data, 760 pairs of apparent recombination events were separated by just one informative marker. If these events were true double-recombination events, then we would observe, in the raw data, a smooth distribution of the numbers of informative

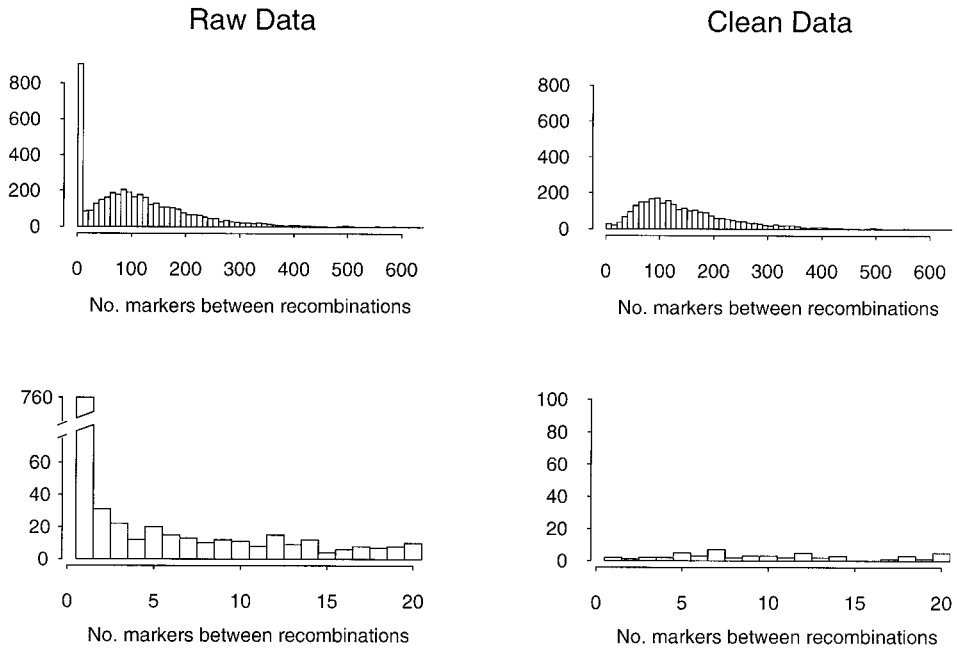

Figure 2 Histograms of the number of informative markers separating every pair of apparent recombination events genomewide, for the raw and clean data. *Top,* Full distributions. *Bottom,* Left tails of the distributions shown in greater detail.

markers between recombinations, as is seen in the clean data (for further discussion, see Broman et al. [1998]).

Number of Crossovers per Chromosome

We compared, for males and females separately, the observed distributions of the number of crossovers per meiotic product, for each chromosome, to five mathematical models of interference: Poisson and truncated Poisson (both of which assume no interference), CL, obligate-chiasma CL, and gamma models (see the Material and Methods section, above). Specific examples of these results, for chromosomes 1 and 4, are presented in tables 1 and 2, respectively. These two chromosomes were chosen because the results for chromosome 1 were typical, whereas those for maternal chromosome 4 were unusual. The crossovers-per-chromosome distributions provided modest evidence for positive crossover interference, insofar as the distributions for five maternal chromosomes (3, 4, 12, 15, and 17) and six paternal chromosomes (3–6, 14, and 21) fit the Poisson and/or truncated Poisson models poorly ($P < .05$). The observed distributions generally showed fewer meiotic products that had very low or high numbers of crossovers than would be predicted under the no-interference model.

In contrast to the Poisson models, the CL, obligate-chiasma CL, and gamma models generally fit the recombination-distribution data well. None of these three models was obviously superior in fit, with regard to this aspect of the data. Several of the smaller paternal chromosomes (19, 21, and 22) showed a poor fit to the models requiring at least one chiasma on the four-strand bundle. For example, of the 92 paternal chromosomes 22, 69 showed no crossovers, 22 showed one crossover, and 1 showed two crossovers. The Poisson model fits these data reasonably well, but under no chromatid interference with an obligate chiasma, no more than 50% of the meiotic products would be expected to have no crossovers. The distribution for maternal chromosome 4 was unusual; it deviated considerably from all five models, apparently because of an excess of meiotic products with exactly two crossovers—an excess that possibly was the result of chromatid interference or of individual variation in recombination.

Note that the average numbers of observed crossovers on the paternal chromosomes were somewhat smaller than those that would be expected, given chiasma frequencies observed cytologically in male sperm (e.g., see table 4 in the study by Laurie and Hultén [1985*a*]). The significance of these differences is unclear, given that we excluded telomeric markers.

Intercrossover Distances

In figure 3, histograms of the intercrossover distances for the maternal and paternal chromosomes 1 and 2 are displayed along with fitted distributions for the gamma, Poisson, and CL models. The fitted curves for the truncated Poisson and obligate-chiasma CL models were vir-

Table 1

Distribution of Observed Numbers of Recombination Events, for Maternal and Paternal Chromosomes 1

Source of Data	No. of Recombinations							χ^2	df	P
	0	1	2	3	4	5	>5			
Maternal:										
Observed	2	7	14	22	16	15	3			
Poisson	3.0	9.8	16.1	17.5	14.3	9.3	8.9	10.1	5	.07
Truncated Poisson	2.9	9.9	16.1	17.5	14.3	9.3	8.9	10.2	5	.07
CL	2.0	6.8	14.4	21.0	19.5	11.1	4.1	2.3		
Obligate-chiasma CL	2.0	6.8	14.3	21.1	19.6	11.1	4.0	2.3		
Gamma	.9	5.3	13.7	20.9	19.8	12.0	6.4	5.2	4	.27
Paternal:										
Observed	8	30	30	19	5	0	0			
Poisson	15.0	27.2	24.7	14.9	6.8	2.5	1.0	6.7	3	.08
Truncated Poisson	13.8	29.0	25.2	14.6	6.3	2.2	.8	8.5	3	.04
CL	9.0	27.3	32.5	18.6	4.5	.1	.0	.6		
Obligate-chiasma CL	8.9	27.4	32.7	18.6	4.4	.0	.0	.6		
Gamma	8.9	27.1	31.4	18.0	5.4	.9	.1	.9	2	.65

tually identical to those for the CL model. These distributions include all intercrossover and crossover-to-terminal-marker distances. The "bumps" at the far right of each histogram correspond to whole chromosomes transmitted intact. Note that the heights of these bumps do not accurately reflect the number of whole chromosomes transmitted without recombination, because of the smoothing of the kernel-density estimate (see the Material and Methods section, above); the expected numbers of meiotic products without crossovers under the gamma and CL models corresponded closely to the numbers observed; the no-interference model generally predicted somewhat more meiotic products without crossovers than was observed.

The Poisson and CL models fit this aspect of the data very poorly, whereas the gamma model fit remarkably well. Similar results were obtained for all chromosomes longer than 150 cM. (A notable exception was maternal chromosome 8, which showed an unusually large number of intercrossover distances that were $\leq$10 cM. Inspection of the maternal chromosome 8 haplotypes in the progeny of the families revealed apparent triple-recombination events within 40 cM in five different individuals.) For the chromosomes shorter than 150 cM, it is difficult to assess the fit of the models, given the small number of observed meiotic products with multiple crossovers.

Levels of Interference

Table 3 contains estimates (and estimated standard errors) of the interference parameter ν from the gamma model. Separate estimates were obtained for each sex and each chromosome; pooled estimates were obtained by the combination of the intercrossover distances across sex and/or chromosome number. The sex-specific estimates are displayed graphically, with ~95% confidence limits, in figure 4. The horizontal line in figure 4 corresponds to the estimate of ν obtained with the use of data from all autosomes plus the maternal X chromosome. Note that there were no paternal chromosomes 21 that exhibited more than one crossover and, therefore, the likelihood for ν continued to increase as $\nu \rightarrow \infty$, so that $\hat{\nu} = \infty$, where $\hat{\nu}$ is the maximum-likelihood estimate of ν.

The results of likelihood-ratio tests indicated that all estimates $\hat{\nu}$ (with the exception of that for paternal chromosome 22) were significantly greater than $\nu = 1$, which is the value corresponding to no interference. In addition, most of the estimates $\hat{\nu}$ were significantly greater than $\nu = 2.6$, which is the value corresponding to a gamma model with a corresponding map function that is approximately equal to the Kosambi map function (Zhao and Speed 1996). For example, for the sex-pooled estimates, only chromosomes 8, 20, and 22 had $\hat{\nu}$ that was not significantly greater than 2.6.

Next, we investigated whether the level of interference, as measured by $\hat{\nu}$, varied between the sexes and among different chromosomes. Significant differences between the sexes were found on chromosomes 6–8, 11, 16, and 21. However, relative to the estimated standard errors for the $\hat{\nu}$, the chromosome-specific differences were small, with the biggest difference, for chromosome 7, being 2.3 times the estimated standard error of the difference. As displayed in figure 4, the sex-specific confidence intervals for each chromosome all overlapped. Also, the sex-specific estimates pooled across chromosomes were not significantly different (table 3).

Tests for heterogeneity across chromosomes were significant for the female and pooled estimates but not for the male estimates. Inspection of the chromosomes' con-

Table 2

Distribution of Observed Numbers of Recombination Events, for Maternal and Paternal Chromosomes 4

	No. of Recombinations									
Source of Data	0	1	2	3	4	5	>5	χ^2	df	*P*
Maternal:										
Observed	1	17	28	13	14	6	0			
Poisson	6.4	16.2	20.2	16.9	10.6	5.3	3.4	9.8	4	.04
Truncated Poisson	6.1	16.6	20.5	16.9	10.5	5.2	3.2	10.4	4	.03
CL	4.0	14.8	22.9	20.2	11.9	4.5	.8	6.8		
Obligate-chiasma CL	4.0	14.8	22.9	20.2	11.9	4.5	.8	6.8		
Gamma	3.0	12.9	23.2	22.7	11.7	4.3	1.1	8.3	3	.04
Paternal:										
Observed	10	39	32	10	1	0	0			
Poisson	20.8	30.9	23.0	11.4	4.3	1.3	.4	11.1	3	.01
Truncated Poisson	19.0	34.6	23.3	10.4	3.5	.9	.3	15.5	3	.01
CL	12.9	35.0	32.3	11.0	.8	.0	.0	1.2		
Obligate-chiasma CL	12.9	35.1	32.3	10.9	.8	.0	.0	1.2		
Gamma	13.7	34.9	30.6	10.9	1.8	.1	.0	2.1	2	.36

tributions to the likelihood-ratio-test statistics indicated that maternal chromosomes 7, 8, and 14 were largely responsible for these differences, with maternal chromosome 8 being by far the largest contributor. These three chromosomes—as well as paternal chromosome 11 and maternal chromosome 20—were the only chromosomes for which the confidence intervals in figure 4 did not overlap the pooled estimate of ν. Once again, it is important to note that maternal chromosome 8 was the one chromosome for which the gamma model gave a poor fit to the distribution of the intercrossover distances—largely as a result of several apparent triple-recombination events on chromosome 8p.

To study possible individual variation in the level of interference, we estimated the parameter ν separately for each of the 16 parents in the eight families, by use of data pooled across the chromosomes (chromosomes 1–X for the mothers and chromosomes 1–22 for the fathers). The individual-specific estimates are displayed in figure 5. Large differences in the level of interference are seen in the eight mothers: the mothers in families 1331 and 1416 show more interference, whereas the mother in family 1413 shows less interference. There appears to be little variation among the fathers, although the reduced recombination in males results in wider confidence intervals for ν and, thus, in lower power to detect differences. Likelihood-ratio tests for homogeneity in interference between individuals gave *P* values of .002 and .55 for the mothers and fathers, respectively.

Interference across the Centromere

Following Mather (1938), it has generally been assumed that the centromere plays a special role in the recombination process and that the centromere acts as a barrier to interference. The estimates of the interference parameters ν for the gamma model were obtained under the assumption of a constant level of interference across the chromosome, with the centromere playing no special role in the recombination process. The parameters were reestimated under the assumption of no interference across the centromere (i.e., the processes on the two arms of each chromosome were treated as independent processes). The new estimates of ν were only slightly changed. In most cases, $\hat{\nu}$ increased (as would be expected if there were no interference across the centromere), but the increases were generally smaller than one standard error. The major exceptions were paternal chromosomes 16–20, for which very few chromosome arms exhibited more than one crossover. Under the assumption of no interference across the centromere, paternal chromosomes 17, 19, and 20 had $\hat{\nu} = \infty$, whereas, for paternal chromosomes 16 and 18, $\hat{\nu}$ was finite but was greater than 15, with a standard error >8.

The appropriateness of the assumption of no interference across the centromere was studied via the correlation function $\rho(d) = \text{corr}\,(x_p, x_q \mid \max\{x_p, x_q\} \leqslant d, \min\{L_p, L_q\} \geqslant d)$ for d = 20–100 cM, where x_p and x_q are the distances from the centromere to the nearest crossovers on the p and q arms, respectively, and where L_p and L_q are the lengths of the arms. In other words, $\rho(d)$ is the correlation between the distances from the centromere to the nearest crossovers on the p and q arms, among chromosomes for which the crossovers on the two arms were each no more than d cM away from the centromere. We restricted analysis to those chromosomes for which both arms are at least d cM in length, to avoid a spurious correlation resulting from the variation in the lengths of the arms. In the case of interference across the centromere, we expect $\rho(d) < 0$, which means that, if a crossover oc-

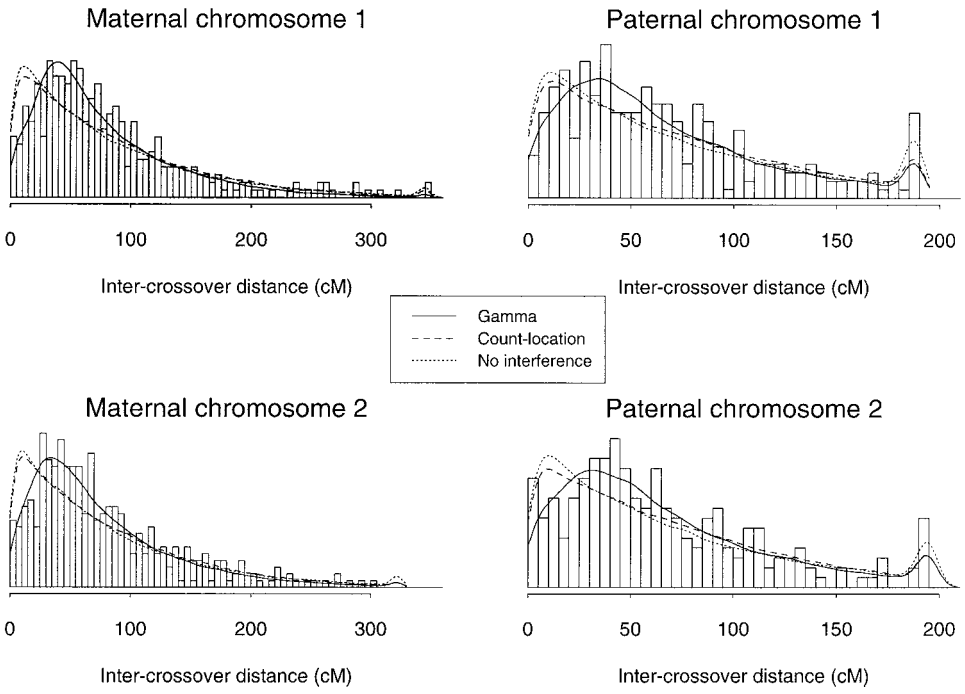

Figure 3 Histograms of intercrossover distances for the maternal and paternal chromosomes 1 and 2 (including telomere-to-crossover distances), with fitted distributions for the gamma, CL, and no-interference models. Note that the fitted distributions for the truncated Poisson and obligate-chiasma CL models were virtually identical to those for the CL model. The "bumps" to the far right in each histogram correspond to whole chromosomes transmitted without crossovers.

curs near the centromere on one arm, the nearest crossover on the other arm will tend to be farther away from the centromere.

We considered data for all chromosomes showing at least one crossover on each arm, pooled across sex and chromosome number. The estimate of $\rho(d)$ is displayed in figure 6, along with approximate pointwise confidence intervals. As shown in figure 6, $\hat{\rho}(d) \approx -0.19$, and its upper confidence limit is < 0, for d = 45–80 cM. For $d < 40$ cM, the confidence intervals for $\rho(d)$ increase in width, as a result of the small number of chromosomes with crossovers close to the centromere on each arm. For $d > 80$ cM, $\hat{\rho}(d)$ begins to increase toward 0. We conclude that the processes on the two arms are not independent; in other words, interference does act across the centromere.

Discussion

The results of the present study provide strong support for a high level of positive crossover interference in human meiosis genomewide. Zhao and Speed (1996) showed that the map functions derived from a gamma model with $\nu \approx 2.6$ and 7.6 correspond closely to the Kosambi and Carter-Falconer map functions, respectively. In the present study, the $\hat{\nu}$ value of 4.3, which was obtained when the data were pooled across both sexes and all chromosomes (table 3), indicates that the appropriate map function for human meiosis is between these two map functions.

We found little evidence for variation in the level of interference among chromosomes and between sexes. The sex- and chromosome-specific confidence intervals for ν failed to overlap the pooled estimate of ν only for a few chromosomes (fig. 4). Kaback et al. (1992, 1999) showed, in yeast, that the size of a chromosome may have a causal effect on its level of recombination and that this may be the result of size-dependent control in the level of interference. Kaback (1996) showed that, for human chromosomes, there is a negative relationship between the physical length of a chromosome and the recombination rate per unit of physical length; this leads to the hypothesis that smaller human chromosomes exhibit a lower level of crossover interference. We did not uncover evidence to support this hypothesis, but this may have been because of the paucity of meioses considered here. When additional data become available, variation in interference among chromosomes and sexes may become apparent.

We detected individual variation in interference among the eight mothers of these families, although we saw little variation among the fathers (fig. 5). This

Table 3

Estimates (and Estimated Standard Errors) of the Parameter ν in the Gamma Model, by Chromosome, for Female and Male Meiosis Separately and Pooled

	Estimates (Estimated Standard Errors) of ν in the Gamma Model, for Meiosis		
Chromosome	Female	Male	Pooled
1	5.5 (.9)	4.5 (.9)	5.0 (.6)
2	3.7 (.5)	4.6 (.9)	4.0 (.5)
3	3.7 (.6)	6.4 (1.6)	4.2 (.5)
4	4.2 (.7)	5.3 (1.2)	4.6 (.6)
5	4.3 (.8)	3.4 (.8)	3.9 (.6)
6	4.0 (.7)	8.3 (2.2)	4.7 (.7)
7	7.9 (1.4)	3.9 (1.0)	5.9 (.8)
8	2.2 (.4)	4.9 (1.4)	2.6 (.4)
9	4.3 (.9)	3.2 (.8)	3.8 (.6)
10	3.7 (.7)	4.3 (1.0)	3.9 (.6)
11	5.5 (1.1)	2.9 (.7)	4.1 (.7)
12	6.7 (1.3)	4.3 (1.1)	5.5 (.9)
13	4.7 (1.2)	3.8 (1.2)	4.3 (.9)
14	8.9 (2.0)	5.3 (1.6)	7.3 (1.4)
15	5.6 (1.4)	4.2 (1.4)	5.0 (1.0)
16	3.5 (.7)	10.1 (3.5)	4.2 (.7)
17	4.6 (1.0)	8.4 (3.0)	5.3 (1.0)
18	4.7 (1.1)	5.7 (1.9)	5.0 (1.0)
19	7.9 (2.2)	5.7 (2.0)	6.9 (1.5)
20	2.6 (.6)	3.0 (1.1)	2.7 (.5)
21	4.3 (1.7)	∞	5.4 (1.8)
22	3.6 (1.4)	2.9 (1.7)	3.4 (1.1)
X	7.0 (1.4)		
Pooled	4.2 (.2)	4.5 (.3)	4.3 (.1)

should be considered as a preliminary finding, since the present analysis required pooling of data across chromosomes and since it did not account for known individual variation in recombination. We have previously found that the mothers in these families vary in terms of the total number of recombination events per meiosis (see table 2 in the study by Broman et al. [1998]). One might imagine that mothers that show a greater degree of recombination would display a decreased level of interference, and vice versa. However, the mothers in families 1413 and 1416, which showed low and high levels of interference, respectively, displayed approximately equal, somewhat elevated levels of recombination. Thus, the individual differences in recombination do not explain the observed differences in interference.

With regard to interference across the centromere, we have shown that there is a negative correlation between the distances from the centromere to the nearest crossovers on the p and q arms, among chromosomes exhibiting at least one crossover on each arm. Interference therefore continues to operate across the centromere. The data from the present study support the conclusion of a study by Colombo and Jones (1997), which indicates that the centromere is not a barrier to interference.

In a study of the relationship between nondisjunction and the obligate-chiasma hypothesis, Bugge et al. (1988) used a portion of the data considered in the present study (family 102 was excluded). They found that, in addition to paternal chromosomes 21 and 22 (as seen here), paternal chromosome 6 did not fit the obligate-chiasma CL model. However, their finding of an excess of paternal chromosomes 6 with no chiasmata likely was the result of their use of family 884, for which a long autozygous segment in the father (Broman and Weber 1999) likely led to several double crossovers being missed. The most-reasonable explanation for the observations on paternal chromosomes 21 and 22 is that the genetic markers did not provide sufficiently complete coverage of these two small chromosomes, rather than that there was not an obligate chiasma in these meioses.

Because of the difficulty in determination of marker order and identification of true recombination events at telomeres, we excluded from our analysis the five terminal markers from each end of each chromosome (~3% of the total markers). As a result, we cannot draw conclusions about interference at telomeres. Although this approach may affect our inference on the obligate-chiasma hypothesis, it does not bias our conclusions about the presence and level of interference in the regions studied. It is possible that interference is different in the telomeres than in the interiors of the chromosomes—especially, perhaps, in males, who generally have an excess of recombination at chromosome ends—and, therefore, the estimated level of interference obtained here may not apply to those regions. It is of

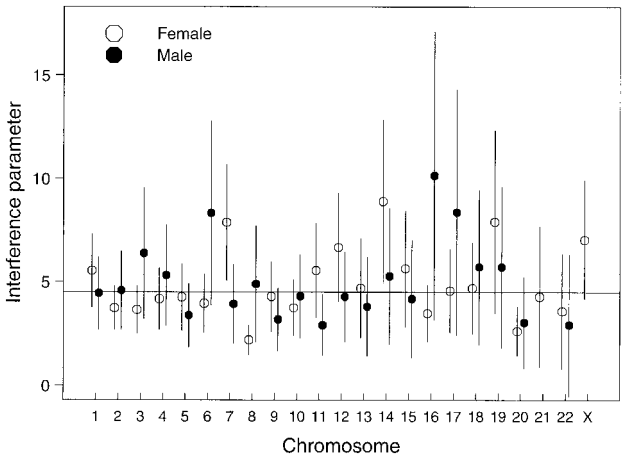

Figure 4 Sex-specific estimates of the parameter ν from the gamma model, for each chromosome. The intervals correspond to $\pm$ twice the estimated standard error. A horizontal line is plotted at the pooled estimate of ν obtained with the use of data from all autosomes plus the maternal X chromosome. Paternal chromosome 21 gave $\hat{\nu} = \infty$ and thus was not plotted.

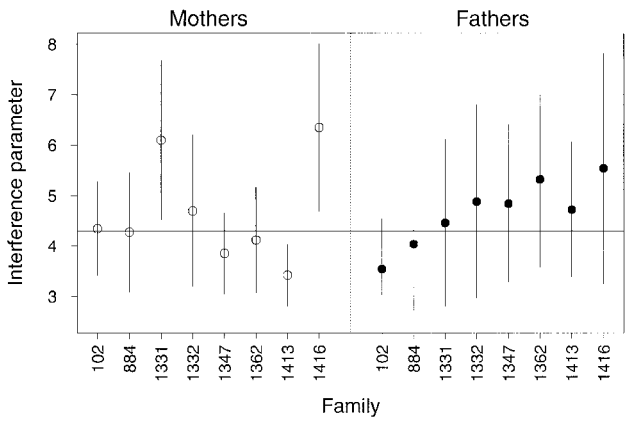

Figure 5 Individual-specific estimates of the parameter ν from the gamma model, for each of the 16 parents, with data pooled across chromosomes. The intervals correspond to ± twice the estimated standard error. A horizontal line is plotted at the pooled estimate of ν obtained with the use of data from all autosomes plus the maternal X chromosome.

interest to note that, when we included the telomeric data in the fit of the gamma model, the estimates of the interference parameter ν were largely unchanged. On average, the relative change in $\hat{\nu}$ was only 10%; the overall estimate of ν dropped from 4.3 to 4.1.

Gene conversion could be responsible for a small fraction of the apparent tight double-recombination events involving only a single marker (as is shown for the raw data in fig. 2). From the results of previous work, we know that most of the apparent tight double-recombination events are caused by genotyping errors and mutations (see, for example, Haines et al. [1993], Weber et al. [1993], and Zahn and Kwiatkowski [1995]). A small fraction, however, could also be the result of gene conversion. Gene conversion would slightly increase the probabilities (shown in fig. 7) for the presence of an allele of the opposite phase between two flanking, nonrecombinant markers.

Several other approximations and assumptions, although they are unlikely to alter the primary conclusions of the present study, should still be considered. We assumed that the markers were in the correct order and that the intercrossover distances were known exactly. However, the estimated genetic distances between markers, based on 92 meioses, were subject to considerable sampling variation. For example, an estimated distance of 5 cM had a standard error of ~2 cM. We assumed that there was no individual variation in recombination or interference, although several studies have shown individual variation in recombination (Tanzi et al. 1992; Yu et al. 1996; Broman et al. 1998). Except for interference near the centromere, we assumed that there was no intrachromosomal variation in interference. Chromatid interference was ignored. The interplay between crossover interference and chromatid interference (if it exists) cannot be resolved through analysis of recombinations within single meiotic products. Finally, we neglected the effect of interval censoring of the crossover locations (i.e., that the crossovers could be localized only to the interval between flanking informative markers). An analysis that made account of the interval censoring (data not shown) gave very similar results; this may have been expected, given the precision (see fig. 1) with which the crossovers could be localized. It was crucial, however, to exclude chromosomes for which there were large gaps between informative markers, the majority of which were the result of autozygosity in the parents in family 884 (Broman and Weber 1999).

The gamma model provided a much better fit to the data than did the CL models, as has previously been seen in data on experimental organisms (McPeek and Speed 1995). The CL models allow crossovers to occur more closely together than was seen in the data from the present study. A modified version of the CL model described by Goldgar and Fain (1988), in which chiasmata are prevented from occurring close together, may provide similar improvements on the CL model. We found the gamma model to be more natural and mathematically tractable, and, therefore, we chose not to fit the model of Goldgar and Fain (1988).

The distribution of the distances between recombination events (fig. 3) was much more valuable than were the distributions of the numbers of recombination events per chromosome (tables 1 and 2), both in the demonstration of interference and in discrimination between the gamma and CL models. This could be anticipated, given the experience with experimental organ-

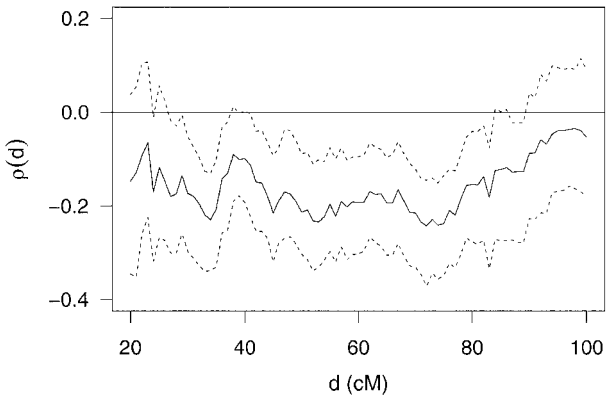

Figure 6 Estimate of the correlation function $\rho(d) = \mathrm{corr}(x_p, x_q \mid \max\{x_p, x_q\} \leq d, \min\{L_p, L_q\} \geq d)$ obtained with the use of all meiotic products that had at least one crossover on each arm, pooled across sex and chromosome number. The dashed curves correspond to ~95% pointwise confidence limits.

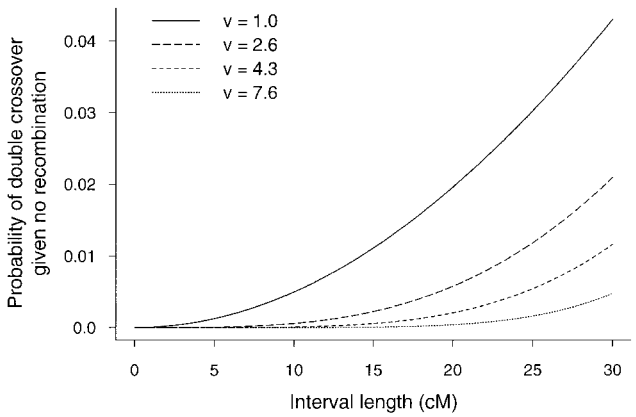

Figure 7 Probability of a double crossover in an interval, given that no recombination has occurred, as a function of interval length, for the gamma model with different values of the parameter ν. The value $\nu = 1.0$ corresponds to no interference, and $\nu = 4.3$ corresponds to the pooled estimate for the present data. The values $\nu = 2.6$ and 7.6 correspond to gamma models for which the corresponding map functions are, approximately, the Kosambi and Carter-Falconer map functions, respectively.

isms (see McPeek and Speed 1995). The distance between recombination events provides a more direct view of crossover interference, in which each chiasma inhibits the formation of other chiasmata nearby.

Lin and Speed (1996) fitted the χ^2 model (equivalent to the gamma model with $v = m + 1$, where m is an integer) to data on six linked loci from chromosome 10 that covered ~60 cM. They obtained the estimate $\hat{m} = 4$, which corresponds to $\hat{\nu} = 5$, a finding that is quite close to the results of the present study.

Collins et al. (1996) provided estimates of the level of interference on each chromosome genomewide. They estimated the Rao map-function parameter p, by combining pairwise recombination fractions separately for each chromosome. Note that $p = 1$, $p = 0.5$, $p = 0.25$, and $p = 0$ correspond to the Haldane, Kosambi, Carter-Falconer, and Morgan (i.e., $M(d) = d$) map functions, respectively, so that smaller values of p correspond to greater levels of interference. Collins et al. (1996) obtained estimates with a range of 0.46 (chromosome 6) to 0.22 (chromosomes 18 and 19). Thus, their conclusions regarding interference were similar to those obtained here: the level of interference in human meiosis is between the levels of interference implied by the Kosambi and Carter-Falconer map functions. The results of the present study improve on the those of Collins et al. (1996), in that (*a*) we provide standard errors for the estimated levels of interference and thus are able to show that the observed variation between chromosomes and between sexes could reasonably be due to chance and (*b*) the present study results in a mathematical model for the recombination process that is valuable both for computer simulations and for incorporation into new methods of mapping genes (see below).

The findings of the present study have at least three important implications. First, many authors have shown that a proper account of crossover interference in methods for gene mapping may increase the efficiency of such methods (Bishop and Thompson 1988; Goldgar et al. 1989; Goldstein et al. 1995; Lin and Speed 1999). Although statistical genetic analyses are considerably more simple computationally when calculations are performed with the assumption of no interference, significant improvement in efficiency may be gained—at least for the problems of exclusion mapping and the ordering of markers—with the use of an appropriate mathematical model of crossover interference.

Second, the gamma model for the recombination process, allowing for interference, may be useful in computer simulations of meiosis in families. Such simulations, which are generally performed with the assumption of no interference, are valuable in the study of the statistical properties of methods for mapping genes. Insofar as the gamma model improves on the no-interference model, simulations that make use of the gamma model may provide an improved understanding of the properties of statistical methods for gene mapping.

Finally—and, perhaps, most importantly—the present study has valuable clinical application. When studying the transmission of a deleterious mutation through a pedigree with data on two informative flanking markers, a clinician must consider the chance of a double crossover, given that no recombination occurred between the flanking markers. Figure 7 shows the probability of a double crossover between markers separated by d cM, given that there is no recombination between the markers, calculated under the gamma model with four different values for the interference parameter ν. The value $\nu = 1$ corresponds to no interference, $\nu = 4.3$ corresponds to the pooled estimate in the present study, and $\nu = 2.6$ and 7.6 correspond to gamma models whose map functions are, approximately, the Kosambi and Carter-Falconer map functions, respectively. At 20 cM, this probability is 2%, .6%, .2%, and .04% for $\nu = 1$, 2.6, 4.3, and 7.6, respectively. Although calculation of the curves in figure 7 involved rather complicated numerical integrals, the equation $(0.0114d - 0.0154)^4$, in which d is measured in cM, provides a good estimate of the probability plotted (accurate to within .1%), for our pooled estimate of interference ($\nu = 4.3$). This equation provides a reasonable guide for the clinical geneticist (or gene mapper), although the results should be treated with care, given the common observation of individual variation in recombination,

the possibility of gene conversion, and imperfect estimates of the genetic distances between markers.

The CEPH-family genotyping data are likely the best set of human data that has been produced, to date, for the analysis of genomewide interference (and other meiotic properties). The data permit a strong conclusion of positive crossover interference and a reasonable estimation of the average strength of interference over the entire genome. However, these data are still insufficient to allow us to detect subtle sex or chromosome differences in interference levels. The use of genotype data for dense collections of polymorphisms from many more families will be necessary to resolve these issues.

Acknowledgments

Our work on interference benefited greatly from discussions we had with Mary Sara McPeek, Terry Speed, and Hongyu Zhao. Sharon Browning, Penny Chua, Eleanor Feingold, Mark Neff, and Terry Speed generously provided comments for improvement of the manuscript. We further thank an anonymous reviewer who pointed out a flaw in our original analysis of interference across the centromere. This work was supported in part by a John Wasmuth Fellowship in Genomic Analysis, National Human Genome Research Institute grant F32-HG000198 (to K.W.B.), and National Heart, Lung, and Blood Institute contract N01-HV48141 for the Mammalian Genotyping Service (to J.L.W.).

Electronic-Database Information

URLs for data in this article are as follows:

Center for Medical Genetics, Marshfield Medical Research Foundation, http://www.marshmed.org/genetics/ (for STRPs and genetic maps)

Cooperative Human Linkage Center, The, http://lpg.nci.nih.gov/CHLC/ (for STRPs)

Généthon, http://www.genethon.fr/ (for STRPs)

Whitehead Institute for Biomedical Research/MIT Center for Genome Research, http://www.genome.wi.mit.edu/ (for radiation-hybrid maps)

References

Attwood J, Chiano M, Collins A, Donis-Keller H, Dracopoli N, Fountain J, Falk C, et al (1994) CEPH Consortium map of chromosome 9. Genomics 19:203–214

Bishop DT, Thompson EA (1988) Linkage information and bias in the presence of interference. Genet Epidemiol 5: 107–119

Blank RD, Campbell GR, Calabro A, D'Eustachio P (1988) A linkage map of mouse chromosome 12: localization of Igh and effects of sex and interference on recombination. Genetics 120:1073–1083

Brent RP (1973) Algorithms for minimization without derivatives. Prentice-Hall, Englewood Cliffs, NJ

Broman KW, Murray JC, Sheffield VC, White RL, Weber JL (1998) Comprehensive human genetic maps: individual and sex-specific variation in recombination. Am J Hum Genet 63:861–869

Broman KW, Weber JL (1999) Long homozygous chromosomal segments in reference families from the Centre d'Étude du Polymorphisme Humain. Am J Hum Genet 65:1493–1500

Bugge M, Collins A, Petersen MB, Fisher J, Brandt C, Hertz JM, Tranebjærg L, et al (1998) Non-disjunction of chromosome 18. Hum Mol Genet 7:661–669

Carter TC, Falconer DS (1951) Stocks for detecting linkage in the mouse, and the theory of their design. J Genet 50:307–323

Ceci JD, Siracusa LD, Jenkins NA, Copeland NG (1989) A molecular genetic linkage map of mouse chromosome 4 including the localization of several proto-oncogenes. Genomics 5:699–709

Chua PR, Roeder GS (1997) Tam1, a telomere-associated meiotic protein, functions in chromosome synapsis and crossover interference. Genes Dev 11:1786–1800

Collins A, Forabosco P, Lawrence S, Morton NE (1995) An integrated map of chromosome 9. Ann Hum Genet 59:393–402

Collins A, Frezal J, Teague J, Morton NE (1996) A metric map of humans: 23,500 loci in 850 bands. Proc Natl Acad Sci USA 93:14771–14775

Collins A, Keats BJ, Dracopoli N, Shields DC, Morton NE (1992) Integration of gene maps: chromosome 1. Proc Natl Acad Sci USA 89:4598–4602

Colombo PC, Jones GH (1997) Chiasma interference is blind to centromeres. Heredity 79:214–227

Cox DR (1962) Renewal theory. Methuen, London

Cox S, Bryant SP, Collins A, Weissenbach J, Donis-Keller H, Koeleman BP, Steinkasserer A, et al (1995) Integrated genetic map of human chromosome 2. Ann Hum Genet 59:413–434

Dempster AP, Laird NM, Rubin DB (1977) Maximum likelihood from incomplete data via the EM algorithm. J R Stat Soc B 39:1–38

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

Egel R (1995) The synaptonemal complex and the distribution of meiotic recombination events. Trends Genet 11:206–208

Fisher RA, Lyon MF, Owen ARG (1947) The sex chromosome in the house mouse. Heredity 1:335–365

Forabosco P, Collins A, Morton NE (1995) Integration of gene maps: updating chromosome 1. Ann Hum Genet 59:291–305

Foss EJ, Lande R, Stahl FW, Steinberg CM (1993) Chiasma interference as a function of genetic distance. Genetics 133: 681–691

Foss EJ, Stahl FW (1995) A test of a counting model for chiasma interference. Genetics 139:1201–1209

Goldgar DE, Fain PR (1988) Models of multilocus recombination: nonrandomness in chiasma number and crossover positions. Am J Hum Genet 43:38–45

Goldgar DE, Fain PR, Kimerling WJ (1989) Chiasma-based models of multilocus recombination: increased power for exclusion mapping and gene ordering. Genomics 5:283–290

Goldstein DR, Zhao H, Speed TP (1995) Relative efficiencies of χ^2 models of recombination for exclusion mapping and gene ordering. Genomics 27:265–273

Haines JL, Guillemette W, Rosen D, Brown R Jr, Donaldson D, Patterson D (1993) A genetic linkage map of chromosome 21: a look at meiotic phenomena. Prog Clin Biol Res 384: 51–61

Haldane JBS (1919) The combination of linkage values, and the calculation of distances between the loci of linked factors. J Genet 8:299–309

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

Hultén M (1974) Chiasma distribution at diakinesis in the normal human male. Hereditas 76:55–78

Hultén M, Lawrie NM, Laurie DA (1990) Chiasma-based genetic maps of chromosome 21. Am J Med Genet Suppl 7: 148–154

Kaback DB (1996) Chromosome-size dependent control of meiotic recombination in humans. Nat Genet 13:20–21

Kaback DB, Barber D, Mahon J, Lamb J, You J (1999) Chromosome size-dependent control of meiotic reciprocal recombination in *Saccharomyces cerevisiae:* the role of crossover interference. Genetics 152:1475–1486

Kaback DB, Guacci V, Barber D, Mahon JW (1992) Chromosome size-dependent control of meiotic recombination. Science 256:228–232

Karlin S, Liberman U (1978) Classifications and comparisons of multilocus recombination distributions. Proc Natl Acad Sci USA 75:6332–6336

Karlin S, Liberman U (1979) A natural class of multilocus recombination processes and related measures of crossover interference. Adv Appl Prob 11:479–501

Karlin S, Liberman U (1994) Theoretical recombination processes incorporating interference effects. Theor Popul Biol 46:198–231

Kingsley DM, Jenkins NA, Copeland NG (1989) A molecular genetic linkage map of mouse chromosome 9 with regional localizations for the Gsta, T3g, Ets-1 and Ldlr loci. Genetics 123:165–172

Kosambi DD (1944) The estimation of map distances from recombination values. Ann Eugen 12:172–175

Kwiatkowski DJ, Dib C, Slaugenhaupt SA, Povey S, Gusella JF, Haines JL (1993) An index marker map of chromosome 9 provides strong evidence for positive interference. Am J Hum Genet 53:1279–1288

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

Laurie DA, Hultén MA (1985*a*) Further studies on bivalent chiasma frequency in human males with normal karyotypes. Ann Hum Genet 49:189-201

Laurie DA, Hultén MA (1985*b*) Further studies on chiasma distribution and interference in the human male. Ann Hum Genet 49:203–214

Lawrence S, Collins A, Keats BJ, Hultén M, Morton NE (1993) Integration of gene maps: chromosome 21. Proc Natl Acad Sci USA 90:7210–7214

Lin S, Speed TP (1996) Incorporating crossover interference into pedigree analysis using the χ^2 model. Hum Hered 46: 315–322

Lin S, Speed TP (1999) Relative efficiencies of the χ^2 recombination models for gene mapping with human pedigree data. Ann Hum Genet 63:81–95

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

Mather K (1938) Crossing-over. Biol Rev 13:252–292

McInnis MG, Chakravarti A, Blaschak J, Petersen MB, Sharma V, Avramopoulos D, Blouin J-L, et al (1993) A linkage map of human chromosome 21: 43 PCR markers at average intervals of 2.5 cM. Genomics 16:562–571

McPeek MS, Speed TP (1995) Modeling interference in genetic recombination. Genetics 139:1031–1044

Muller HJ (1916) The mechanism of crossing-over. Amer Nat 50:193–221, 284–305, 350–366, 421–434

Munz P (1994) An analysis of interference in the fission yeast *Schizosaccharomyces pombe.* Genetics 137:701–707

Ott J (1996) Estimating crossover frequencies and testing for numerical interference with highly polymorphic markers. In: Speed T, Waterman MS (eds) Genetic mapping and DNA sequencing: IMA volumes in mathematics and its applications. Vol 81. Springer-Verlag, New York, pp 49–63

Povey S, Smith M, Haines J, Kwiatkowski D, Fountain J, Bale A, Abbott C, et al (1992) Report and abstracts on the First International Workshop on chromosome 9. Ann Hum Genet 56:167–221

Press WH, Teukolsky SA, Vetterling WT, Flannery BP (1992) Numerical recipes in C: the art of scientific computing, 2d ed. Cambridge University Press, Cambridge

Rao DC, Morton NE, Lindsten J, Hultén M, Yee S (1977) A mapping function for man. Hum Hered 27:99–104

Risch N, Lange K (1979) An alternative model of recombination and interference. Ann Hum Genet 43:61–70

Roeder GS (1997) Meiotic chromosomes: it takes two to tango. Genes Dev 11:2600–2621

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

Schmitt K, Lazzeroni LC, Foote S, Vollrath D, Fisher EMC, Goradia TM, Lange K, et al (1994) Multipoint linkage map of the human pseudoautosomal region, based on single-sperm typing: do double crossovers occur during male meiosis? Am J Hum Genet 55:423–430

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

Shields DC, Collins A, Buetow KH, Morton NE (1991) Error filtration, interference, and the human linkage map. Proc Natl Acad Sci USA 88:6501–6505

Silverman BW (1996) Density estimation for statistics and data analysis. Chapman and Hall, London

Speed TP (1996) What is a genetic map function? In: Speed

T, Waterman MS (eds) Genetic mapping and DNA sequencing: IMA volumes in mathematics and its applications. Vol 81. Springer-Verlag, New York, pp 65–88

Sturt E (1976) A mapping function for human chromosomes. Ann Hum Genet 40:147–163.

Sturtevant AH (1915) The behavior of the chromosomes as studied through linkage. Z induktive Abstammungs- Vererbungslehre 13:234–287

Sunden SLF, 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

Sym M, Roeder GS (1994) Crossover interference is abolished in the absence of a synaptonemal complex protein. Cell 79: 283–292

Tanzi RE, Watkins PC, Stewart GD, Wexler NS, Gusella JF, et al (1992) A genetic linkage map of human chromosome 21: analysis of recombination as a function of sex and age. Am J Hum Genet 50:551–558

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

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

Weeks DE, Ott J, Lathrop GM (1994) Detection of genetic interference: simulation studies and mouse data. Genetics 136:1217–1226

Weinstein A (1936) The theory of multiple-strand crossing over. Genetics 21:155–199

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

Yu K, Feingold E (1998) Statistical methods for estimating tetrad crossover frequencies. Am J Hum Genet Suppl 63: A225

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

Zhao H, McPeek MS, Speed TP (1995*a*) Statistical analysis of chromatid interference. Genetics 139:1057–1065

Zhao H, Speed TP (1996) On genetic map functions. Genetics 142:1369–1377

Zhao H, Speed TP, McPeek MS (1995*b*) Statistical analysis of crossover interference using the chi-square model. Genetics 139:1045–1056