GENOMICS 21, 92–103 (1994)

Multipoint Radiation Hybrid Mapping: Comparison of Methods, Sample Size Requirements, and Optimal Study Characteristics

KATHRYN L. LUNETTA AND MICHAEL BOEHNKE[1]

Department of Biostatistics, University of Michigan School of Public Health, Ann Arbor, Michigan 48109

Received September 13, 1993; revised February 14, 1994

There are several statistical methods available for analyzing radiation hybrid (RH) data, but little is known about the ordering accuracy we can expect under common study conditions. Using analytic methods and computer simulation, we compared the ordering accuracy of three multipoint statistical methods: minimum breaks (MB), maximum likelihood (ML), and *maximum posterior probability* (PP). For 8, 12, and 16 markers and all combinations of numbers of hybrids, retention patterns, and marker spacings considered, the probabilities that the true order is identified as the best order were considerably higher with the ML and PP methods than with the MB method. ML and PP performed similarly, but PP tended to give slightly greater support for the best order than did ML. Our results can be used as guidelines for determining sample size requirements and optimal marker spacing for future RH mapping experiments. For equally spaced markers, intermarker spacing of 30 to 50 cR gave the highest probability of correctly ordering all the markers. For randomly spaced markers, 10–20 cR average intermarker spacing resulted in the highest proportion of markers being placed in a 1000:1 framework map. Assuming equal retention in the analysis when a centromeric model would be more appropriate did not affect the ability of the ML method to accurately order the markers, but did influence the distance estimates obtained.

INTRODUCTION

Radiation hybrid (RH) mapping, developed in the early 1970s by Goss and Harris (1975, 1977) and adapted and refined by Cox and colleagues (Cox *et al.*, 1990; Burmeister *et al.*, 1991), has proven to be a valuable tool for ordering loci along human chromosomes. In RH mapping, a lethal dose of radiation is used to break a human chromosome in rodent–human somatic *cell hybrids into multiple fragments. The cells con*taining the fragmented chromosomes are rescued by fusion with a normal rodent cell line that is deficient in HPRT. Growth in HAT medium then selects for the fused rodent cells containing the chromosomal fragments from the rodent–human hybrid. Each resulting hybrid clone contains a unique set of human chromo*some fragments. The clones can be screened for the* presence or absence of markers present on the human chromosome. RH mapping makes use of the principle that the closer two markers are on a chromosome, the less likely it is that they will be separated by a radiation-induced break. By analyzing the patterns of presence and absence of the various loci in the hybrid clones, the order of the markers can be inferred.

RH mapping is now a widely used method of gene mapping, and multipoint statistical methods are becoming a standard analytic approach. However, few attempts have been made (1) to compare the multipoint methods for ordering markers using RH data or (2) to determine the accuracy of these methods under typical study conditions: number of hybrids, number of markers, spacing of markers, and marker retention patterns.

In this paper, we compare the ordering accuracy of the nonparametric minimum breaks method (Boehnke *et al.*, 1991; Bishop and Crockford, 1992; Boehnke, 1992; Weeks *et al.*, 1992) and two model-based methods: maximum likelihood (Boehnke *et al.*, 1991; Chakravarti and Reefer, 1992; Green, 1992) and maximum posterior probability (Lange and Boehnke, 1992). We use analytic methods to obtain results for three markers, assuming all fragments are retained with probability 0.5, and computer simulation to obtain results for more realistic situations. We examine the accuracy of the three multipoint methods under a variety of conditions characteristic of RH data found in the literature. In so doing, we provide sample size guidelines that may be useful in designing RH mapping studies.

MATERIALS AND METHODS

Notation

Suppose the markers $A_1, A_2, \ldots, A_M$ are typed on H radiation hybrids. The observation vector for a hybrid given a specific locus order is $\mathbf{x} = (x_1, x_2, \ldots, x_M)$, where

[1] To whom correspondence should be addressed at Department of Biostatistics, School of Public Health, University of Michigan, 1410 Washington Heights, Ann Arbor, MI 48109-2029. Telephone: (313) 936-1001. Fax: (313) 763-2215. E-mail: Michael.Boehnke@um.cc.umich.edu.

0888-7543/94 $6.00

$$x_i = \begin{cases} 1 \text{ if marker } i \text{ is typed and retained} \\ 0 \text{ if marker } i \text{ is typed and not retained} \\ ? \text{ if marker } i \text{ is untyped.} \end{cases}$$

We define the breakage probability $\theta_i (i = 1, \ldots, M - 1)$ as the probability of at least one chromosome break between markers A_i and A_{i+1}. We assume that breakage occurs at random along the chromosome, so that breakage can be modeled as a Poisson process. In this situation, a breakage probability θ can be converted to an additive distance d using the formula $d = -\log(1 - \theta)$ (Cox *et al.*, 1990). This function is analogous to Haldane's (1919) no-interference mapping function used in linkage analysis. The distance d can also be interpreted as the expected number of chromosome breaks per hybrid between the two loci.

We define the retention probability $r_{ij} (i \leq j)$ as the probability that a fragment including markers $A_i, A_{i+1}, \ldots, A_j$ is retained in a hybrid. Here we consider two fragment retention models: equal retention, where $r_{ij} = r$ for all i, j; and centromeric retention, where $r_{ij} = r_1$ for all j, $r_{ij} = r_2$ for $2 \leq i \leq j$. The centromeric model with $r_1 > r_2$ results in a gradient of marker retention along the chromosome, with highest retention r_1 near the centromere and retention decreasing to r_2 far from the centromere.

Methods of Multipoint Analysis

A number of multipoint methods for ordering radiation hybrids have been developed; we compare three: minimum breaks (Boehnke *et al.*, 1991; Bishop and Crockford, 1992; Boehnke, 1992; Weeks *et al.*, 1992), maximum likelihood (Boehnke *et al.*, 1991; Chakravarti and Reefer, 1992; Green, 1992), and maximum posterior probability (Lange and Boehnke, 1992).

Minimum breaks. The minimum breaks (MB) method selects as best the marker order that requires the fewest obligate chromosome breaks and is analogous to inferring order by minimizing the number of obligate recombinants in linkage mapping (Thompson, 1987). MB is based on the idea that the farther apart two loci are on a chromosome, the more likely it is that the radiation will cause a break between them. The only assumption required is that the markers are ordered linearly along the chromosome. For any particular order, the number of obligate breaks is summed over all hybrids; the best orders are those that require the fewest obligate breaks. As a simple example, if we have three markers typed in one hybrid, then the order (1,0,1) implies two obligate breaks—the first between markers 1 and 2, the second between markers 2 and 3. Two other possible orders, (1,1,0) and (0,1,1), require only one break each. Barrett (1992) showed that the MB method is consistent, in the sense that as the number of hybrids increases, the probability of inferring the correct order converges to one.

Maximum likelihood. For the maximum likelihood (ML) method, we construct a model for the RH data and estimate model parameters by maximizing the likelihood for each possible order. The best orders are those that have the largest maximum likelihood. The ML method provides the advantage of allowing estimation of distances between loci, but requires a model for chromosome breakage and fragment retention. We assume random breakage along the chromosome and that fragments are retained independently within the hybrids. In most of the analyses in this paper, we assume that retention probabilities are equal for all fragments. We also examine the effect of assuming equal retention when retention actually is greater near the centromere, an effect observed by several investigators (Cox *et al.*, 1990; Burmeister *et al.*, 1991; Ceccherini *et al.*, 1992; Gorski *et al.*, 1992).

Since different locus orders are not nested models, significance levels for comparisons among orders cannot be obtained. In practice, a difference in log likelihood of 3 between the best order and all other orders (i.e., the best order has maximum likelihood 1000 times that of any other order) is taken as strong evidence that the maximum likelihood order is correct.

Maximum posterior probability. The Bayesian approach of maximum posterior probability (PP) provides a direct means for comparison of orders by answering the question: what is the probability that a particular locus order is correct? We compute an approximation of the posterior probability developed by Lange and Boehnke (1992), which assumes random marker distribution on the chromosome, equal fragment retention probability, and equal prior probability for each order of $2/M!$. Under these assumptions, the posterior probability that a specific order i is the true order given the observed hybrid data is

$$P(\text{order } i \,|\, \text{data}) = \frac{P(\text{data} \,|\, \text{order } i)}{\sum_{j=1}^{M!/2} P(\text{data} \,|\, \text{order } j)}. \quad [1]$$

The approximation to [1] that we use requires two steps. First, we approximate $P(\text{data} \,|\, \text{order } i)$ by approximating the beta distributions of the marker spacing by exponential distributions with the same mean. This approximation also relies on the assumption that the distribution of the spacings between adjacent loci is approximately independent; this assumption is reasonable if M is large enough. See Lange and Boehnke (1992) for a detailed description of this computation. When this first approximation is used, computing $P(\text{data} \,|\, \text{order } i)$ for all $M!/2$ possible orders would be computationally impractical even for modest M. Therefore, we further approximate the denominator in [1] by summing only over the set of all orders with maximum likelihood within 10^5 of the ML order, rather than all $M!/2$ possible orders in the denominator. We obtain this set of orders from the ML analysis using the branch-and-bound or stepwise locus ordering algorithms described below. In cases in which we used stepwise locus ordering, we cannot be sure that all orders within 10^5 of the best order were identified. However, our experience with this algorithm leads us to believe that we have identified nearly all of these orders. Since there is no obvious prior distribution for the retention probability r, and since it tends to be well estimated by ML, we hold the retention probability fixed at its ML estimate in our calculations (Lange and Boehnke, 1992).

Methods for Identifying Best Locus Orders

Due to the time-intensive computational efforts required, it is not practical to evaluate the number of obligate breaks, maximum likelihood, and posterior probability for all $M!/2$ possible locus orders unless the number of loci M is small. For MB and ML with $M = 8$, we used the branch-and-bound strategy to identify the best locus orders (e.g., see Nijenhuis and Wilf, 1978). In our implementation, branch-and-bound requires a candidate locus order and then builds locus orders one locus at a time, keeping under consideration only those partial locus orders that are within K units (of minimum breaks or log likelihood) of the candidate order. All complete orders consistent with a partial order that are not within K units of the candidate order are eliminated. Branch-and-bound guarantees that the best order and all orders within K units of the best order will be identified and substantially decreases the number of orders that must be evaluated. However, the number of orders to be evaluated still may scale exponentially with the number of markers M (Boehnke *et al.*, 1991).

For $M \geq 12$, branch-and-bound often requires the evaluation of an impractical number of orders; in these situations we used the related strategy of stepwise locus ordering (Boehnke *et al.*, 1991). Stepwise locus ordering builds locus orders one marker at a time and at each stage keeps under consideration only those partial orders that are within K units of the current best partial locus order. If a partial order of the same length as the current best partial order is eliminated, all complete orders consistent with that partial order are eliminated as well. This strategy can result in the elimination of many more orders than branch-and-bound, but does not absolutely guarantee that the best locus order will be identified. Larger values of K increase the probability that the best order will be found, but also increase the number of orders that must be evaluated. In most of our analyses, we use the criterion $K = 8$ breaks or $\log_{10}$ likelihood units. In a few cases $K = 8$ did not allow identification of the true MB order. In these cases, we used $K = 20$. Our strategy here is possible only because we already knew the correct order. For a real RH mapping problem, we might first order the markers using the

faster criterion $K = 8$ and then rerun the analysis with a larger K to ensure that the same results were obtained. These strategies are described in detail in Boehnke *et al.* (1991) and implemented in the RH mapping software package RHMAP (Boehnke *et al.*, 1992).

Study Variables

The three ordering methods we examined can make use of partially typed hybrids. However, to simplify comparisons, we simulated RH data in which all markers were typed on all hybrids; hence, $x_i = 0$ or 1. For each combination of study variables considered, we generated 500 replicate data sets. In our simulated data, we varied four characteristics to study the ability of the methods to order markers under a wide range of circumstances that may be found in actual RH mapping studies.

Typical RH studies have typed 10 to 20 markers in 100 or fewer hybrids (e.g., Cox *et al.*, 1990; Burmeister *et al.*, 1991; Richard *et al.*, 1991, 1993; Warrington *et al.*, 1991; Frazer *et al.*, 1992; Gorski *et al.*, 1992). To reflect this range while dealing with manageable numbers of orders, we simulated $M = 8$, 12, or 16 markers in data sets of $H = 50$ or 100 hybrids. We also simulated data sets of $H = 200$ hybrids under some conditions to determine the benefit of greater sample sizes. In some cases, observed retention probabilities are well-approximated by an equal-retention model, but often fragment retention follows a pattern. For example, investigators have observed greater retention near the centromere on chromosomes 21 (Cox *et al.*, 1990; Burmeister *et al.*, 1991), 16 (Ceccherini *et al.*, 1992), and X (Gorski *et al.*, 1992). We simulated fragment retention either under the equal retention probability model with retention probability $r = 0.2$ or 0.5 or under the centromeric retention model with retention probabilities $r_1 = 0.8$ or 0.5 and $r_2 = 0.2$. When a RH mapping study begins, the relative spacing of markers within a length of chromosome is often unknown and can be assumed to be approximately random; if we wish to build a framework map, we want to have spacing as nearly equal as possible between markers. To represent these two situations, we simulated either equal spacing of markers, with interlocus distances of $d = 20$, 30, 40, or 50 cR, or random spacing of markers along a section of chromosome $D = d \times (M + 1)$ cR long, so that the average interlocus distance was $d = 10$, 20, 30, 40, or 50 cR. Our choice of range of interlocus distances reflects typical interlocus distances found in RH studies and our beliefs about what the optimal marker spacing might be.

Criteria for Evaluation

We employed three main criteria for comparing the three ordering methods and for determining optimal circumstances for accurately ordering markers. First, we compared the three methods by estimating the probability that the true order is identified as the unique best order. In situations in which that probability was low, we used a second criterion: the estimated probability that the true order is equivalent to the best order, that is, that the true order is one of one or more equivalently good best orders. A large difference between these two probabilities implies that the ordering method is not capable of discriminating between a number of equally good orders, often because no obligate breaks are present between pairs of markers in the data. For all situations that we considered, the ML and PP methods performed similarly under these criteria (see Results). Thus, for these methods we also compared the maximum likelihood ratio of the two best ML locus orders to the posterior probability ratio for these orders.

For our study of sample size and optimal study characteristics for RH mapping, we concentrated on results for the ML method. To examine the effects of the number of hybrids H, the number of markers M, the retention probability r, and the intermarker distance d, we used the two criteria described above and also compared methods based on a third criterion: the probability that the true order has maximum likelihood at least 1000 times the maximum likelihood of the next best order. For randomly spaced markers, we examined the distribution of the number of markers that could be placed in a 1000:1 framework map. A 1000:1 framework map is a set of loci that can be ordered with respect to each other at 1000:1 relative maximum likelihood; that is, the best order has maximum likelihood at least 1000 times that of any other order. For $H = 100$ and 200 and $r = 0.2$ and 0.5, we simulated $M = 16$ markers spaced randomly on chromosome segments with total distances of $D = 170$, 340, 510, 680, and 850 cR, corresponding to average marker spacings of $d = 10$, 20, 30, 40, and 50 cR. We obtained framework maps using stepwise locus ordering, adding a new locus to a partial order only if it could be ordered with 1000:1 relative likelihood. This method does not guarantee that the resulting framework map will contain the largest possible number of markers. However, limited experience suggests that the approach often generates nearly maximal framework maps (Boehnke *et al.*, 1992).

RESULTS

Comparison of Methods

Three locus results. For $M = 3$ markers, complete typing, and retention probability r fixed at $r = 0.5$, the ML and MB methods are equivalent. This fact is directly analogous to Thompson's (1987) result for the equivalence of the ML and minimum obligate recombinant methods for ordering three loci in linkage analysis. Following Thompson's argument, if $M = 3$, then there are only three possible orders: ABC, ACB, and BAC. The maximum log likelihood for the order ABC, given H completely typed hybrids, is

$$\log[L(\hat{\theta}, r = 0.5)] = \frac{H}{2}[g(\hat{\theta}_{AB}) + g(\hat{\theta}_{BC})] - H[M \log(2)], \quad [2]$$

where $g(\theta) = (2 - \theta)\log(2 - \theta) + \theta \log(\theta)$ and $\hat{\theta}_{ij}$ is twice the observed proportion of obligate breaks per hybrid between markers i and j. For orders ACB and BAC, the maximum log likelihoods are analogous, with $\hat{\theta}_{AC}$ substituted for $\hat{\theta}_{AB}$ and $\hat{\theta}_{BC}$, respectively. Hence, the maximum log likelihoods for each pair of orders have one breakage probability that is not in common. Since $g(\theta)$ is monotonic decreasing for increasing $0 < \theta \leq 1.0$, the order with the smallest pair of breakage probability estimates, and hence the fewest obligate breaks, will be the order with maximum likelihood. If $r \neq 0.5$ or $M > 3$, ML and MB are not equivalent (see Appendix).

Figure 1 displays analytically calculated probabilities that the true order is the unique best order for $M = 3$ equally spaced markers and fixed retention $r = 0.5$ for a range of interlocus distances d and number of hybrids H. When calculating these probabilities, we assume that there is evidence for linkage, i.e., that $\hat{\theta}_{AB}$ and $\hat{\theta}_{BC}$ are less than 1.0. Except when $H < 20$, the highest probability of correctly ordering the markers occurred for $d \approx 40$ cR, but remained close to its peak from 30 to 70 cR.

Figure 2 shows the effect of unequal marker spacing for $H = 50$ hybrids and $M = 3$ markers. Given a fixed total distance $D = d_1 + d_2$, the probability that the true order is the unique best order increased as d_1, the distance between the first two markers, approached $D/2$. That is, for $M = 3$ markers, the more nearly equal

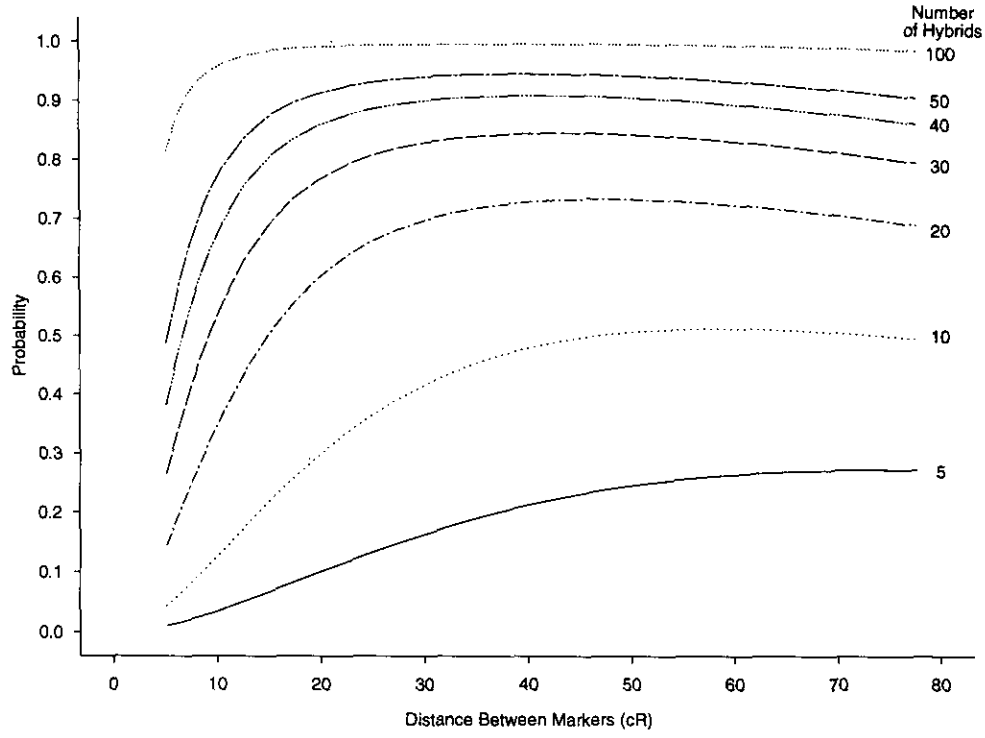

FIG. 1. Analytically calculated probabilities under the ML and MB methods that the true order is the unique best order by distance d between equally spaced markers and by number of hybrids H. $M = 3$ markers and fixed $r = 0.5$ retention.

the intermarker spacing, the higher the probability that the markers could be ordered correctly.

Multilocus results: Equally spaced markers. The probabilities that ML and PP identify the true order as the unique best order for equally spaced markers were virtually identical for every combination of H, M, r, and d considered; these probabilities were generally higher than those with MB (Table 1), although the dif-

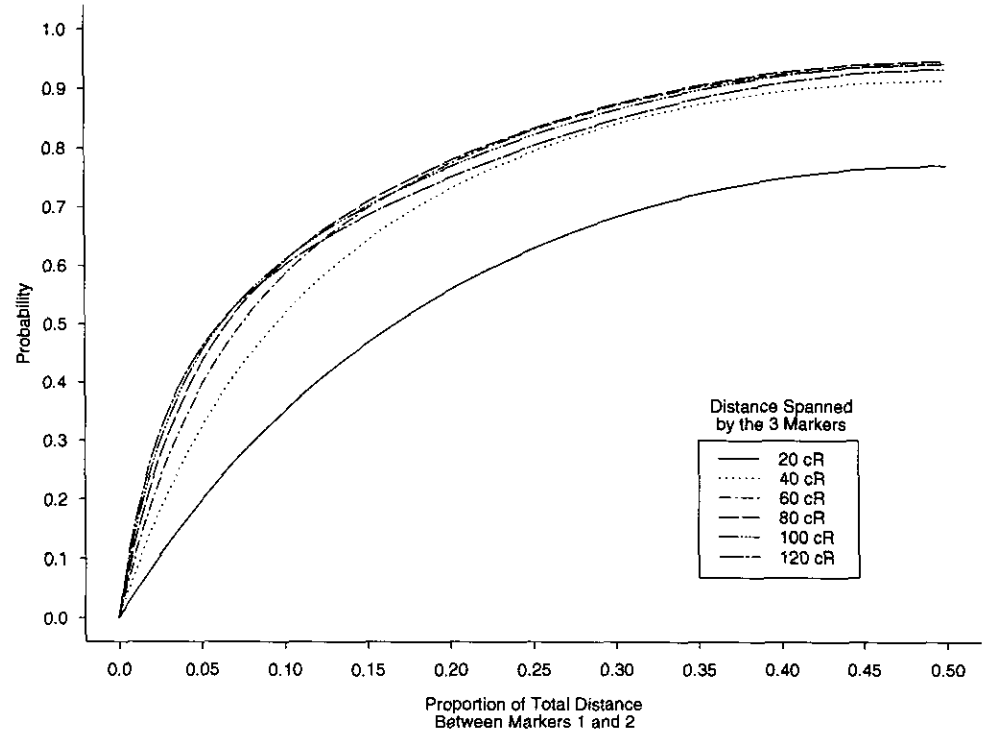

FIG. 2. Analytically calculated probabilities under the ML and MB methods that the true order is the unique best order by the proportion of the total map distance between the first two markers and total map distance spanned by the three markers. $H = 50$ hybrids, $M = 3$ markers, and fixed $r = 0.5$ retention.

TABLE 1

Estimated Probability That the True Order Is the Unique Best Order and That the True Order Has Maximum Likelihood at Least 1000 Times That of All Other Orders (Equally Spaced Markers)

				P(true order is unique best)			
H	*M*	*r*	*d* (cR)	MB	ML	PP	*P*(true order has ML 1000× any other order)
50	8	0.2	20	0.458	0.506	0.508	0.010
			30	0.626	0.710	0.712	0.010
			40	0.644	0.760	0.760	0.004
			50	0.586	0.716	0.720	0.000
		0.5	20	0.768	0.810	0.814	0.092
			30	0.880	0.916	0.916	0.094
			40	0.906	0.922	0.924	0.042
			50	0.868	0.918	0.918	0.020
100	8	0.2	20	0.884	0.920	0.920	0.402
			30	0.958	0.968	0.968	0.516
			40	0.956	0.986	0.986	0.454
			50	0.964	0.980	0.980	0.328
		0.5	20	0.984	0.994	0.994	0.814
			30	0.996	0.998	0.998	0.878
			40	0.998	1.000	1.000	0.824
			50	0.996	0.998	0.998	0.726
	12	0.2	20	0.862	0.896	0.896	0.330
			30	0.922	0.960	0.928	0.432
			40	0.940	0.966	0.966	0.384
			50	0.934	0.966	0.966	0.224
		0.5	20	0.996	0.990	0.990	0.802
			30	1.000	1.000	1.000	0.866
			40	0.998	1.000	1.000	0.842
			50	0.994	0.996	0.996	0.686
	16	0.2	20	0.854	0.892	0.892	0.252
			30	0.932	0.966	0.966	0.414
			40	0.944	0.954	0.952	0.348
			50	0.930	0.964	0.962	0.174
		0.5	20	0.990	0.994	0.994	0.776
			30	0.994	0.996	0.996	0.834
			40	0.994	0.996	0.990	0.808
			50	0.998	1.000	1.000	0.698

ferences were not always statistically significant, particularly for $r = 0.5$.

Since ML and PP appeared to correctly order equally spaced markers with essentially the same probability, we compared the level of support for the best order provided by the two methods by comparing the maximum likelihood ratio (MLR) and posterior probability ratio (PPR). The MLR is the ratio of the maximum likelihoods of the best ML order and the next best order. The PPR is the ratio of the posterior probabilities of these two best orders. ML and PP always identified the same two orders as the best two orders for equally spaced markers using the combinations of study variables we considered. Log(MLR) and log(PPR) had a nearly perfect linear relationship for each combination of H, M, r, and d considered; the Pearson correlations ranged from 0.942 to 0.999. The linear regression line with zero intercept, $\log(\text{PPR}) = b \times \log(\text{MLR}) + \text{error}$, fit the data very well for all combinations of H, M, r, and equally spaced d that we considered. The slope b was always slightly greater than 1.0, ranging from 1.03 to 1.23. Randomly spaced markers result in similar conclusions.

Since for all cases considered the correlations were near 1.0 and the regression slopes were always slightly greater than 1.0, it appears that the log(MLR) for the best and next best order can be used as an accurate, slightly conservative approximation for the log(PPR). Hence, for a $\log_{10}$(MLR) of 3 (or 1000:1 MLR), we can reasonably predict that the posterior probability of the best order is at least 1000 times the posterior probability of the next best order (see Discussion).

Effects of Study Variables on Accuracy of Ordering

Numbers of hybrids and markers. Typing $H = 100$ hybrids for $M = 8$, 12, or 16 markers equally spaced at $d = 20$–50 cR intervals with retention $r = 0.5$ resulted in very high probability that the true order was the unique best order (>0.98) for all three methods (Table 1). For retention $r = 0.2$ and $d = 30$–50 cR, this probability was still always greater than 0.92 (Table 1). Typing $H = 50$ hybrids resulted in substantially lower probabilities, even for just $M = 8$, particularly for retention $r = 0.2$, in which case even for $d = 40$ cR the probability for ML was only 0.76. In general, for fixed r and d, $M = 8$ markers resulted in the highest probabilities, followed by $M = 12$ and $M = 16$. The decrease in probability from $M = 8$ to $M = 12$ generally was greater than that from $M = 12$ to $M = 16$.

TABLE 2

Estimated Probability of Correctly Ordering *M* Randomly Spaced Markers with Average *d* = 40 cR (*H* = 100, *r* = 0.5)

	P(true order is unique best)			*P*(true order is a best order)			Average number of orders equivalent to true order[a]		
M	MB	ML	PP	MB	ML	PP	MB	ML	PP
8	0.546	0.602	0.602	0.892	0.826	0.826	1.574	1.329	1.131
12	0.392	0.462	0.462	0.870	0.808	0.804	2.264	1.829	1.231
16	0.274	0.334	0.334	0.790	0.710	0.708	3.499	2.335	1.285

[a] Average listed for when the true order is equivalent to the best order. The overall average number of orders equivalent to the true order is higher.

Equal marker spacing. For all three methods and for all combinations of H, M, and r, equal between-marker distances of $d = 30$–50 cR produced the highest probability that the true order is the unique best order. This probability increased the most from 20 to 30 cR and remained more constant between 30 and 50 cR. Forty centirays most often resulted in the highest probability for most combinations of H, M, and r considered (Table 1). When we used the more stringent requirement that the maximum likelihood of the true order must be at least 1000 times the maximum likelihood of any other order, spacing of $d = 30$ cR produced the highest probabilities of confidently ordering the markers for the cases considered (Table 1). These results suggest an optimal marker spacing that is somewhat more dense than the 55 cR suggested by minimizing the average coefficient of variation or standard error of the maximum likelihood estimate of the distance between two markers (Lange and Boehnke, 1992).

Retention probability. Consistent with Lange and Boehnke (1992), who showed that $r = 0.5$ results in the greatest amount of mapping information, we found that $r = 0.5$ gave substantially greater probability than $r = 0.2$ that the true order is the unique best in all situations tested (Table 1). The difference between $r = 0.5$ and $r = 0.2$ for the probability that the true order is equivalent to the best order was less substantial (data not shown).

Random marker spacing. When markers were randomly spaced, the probability that the true order is the unique best order was substantially decreased compared to that when markers were equally spaced given the same H, M, r, and average spacing d. $H = 100$ and $r = 0.5$ were insufficient to order correctly with high probability even $M = 8$ markers spaced at an average distance $d = 40$ cR (Table 2). While the probability of identifying the true order as the unique best order was relatively low even for $M = 8$, the probability that the true order was one of the equally good best orders remained greater than 0.7, even for $M = 16$. ML and PP produced slightly lower probabilities than MB that the true order was equivalent to the best order, but the average number of orders equivalent to the true order when the true order was the best order was substantially greater for MB than for ML or PP (Table 2). This result is explained by the fact that the MB method produces discrete-valued scores that can range from 0 to $H \times (M - 1)$ obligate breaks, while ML and PP are continuous-valued scores with infinitely many possible values. Since there are fewer possible values for orders under MB, the true order is more likely to have the same number of obligate breaks as the "best" MB order, particularly when the number of obligate breaks is small.

Building a framework map. Since the three methods appeared to have the same relative merits for ordering randomly spaced markers as for ordering equally spaced markers, we concentrated on the ML method for randomly spaced markers. Consider a situation in which we have a number of new markers to be ordered, and we have some knowledge about the total length of the chromosome segment on which the markers are located. We may wish to know how many markers we can reasonably hope to place in a framework map using ML so that all the markers are mapped with 1000:1 relative maximum likelihood. Figure 3 displays the estimated probabilities of ordering at least n markers ($3 \leq M \leq 16$) at MLR of at least 1000 for $H = 100$ and 200, $r = 0.2$ and 0.5, and a range of total distances D corresponding to average marker spacings of $d = 10$, 20, 30, 40, and 50 cR. Total distances of $D = 170$ and 340 cR, corresponding to average spacings of 10 and 20 cR, appeared to give the highest probabilities of ordering at least 13 markers in a 1000:1 framework map. With $H = 100$ hybrids, retention $r = 0.5$, and total distances of 170 to 340 cR, we can expect to be able to include at least 9 markers in a 1000:1 framework map with probability near 0.90 and at least 10 markers with probability greater than 0.70. With $H = 200$ hybrids, we can expect to be able to include at least 12 markers in a 1000:1 framework map with probability greater than 0.90. Even with $H = 200$ and $r = 0.5$, the probability of ordering all 16 markers in a 1000:1 framework map is low.

Maximum Likelihood Breakage Probability Estimates

One of the benefits of the ML approach to ordering RH marker data is that it provides estimates of the breakage probabilities, and hence distances, between the markers. Table 3 lists the actual and ML-estimated total map lengths (distance between most distant markers) for the true order for the equal spacing case. The map lengths were consistently overestimated for

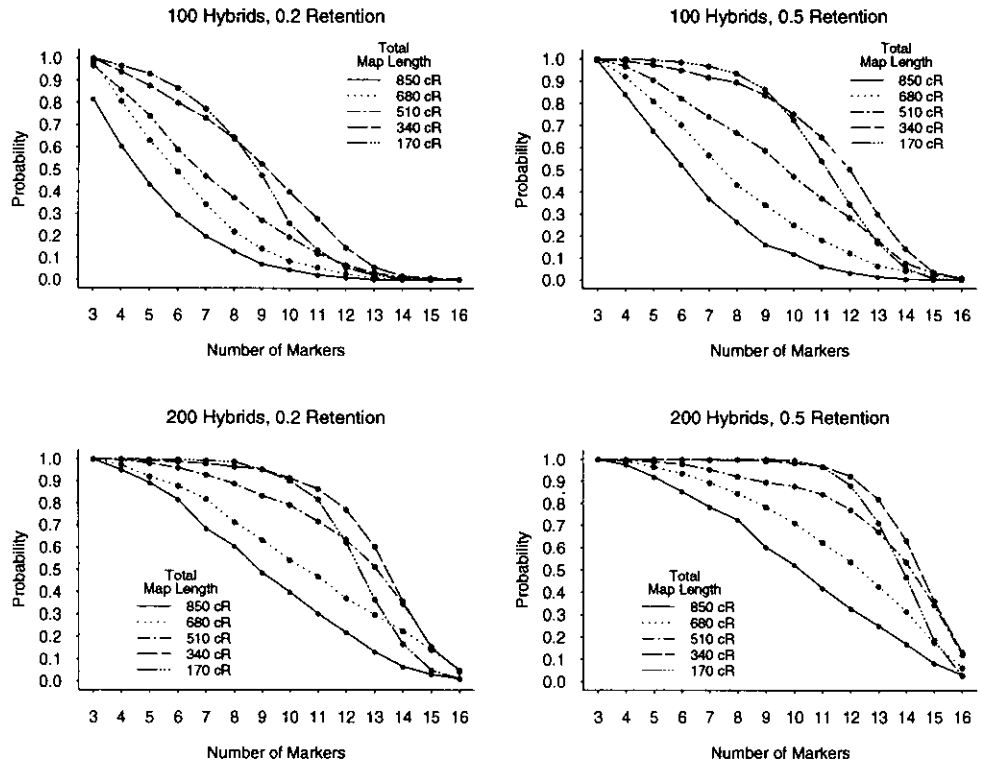

FIG. 3. Estimated probabilities of ordering at least M (≤ 16) randomly spaced markers in a 1000:1 maximum likelihood framework map by total map length. $H = 100$ and $H = 200$ hybrids, $r = 0.2$ and $r = 0.5$ retention.

the conditions we considered, with proportional bias $[(\hat{\theta} - \theta)/\theta]$ ranging from 0.007 to 0.098. Holding all other study variables constant, the bias generally was more severe for $H = 50$ than for $H = 100$, for $r = 0.2$ than for $r = 0.5$, and for larger d than for smaller d. Bias also appeared to be more severe for larger M, particularly when $r = 0.2$. Limited simulations and analyses of data sets with $H = 1000$ hybrids suggest that the bias observed in Table 3 is a consequence of relatively small hybrid sample size.

Centromeric Retention Effect

Often investigators observe that marker retention is not uniform, but shows a gradient with higher values near the centromere. In this situation we can attempt

TABLE 3

Equal Retention Data: ML Map Length Estimates for Equally Spaced Markers

				$r = 0.2$		$r = 0.5$	
H	M	d (cR)	Actual map length (cR)	Map length estimate (cR)	SE (cR)	Map length estimate (cR)	SE (cR)
50	8	20	140	150.0	1.8	145.2	1.2
		30	210	225.9	2.4	220.4	1.7
		40	280	307.4	3.1	291.3	2.0
		50	350	381.4	3.3	367.5	2.3
100	8	20	140	144.3	1.2	143.9	0.9
		30	210	216.3	1.5	214.6	1.1
		40	280	290.8	1.9	284.7	1.3
		50	350	362.2	2.4	360.5	1.6
100	12	20	220	225.1	1.4	221.5	1.1
		30	330	344.2	2.0	339.3	1.5
		40	440	455.2	2.4	450.9	1.7
		50	550	572.7	2.8	560.4	2.0
100	16	20	300	308.6	1.5	303.6	1.2
		30	450	465.5	2.3	458.6	1.7
		40	600	620.2	2.8	613.2	2.0
		50	750	778.0	3.3	763.5	2.3

TABLE 4

Centromeric Retention Data: Comparison of Probabilities of Correct Ordering for Analysis under the Equal Retention Model and under the Centromeric Retention Model (H = 100, M = 12, r_2 = 0.2)[a]

r_1	d (cR)	Analysis model	P(true order is unique best)	P(true order has ML 1000× any other order)
0.8	20	Centromeric	0.994	0.818
		Equal	0.990	0.774
	40	Centromeric	0.994	0.674
		Equal	0.988	0.740
0.5	20	Centromeric	0.992	0.660
		Equal	0.992	0.640
	40	Centromeric	0.992	0.650
		Equal	0.992	0.656

[a] For the centromeric model, if the same order is present in both orientations to the centromere, then only the order that is most likely is considered.

to model the retention gradient in the analysis, or we can incorrectly assume equal retention. Recall that under the centromeric retention model, fragments that included the most centromeric marker have retention probability r_1, while all other fragments have retention probability r_2 ($r_1 \geq r_2$). This results in a gradient of marker retention probabilities, with the highest retention nearest the centromere, as seen in some RH data. The assumption of equal retention allows the consideration of (1) only half as many orders, since orientation along the chromosome is not important in the equal retention case, and (2) one less retention parameter, simplifying the analysis for each order. For the cases we considered for centromeric data (r_1 = 0.5 or 0.8, r_2 = 0.2, equally spaced markers, d = 20 or 40 cR), the probability that the true order is the unique best order given analysis under the centromeric model was virtually equal to the probability given analysis under the equal retention model. The differences between analysis under the centromeric model and analysis under the equal retention model in the probability that the true order has maximum likelihood at least 1000 times greater than the next best order were not statistically significant, although it is interesting to note that for d = 40 cR, the probability appeared to be slightly lower for the centromeric model than for the equal retention model (Table 4), while for d = 20 cR, the probability appeared to be higher for the centromeric model than for the equal retention model.

The assumption of equal retention for centromeric model data affected the estimates of the breakage probabilities of centromeric data more severely than the probability of ordering the markers correctly (Table 5). As for the equal retention data (Table 3), analysis of the centromeric retention data under the centromeric model resulted in overestimates of the actual map length for the true order. Analyzing the centromeric model data under the equal retention model resulted in underestimation of the map length for all cases considered. In each case considered, the bias was of substantially greater magnitude for analysis under the equal retention model than for analysis under the correct centromeric model.

For the centromeric data, the estimates of the breakage probabilities assuming equal retention formed a gradient, with the larger estimates near the centromere. In general, the first few breakage probability estimates were overestimates of the true probability, and the rest were underestimates. This phenomenon can be explained by examining the expected number of obligate breaks under the two models.

Under any analysis model, the expected number of obligate breaks between two markers should approximate the observed number of obligate breaks. Let R_i be the probability of retaining locus i. If we assume equal breakage probabilities θ, under the centromeric model, $R_i = (r_1 - r_2)(1 - \theta)^{i-1} + r_2$. The expected number of obligate breaks between markers i and $i + 1$ for centromeric data and for equal retention data are

$$B_{c_i} = \theta[R_i(1 - R_{i+1}) + R_{i+1}(1 - R_i)]$$
$$= \theta[(r_1 - r_2)(1 - \theta)^{i-1}(2 - \theta)(1 - 2r_2) - 2(r_1 - r_2)^2(1 - \theta)^{2i-1} + 2r_2(1 - r_2)]$$

and

$$B_{e_i} = \theta[2r(1 - r)], \qquad [3]$$

respectively. Substituting the average retention $r^* = (1/M)\ \Sigma_{i=1}^{M} R_i = r_2 + \{(r_1 - r_2)[1 - (1 - \theta)^M]/M\theta\}$ for r in [3] gives an estimate for the expected number of breaks for centromeric data analyzed under the equal retention model. A comparison of the retention R_i versus the average retention r^* shows that for the first few markers, R_i is greater than r^*. For B_c and B_e to agree, the estimate of θ between these markers would have to be inflated by the equal retention analysis. For the remaining intervals, r^* is greater than R_i, so that the estimate of θ by the equal retention analysis would have to be less than the true θ.

Evaluating a Possible Sample Size Approximation

It is tempting to estimate hybrid sample size requirements for RH mapping by using the fact that to distin-

TABLE 5

Centromeric Retention Data: Map Length Estimates under Centromeric and Equal Retention Analysis Models, Equally Spaced Markers (H = 100, M = 12, r_2 = 0.2)

			Centromeric model		Equal retention model	
r_1	d (cR)	Actual map length (cR)	Map length estimate (cR)	SE (cR)	Map length estimate (cR)	SE (cR)
0.8	20	220	222.5	1.1	214.2	0.9
	40	440	450.1	1.9	384.5	1.5
0.5	20	220	226.1	1.3	199.8	1.1
	40	440	453.7	2.1	400.6	1.8

guish the possible marker orders with any analysis method, at least one obligate break must be observed between each pair of adjacent markers. Given H hybrids and retention probability r, the number of obligate breaks between two loci separated by breakage probability θ follows a binomial distribution on H trials with probability of success $p = 2\theta r(1 - r)$. Hence, the probability of observing at least B obligate breaks between each of the $M - 1$ adjacent marker pairs is

$$[\sum_{y=B}^{H} \binom{H}{y} p^y (1 - p)^{H-y}]^{M-1}.$$

Table 6 lists the probability of at least B = 1, 2, or 3 obligate breaks between each pair of markers and the estimated probability that the true order is the unique best ML order. As expected, the probability of observing at least one obligate break between each pair of markers is always greater than the ML probability of correct ordering. However, this is often a very crude upper bound, and there does not appear to be a clear relationship between the probability of identifying the true order as the unique best order using ML and the probability of observing at least B = 1, 2, or 3 breaks. Hence, the probability of at least one obligate break between each pair of adjacent markers does not appear to be a reliable substitute for the probability that the true order is the unique best order for estimating sample size.

DISCUSSION

Multipoint methods for RH mapping make efficient use of information from all loci simultaneously and take advantage of information contained in incompletely typed radiation hybrids. We compared three multipoint methods for ordering genetic loci using radiation hybrid data and determined the limits of ordering accuracy of RH mapping under conditions characteristic of recent RH mapping experiments. A variety of other strategies have been developed, including the methods of Falk (1991, 1992), Wilson (1992), and Cox *et al.* (1990), which combine two-locus results to build a multipoint map, and the multipoint methods of Green (1992) and Lawrence and Morton (1992), which are very similar to the maximum likelihood approach of Boehnke *et al.* (1991). Guerra *et al.* (1992) developed

TABLE 6

Probability of Observing at Least B Obligate Breaks between Each of $M - 1$ Adjacent Marker Pairs Compared to the ML Estimated Probability That the True Order Is the Unique Best Order

				P(at least B obligate breaks)			
H	M	r	d (cR)	B = 1	B = 2	B = 3	P(true order is unique best)
50	8	0.2	20	0.6963	0.1997	0.0173	0.5060
			30	0.9113	0.5892	0.2010	0.7100
			40	0.9737	0.8306	0.5135	0.7600
			50	0.9917	0.9333	0.7503	0.7160
		0.5	20	0.9410	0.6894	0.3026	0.8100
			30	0.9932	0.9441	0.7822	0.9160
			40	0.9991	0.9907	0.9508	0.9220
			50	0.9999	0.9984	0.9894	0.9180
100	16	0.2	20	0.9626	0.7594	0.3599	0.8920
			30	0.9974	0.9741	0.8761	0.9660
			40	0.9998	0.9972	0.9825	0.9540
			50	1.0000	0.9997	0.9975	0.9640
		0.5	20	0.9989	0.9878	0.9346	0.9940
			30	1.0000	0.9998	0.9982	0.9960
			40	1.0000	1.0000	1.0000	0.9960
			50	1.0000	1.0000	1.0000	1.0000

an alternative Bayesian approach, but it is apparently limited to the analysis of three loci at a time.

The three multipoint methods that we compared each have advantages and disadvantages. The minimum breaks method demands only minimal assumptions about the data and requires the simplest computations; it is the fastest of the three methods to use. However, it cannot supply estimates of intermarker distances and does not provide a meaningful way to compare the relative probabilities of competing orders. The maximum likelihood method provides distance estimates between markers, but requires both the assumption that fragments are retained independently and the specification of a model for fragment retention. In addition, ML demands more complex computation than the minimum breaks method and is therefore considerably slower. ML allows the comparison of orders via relative maximum likelihoods, but there can be no formal tests to determine the probability that an order is the true order. The Bayesian maximum posterior probability method allows direct comparison of orders through their posterior probabilities and can also be used to estimate distances between markers (Lange and Boehnke, 1992). Its main failings are that it requires a set of predetermined best orders, and therefore requires the use of an alternative method to determine the best orders, and that it is computationally complex and therefore slow.

The positive and negative aspects of the ML and PP methods are partially reconciled by our finding that we can use the log(MLR) as a slightly conservative approximation of the log(PPR). Our approximation for the denominator of the posterior probability will cause the posterior probabilities to be slightly overestimated. However, the denominator approximation does not affect the probability ratio. Therefore, our result that the slope of the regression line for log(PPR) and log(MLR) is always slightly greater than 1.0 is not an artifact of this part of the approximation of the posterior probability.

The advantage of the PPR over the MLR is its simple, direct interpretation: PPR = 1000 means that the posterior probability of the best order has posterior probability 1000 times greater than that of any other order. Using the MLR as an approximation for the PPR eliminates the need for the additional complex computations of the PP method, but allows us to make a simple, direct comparison of the best order to the other orders. In addition, since the MLR can be used to approximate the PPR, it would be reasonable to approximate the posterior probability of a particular order by the maximum likelihood of that order divided by the sum of the maximum likelihoods of the set of best ML orders. Hence, we can approximate the actual posterior probabilities using the maximum likelihoods. Rogatko and Zacks (1993) demonstrate an analogous result in the context of linkage analysis.

In most RH mapping projects to date, the number of hybrids H has been near 100. We simulated samples of sizes $H = 50$, 100, and 200 in different situations to determine the range of accuracy that can be expected. The retention models and ranges of retention values we have simulated approximate the values seen in data from many RH mapping projects. We chose to simulate equally spaced markers because this represents a best-case scenario and because when we build a framework map, we generally want to produce a map with approximately equally spaced markers. If there are enough markers, it may be possible to choose a subset of markers that are approximately equally spaced; our results suggest that these markers will be ordered correctly with high probability under many conditions common to recent RH mapping experiments. Generally, markers are not equally spaced. Our random-spacing results should serve as an indication of what can be expected in these situations.

We have assumed complete marker typing in all of our simulations. While incomplete marker typing may affect the performance of the ordering methods we compared and will lessen the power to determine marker order accurately, results for complete marker typing should be similar when the percentage of untyped loci is low, as has generally been the case in recent studies (e.g., Warrington *et al.*, 1991; Gorski *et al.*, 1992; Richard *et al.*, 1993).

For randomly spaced markers, total distances corresponding to average marker spacings of 10 to 20 cR gave the highest probability of mapping the largest proportion of markers in a 1000:1 framework map (Fig. 3). One of the attractive features of RH mapping is that some elements of experimental design can be used to increase the probability that the correct order will be inferred (Goss and Harris, 1975, 1977; Lange and Boehnke, 1992). In theory, the dose of radiation applied to the chromosome can be adjusted to increase or decrease the average number of breaks per chromosome and hence between markers. For example, suppose we want to order 16 markers that lie within an interval approximately 34 Mb in length so that the average distance between markers is approximately 2 Mb. If previous RH experiments in similar situations suggest that 50 kb $\approx$ 1 cR under 8000-rad radiation, then we would expect an average of 0.40 breaks or 40 cR between markers. We would want to decrease the radiation dose somewhat for our experiment to decrease the average number of breaks between markers and hence increase the probability of being able to map a large proportion of the markers in a 1000:1 map.

Since in most cases we simulated RH data and then analyzed the data under the model we knew to be correct, our results may tend to overestimate the ordering accuracy of the three methods we compared. However, our results concerning data with a gradient of retention probabilities confirm that, as a number of researchers have already noted (Boehnke *et al.*, 1991; Chakravarti and Reefer, 1992; Lange and Boehnke, 1992), the retention model that is assumed does not strongly affect which orders are inferred to be the best. It is important to note that the map length and intermarker distance estimates may be more strongly influenced by the re-

tention model assumed, particularly when a strong centromeric effect is present.

We have shown that typing 100 hybrids results in a high probability of accurate ordering of up to 16 markers, given fragment retention probabilities in the range 0.2 to 0.5 and approximately equal marker spacing with intermarker distances between 30 and 50 cR. If markers are not approximately equally spaced, 100 hybrids still give high probability that a fairly large proportion of the loci can be placed into a 1000:1 maximum likelihood framework map, provided that the fragment retention is near 0.5 and the markers are spaced at average distances of 10 to 20 cR. These findings should be useful in planning future RH mapping studies.

APPENDIX

We demonstrate in the text that if the number of markers $M = 3$ and the retention probability r is fixed at 0.5, then the minimum breaks and maximum likelihood methods are equivalent. We now show that both of these conditions are necessary to guarantee this equivalence. If $M = 3$ but $r \neq 0.5$, then the likelihood does not have the simple form of [2], and the breakage parameter estimates are not a simple function of the number of obligate breaks. The RH data

A	B	C	Number of hybrids observed
0	1	1	1
1	0	1	1
1	0	0	1
0	0	1	2
0	0	0	5
			10

yield the following results under the maximum likelihood estimates $\hat{r}$:

Order	Obligate breaks	Maximum log likelihood
ABC	6	−6.995
ACB	7	−6.985
BAC	7	−7.068

For MB, the best order is ABC, followed by ACB and BAC, which are equally good. But for ML evaluated at the maximum likelihood estimate $\hat{r}$, the order ACB is best, followed by ABC and BAC.

When $r = 0.5$ but $M \geq 4$, the log likelihood has a form analogous to [2]. For example, for the order ABCD, the maximum log likelihood is

$$\log[L(\hat{\theta}, r = 0.5)] = \frac{H}{2}[g(\hat{\theta}_{AB}) + g(\hat{\theta}_{BC}) + g(\hat{\theta}_{CD})] - H[M \log(2)].$$

However, there are cases where $M \geq 4$ for which comparisons must be made between orders having zero or having one breakage probability in common. For example, orders ABCD and ADBC have only θ_{BC} in common; orders ABCD and BDAC have no breakage probabilities in common. When there is one probability in common, the log likelihood will depend on the two estimates $\hat{\theta}_i$ and $\hat{\theta}_j$ not held in common through $[g(\hat{\theta}_i) + g(\hat{\theta}_j)]$, while the number of obligate breaks will depend on $(\hat{\theta}_i + \hat{\theta}_j)$. Since $[g(\hat{\theta}_i) + g(\hat{\theta}_j)]$ is not monotonic in $(\hat{\theta}_i + \hat{\theta}_j)$, situations exist in which the MB and ML ranking of orders will not agree. A similar argument holds if no estimates are in common. For example, the RH data

A	B	C	D	Number of hybrids observed
0	0	0	0	1
1	0	0	0	2
1	1	0	0	1
1	1	1	0	1
1	1	1	1	5
				10

give the following results for the three best orders:

		Maximum log likelihood	
Order	Obligate breaks	$r = \hat{r}$	$r = 0.50$
ABCD	4	−7.434	−8.007
BCDA	6	−7.731	−8.757
ABDC	5	−8.025	−8.769

Using MB, we would rank ABCD as the best order, ABDC as second best, and BCDA as third best. Using ML with $r = 0.5$ or the maximum likelihood estimate $\hat{r}$, we would reverse the rank of the second and third orders.

ACKNOWLEDGMENTS

This research was supported by National Institutes of Health Grants HG00376 and HG00209 (M.B.) and a Regents predoctoral fellowship from the University of Michigan (K.L.L.).

REFERENCES

Barrett, J. H. (1992). Genetic mapping based on radiation hybrids. *Genomics* **13:** 95–103.

Bishop, D. T., and Crockford, G. P. (1992). Comparisons of radiation hybrid mapping and linkage mapping. *Cytogenet. Cell Genet.* **59:** 93–95.

Boehnke, M. (1992). Radiation hybrid mapping by minimization of the number of obligate chromosome breaks. *Cytogenet. Cell Genet.* **59:** 96–98.

Boehnke, M., Hauser, E., Lange, K., Uro, J., and Vanderstoep, J. (1992). "RHMAP: Statistical package for multipoint radiation hybrid mapping" (FORTRAN), unpublished computer program documentation.

Boehnke, M., Lange, K., and Cox, D. R. (1991). Statistical methods for multipoint radiation hybrid mapping. *Am. J. Hum. Genet.* **49:** 1174–1188.

Burmeister, M., Kim, S., Price, E. R., de Lange, T., Tantravahi, U., Myers, R. M., and Cox, D. R. (1991). A map of the distal region of the long arm of human chromosome 21 constructed by radiation hybrid mapping and pulsed-field gel electrophoresis. *Genomics* **9:** 19–30.

Ceccherini, I., Romeo, G., Lawrence, S., Breuning, M. H., Harris,

P. C., Himmelbauer, H., Frischauf, A. M., Sutherland, G. R., Germino, G. G., Reeders, S. T., and Morton, N. E. (1992). Construction of a map of chromosome 16 by using radiation hybrids. *Proc. Natl. Acad. Sci. USA* **89:** 104–108.

Chakravarti, A., and Reefer, J. E. (1992). A theory for radiation hybrid (Goss–Harris) mapping: Application to proximal 21q markers. *Cytogenet. Cell Genet.* **59:** 99–101.

Cox, D. R., Burmeister, M., Price, E. R., Kim, S., and Myers, R. M. (1990). Radiation hybrid mapping: A somatic cell genetic method for constructing high resolution maps of mammalian chromosomes. *Science* **250:** 245–250.

Falk, C. T. (1991). A simple method for ordering loci using data from radiation hybrids. *Genomics* **9:** 120–123.

Falk, C. T. (1992). Multilocus mapping strategies on chromosome 21 data sets: Comparison of results from family data, radiation hybrids and somatic cell hybrids. *Cytogenet. Cell Genet.* **59:** 119–121.

Frazer, K. A., Boehnke, M., Budarf, M. L., Wolff, R. K., Emanuel, B. S., Myers, R. M., and Cox, D. R. (1992). A radiation hybrid map of the region on human chromosome 22 containing the neurofibromatosis type 2 locus. *Genomics* **14:** 574–584.

Gorski, J. L., Boehnke, M., Reyner, E. L., and Burright, E. N. (1992). A radiation hybrid map of the proximal short arm of the human X chromosome spanning incontinentia pigmenti 1 (IP1) translocation breakpoints. *Genomics* **14:** 657–665.

Goss, S. J., and Harris, H. (1975). New method for mapping genes in human chromosomes. *Nature* **255:** 680–684.

Goss, S. J., and Harris, H. (1977). Gene transfer by means of cell fusion. *J. Cell Sci.* **25:** 17–57.

Green, P. (1992). Construction and comparison of chromosome 21 radiation hybrid and linkage maps using CRI-MAP. *Cytogenet. Cell Genet.* **59:** 122–124.

Guerra, R., McPeek, M. S., Speed, T. P., and Sewart, P. M. (1992). A Bayesian analysis for mapping from radiation hybrid data. *Cytogenet. Cell Genet.* **59:** 104–106.

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

Lange, K., and Boehnke, M. (1992). Bayesian methods and optimal experimental design for gene mapping by radiation hybrids. *Ann. Hum. Genet.* **56:** 119–144.

Lawrence, S., and Morton, N. (1992). Physical mapping by multiple pairwise analysis. *Cytogenet. Cell Genet.* **59:** 107–109.

Nijenhuis, A., and Wilf, H. S. (1978). "Combinatorial Algorithms," pp. 240–246, Academic Press, New York.

Richard, C. W., Cox, D. R., Kapp, L., Murnane, J., Cornelis, F., Julier, C., Lathrop, G. M., and James, M. R. (1993). A radiation hybrid map of human chromosome 11q22–q23 containing the ataxia-telangiectasia disease locus. *Genomics* **17:** 1–5.

Richard, C. W., Withers, D. A., Meeker, T. C., Maurer, S., Evans, G. A., Myers, R. M., and Cox, D. R. (1991). A radiation hybrid map of the proximal long arm of human chromosome 11 containing the multiple endocrine neoplasia type 1 (MEN-I) and *bcl*-I disease loci. *Am. J. Hum. Genet.* **49:** 1189–1196.

Rogatko, A., and Zacks, S. (1993). Ordering genes: Controlling the decision-error probabilities. *Am. J. Hum. Genet.* **52:** 947–957.

Thompson, E. A. (1987). Crossover counts and likelihood in multipoint linkage analysis. *IMA J. Math. Appl. Med. Biol.* **4:** 93–108.

Warrington, J. A., Hall, L., Hinton, L. M., Miller, J. N., Wasmuth, J. J., and Lovett, M. (1991). Radiation hybrid map of 13 loci on the long arm of chromosome 5. *Genomics* **11:** 701–708.

Weeks, D. E., Lehner, T., and Ott, J. (1992). Preliminary ranking procedures for multilocus ordering based on radiation hybrid data. *Cytogenet. Cell Genet.* **59:** 125–127.

Wilson, S. R. (1992). A statistical and numerical view of the chromosome 21 physical map data. *Cytogenet. Cell Genet.* **59:** 110–111.