

# Confidence limits for estimates of gene linkage based on analysis of recombinant inbred strains

Jonathan Silver

**ABSTRACT:** Recombinant inbred (RI) strains are extremely useful for genetic mapping. This paper presents a simple method for determining confidence intervals for linkage estimates based on analysis of RI strains. The results show that such confidence intervals are usually large with the currently available numbers of RI strains. Therefore, map positions based only on analysis of RI strains should be interpreted with caution. To facilitate interpretation of linkage data derived from RI strains, a table is presented giving the 95 percent and 99 percent confidence intervals for all possible linkages detected with up to 45 RI strains.

RECOMBINANT INBRED (RI) strains<sup>1,3,6</sup> are used more and more frequently for genetic mapping. Standard formulas are available for estimating the recombination fraction,  $r$ , between two loci based on the analysis of RI strains<sup>1,3,4,6</sup>. However, the calculation of appropriate confidence limits for  $r$  poses a problem that has not been adequately addressed in the literature. This communication presents a simple method for calculating such confidence limits, along with a table for the most commonly encountered situations. The results indicate that most linkage relationships detected by analysis of RI strains will have large confidence intervals, unless the number of recombinants is very small and the number of strains analyzed is large.

RI strains are families of inbred strains derived by inbreeding the progeny of a cross between two parental strains. Closely linked genes tend to be inherited together during the crosses used to establish RI strains, so that parental combinations of alleles at these loci tend to be maintained. When brother-sister matings are used to construct RI strains, the relationship between the probability,  $r$ , of recombination between two loci in a single meiosis and the probability,  $R$ , that an RI strain will carry a recombinant (nonparental) combination of alleles at these loci is given by<sup>4</sup>:

$$R = 4r/(1 + 6r) \quad (1)$$

$R$  also can be interpreted as the expected proportion of RI strains that carry recombinant

combinations of alleles at the two loci, and  $r$  as the map distance (for small values of  $r$ ) between the two loci. In formula (1),  $r$  is considered to lie in the range  $0 \leq r \leq 1/2$ , which implies that  $0 \leq R \leq 1/2$ .

Formula (1) gives the relationship between the probabilities  $r$  and  $R$ . These probabilities are not measured directly, but are estimated from observations on a set of RI strains. Suppose that the number of RI strains tested is  $N$  and that  $i$  of these strains are found to carry recombinant combinations of alleles at the two loci. Then the observed proportion  $\hat{R} = i/N$  provides an estimate for the "true" probability,  $R$ , that an RI strain will carry a recombinant chromosome for these two loci. The corresponding estimate,  $\hat{r}$ , for the recombination fraction between the two loci is obtained by solving (1) for  $r$  and substituting the estimate,  $\hat{R}$ , for  $R$ . This gives:

$$r = R/(4 - 6R) \quad (2)$$

and

$$\hat{r} = \hat{R}/(4 - 6\hat{R}). \quad (3)$$

While formulas (2) and (3) have singularities for  $R = 2/3$  and  $\hat{R} = 2/3$ , these values lie outside the range of biological interest ( $0 \leq R \leq 1/2$ ) and do not affect the subsequent analysis.

How would the estimates  $\hat{R}$  and  $\hat{r}$  vary if the experiment were repeated many times, i.e., if many independent sets of RI strains, each containing  $N$  strains, were examined for the same loci? If the "true" probability of an RI strain being recombinant for the two loci is  $R$ , then

The author is affiliated with the Laboratory of Molecular Microbiology, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, Maryland 20205. He thanks Dr. David Alling and Dr. Earl Green for helpful discussions and detailed review of this manuscript.

© 1985, American Genetic Association.

the probability that  $i$  strains out of  $N$  would be found to be recombinant is given by the binomial proportion:

$$P_i = \binom{N}{i} (R)^i (1-R)^{N-i} \quad (4)$$

where  $\binom{N}{i} = \frac{N!}{i!(N-i)!}$ . Since  $P_i$  also is the probability of obtaining an estimate  $\hat{R} = i/N$ , the set of estimates  $\{\hat{R}\}$  follows a binomial distribution and has a standard deviation (STD) equal to  $[R(1-R)/N]^{1/2}$  (see Snedecar and Cochran<sup>5</sup>). The standard deviation of the set of estimates  $\{\hat{r}\}$  can then be approximated (see Green<sup>3</sup>) using the "propagation of error" rule that the standard deviation of a function,  $f(x)$ , is approximately equal to  $df/dx$  times the standard deviation of  $x$ . This gives  $\text{STD}(\hat{r}) \cong [d\hat{r}/d\hat{R}] \text{STD}(\hat{R})$ . Evaluating  $d\hat{r}/d\hat{R}$  from (3), substituting the  $\text{STD}(\hat{R})$ , and expressing the result in terms of the estimate  $\hat{r}$  gives:

$$\text{STD}(\hat{r}) \cong [(1 + 6\hat{r})/2] [\hat{r}(1 + 2\hat{r})/N]^{1/2}. \quad (5)$$

While formula (5) is technically correct, it is not very useful for estimating confidence limits for  $\hat{r}$ . If the set of estimates  $\{\hat{r}\}$  followed a normal distribution, then 95 percent of the estimates would be expected to lie within 1.96 standard deviations of the mean, giving 95 percent confidence limits of  $\hat{r} \pm 1.96 \text{STD}(\hat{r})$ . This normal approximation has been used frequently in the literature. However, the assumption of normality is seriously inaccurate. The number of RI strains tested is usually small and the proportion of strains that are recombinant has to be small for the data to indicate linkage. In this case the binomial distribution of the estimates  $\{\hat{R}\}$  is quite skewed and therefore quite different from the normal distribution, as is the distribution of the estimates  $\{\hat{r}\}$ . Use of the normal approximation in this situation often leads to meaningless results. For example, the range  $\hat{r} \pm 1.96 \text{STD}(\hat{r})$  frequently includes negative numbers and zero. But  $r < 0$  is meaningless, and  $r = 0$  is clearly wrong if one has already observed recombinants among the RI strains. If no recombinants are observed in a set of RI strains, formula (3) gives  $\hat{r} = 0$  and (5) gives  $\text{STD}(\hat{r}) = 0$ . This means that confidence limits based on multiples of  $\text{STD}(\hat{r})$  also will be zero. But the upper confidence limit for  $\hat{r}$  should not be zero, because it is intuitively clear that a small distance between two loci ( $r > 0$ ) is perfectly compatible with observing no recombinants in a limited number of RI strains.

More accurate and meaningful confidence limits for  $\hat{r}$  can be derived based on the binomial distribution rather than the normal distribution. Suppose that in a set of  $N$  RI strains one observes  $i$  recombinants for two loci of

interest. We want to know the range of possible recombination fractions between the two loci for which the observed number of recombinants is not too unlikely. This range of possible recombination fractions forms a confidence interval for  $\hat{r}$ , as illustrated in the following example. Suppose one observes five recombinants in a set of 23 RI strains. If the true recombination percentage were greater than 32, the chance of seeing five or fewer recombinants would be less than 2.5 percent; and if the true recombination percentage were less than 2, the chance of seeing five or more recombinants would be less than 2.5 percent. In this case, the range 2-32 would form the 95 percent confidence limits for  $\hat{r}$ . The method of calculating these confidence limits follows.

The probability of observing  $i$  or more recombinants in a set of  $N$  RI strains, designated  $P_{i,N}$  is given by the sum  $P_i + P_{i+1} + \dots + P_N$ . Using (4) and (1) to express  $P_{i,N}$  in terms of  $r$  gives:

$$P_{i,N} = \sum_{j=i}^N \binom{N}{j} [4r/(1+6r)]^j \{1-[4r/(1+6r)]\}^{N-j}. \quad (6)$$

It is clear from (6) that for  $i \neq 0$ ,  $P_{i,N}$  approaches zero as  $r$  approaches zero. Furthermore,  $P_{i,N}$  decreases monotonically when  $r < (i/N)/(4 - 6i/N)$ , since  $dP_{i,N}/dr$  is positive for this range of values of  $r$ . Therefore, for any small value  $\alpha/2$ , there is a unique value of  $r$ , designated  $\underline{r}$ , such that  $P_{i,N}$  is less than  $\alpha/2$  when  $r < \underline{r}$ . Setting  $P_{i,N} = \alpha/2$  for  $r = \underline{r}$  gives an equation for  $\underline{r}$  in terms of  $\alpha$ ,  $i$  and  $N$ :

$$\sum_{j=i}^N \binom{N}{j} [4\underline{r}/(1+6\underline{r})]^j \{1-[4\underline{r}/(1+6\underline{r})]\}^{N-j} = \alpha/2 \quad (7)$$

$\underline{r}$  is the lower confidence limit for  $\hat{r}$ . In words,  $\underline{r}$  is the closest two loci can be to one another before the chance of observing  $i$  or more recombinants between them falls to less than the small value  $\alpha/2$ . The lower confidence limit,  $\underline{r}$ , is not strictly defined when  $i = 0$ . However, since one thinks of confidence limits as defining an interval of reasonably likely values of  $r$ , it is not unreasonable to let  $\underline{r} = 0$  when  $i = 0$ .

The upper confidence limit can be defined analogously. The probability of observing  $i$  or fewer recombinants, designated  $P_{0,i}$ , is given by the sum of  $P_0 + P_1 + \dots + P_i$ . Expressing  $P_{0,i}$  in terms of  $r$ , it is easy to show that  $P_{0,i}$  decreases monotonically as  $r$  increases, provided that  $r > (i/N)/(4 - 6i/N)$ . The upper confidence limit, designated  $\bar{r}$ , is defined by the requirement that  $P_{0,i}$  be less than some

small value  $\alpha/2$  when  $r > \bar{r}$ . It is important to realize that there may be no value  $\bar{r}$  in the biologically interpretable range,  $0 \leq r \leq 1/2$ , for which  $P_{0,i}$  is less than  $\alpha/2$  when  $r > \bar{r}$ . In this case a biologically interpretable upper confidence limit does not exist. The existence of an upper confidence limit less than 1/2 requires that  $i/N$  be sufficiently small. In words, the upper confidence limit,  $\bar{r}$ , is the farthest apart two loci can be before the chance of observing  $i$  or fewer recombinants becomes smaller than the small value  $\alpha/2$ . If there is no upper confidence limit less than 1/2, then the observed results would be reasonably likely ( $p$  greater than  $\alpha/2$ ) if the loci were not linked. Setting  $P_{0,i}$  equal to  $\alpha/2$  for  $r = \bar{r}$  gives an equation for  $\bar{r}$  in terms of  $\alpha$ ,  $i$  and  $N$ :

$$\sum_{j=0}^i \binom{N}{j} [4\bar{r}/(1+6\bar{r})]^j \{1-[4\bar{r}/(1+6\bar{r})]\}^{N-j} = \alpha/2. \quad (8)$$

The confidence limits  $\underline{r}$  and  $\bar{r}$  could be determined by solving (7) and (8) numerically, but it is convenient to note that substituting:

$$\underline{r} = \underline{x}/(4 - 6\underline{x}) \quad (9)$$

$$\text{and } \bar{r} = \bar{x}/(4 - 6\bar{x}) \quad (10)$$

in (7) and (8), respectively, leads to equations for  $\underline{x}$  and  $\bar{x}$  which define the "exact" confidence limits for a binomial distribution. Tables of values of  $\underline{x}$  and  $\bar{x}$  for various values of  $\alpha/2$  can be found in books of statistical tables (e.g., see Diem and Lentner<sup>2</sup>). The values of  $\underline{r}$  and  $\bar{r}$  can be derived from such tables by using formulas (9) and (10).

For convenience, an abridged table of 95 percent and 99 percent confidence limits for  $\hat{r}$  is given in Table I. This table was constructed by converting values of  $\underline{x}$  and  $\bar{x}$  from existing tables<sup>2</sup> to  $\underline{r}$  and  $\bar{r}$  using formulas (9) and (10). Table I gives values of  $\hat{r}$ ,  $\underline{r}$  and  $\bar{r}$ , for up to 45 RI strains, for all possible numbers of recombinants that might reasonably be interpreted as indicating linkage. The values of  $i$  excluded from Table I are those for which the upper 95 percent confidence limit is greater than 50. An upper 95 percent confidence limit larger than 50 means that results as or more extreme than those observed ( $i$  or fewer recombinants) would not be very unlikely ( $P > 0.025$ ) if the loci were unlinked. In other words, for a given  $N$ , values of  $i$  larger than those in Table I would not provide strong evidence for linkage. The more stringent 99 percent confidence limits for  $r$  were omitted from Table I when the upper 99 percent confidence limit was greater than 50.

To illustrate the use of Table I, suppose one observes five recombinants in a set of 23 RI strains. For  $N = 23$  and  $i = 5$ , Table I states

that the estimate for the percent recombination between the loci in a single meiosis is  $\hat{r} = 8.1$ , and the 95 percent confidence limits are 2.1 to 31.7. (The last decimal place has been

dropped for lack of biological significance). This means that the observed results would be unlikely ( $P < 0.025$ ) if the true recombination percentage were greater than 31.7, and

similarly unlikely ( $P < 0.025$ ) if the recombination percentage were less than 2.1. The upper and lower confidence intervals are asymmetric about the estimate  $\hat{r} = 8.1$  because they

**Table 1. Estimates of the percent recombination in a single meiosis,  $\hat{r}$ , between two loci and the upper,  $\bar{r}$ , and lower,  $r$ , 95 percent and 99 percent confidence limits for  $\hat{r}$ , given  $i$  recombinants in  $N$  RI strains**

$N$	$i$	$\hat{r}$	95% limits		99% limits		$N$	$i$	$\hat{r}$	95% limits		99% limits		$N$	$i$	$\hat{r}$	95% limits		99% limits	
			$r$	$\bar{r}$	$r$	$\bar{r}$				$r$	$\bar{r}$	$r$	$\bar{r}$				$r$	$\bar{r}$		
6	0	0.00	0.00	36.92			21	3	4.55	0.80	19.97	0.43	30.72	28	6	7.89	2.37	26.54	1.61	39.46
7	0	0.00	0.00	26.56			21	4	6.67	1.48	28.21	0.89	45.38	28	7	10.00	3.18	34.31		
8	0	0.00	0.00	20.71	0.00	44.26	21	5	9.26	2.34	40.32			28	8	12.50	4.12	45.07		
9	0	0.00	0.00	16.97	0.00	33.46	22	0	0.00	0.00	5.02	0.00	7.88	29	0	0.00	0.00	3.64	0.00	5.57
9	1	3.33	0.07	43.67			22	1	1.22	0.03	8.69	0.01	13.02	29	1	0.91	0.02	6.05	0.01	8.76
10	0	0.00	0.00	14.36	0.00	26.84	22	2	2.63	0.28	12.96	0.12	19.30	29	2	1.92	0.22	8.65	0.09	12.24
10	1	2.94	0.06	33.46			22	3	4.29	0.76	18.32	0.41	27.68	29	3	3.06	0.57	11.59	0.31	16.32
11	0	0.00	0.00	12.44	0.00	22.39	22	4	6.25	1.41	25.44	0.85	39.80	29	4	4.35	1.03	15.07	0.63	21.30
11	1	2.63	0.06	27.10			22	5	8.62	2.21	35.51			29	5	5.81	1.60	19.30	1.04	27.61
12	0	0.00	0.00	10.97	0.00	19.20	23	0	0.00	0.00	4.76	0.00	7.44	29	6	7.50	2.27	24.57	1.54	35.93
12	1	2.38	0.05	22.75	0.01	41.92	23	1	1.16	0.03	8.18	0.01	12.17	29	7	9.46	3.05	31.38	2.13	47.42
12	2	5.56	0.54	44.19			23	2	2.50	0.27	12.10	0.12	17.83	29	8	11.76	3.93	40.53		
13	0	0.00	0.00	9.82	0.00	16.80	23	3	4.05	0.73	16.93	0.39	25.19	30	0	0.00	0.00	3.50	0.00	5.35
13	1	2.17	0.05	19.60	0.01	34.38	23	4	5.88	1.34	23.18	0.81	35.43	30	1	0.88	0.02	5.80	0.01	8.36
13	2	5.00	0.49	35.70			23	5	8.06	2.10	31.71			30	2	1.85	0.21	8.25	0.09	11.63
14	0	0.00	0.00	8.87	0.00	14.94	23	6	10.71	3.02	44.19			30	3	2.94	0.54	11.02	0.30	15.41
14	1	2.00	0.05	17.21	0.01	29.12	24	0	0.00	0.00	4.53	0.00	7.05	30	4	4.17	1.00	14.24	0.60	19.97
14	2	4.55	0.46	29.91			24	1	1.11	0.03	7.73	0.01	11.44	30	5	5.56	1.54	18.11	1.00	25.63
15	0	0.00	0.00	8.10	0.00	13.44	24	2	2.38	0.26	11.34	0.11	16.57	30	6	7.14	2.18	22.88	1.48	32.97
15	1	1.85	0.04	15.34	0.01	25.25	24	3	3.85	0.69	15.72	0.37	23.11	30	7	8.97	2.92	28.90	2.05	42.83
15	2	4.17	0.43	25.73	0.18	44.94	24	4	5.56	1.28	21.27	0.77	31.90	30	8	11.11	3.76	36.81		
15	3	7.14	1.16	43.15			24	5	7.58	2.00	28.65	1.29	44.66	30	9	13.64	4.73	47.68		
16	0	0.00	0.00	7.45	0.00	12.21	24	6	10.00	2.86	39.01			31	0	0.00	0.00	3.37	0.00	5.14
16	1	1.72	0.04	13.83	0.01	22.28	25	0	0.00	0.00	4.32	0.00	6.69	31	1	0.85	0.02	5.57	0.01	8.00
16	2	3.85	0.40	22.57	0.17	37.84	25	1	1.06	0.03	7.32	0.01	10.78	31	2	1.79	0.20	7.89	0.09	11.08
16	3	6.52	1.08	36.20			25	2	2.27	0.25	10.68	0.11	15.48	31	3	2.83	0.53	10.49	0.28	14.60
17	0	0.00	0.00	6.90	0.00	11.19	25	3	3.66	0.66	14.68	0.36	21.34	31	4	4.00	0.96	13.50	0.58	18.79
17	1	1.61	0.37	12.59	0.01	19.92	25	4	5.26	1.22	19.66	0.74	29.03	31	5	5.32	1.48	17.07	0.97	23.93
17	2	3.57	0.37	20.09	0.16	32.64	25	5	7.14	1.90	26.12	1.23	39.77	31	6	6.82	2.10	21.39	1.43	30.44
17	3	6.00	1.01	31.15			25	6	9.38	2.72	34.92			31	7	8.54	2.80	26.79	1.97	39.01
17	4	9.09	1.90	49.60			25	7	12.07	3.68	47.65			31	8	10.53	3.61	33.71		
18	0	0.00	0.00	6.42	0.00	10.32	26	0	0.00	0.00	4.13	0.00	6.37	31	9	12.86	4.52	42.98		
18	1	1.52	0.04	11.55	0.01	18.02	26	1	1.02	0.03	6.96	0.01	10.19	32	0	0.00	0.00	3.25	0.00	4.95
18	2	3.33	0.35	18.10	0.15	28.69	26	2	2.17	0.24	10.08	0.10	14.52	32	1	0.82	0.02	5.36	0.01	7.67
18	3	5.56	0.95	27.34	0.51	45.66	26	3	3.49	0.64	13.76	0.34	19.82	32	2	1.72	0.19	7.56	0.08	10.58
18	4	8.33	1.77	41.73			26	4	5.00	1.17	18.28	0.71	26.62	32	3	2.73	0.51	10.01	0.28	13.87
19	0	0.00	0.00	6.00	0.00	9.58	26	5	6.76	1.82	24.01	1.18	35.83	32	4	3.85	0.93	12.82	0.56	17.75
19	1	1.43	0.03	10.68	0.01	16.44	26	6	8.82	2.59	31.61	1.75	49.09	32	5	5.10	1.43	16.13	0.93	22.43
19	2	3.13	0.33	16.47	0.14	25.59	26	7	11.29	3.50	42.19			32	6	6.52	2.02	20.09	1.38	28.29
19	3	5.17	0.89	24.35	0.48	39.32	27	0	0.00	0.00	3.95	0.00	6.08	32	7	8.14	2.70	24.95	1.89	35.83
19	4	7.69	1.66	36.00			27	1	0.98	0.02	6.63	0.01	9.66	32	8	10.00	3.46	31.09	2.49	45.94
20	0	0.00	0.00	5.63	0.00	8.94	27	2	2.08	0.23	9.55	0.10	13.67	32	9	12.17	4.33	39.12		
20	1	1.35	0.03	9.92	0.01	15.12	27	3	3.33	0.61	12.96	0.33	18.50	33	0	0.00	0.00	3.14	0.00	4.77
20	2	2.94	0.31	15.11	0.13	23.08	27	4	4.76	1.12	17.07	0.68	24.58	33	1	0.79	0.02	5.16	0.01	7.37
20	3	4.84	0.84	21.94	0.45	34.50	27	5	6.41	1.74	22.20	1.13	32.59	33	2	1.67	0.19	7.26	0.08	10.12
20	4	7.14	1.57	31.63			27	6	8.33	2.48	28.86	1.68	43.76	33	3	2.63	0.49	9.58	0.27	13.20
20	5	10.00	2.49	46.58			27	7	10.61	3.33	37.84			33	4	3.70	0.90	12.22	0.54	16.80
21	0	0.00	0.00	5.31	0.00	8.38	28	0	0.00	0.00	3.79	0.00	5.81	33	5	4.90	1.38	15.29	0.90	21.12
21	1	1.28	0.03	9.27	0.01	14.00	28	1	0.94	0.02	6.33	0.01	9.19	33	6	6.25	1.95	18.94	1.33	26.41
21	2	2.78	0.30	13.95	0.13	21.02	28	2	2.00	0.22	9.07	0.10	12.92	33	7	7.78	2.59	23.36	1.83	33.10
							28	3	3.19	0.59	12.24	0.32	17.34	33	8	9.52	3.33	28.86	2.40	41.88
							28	4	4.55	1.07	16.02	0.65	22.82	33	9	11.54	4.15	35.88		
							28	5	6.10	1.67	20.65	1.08	29.89	33	10	13.89	5.09	45.21		

are related to the confidence limits for a binomial distribution rather than a normal distribution. Since no 99 percent confidence limits are given for  $N = 23$ ,  $i = 5$ , the upper 99 percent

confidence limit is not less than 50. This means that the observed results are not inconsistent at the 0.005 level with the loci being unlinked. In contrast, if one tries to apply for-

mula (5) and the normal approximation to this example, one gets 95 percent confidence limits ( $\hat{p} \pm 1.96\text{STD}(\hat{p})$ ) of -1.2 to 17.4. The lower limit from this method is meaningless,

Table I Continued

N	i	$\hat{p}$	95% limits		99% limits		N	i	$\hat{p}$	95% limits		99% limits		N	i	$\hat{p}$	95% limits		99% limits	
			L	F	L	F				L	F	L	F				L	F		
34	0	0.00	0.00	3.04	0.00	4.60	38	6	5.17	1.65	14.71	1.13	19.82	42	7	5.56	1.95	14.80	0.00	19.65
34	1	0.77	0.02	4.98	0.00	7.09	38	7	6.36	2.19	17.69	1.55	23.99	42	8	6.67	2.47	17.47	1.79	23.31
34	2	1.61	0.18	6.98	0.08	9.70	38	8	7.69	2.79	21.19	2.02	29.05	42	9	7.90	3.05	20.55	2.25	27.64
34	3	2.54	0.48	9.18	0.26	12.61	38	9	9.18	3.45	25.38	2.54	35.34	42	10	9.26	3.68	24.16	2.76	32.90
34	4	3.57	0.87	11.67	0.53	15.97	38	10	10.87	4.19	30.48	3.13	43.34	42	11	10.78	4.37	28.45	3.32	39.37
34	5	4.72	1.34	14.54	0.87	19.94	38	11	12.79	5.01	36.84			42	12	12.50	5.14	33.64	3.94	47.61
34	6	6.00	1.88	17.91	1.28	24.77	38	12	15.00	5.93	45.00			42	13	14.44	5.99	40.09		
34	7	7.45	2.50	21.96	1.76	30.76	39	0	0.00	0.00	2.61	0.00	3.92	42	14	16.66	6.93	48.25		
34	8	9.09	3.20	26.91	2.31	37.76	39	1	0.67	0.02	4.22	0.00	5.96	43	0	0.00	2.34	0.00	0.00	3.51
34	9	10.98	3.99	33.14	2.93	48.82	39	2	1.39	0.16	5.85	0.07	8.03	43	1	0.60	0.02	3.77	0.00	5.28
34	10	13.16	4.88	41.24			39	3	2.17	0.42	7.60	0.23	10.27	43	2	1.25	0.14	5.18	0.06	7.05
35	0	0.00	0.00	2.94	0.00	4.45	39	4	3.03	0.75	9.51	0.46	12.75	43	3	1.95	0.37	6.67	0.00	8.94
35	1	0.75	0.02	4.81	0.00	6.83	39	5	3.97	1.15	11.65	0.75	15.59	43	4	2.70	0.67	8.29	0.41	10.99
35	2	1.56	0.18	6.72	0.08	9.31	39	6	5.00	1.61	14.08	1.10	18.87	43	5	3.52	1.03	10.05	0.68	13.27
35	3	2.46	0.46	8.81	0.25	12.05	39	7	6.14	2.13	16.87	1.50	22.74	43	6	4.41	1.44	12.02	0.99	15.84
35	4	3.45	0.84	11.16	0.51	15.20	39	8	7.41	2.70	20.12	1.96	27.36	43	7	5.39	1.90	14.23	1.34	18.79
35	5	4.55	1.30	13.85	0.85	18.89	39	9	8.82	3.34	23.98	2.47	33.03	43	8	6.45	2.73	16.73	1.75	22.20
35	6	5.77	1.82	16.99	1.24	23.31	39	10	10.42	4.05	28.62	3.03	40.18	43	9	7.63	2.96	19.61	2.19	26.22
35	7	7.14	2.42	20.71	1.70	28.75	39	11	12.22	4.84	34.31	3.66	49.41	43	10	8.93	3.57	22.96	2.68	31.03
35	8	8.70	3.09	25.22	2.23	35.60	39	12	14.29	5.71	41.52			43	11	10.38	4.24	26.91	3.22	36.89
35	9	10.47	3.84	30.80	2.82	44.56	40	0	0.00	0.00	2.54	0.00	3.81	43	12	12.00	4.98	31.65	3.82	44.19
35	10	12.50	4.69	37.89			40	1	0.65	0.02	4.10	0.00	5.77	43	13	13.83	5.79	37.44		
35	11	14.86	5.64	47.28			40	2	1.35	0.15	5.67	0.07	7.76	43	14	15.91	6.68	44.63		
36	0	0.00	0.00	2.85	0.00	4.31	40	3	2.11	0.40	7.34	0.22	9.90	44	0	0.00	0.00	2.29	0.00	3.42
36	1	0.72	0.02	4.64	0.00	6.59	40	4	2.94	0.73	9.17	0.44	12.26	44	1	0.59	0.02	3.67	0.00	5.13
36	2	1.52	0.17	6.48	0.07	8.95	40	5	3.85	1.12	11.20	0.73	14.94	44	2	1.22	0.14	5.04	0.06	6.85
36	3	2.38	0.45	8.47	0.24	11.55	40	6	4.84	1.56	13.50	1.07	18.02	44	3	1.90	0.37	6.48	0.20	8.65
36	4	3.33	0.82	10.70	0.50	14.50	40	7	5.93	2.06	16.12	1.46	21.60	44	4	2.63	0.66	8.03	0.40	10.62
36	5	4.39	1.26	13.23	0.82	17.94	40	8	7.14	2.62	19.16	1.90	25.86	44	5	3.42	1.00	9.72	0.66	12.79
36	6	5.56	1.76	16.15	1.20	22.01	40	9	8.49	3.24	22.71	2.39	31.03	44	6	4.29	1.40	11.59	0.96	15.24
36	7	6.86	2.33	19.59	1.65	26.96	40	10	10.00	3.92	26.96	2.93	37.41	44	7	5.22	1.84	13.69	1.31	18.00
36	8	8.33	2.98	23.71	2.15	33.12	40	11	11.70	4.67	32.12	3.54	45.56	44	8	6.25	2.33	16.05	1.70	21.21
36	9	10.00	3.70	28.75	2.72	41.00	40	12	13.64	5.51	38.51			44	9	7.37	2.87	18.76	2.13	24.94
36	10	11.90	4.51	35.07			40	13	15.85	6.43	46.69			44	10	8.62	3.46	21.88	2.60	29.35
36	11	14.10	5.42	43.21			41	0	0.00	0.00	2.47	0.00	3.70	44	11	10.00	4.11	25.54	3.13	34.69
37	0	0.00	0.00	2.77	0.00	4.17	41	1	0.63	0.02	3.98	0.00	5.60	44	12	11.54	4.82	29.87	3.70	41.24
37	1	0.70	0.02	4.49	0.00	6.36	41	2	1.32	0.15	5.49	0.07	7.51	44	13	13.27	5.60	35.09	4.33	49.52
37	2	1.47	0.17	6.25	0.07	8.62	41	3	2.06	0.39	7.10	0.21	9.55	44	14	15.22	6.45	41.55		
37	3	2.31	0.44	8.16	0.24	11.09	41	4	2.86	0.71	8.85	0.43	11.80	44	15	17.44	7.40	49.68		
37	4	3.23	0.79	10.27	0.48	13.87	41	5	3.73	1.09	10.79	0.71	14.34	45	0	0.00	0.00	2.23	0.00	3.33
37	5	4.24	1.22	12.65	0.80	17.09	41	6	4.69	1.52	12.97	1.04	17.23	45	1	0.57	0.02	3.57	0.00	5.00
37	6	5.36	1.71	15.39	1.17	20.86	41	7	5.74	2.00	15.44	1.42	20.57	45	2	1.19	0.14	4.90	0.06	6.65
37	7	6.60	2.26	18.60	1.60	25.39	41	8	6.90	2.54	18.28	1.84	24.52	45	3	1.85	0.36	6.29	0.19	8.39
37	8	8.00	2.88	22.38	2.08	30.95	41	9	8.18	3.14	21.57	2.32	29.23	45	4	2.56	0.64	7.78	0.39	10.28
37	9	9.57	3.57	26.96	2.63	37.94	41	10	9.62	3.79	25.47	2.84	35.02	45	5	3.33	0.98	9.41	0.64	12.35
37	10	11.36	4.35	32.61	3.24	47.09	41	11	11.23	4.52	30.16	3.43	42.26	45	6	4.17	1.37	11.20	0.94	14.67
37	11	13.41	5.21	39.77			41	12	13.04	5.32	35.93			45	7	5.07	1.80	13.20	1.28	17.29
37	12	15.79	6.17	49.17			41	13	15.12	6.20	43.15			45	8	6.06	2.27	15.43	1.65	20.29
38	0	0.00	0.00	2.69	0.00	4.04	42	0	0.00	0.00	2.41	0.00	3.60	45	9	7.14	2.80	17.98	2.07	23.76
38	1	0.68	0.02	4.35	0.00	6.16	42	1	0.62	0.02	3.87	0.00	5.44	45	10	8.33	3.37	20.90	2.53	27.85
38	2	1.43	0.16	6.05	0.07	8.32	42	2	1.28	0.15	5.33	0.06	9.24	45	11	9.65	3.99	24.29	3.04	32.72
38	3	2.24	0.43	7.87	0.23	10.66	42	3	2.00	0.38	6.88	0.21	9.24	45	12	11.11	4.67	28.27	3.59	38.65
38	4	3.13	0.77	9.87	0.47	13.29	42	4	2.78	0.69	8.56	0.42	11.38	45	13	12.75	5.42	33.03	4.20	46.05
38	5	4.10	1.18	12.14	0.77	16.31	42	5	3.62	1.06	10.41	0.69	13.79	45	14	14.58	6.24	38.84		
							42	6	4.55	1.48	12.48	1.01	16.51	45	15	16.66	7.14	46.05		

being less than zero, and the upper limit seriously underestimates the "exact" upper 95 percent confidence limit (17.4 compared to 31.7).

Inspection of Table I leads to two important conclusions. First, the 95 percent confidence limits are quite large except for the smallest values of  $i$  and the largest values of  $N$ . The large confidence interval for most of the entries in Table I indicates that analysis of RI strains will very rarely localize a gene with confidence to within a small region. When linkage is detected with RI strains, a backcross or other cross is usually indicated to confirm the result and to provide an independent estimate of map position. Second, all of the values of  $\hat{r}$  in Table I are small ( $< 17.5$ ). This is a consequence of the large upper confidence intervals and the fact that  $\bar{r}$  must be less than 50 for the data to be reasonably interpreted as indicating linkage. The fact that only

small values of  $\hat{r}$  are consistent with a high probability of linkage may lead to a bias toward reporting underestimates of map distances using RI strains, since overestimates will frequently go unreported for want of compelling evidence of linkage. Therefore, it is not surprising that analysis of backcross mice has in the past led to upward revision of the estimate of  $r$  derived from analysis of RI strains<sup>7</sup>. While RI strains are extremely useful for screening for linked loci, a clear appreciation of the quantitative limitations of the method and the probabilistic meaning of the results is important for proper interpretation of linkage data.

### References

1. BAILEY D. W. Recombinant inbred strains and bilineal congenic strains. *In* The Mouse in Biomed-

cal Research: vol. I. H. L. Foster, J. D. Small, and J. G. Fox, Eds., p. 223-239. Academic Press, NY, 1981.

2. DIEM, K. and C. LENTNER, Eds. Scientific Tables. Geigy Pharmaceuticals, Ardsley, NY. p. 85-103. 1975.
3. GREEN, E. L. Genetics and Probability in Animal Breeding Experiments. Oxford University Press, NY. p. 137. 1981.
4. HALDANE, J. B. S. and C. H. WADDINGTON. Inbreeding and Linkage. *Genetics* 16:357-374. 1931.
5. SNEDECOR, G. W. and W. G. COCHRAN. Statistical Methods. Iowa State University Press, Ames. p. 207. 1972.
6. TAYLOR, B. A. Recombinant inbred strains: use in gene mapping. *In* Origins of Inbred Mice, H. C. Morse, III, Ed. The Jackson Laboratory, Bar Harbor, Maine. p. 423-438. 1978.
7. WOMACK, J. E., M. A. LYNES, and B. A. TAYLOR. Genetic variation of an intestinal leucine arylamidopeptidase (*Lap-1*) in the mouse and its location on chromosome 9. *Biochem. Genet.* 13:511-513. 1975.