Behavior Genetics, Vol. 22, No. 6, 1992

Empirical Estimates of Bonferroni Corrections for Use in Chromosome Mapping Studies with the BXD Recombinant Inbred Strains

J. K. Belknap[1]

Most chromosome mapping efforts with the BXD recombinant inbred (RI) strains involve comparisons between a trait of interest and each of a large number of marker loci for evidence of linkage. Such multiple tests or comparisons greatly increase the Type I error rate compared to the single-test situation. Perhaps the most direct way to obtain multiple-test error rates is to employ a Bonferroni correction, where the single-test alpha is multiplied by the number of independent (nonredundant) comparisons (k) *to yield a multiple-test alpha that protects against even one fortuitous association with any of the markers. Several empirical estimates of* k *are discussed for the published BXD marker set of 142 mapped loci and the newer unpublished marker set (October 1991) comprised of 352 marker loci. Reasonable estimates of* k *appear to be roughly 40 and 65 for these two marker sets, respectively.*

KEY WORDS: BXD recombinant inbred strains; chromosome mapping.

The BXD RI strains were originally developed as a tool for linkage detection in gene mapping (Bailey, 1981; Taylor, 1978). When RI strain means on a given trait are found to fall in a bimodal distribution (i.e., some RI strains resemble the B6 progenitor and some resemble the D2 progenitor and none are intermediate), this is presumptive evidence for control of that trait by a single major gene locus. Comparison of the strain distribution pattern (SDP) for that trait (i.e., a listing of which

This work was supported by PHS Grants AA08621, AA06243, and DA05228, NIDA Contract 271-90-7405, and a Merit Review Grant from the Department of Veterans Affairs.

[1] Research Service (151W), VA Medical Center, and Department of Medical Psychology, Oregon Health Sciences University, Portland, Oregon 97201.

0001-8244/92/1100-0677$06.50/0

strains are "B6-like" and which are "D2-like") can be made with SDPs for known marker loci previously mapped to a particular chromosome region. A close match in SDPs between the unknown (test) locus and a marker locus allows provisional mapping to the chromosome region of the marker (Bailey, 1981; Taylor, 1978). The usefulness of RI strains for gene mapping is largely due to the substantial number of marker loci previously mapped through many years of cumulative effort, which allows a large proportion of the mouse genome to be readily scanned for possible linkage with a new test SDP (Bailey, 1981; Taylor, 1978; 1989).

In the analysis that follow, we used two marker sets in the BXD series. The first was that published by Taylor (1989), comprised of 142 mapped marker loci. The second (unpublished) was kindly provided to us by Dr. Benjamin Taylor of The Jackson Laboratory in October 1991. This newer marker set was comprised of 352 mapped loci, or 2.5-fold the number contained in the published set. This proliferation of markers will undoubtedly continue as a result of new advances in molecular genetics, particularly microsatellite sequences (simple sequence repeats), which have proven to be highly polymorphic. The proportion of the genome within 10 cM of any one marker was 808 cM for the published marker set and 1192 cM for the unpublished one, for a gain of 48%. Taylor (1978) has estimated that a $\pm$10-cM span from each marker yields a power (1 $-$ Type II error rate) for linkage detection of about 60% at $p < .01$ for $n = 25$ strains; larger spans would yield lower power estimates. In these calculations, the map locations of the markers was assumed to be that given in the consensus linkage map published by Hillyard *et al.* (1992).

When comparing a test SDP with a marker SDP for n RI strains, linkage is suggested whenever the number of recombinant strains (SDP mismatches), or i, is significantly fewer than the $0.5n$ expected with nonlinkage. Thus, with $n = 22$ RI strains, the expectation for the null hypothesis (H_o; nonlinkage) is for 11 strains to be recombinant (discordant in their SDPs). With a preselected alpha level of $p < .05$, and a simple application of the binomial distribution ($p = q = 0.5$, $n = 22$, mean $= 0.5n$, variance $= 0.25n$), the null hypothesis will be rejected (linkage is indicated) whenever $i \leq 6$ strains (one-tailed) between two SDPs (Green, 1981). This reasoning is warranted whenever a test (i.e., unmapped) SDP is matched against a single marker SDP, i.e., a single significance test is made. However, a serious problem arises whenever the test SDP is matched against a large number of marker loci, as is the case with most gene mapping efforts. Such multiple tests of significance greatly increase the probability of fortuitous linkage indications, i.e., rejections of H_o when linkage does not exist, or Type I errors. For

example, if a test SDP is compared to 100 marker SDPs (a total of 100 individual tests of significance), 5 of the comparisons will be expected to be significant at the $p < .05$ level due to chance alone. Therefore, the preset alpha levels for the single or individual test will not provide adequate protection against multiple-test Type I errors (Miller, 1981; Rice, 1989). The latter has also been referred to as the experimentwise, groupwise, simultaneous comparison, or family error rate (Miller, 1981).

There are two primary avenues for dealing with this problem. First, confirmation of the BXD data can (and should) be sought by using other genetic models to determine which candidate chromosome map sites from the BXD data can be independently supported. The BXD data would serve as a screen to identify a handful of candidate map locations, which then would be specifically tested in other RI series, standard inbred (non-RI) strains (Goldman *et al.*, 1987), or congenic lines (Bailey, 1981) or by linkage analysis in F_2 or backcross populations (Lander and Botstein, 1989; Rise *et al.*, 1991; Neumann and Collins, 1991). If such confirmation testing is to be done, then $p < .05$ (single test) seems appropriate for the BXD results, since it yields lower Type II error rates (failing to detect important map sites) compared to more stringent alpha values. In this case, the primary protection against Type I errors would reside in the confirmation test.

Second, if the BXD data are to stand alone, without confirmation from other genetic models, then a correction is needed for the multiple comparisons calculated for the marker loci. Perhaps the simplest way is to apply a Bonferroni correction to the single-test alpha level ($alpha_s$) to obtain a multiple-test estimate of Type I error rates ($alpha_m$) (Miller, 1981; Rice, 1989). The correction involves multiplying the single-test alpha by the number of nonredundant (independent) tests of significance made, or k. This derives from the equation $\alpha_m = k\,\alpha_s$ (Rice, 1989). For example, with a single-test alpha of $p < .001$, and $k = 50$ nonredundant marker loci, then the estimated multiple-test alpha is $.001 \times 50 = .05$. The latter protects against *even one* fortuitous linkage indication (rejection of H_o) arising anywhere among the 50 marker loci (Lander and Botstein, 1989). It does not matter what statistic (e.g., r, χ^2, t, or F) is appropriately used for testing significance—only the number of independent (orthogonal) tests. This correction is considered to be conservative (Miller, 1981). Since most statistical software packages routinely provide only single-test p values, a Bonferroni correction can be readily applied to such output using a calculator.

While nonredundancy is not a requirement for the Bonferroni correction, an increasingly conservative test will result if some of the comparisons between test and marker loci are largely redundant with regard

to chromosome location because of close linkage among the markers, resulting in many similar or identical SDPs per chromosome region. It is almost certain that any large collection of marker loci will have a considerable degree of such redundancy. What is desirable is an estimate of the number of nonredundant (orthogonal) linkage groups among the marker loci for use in Bonferroni corrections, which is approximated when marker loci are located just sufficiently far apart that they are not significantly associated with each other beyond that due to chance, i.e., their mean correlation with each other is essentially zero. The purpose of this paper is to discuss several options in deriving such estimates.

There are several approaches to estimating an appropriate value of k. Taylor (1978) has estimated that an average of 4 crossovers per 100-cM chromosome length has occurred in the derivation of each of the BXD RI strains, or about 64 crossovers for the entire genome of about 1600 cM. Since the number of chromosome segments remaining intact (not recombined) on each chromosome will be one more than the number of crossovers, a total of 84 linked regions will result (i.e., 64 + 20) throughout the genome, or $k = 84$. Neumann (1992) recently calculated a value of $k = 84$ assuming 21 RI strains, a swept radius per marker locus of 12.5 cM, and a complete marker set encompassing the entire genome. However, the marker sets available for any existing RI series are not complete in that sizable segments of the mouse genome are not represented by any known marker. Thus, these k estimates for the entire genome are probably too large for the marker sets presently available. For example, the published BXD markers (Taylor, 1989), assuming a linkage detection span of 10 cM on either side of each marker, encompassed only 808 cM in the aggregate, or half the estimated genome of 1600 cM. The newer set of 352 mapped loci covered 1192 cM, or about three-fourths of the genome. In the newer marker set, the most obvious portions of the genome where polymorphic markers are absent were the middle portions of chromosomes 3 and 13, the distal portions of chromosomes 5 and 14, and most of the X chromosome.

There are several possible explanations for the relatively low proportion of the genome covered by the existing BXD markers. First, the present markers represent an incomplete sample of the genome (sampling error). However, if this were the only explanation, 142 markers would be expected to cover 83% of the genome based on a random sampling from the uniform probability distribution, and 99% for 352 markers, assuming ± 10 cM from each locus. The fact that the two marker sets cover much less than expected with random sampling (51 and 74%, respectively) indicates that the sampling of markers has been biased, i.e., it has been confined largely to a subset of the genome, especially for the

published markers. While the reasons for this are not fully known, one possibility is that linkage disequilibrium fixed by inbreeding may exist from the common ancestor to the two progenitors (B6 and D2), resulting in several chromosome segments of 20 cM or more in the BXD series that show no allelic variation. Of course, many of these regions may ultimately prove to be polymorphic in the BXDs with the discovery of new markers. However, it is a distinct possibility that the polymorphic genome in the BXDs may ultimately fall short of the entire genome.

We took an empirical approach to estimating appropriate values of k by determining the number of linked regions inherited intact (not recombined) from one or the other progenitor strain in the BXD set. A linked region was defined as a series of consecutive marker loci along the length of each chromosome originating from a single progenitor strain. The mean value ($\pm$SD) across all 24 existing BXD strains was 34 ± 4 for the published marker set and 56 ± 4 for the newer marker set. This provides one estimate of the number of independent linkage groups involved in these marker sets. We had previously reported a value of 52 ± 4 for an earlier, slightly smaller marker set (Belknap *et al.*, 1992). When more complete marker sets become available in the BXDs, this number will undoubtedly increase somewhat.

Another empirical approach to estimating k was to determine the number of nonredundant linkage groups involved in the two marker sets. This was done by determining the maximum number of marker loci for each chromosome that were not significantly correlated with each other at $p < .05$ (single test), i.e., that were roughly statistically independent (uncorrelated). Starting with the marker closest to the proximal end of each chromosome, this SDP was matched in linear order with more distal markers until a marker was found that was no longer significantly associated at $p < .05$. This second marker was then recorded, and its SDP, in turn, was matched in linear order with more distal markers until statistical significance was again no longer observed. This process was repeated until the most distal marker was reached. The linear order used was that given by Taylor (1989) or Taylor (unpublished) for the two marker sets, respectively. When summed over all chromosomes, this number of roughly independent markers (k) was 42 for the published marker set and 73 for the newer marker set, of 142 and 352 loci, respectively. These numbers are likely overestimates, since many of the "independent" markers on each chromosome remained positively associated (correlated) with each other (although below the $p < .05$ critical value) rather than showing a mean correlation of zero expected with full independence.

A third empirical estimate was based on a simulation utilizing the

published (Taylor, 1989) marker set. Of 142 marker loci of known chromosome location, 25 were chosen at random to serve as test loci, i.e., to simulate mapping of new loci. For each of these 25 "test" loci, the concordances (proportion of matches between strain distribution patterns) were calculated with all other marker loci, and tests of significance carried out for each pair of loci for evidence of linkage. This resulted in 3525 individual tests of significance. Those meeting the $p < .05$, .01, and .001 alpha levels (single-test Type I error rates) were examined for the number of correct and incorrect linkage indications (rejections of the null hypothesis) to determine the multiple-test error rate. Dividing the observed multiple-test Type I error rate by the single-test Type I error rate yields an empirical estimate of k, which was 33. The full results of this study are to be reported separately.

Based on the various estimates of k described above, values of 40

Table I. Critical Values of i (Number of Recombinant Strains or SDP Mismatches Between Two SDPs) Out of n RI Strains Required to Obtain a Multiple-Test Alpha of $p < .05$ for Linkage Detection After a Bonferroni Correction (k) Has Been Applied (See Text)[a]

n RI strains	$k = 40$	$k = 65$	$k = 84$
10	0	NA	NA
11	0	0	0
12	0	0	0
13	0	0	0
14	1	0	0
15	1	1	1
16	1	1	1
17	2	1	1
18	2	2	1
19	2	2	2
20	3	2	2
21	3	3	2
22	3	3	3
23	3	3	3
24	4	4	3
25	4	4	4
26	5	4	4

[a] In each case, the single-test p value calculated from the binomial distribution (one-tailed) was multiplied by one of three different values of k (40, 65, or 84) as tabulated. The values of i presented assume that an SDP for a trait of interest is matched against each of k statistically independent (nonredundant) marker loci and that the BXD data are to stand alone, without confirmation testing with other genetic models. Values of $k = 40$ are suggested for the Taylor (1989) published markers, $k = 65$ for the newer unpublished marker set described in the text, and $k = 84$ for a complete marker set. For example, with $n = 20$ BXD RI strains, and using the published markers ($k = 40$), 3 or fewer SDP mismatches are required to obtain a multiple-test alpha of $p < .05$.

and 65 seem appropriate for the 142 and 352 loci marker sets, respectively. When $k = 65$, a $p < .0008$ (single-test) alpha becomes $p < .05$ (multiple test) after correction. Higher values will be needed for more complete marker sets and lower values for less complete ones. Thus, the most appropriate values of k should be determined for each marker set employed. In the absence of such determinations, Neumann's (1992) estimate of $k = 84$ is a reasonable choice for all RI series with relatively complete marker sets. In this case, a single-test $p < .0006$ is required to obtain a multiple-test $p < .05$. Table I presents the critical values of i (SDP mismatches) out of n RI strains to obtain a multiple-test alpha of $p < .05$ when $k = 40$, 65, or 84.

Recent assessments of the error rates associated with linkage detection in existing mouse RI sets have shown them to be greater than previously believed (Neumann, 1990, 1991, 1992; Silver and Buckler, 1986). The present study provides further support for this trend. Neumann (1990, 1991) provides some useful suggestions for reducing error rates. However, these recent developments do not diminish the usefulness of RI strains for gene mapping efforts; they merely signal the need for more caution when the only data available are from one RI set. Over the past two decades, the tremendous value of RI strains has been demonstrated frequently for quickly and efficiently screening large portions of the mouse genome for possible linkage with a test locus. This often results in a small number of candidate chromosome sites suitable for further testing in other genetic models.

NOTE ADDED IN PROOF

Assuming a random distribution of markers throughout the genome, the expected proportion of the genome (P) within d cM of each of N marker loci is given by the expression: $P = 1 - e^{-2Nd/D}$, where D is the map length of the entire genome, about 1600 cM in the mouse, d is the span in cM on either side of each marker locus, here assumed to be ± 10 cM, and e is 2.71. For example, with $N = 142$, $P = 0.83$, and for $N = 352$, $P = 0.99$. No correction is employed for ends of chromosomes. This expression is from Jacob, H.J., Lindpaintner, K., Lincoln, S.E., Kusumi, K., Bunker, R.K., Mao, Y-P., Ganten, D., Dzau, V.J. and Lander, E.S. (1991). Genetic mapping of a gene causing hypertension in the stroke-prone spontaneously hypertensive rat. Cell 67:213–224.

REFERENCES

Bailey, D. W. (1981). Recombinant inbred strains and bilineal congenic strains. In Foster, H. L., Small, J. D., and Fox, J. G. (eds.), *The Mouse in Biomedical Research, Vol. I*, Academic Press, New York, pp. 223–239.

Belknap, J. K., Crabbe, J. C., Plomin, R., McClearn, G. E., Sampson, K. E., O'Toole, L. A., and Gora-Maslak, G. (1992). Single locus control of saccharin intake in BXD/Ty recombinant inbred mice: Some methodological implications for RI strain analysis. *Behav. Genet.* **22:**81–100.

Goldman, D., Lister, R. G., and Crabbe, J. C. (1987). Mapping of a putative genetic locus determining ethanol intake in the mouse. *Brain Res.* **420:**220–226.

Green, E. L. (1981). *Genetics and Probability in Animal Breeding Experiments*, Oxford, New York.

Hillyard, A. L., Doolittle, D. P., Davisson, M. T., and Roderick, T. H. (1992). Locus map of mouse. *Mouse Genome* **90:**8–21.

Lander, E. S., and Botstein, D. (1989). Mapping mendelian factors underlying quantitative traits using RFLP linkage maps. *Genetics* **121:**185–199.

Miller, R. G., Jr. (1981). *Simultaneous Statistical Inference*, McGraw–Hill, New York.

Neumann, P. E. (1990). Two-locus linkage analysis using recombinant inbred strains and Bayes' theorem. *Genetics* **126:**277–284.

Neumann, P. E. (1991). Three-locus linkage analysis using recombinant inbred strains and Bayes' theorem. *Genetics* **128:**631–638.

Neumann, P. E. (1992). Inference in linkage analysis of multifactorial traits using recombinant inbred strains of mice. *Behav. Genet.* **22:**665–676.

Neumann, P. E., and Collins, R. L. (1991). Genetic dissection of susceptibility to audiogenic seizures in inbred mice. *Proc. Natl. Acad. Sci. USA* **88:**5408–5412.

Rice, W. R. (1989). Analyzing tables of statistical tests. *Evolution* **43:**223–225.

Rise, M. T., Frankel, W. N., Coffin, J. M., and Seyfried, T. N. (1991). Genes for epilepsy mapped in the mouse. *Science* **253:**669–673.

Roderick, T. H., and Guidi, J. (1989). Strain distribution of polymorphic variants. In Lyon, M. F., and Searle, A. G. (eds.), *Genetic Variants and Strains of the Laboratory Mouse*, 2nd ed., Oxford University Press, Oxford, pp. 663–772.

Silver, J., and Buckler, C. E. (1986). Statistical considerations for linkage analysis using recombinant inbred strains and backcrosses. *Proc. Natl. Acad. Sci. USA* **83:**1423–1427.

Sokal, R., and Rohlf, J. (1981). *Biometry*, Freeman, San Francisco.

Taylor, B. A. (1978). Recombinant inbred strains: Use in gene mapping. In Morse, H. C. (ed.), *Origins of Inbred Mice*, Academic Press, New York, pp. 423–438.

Taylor, B. A. (1989). Recombinant inbred strains. In Lyon, M. F., and Searle, A. G. (eds.), *Genetics Variants and Strains of the Laboratory Mouse*, 2nd ed., Oxford University Press, Oxford, pp. 773–789.