Amati P, Chomel JC, Nivelon-Chevalier A, Gigelkrantz S, Kitzis A, Kaplan J, Bonneau D (1995) A gene for blepharophimosis, ptosis, epicanthus inversus maps to chromosome 3q23. Hum Genet 96:213–215

Jones C, Collins JRO (1984) Blepharophimosis and its association with female infertility. Br J Ophthalmol 68:533–534

Lawson CT, Toomes C, Fryer A, Carette MJM, Taylor GM, Fukushima Y, Dixon M (1995) Definition of the blepharophimosis, ptosis, based on the analysis of chromosomal anomalies. Hum Mol Genet 4:963–967

McKusick VA (1992) Mendelian inheritance in man, 10th ed. Johns Hopkins University Press, Baltimore

Nicolino M, Bost M, David M, Chaussain JL (1995) Familial blepharophimosis: an uncommon marker of ovarian dysgenesis. J Pediatr Endocrinol Metabol 8:127–133

Oley C, Baraitser M (1988) Blepharophimosis, ptosis, epicanthus inversus syndrome (BPES syndrome). J Med Genet 25: 47–51

Small KW, Stalvey M, Fisher L, Mullen L, Dickel C, Beadles K, Reimer R, et al (1995) Blepharophimosis syndrome is linked to chromosome 3q. Hum Mol Genet 4:443–448

Smith A, Fraser IS, Shearman RP, Russell P (1989) Blepharophimosis plus ovarian failure: a likely candidate for a contiguous gene syndrome. J Med Genet 26:434–438

Townes PL, Muechler EK (1978) Blepharophimosis, ptosis, epicanthus inversus and primary amenorrhea. Arch Ophthalmol 97:1664–1665

Vassilli A, Matzuk MM, Gardner HAR, Lee KF, Jaenisch R (1994) Activin/inhibin βB subunit gene disruption lead to defects in eyelid development and female reproduction. Genes Dev 8:414–427

Zlotogora J, Sagi M, Cohen T (1983) The blepharophimosis, ptosis, and epicanthus inversus syndrome: delineation of two types. Am J Hum Genet 35:1020–1027

Address for correspondence and reprints: Dr. Dominique Bonneau, Unité de Génétique Médicale, Centre Hospitalier Universitaire, BP 577, 86021 Poitiers Cedex, France. E-mail: bonneau@univ-poitiers.fr
© 1996 by The American Society of Human Genetics. All rights reserved.
0002-9297/96/5805-0022$02.00

Am. J. Hum. Genet. 58:1092–1093, 1996

Limits on Fine Mapping of Complex Traits

To the Editor:

We recently published a paper in the *Journal* on high-resolution genetic mapping of complex traits (Kruglyak and Lander 1995). In that paper, we considered the confidence region for the position of a gene localized by genetic mapping. We showed that the size of this confidence region increases rapidly as the trait becomes more complex (that is, as the relative risk or the proportion of alleles shared by affected relatives decreases). We concluded that using affected-relative-pair analysis to localize a gene conferring a modest increased risk to a region suitable for positional cloning (e.g., 1 cM) requires a large number of relative pairs (or, more generally, meioses).

Since the paper appeared, we have discovered an intuitive way to understand the difficulty of fine mapping of susceptibility genes for complex traits. The insight was motivated by a colleague who inquired why one could not simply confine the search for a susceptibility gene to the region of maximum allele sharing in a sib-pair (or other relative pair) data set.

The key question is thus: what is the chance that a susceptibility gene will *not* lie in the region of maximum allele sharing? The answer is easily obtained by using the methods described in our previous paper (Kruglyak and Lander 1995).

Proposition

Consider a susceptibility locus at which affected sibs share a proportion of alleles $z > 1/2$. (This proportion is given by $z = (z_1 + 2z_2)/2$, where z_1 and z_2 are the proportions of affected sib pairs sharing 1 and 2 alleles identical by descent at the susceptibility locus.) The probability that the gene will not lie in the region of maximum allele sharing in an affected-sib-pair study is $(1 - z)(3z - 1)/z^2$.

Proof

In fact, one can easily show a stronger result. Consider an affected-relative-pair study involving relative pairs with allele-sharing proportion α at random loci and allele-sharing proportion z at a susceptibility locus. The chance that the number of pairs sharing alleles at the true susceptibility locus is lower by at least Δ than the maximum observed number is $q^{\Delta}(2 - q^{\Delta})$, where the quantity $q = \alpha(1 - z)/[z(1 - \alpha)]$. The proof follows from appendices D and E of Kruglyak and Lander (1995). The special case above corresponds to $\alpha = 1/2$ and $\Delta = 1$. The proof implicitly assumes that a large number of relative pairs has been studied; this is a realistic assumption in the context of fine mapping.

Consider the consequences of this relation for positional cloning based on sib-pair data. Let λ_O, λ_S, and λ_M denote the relative risk ratios for an offspring, a sibling, and a monozygotic twin of an affected individual, respectively. Then, $z = (\lambda_O/\lambda_S + \lambda_M/\lambda_S)/4$ for a single-locus trait and $z = (3\lambda_S - 1)/4\lambda_S$ in the special case of an additive single-locus trait for which $\lambda_O = \lambda_S$ and $\lambda_M = \lambda_S - 1$ (Risch 1990*a*, 1990*b*). Thus, for an additive trait with $\lambda_S = 40$, 6, 3, 2, and 1.5, the sharing proportion $z = 0.74$, 0.71, 0.67, 0.63, and 0.58, respectively. The corresponding chance that the gene lies outside the region of maximal sharing is 0.57, 0.65, 0.74, 0.84, and 0.92. By looking only in the region of maximum sharing, one will thus miss a gene conferring sixfold increased risk ~2/3 of the time, a gene conferring threefold increased risk ~3/4 of the time, and a gene conferring

twofold increased risk nearly 85% of the time. This is clearly unacceptable in a positional cloning project. While the precise numbers change, the general conclusion applies to all types of meiotic mapping data.

The simple argument presented above underscores the difficulty of finely mapping genes underlying complex traits. This situation is in contrast to that of a rare simple Mendelian trait, for which the gene *always* lies in the region of maximal sharing delimited by the closest flanking recombinants. Complex traits are different because a single recombinant cannot be trusted to rule out a region as the gene's location—the observed lack of allele sharing may instead reflect the fact that an affected individual happens not to carry the susceptibility gene.

LEONID KRUGLYAK AND ERIC S. LANDER
Whitehead Institute for Biomedical Research, Cambridge, MA

References

Kruglyak L, Lander ES (1995) High-resolution genetic mapping of complex traits. Am J Hum Genet 57:1212–1223

Risch N (1990*a*) Linkage strategies for genetically complex traits. I. Multilocus models. Am J Hum Genet 46:222–228

——— (1990*b*) Linkage strategies for genetically complex traits. II. The power of affected relative pairs. Am J Hum Genet 46:229–241

Address for correspondence and reprints: Dr. Leonid Kruglyak, Whitehead Institute, Center for Genome Research, 1 Kendall Square, Building 300, Cambridge, MA 02139. E-mail: leonid@genome.wi.mit.edu.
© 1996 by The American Society of Human Genetics. All rights reserved.
0002-9297/96/5805-0023$02.00

Am. J. Hum. Genet. 58:1093–1095, 1996

Likelihood Ratio Tests for Linkage and Linkage Disequilibrium: Asymptotic Distribution and Power

To the Editor:

Terwilliger (1995) proposes an interesting likelihood ratio test for linkage disequilibrium that appears conservative under the null hypothesis and powerful when one of several alleles is positively associated with the disease. In a model where p_j is the population frequency of marker allele j, p_D is the population frequency of the disease allele, and λ is a parameter specifying the magnitude of the linkage disequilibrium, he defines the log-likelihood of the data conditional on allele i being positively associated with the disease to be $\ln[L_i(\lambda)] = \Sigma_j [X_j \ln(q_j) + Y_j \ln(r_j)]$, where the observed counts of marker allele j on disease and control chromosomes are X_j and Y_j and where the predicted allele frequencies are $q_j = p_j + \lambda(1 - p_j)$ and $r_j = p_j - \lambda(1 - p_j)p_D/(1 - p_D)$ when $j \neq i$ (i.e., the associated allele); and $q_j = p_j - \lambda p_j$ and $r_j = p_j + \lambda p_j p_D/(1 - p_D)$ when $j \neq i$ (i.e., the nonassociated alleles). (Incidentally, the likelihood function in eq. [1] in the paper should be a product rather than a sum, although the correct formula was used in the computer program that implemented the test.) He then defines the overall likelihood to be a weighted sum of the conditional likelihoods over all marker alleles; that is, $L(\lambda) = \Sigma_i p_i L_i(\lambda)$. A likelihood ratio statistic is then

$$\Lambda = 2\{\mathrm{Max}_\lambda[\ln[L(\lambda)]]/\{\ln[L(\lambda = 0)]\} . \qquad (1)$$

This statistic assumes that allele frequencies are known (as in standard linkage analysis); when allele frequencies are uncertain, a similar likelihood ratio statistic can be defined by maximizing the numerator likelihood with respect to λ and allele frequencies jointly and maximizing the denominator likelihood with respect to allele frequencies only. In either case Λ is assumed to be asymptotically distributed as a 50:50 mixture of 0 and χ_1^2 under the null hypothesis (H_0, $\lambda = 0$). The reasoning given for the 50% point mass at 0 is that the test is "one-sided" (that is, H_0, $\lambda = 0$, is tested against H_1, $\lambda > 0$). Thus, the numerator likelihood will maximize at $\lambda = 0$ (giving $\Lambda = 0$) whenever the unrestricted maximum falls in the inadmissible region $\lambda < 0$, and this occurs with probability 0.5 under H_0.

Using this null distribution for Λ, however, Terwilliger found that the test tended to be conservative. This finding suggests that this distribution is incorrect and that the standard argument for a "one-sided" test does not apply to Λ. To simplify matters in order to gain insight into the apparent "anomalous" behavior of Λ, it is helpful to consider a particular situation under which Λ has some properties similar to LOD scores for phase-unknown sibships. The situation is when both the disease and marker loci are biallelic with known allele frequencies $p_1 = p_2 = p_D = \frac{1}{2}$. In this special case, given the observed data (X_1, X_2, Y_1, Y_2), the overall likelihood can be written as

$$L = (\tfrac{1}{2})\theta^R(1 - \theta)^{N-R} + (\tfrac{1}{2})\theta^{N-R}(1 - \theta)^R , \qquad (2)$$

where $R = X_1 + Y_2$, $N = X_1 + X_2 + Y_1 + Y_2$, and $\theta = (1 - \lambda)/2$. This likelihood function is identical in form to that of a phase-unknown sibship in which there are R gametes of one type and $N - R$ gametes of the other, and where θ is the recombination fraction. In both cases H_0 corresponds to $\theta = \frac{1}{2}$, so that R is a binomial random variable with parameters $(N, \frac{1}{2})$. It is clear that R and $N - R$ are interchangeable without affecting the value of L, so we can set $R \leq N - R$ and define $K = N - 2R$. Thus,

$$L = \theta^R(1 - \theta)^R[\theta^K + (1 - \theta)^K]/2 . \qquad (3)$$