quences than in A/T-rich R-band sequences ($P = .009$). Thus, the flexibility pattern is one of the features that differentiate R- and G-bands.

As mentioned, common fragile sites were found to be A/T rich. The pattern of high-flexibility clusters found in the identified common fragile sites (see table 1) (Mishmar et al. 1998) was significantly different ($P = .02$) from that of A/T-rich control sequences mapped to R-bands. This pattern was not different from that of A/T-rich control sequences mapped to G-bands ($P = .85$). These results might indicate that common fragile sites mapped to R-bands have the flexibility patterns characteristic of G-bands with the same A/T content.

Our previous analysis of potential unusual DNA structures in FRA7H revealed a cluster of regions with potential to form triple helixes (Mishmar et al. 1998). Previous studies, using monoclonal antibodies to triple-helix DNA, showed that G-bands are rich in triple-helix DNA (Burkholder et al. 1991). Thus, clusters of regions with potential to form triple-helix DNA might be added to the G-band characteristics found in common fragile sites.

Together, all the known molecular features of common fragile sites indicate that they might consist of DNA sequences with characteristics of G-bands embedded within R-bands. Of what significance could this feature be to the mechanism of fragility? We think that delayed replication and aberrant condensation of fragile sites might be involved. Chromosomal bands apparently represent regions with several origins of replication that are coordinately controlled to initiate the replication process. The presence of a relatively small region consisting of a common fragile site with G-band characteristics might lead to disturbances in the regional control of replication. This might involve inappropriate initiation of replication in the fragile region. The addition of aphidicolin, which inhibits DNA elongation, might further add to the interference in replication at fragile sites, leading to unreplicated sequences that might adopt abnormal chromatin organization, resulting in fragility.

DAN MISHMAR,[1] YAEL MANDEL-GUTFREUND,[2] HANAH MARGALIT,[2] AYELET RAHAT,[1] AND BATSHEVA KEREM[1]

[1]*Department of Genetics, The Hebrew University, and* [2]*Department of Genetics and Biotechnology, The Hebrew University, Hadassah, Jerusalem*

Electronic-Database Information

URLs for data in this article are as follows:

Genome Database, http://www.gdb.org

Hebrew University of Jerusalem, http://leonardo.ls.huji.ac.il/departments/genesite/faculty/bkerem.htm

References

Boldog F, Gemmill RM, West J, Robinson M, Robinson L, Li E, Roche J, et al (1997) Chromosome 3p14 homozygous deletions and sequence analysis of FRA3B. Hum Mol Genet 6:193–203

Burkholder GD, Latimer LJ, Lee JS (1991) Immunofluorescent localization of triplex DNA in polytene chromosomes of Chironomus and Drosophila. Chromosoma 101:11–18

Gardiner K (1995) Human genome organization. Curr Opin Genet Dev 5:315–322

Huang H, Qian J, Proffit J, Wilber K, Jenkins R, Smith DI (1998) FRA7G extends over a broad region: coincidence of human endogenous retroviral sequences (HERV-H) and small polydispersed circular DNAs (spcDNA) and fragile sites. Oncogene 16:2311–2319

Inoue H, Ishii H, Alder H, Synder E, Druck T, Huebner K, Croce CM (1997) Sequence of the FRA3B common fragile region: implications for the mechanism of FHIT deletion. Proc Natl Acad Sci USA 94:14584–14589

Le Beau MM, Rassool FV, Neilly ME, Espinosa III R, Glover TW, Smith DI, McKeithan TW (1998) Replication of a common fragile site, FRA3B, occurs late in S phase and is delayed further upon induction: implications for the mechanism of fragile site induction. Hum Mol Genet 7:755–761

Mishmar D, Rahat A, Scherer SW, Nyakatura G, Hinzmann B, Kohwi Y, Mandel-Gutfreund Y, et al (1998) Molecular characterization of a common fragile site (FRA7H) on human chromosome 7 by the cloning of an SV40 integration site. Proc Natl Acad Sci USA 95:8141–8146

Sarai A, Mazur J, Nussinov R, Jernigan RL (1989) Sequence dependence of DNA conformational flexibility. Biochemistry 28:7842–7849

Selig S, Okumura K, Ward DC, Cedar H (1992) Delineation of DNA replication time zones by fluorescence in situ hybridization. EMBO J 11:1217–1225

Sutherland GR, Richards RI (1995) The molecular basis of fragile sites in human chromosomes. Curr Opin Genet Dev 5:323–327

Address for correspondence and reprints: Dr. Batsheva Kerem, Department of Genetics, The Hebrew University, Jerusalem, Israel, 91904. E-mail: kerem@leonardo.ls.huji.ac.il

0002-9297/99/6403-0032$02.00

Am. J. Hum. Genet. 64:910–915, 1999

Finite-Sample Properties of Family-Based Association Tests

To the Editor:

During the past few years, there has been much interest in the use of family-based association tests to detect linkage between marker and disease loci, since these methods avoid the problems of ascertaining the appropriate pop-

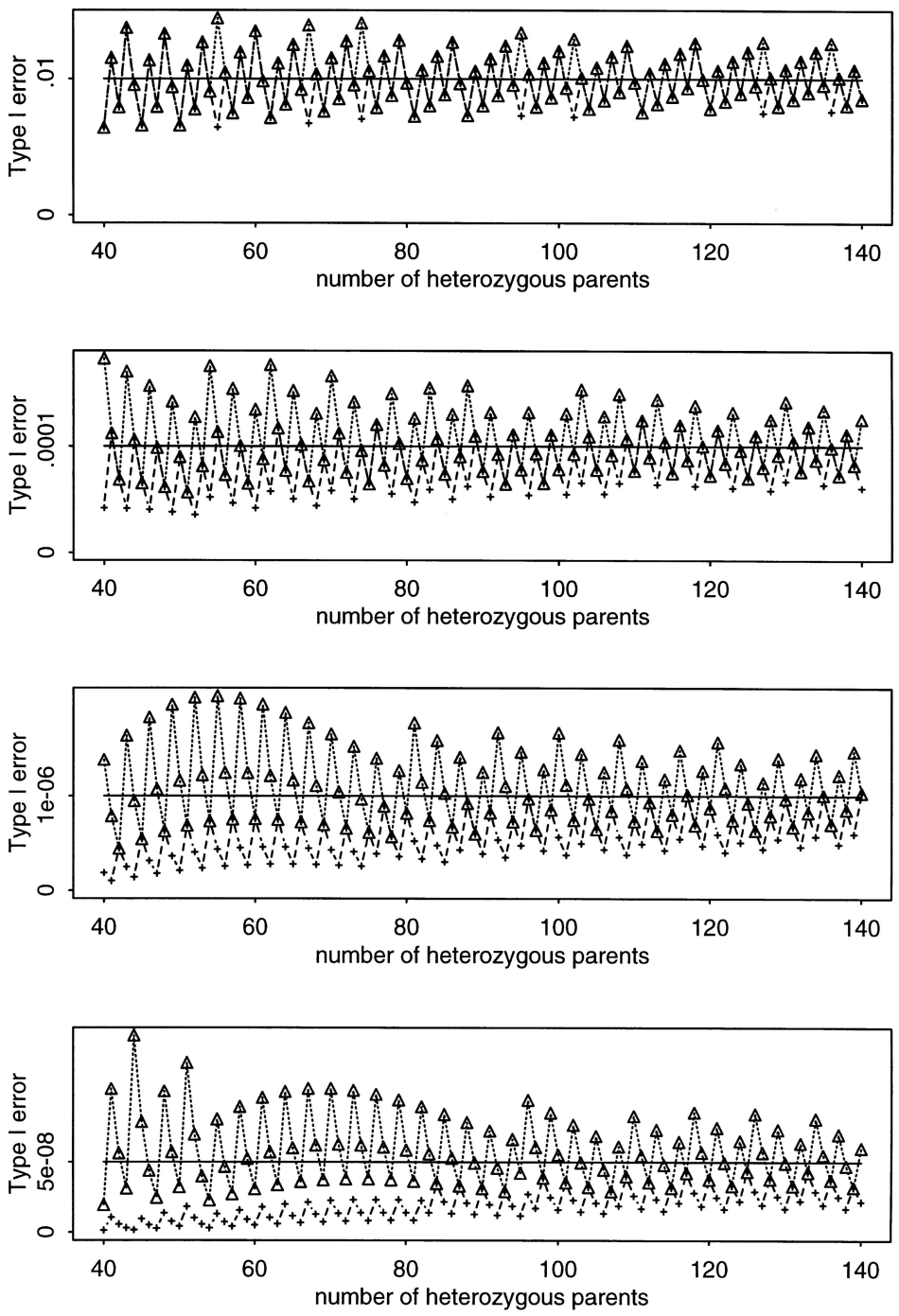

Figure 1 Achieved type I errors for $T_{TDT}(+)$ and Λ (Δ), compared with specified α (*unbroken horizontal line*), for (*from top to bottom*) $\alpha = .01, .0001, 1 \times 10^{-6}$ (1e-06), and 5×10^{-8} (5e-08).

ulations of cases and controls implicit in population association studies. Although these tests were originally developed for candidate-gene studies, the use of such methods in genome scans has recently been proposed (Risch and Merikangas 1996).

Perhaps the best-known family-based association test is the transmission/disequilibrium test (TDT) for diallelic markers, introduced by Spielman and Ewens (1993). A number of similar tests have subsequently been suggested; for reviews, see the work of Spielman and Ewens (1996) and Curnow et al. (1998). However, it has often not been clear how these various tests are related, and

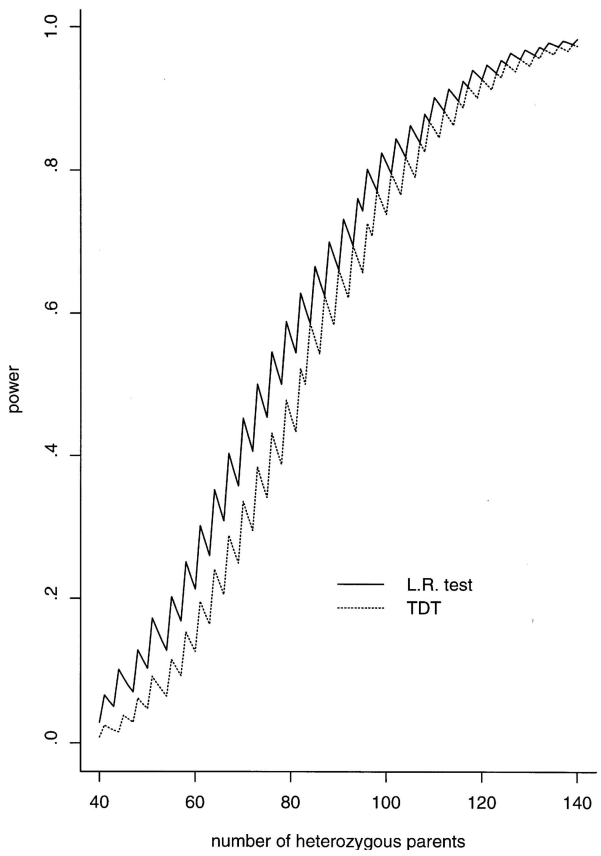

Figure 2 Power versus number of heterozygous parents for T_{TDT} and likelihood-ratio statistic Λ (L.R. test), when $\gamma = 4$ and $\alpha = 5 \times 10^{-8}$.

there has been debate about the advantages and disadvantages of several of the tests (e.g., see Kaplan et al. 1997; Sham 1997). There is a need to investigate the relationships between the suggested tests and to establish which of them should be preferred in a given situation. Determining the relative merits of competing test statistics is often difficult, because the comparisons usually rely on simulation or on asymptotic results that may be of limited relevance to finite-sample data. Here we suggest a way of avoiding these problems for diallelic markers, focusing particularly on the TDT, the extended TDT (ETDT [Sham and Curtis 1995]), the score tests introduced by Schaid (1996), and tests based on the conditional likelihood of the offspring marker types when the parental marker data are given.

We consider a sample of N families, each with a single affected child. All individuals have been genotyped at a marker locus with m alleles, labeled as "M_1," M_2,...,M_m." We wish to use information on the alleles transmitted from the parents to the affected child, to test the null hypothesis of no linkage or no association between the marker and disease. Note that we can only test this compound null hypothesis when we have a single affected child in each family; for families with multiple affected children, the tests discussed below are valid only as tests of linkage, not as tests of association (Spielman and Ewens 1996).

Consider a single family, with parental marker genotypes **g** and **h** and with the genotype of the affected child denoted by **x**. If we use C_A to denote that a child is affected, then the probability of the child's genotype, conditional on the parental genotype is, by Bayes's theorem,

$$P(\mathbf{x}|g,h,C_A) = \frac{P(C_A|\mathbf{x},\mathbf{g},\mathbf{h})P(\mathbf{x}|g,h)P(\mathbf{g},\mathbf{h})}{\sum_{x^* \in G} P(C_A|\mathbf{x}^*,\mathbf{g},\mathbf{h})P(\mathbf{x}^*|g,h)P(\mathbf{g},\mathbf{h})}$$

$$= \frac{P(C_A|\mathbf{x})P(\mathbf{x}|g,h)}{\sum_{x^* \in G} P(C_A|\mathbf{x}^*)P(\mathbf{x}^*|g,h)},$$

where G is the set of possible marker types for the affected child (Schaid 1996). We shall assume normal segregation (which requires, e.g., absence of meiotic drive), so that $P(\mathbf{x}|\mathbf{g}, \mathbf{h})$ is easily calculated, leaving only $P(C_A|\mathbf{x})$, the risk of disease for a particular marker genotype, to be discussed.

We could model the disease locus explicitly, but it is often more convenient (Self et al. 1991; Schaid 1996) to work directly with the marker genotype, by putting $f_{\mathbf{x}} = P(C_A|\mathbf{x})$. The $f_{\mathbf{x}}$ then reflects both the disease-locus penetrances and the strength of allelic association between the marker and disease loci in the affected children. Note that $m(m + 1)/2$ parameters are needed in the general model; this will be large for highly polymorphic markers. Schaid (1996) derives score tests for general $f_{\mathbf{x}}$ and for various special cases representing particular disease models. In particular, Schaid shows that, for a log-additive model—that is, one in which allelic effects combine multiplicatively at the marker, so that $f_{\mathbf{x}} = f_{x_1} f_{x_2}$—the score test for a diallelic marker is the TDT statistic.

Note that we are using this multiplicative model as a convenient approximation, rather than as something that we believe is exactly correct. If the true disease model is not multiplicative, then tests based on the multiplicative model remain valid in the sense of having the correct size, but they may not be optimal, in that there may exist other test statistics with higher power. However, the multiplicative model has the advantage of requiring $(m - 1)$ parameters, where tests derived by means of the general model need $m(m + 1)/2$. This means that tests based on the multiplicative model can be more powerful than general alternatives even when the allelic effects do not combine multiplicatively at the marker (e.g., see Schaid 1996).

Schaid (1996) comments that likelihood-ratio tests

could be used instead of score tests, but he opts for score statistics because of their ease of calculation. Here we derive the likelihood-ratio test for the multiplicative disease model discussed above.

It can be shown (e.g., see Curnow et al. 1998) that, in the presence of allelic association, the marker alleles transmitted to an affected child from the child's two parents are independent if and only if the multiplicative model holds at the marker. Therefore, under the multiplicative model, $P(\mathbf{x}|g,h,C_A) = P(x_1|\mathbf{g},C_A)P(x_2|\mathbf{h},C_A)$, where x_1 and x_2 are the alleles transmitted from parents with genotypes **g** and **h**, respectively. Now,

$$P(x_1 = i|\mathbf{g},C_A) = 0 \quad \text{if } g_1 \neq i \text{ and } g_2 \neq i ,$$

$$P(x_1 = i|\mathbf{g},C_A) = 1 \quad \text{if } g_1 = g_2 = i ,$$

$$P(x_1 = i|\mathbf{g},C_A) = \frac{f_i P(x_1 = i|\mathbf{g})}{f_{g_1}P(x_1 = g_1|\mathbf{g}) + f_{g_2}P(x_1 = g_2|\mathbf{g})} \quad \text{otherwise} .$$

so that, if $\mathbf{g} = (i,j)$ and $x_1 = i$, then

$$P(x_1|\mathbf{g},C_A) = 1 \quad \text{if } i = j ,$$

$$P(x_1|\mathbf{g},C_A) = \frac{f_i}{(f_i + f_j)} \quad \text{if } i \neq j ,$$

because $P(x_1 = i|\mathbf{g}) = P(x_1 = j|\mathbf{g}) = .5$.

Let n_{ij} be the number of transmissions of M_i from M_iM_j parents in our sample of $2N$ parents. Then $L(\mathbf{f})$, the likelihood of the child genotypes, given the parental genotypes, under the multiplicative model, is

$$L(\mathbf{f}) = \prod_{i=1}^{m}\prod_{j<i}\binom{n_{ij}+n_{ji}}{n_{ij}}\left(\frac{f_i}{f_i+f_j}\right)^{n_{ij}}\left(\frac{f_j}{f_i+f_j}\right)^{n_{ji}} ,$$

by derivation from the formula above. This can be maximized over **f**, to give L_A. The null hypothesis is no linkage or no association between marker and disease; in this case, the two parental marker alleles are equally likely to be transmitted, so that the likelihood under the null hypothesis is

$$L_0 = \prod_{i=1}^{m}\prod_{j<i}\binom{n_{ij}+n_{ji}}{n_{ij}}\left(\frac{1}{2}\right)^{n_{ij}} .$$

The likelihood-ratio statistic is $\lambda = L(\mathbf{f})/L_0$, and, by standard theory, $-2\ln(\lambda)$ has an approximate χ^2 distribution with $m - 1$ df, under the null hypothesis.

Note that $L(\mathbf{f})$ is equivalent to the likelihood derived by Sham and Curtis (1995), with our f_i being equivalent to their d_{ii}. Sham and Curtis (1995) made assumptions that, at first sight, seem to be rather different from those which we have made here: they assume that there is no recombination between marker and disease loci and that parental transmissions of marker alleles are independent. However, as we have noted above, parental transmissions are independent if and only if the multiplicative model holds at the marker locus; the two sets of assumptions are therefore directly equivalent, and we should expect to obtain the same likelihoods.

In summary, the score test for the multiplicative model is Schaid's (1996) general TDT statistic, and the likelihood-ratio test is Sham and Curtis's (1995) ETDT statistic. By standard theory (Cox and Hinkley 1974), these tests are asymptotically equivalent; we will now show that, for diallelic markers, a stronger result holds.

Remember that in this case the score test is the TDT of Spielman et al. (1993), so that, for a test of size α, we reject the null if

$$T_{\mathrm{TDT}} = \frac{(n_{12} - n_{21})^2}{n_{12} + n_{21}} > k_{\mathrm{TDT}} ,$$

where the critical value k_{TDT} is chosen to give the required type I error rate α. In most cases, it seems that $L(\mathbf{f})$ must be maximized numerically, but, if $m = 2$, then we have

$$L(\mathbf{f}) \propto \frac{f_1^{n_{12}}f_2^{n_{21}}}{(f_1+f_2)^{n_{12}+n_{21}}} ,$$

and it is easy to show that $L(\mathbf{f})$ is maximized when $f_1/f_2 = n_{12}/n_{21}$. The likelihood-ratio test with size α is therefore likely to reject the null hypothesis if

$$\Lambda = 2\ln\left\{\frac{n_{12}^{n_{12}}n_{21}^{n_{21}}}{[0.5(n_{12}+n_{21})]^{n_{12}+n_{21}}}\right\} > k_\Lambda ,$$

where, again, the critical value k_Λ is chosen to give the required type I error rate α. Usually, asymptotic results are relied on, so that $k_\Lambda = k_{\mathrm{TDT}} = \chi^2_{1,1-\alpha}$.

Now suppose that there are H heterozygote parents in the sample. The values of T_{TDT} and Λ are completely determined by n_{12}, because $n_{21} = H - n_{12}$. Both of the aforementioned tests can be rewritten with rejection region $\{c_l > n_{12}\} \cup \{n_{12} > c_u\}$, where $c_l = H - c_u$ by symmetry; in fact, it is clear that any reasonable test statistic must have this form. Moreover, under the multiplicative disease model discussed above, the probability that a heterozygote parent will transmit the M_1 allele rath-

er than the M_2 is $\nu = f_1/(f_1 + f_2)$, so that, conditional on H heterozygote parents in the sample, $n_{12} \sim \text{binomial}(H,\nu)$. The constant c_u can be chosen to give the approximate type I error rate, by use of this binomial distribution; any c_u corresponds to a particular k_Λ and k_{TDT}, and, clearly, for a particular c_u, the test statistics will be exactly equivalent. It follows that, provided that type I error rates are properly controlled by appropriate choice of the critical values k_Λ and k_{TDT}, Λ and T_{TDT} will have identical power. However, the appropriate critical values will be the same for the two tests only if Λ and T_{TDT} have the same distribution, for then we will have $k_\Lambda = k_{TDT}$ for any c_u; if not, then setting $k_\Lambda = k_{TDT} = \chi^2_{1,1-\alpha}$ will give different rejection regions—and, therefore, different sizes and powers—for Λ and T_{TDT}. We now consider the properties of the test statistics if the χ^2 approximation is used.

It is convenient to put $\gamma = f_1/f_2$ so that $\nu = \gamma/(1 + \gamma)$ and to condition on there being H heterozygote parents, so that the distribution of the tests statistic depends only on the two parameters H and ν. Note that the null hypothesis that the two parental maker alleles are equally likely to be transmitted corresponds to $\gamma = 1$. We can investigate the *actual* size and power of the test directly, because, for any α, it is easy to calculate the probability that T_{TDT} or Λ is greater than $\chi^2_{1,1-\alpha}$, via the above binomial distribution. We consider $\alpha = .01, .0001, 1 \times 10^{-6}$, and 5×10^{-8}. The first two values of α might be appropriate for candidate loci, whereas the last has been suggested, by Risch and Merikangas (1996), for genome scans using the TDT.

First, consider the actual type I error rates, which are shown in figure 1, for $\alpha = .01, .0001, 1 \times 10^{-6}$, and 5×10^{-8}. Achieved type I errors for Λ oscillate about the asymptotic size, α, with the amplitude of the oscillation being relatively greatest for small α. Achieved type I errors for T_{TDT} are less variable, and T_{TDT} is conservative for small α. Overall, the null distributions of T_{TDT} and Λ are well approximated by χ^2_1, for most α, with the approximation much less satisfactory in the extreme tails of the distribution, despite reasonably large sample sizes. This suggests that the χ^2 approximation should not be used in genome scans, unless sample sizes are very large. The oscillation of the type I errors about the asymptotic size is caused by the underlying discreteness of the data; for example, the critical value $\chi^2_{1,1-\alpha}$ is 29.72 when $\alpha = 5 \times 10^{-8}$, and, if $H = 138$ or $H = 139$, this is exceeded by T_{TDT} only if $n_{21} \geq 102$ or $n_{12} \geq 102$. The type I error rate for T_{TDT} for $H = 138$ or $H = 139$ is therefore $P(n_{12} \geq 102) + P(n_{21} \geq 102))$ when $\gamma = 1$, which is greater if $H = 139$ than if $H = 138$. However, if $H = 140$, then T_{TDT} exceeds 29.72 only if $n_{12} \geq 103$ or $n_{21} \geq 103$, so that the type I error rate for T_{TDT} is $P(n_{12} \geq 103) + P(n_{21} \geq 103)$, for $\gamma = 1$, and this is less than the type I error rate for $H = 139$. Note that the effect on the error rate can be large, even for reasonable sample sizes; for example, if $H = 118$, then the achieved type I error rate for Λ is 8.55×10^{-8}, nearly twice the nominal 5×10^{-8}.

We stress again that, if critical values are correctly set, then T_{TDT} and Λ have identical size and power. However, the χ^2 approximation does not provide the correct critical values, and we will now show that this also leads to misleading power results. In particular, the fact that Λ tends to have higher type I error rates than does T_{TDT}, when the χ^2 approximation is used, can result in Λ appearing to have more power than T_{TDT} does; for example, consider figure 2, which plots the power of Λ and T_{TDT} as a function of H, for $\alpha = 5 \times 10^{-8}$ and $\gamma = 4$. We see that, although T_{TDT} and Λ have similar power for small γ, Λ can have considerably higher power for large γ, particularly for small α. The fact that power does not increase monotonically with sample size is, at first sight, surprising but, as with the oscillations in the type I error that have been noted above, is explained by the underlying discreteness of the data.

We have derived the likelihood-ratio test for a sample of families, in each of which there is a single affected child and all individuals have been genotyped at a particular marker locus, under the assumption that allelic effects combine multiplicatively at the marker. This test proves to be the ETDT (Sham and Curtis 1995), because the alleles transmitted to an affected child are independent if and only if the multiplicative model holds. For diallelic markers, the finite-sample properties of this statistic can be investigated by exact calculation, because then, for H heterozygous parents, only $H + 1$ outcomes need be considered, and these are easily enumerated. The computations become more complicated if the marker under consideration has more than two alleles, but, as a way of assessing the properties of test statistics, this type of exact calculation may be preferable to simulation, because it gives the exact sampling distribution of the test statistic under investigation, at any point in the parameter space.

For diallelic markers, the likelihood-ratio test Λ and the corresponding score test T_{TDT} are exactly equivalent, provided that type I error rates are correctly controlled. Exact calculation of the distribution of Λ and T_{TDT} shows that type I error rates are adequately controlled by reliance on asymptotic distributions for α that are appropriate for candidate loci, but not for the very small α required for genome screens. It thus seems that, if Λ or T_{TDT} is to be used in genome scans, then either significance levels must be calculated exactly, by means of the binomial distribution given here, or Monte Carlo approaches (Kaplan et al. 1997; Morris et al. 1997) must be used. Note also the exact test of Cleves et al. (1997).

It is easy to derive corresponding likelihood tests for other models, such as those for recessive or dominant

diseases, or for the general model in which no relationship is assumed between the f_x, and it is also easy to extend the test to cope with multiple affected or unaffected sibs (Thompson 1997). The properties of such tests deserve further study; comparisons with the corresponding score tests (Schaid 1996) would be of particular interest. By analogy with the results given here, reliance on asymptotic null distributions should be avoided for such tests; also see the work of Chapman (1976). Monte Carlo approaches are always available, and they provide a simple alternative to asymptotic approximations.

Finally, note that the results given here illustrate the problems of comparing the test statistics by stochastic simulation, especially when asymptotic distributional results are relied on. When the χ^2 approximation is used, the properties of Λ and T_{TDT} are very sensitive to the value of H, and to exhaustively survey the relevant parameter space by simulation would be very time consuming. On the basis of simulation results produced by the asymptotic χ^2 distribution, it would have been easy to conclude, according to the values of α, H, and γ studied, either that Λ is preferable to T_{TDT}, because it has higher power for some parameter values, or that T_{TDT} is preferable to Λ, because Λ can be anticonservative for certain values of H. For example, if $\alpha = 5 \times 10^{-8}$ and $H = 100$, then T_{TDT} is conservative, whereas the achieved type I error rate for Λ is close to the nominal value, so we might conclude that Λ is preferable. However, for $\alpha = 5 \times 10^{-8}$ and $H = 110$, T_{TDT} is still conservative but Λ is unacceptably anticonservative, and we would prefer T_{TDT}. As we have seen, we would be mistaken in both cases, because the differences are due to failure of the asymptotic χ^2 approximation, rather than to differences between the test statistics.

Acknowledgments

We are grateful to Robert Curnow and Andrew Morris for helpful discussions of this work and to two anonymous referees who made useful comments on this letter.

J. C. Whittaker[1] and D. J. Thompson[2]
[1]*Department of Applied Statistics, University of Reading, Reading, United Kingdom; and* [1]*CRC Genetic Epidemiology Unit, University of Cambridge, Cambridge*

References

Chapman JW (1976) A comparison of the χ^2, $-2\log R$ and multinomial probability criteria for significance tests when expected frequencies are small. J Am Stat Assoc 71:854–863

Cleves MA, Olson JM, Jacobs KB (1997) Exact transmission-disequilibrium tests with multiallelic markers. Genet Epidemiol 14:337–347

Cox DR, Hinkley DV (1974) Theoretical statistics. Chapman & Hall, London

Curnow RN, Morris AP, Whittaker JC (1998) Locating genes involved in human diseases. Appl Stat 47:63–76

Kaplan NL, Martin ER, Weir BS (1997) Power studies for transmission/disequilibrium tests with multiple alleles. Am J Hum Genet 60:691–702

Morris AP, Curnow RN, Whittaker JC (1997) Randomisation tests of disease-marker association. Ann Hum Genet 61: 49–60

Risch N, Merikangas K (1996) The future of genetic studies of complex human diseases. Science 273:1516–1517

Schaid DJ (1996) General score tests for associations of genetic markers with disease using cases and their parents. Genet Epidemiol 13:423–449

Self SG, Longton G, Kopecky KJ, Liang KY (1991) On estimating HLA-disease association with application to a study of aplastic anemia. Biometrics 47:53–61

Sham P (1997) Transmission/disequilibrium tests for multiallelic loci. Am J Hum Genet 61:774–778

Sham PC, Curtis D (1995) An extended transmission/disequilibrium test (TDT) for multiallele marker loci. Ann Hum Genet 59:323–336

Spielman RS, Ewens WJ (1996) The TDT and other family-based tests for linkage disequilibrium and association. Am J Hum Genet 59:983–989

Spielman RS, McGinnis RE, Ewens WJ (1993) Transmission test for linkage disequilibrium: the insulin gene region and insulin-dependent diabetes mellitus (IDDM). Am J Hum Genet 52:506–516

Thompson DJ (1997) Using information from unaffected sibs in genetic association studies. MS diss, Department of Applied Statistics, University of Reading, Reading, UK

Address for correspondence and reprints: Dr. John Whittaker, Department of Applied Statistics, University of Reading, P.O. Box 240, Whiteknights Road, Reading, RG6 2FN, United Kingdom. E-mail: j.c.whittaker@reading.ac.uk

0002-9297/99/6403-0033$02.00