# Linkage Disequilibrium and the Search for Complex Disease Genes

## L.B. Jorde

*Eccles Institute of Human Genetics, University of Utah Health Sciences Center, Salt Lake City, Utah 84112, USA*

During the past two decades, linkage analysis has been phenomenally successful in localizing Mendelian disease genes. Linkage disequilibrium (LD) analysis, which effectively incorporates the effects of many past generations of recombination, has often been instrumental in the final phases of gene localization (Feder et al. 1996; Hästbacka et al. 1994; Kerem et al. 1989). These successes have fueled hopes that similar approaches will be effective in localizing genes underlying susceptibility to common, complex diseases.

With the exception of Mendelian subsets of common diseases (e.g., *BRCA1* and *BRCA2* for breast cancer, APC for colon cancer, the *LDL* receptor gene for heart disease), progress on this front has been limited. Typically, a nonparametric linkage analysis, such as a sib-pair analysis, will implicate several genetic regions as targets for further investigation. These regions, often 10–20 Mb in size, remain intractably large for effective positional cloning. It is now hoped that LD approaches, using hundreds of thousands of new polymorphic markers, will overcome this impasse (Risch and Merikangas 1996).

The rationale underlying LD mapping of complex disease genes is straightforward and similar to the justification for LD mapping of Mendelian disease genes. With both types of disease genes, the primary advantage of LD analysis remains its ability to use the effects of dozens or hundreds of past generations of recombination to achieve fine-scale gene localization (Jorde 1995). An important difficulty, common to both types of disease genes, is that past historical events (admixture, genetic drift, multiple mutations, and natural selection) can disturb the relationship between LD and inter-locus physical distance. A major difference, of course, is that locus heterogeneity complicates the analysis of complex diseases and may be more extensive for these diseases than for most Mendelian diseases. Furthermore, allelic heterogeneity may be present at each locus. This heterogeneity, the scope of which is largely unknown, will limit the strength of association between a given polymorphism and an observable phenotype.

Despite these challenges, LD mapping holds considerable appeal, and there is great demand to resolve the genetics of complex diseases. Consequently, many new techniques have been devised to carry out LD analysis, often with a view toward mapping complex disease loci. The purpose of this review is to summarize these techniques and some of the issues surrounding their application. In particular, the evolutionary factors that can confound or enhance disequilibrium analysis will be discussed, and some thoughts will be offered on the optimal choice of markers and populations for LD analysis.

## Linkage Disequilibrium Measures

### Two-Locus Methods

LD, simply defined, is the nonrandom association of alleles at linked loci. Although concepts of LD date to the early part of the twentieth century (Jennings 1917), the first commonly used LD measure, *D,* was developed by Richard Lewontin some 35 years ago (Lewontin 1964). For a pair of diallelic loci, *A* and *B,* this statistic measures the difference between two quantities: (1) the observed frequency of co-occurrence of an allele of *A* ($A_1$) and an allele of *B* ($B_1$) on the same chromosome and (2) the expected frequency of co-occurrence under linkage equilibrium. The observed frequency, denoted by $P_{11}$, is the proportion of chromosomes on which alleles $A_1$ and $B_1$ co-occur in a population. The expected value of $P_{11}$ under linkage equilibrium is the product of the allele frequencies of $A_1$ and $B_1$ in the population. Thus, $D = P_{11} - p_1q_1$, where the allele frequencies are symbolized as follows: $p_1 = f(A_1)$; $p_2 = 1 - p_1 = f(A_2)$; $q_1 = f(B_1)$; $q_2 = 1 - q_1 = f(B_2)$. If $D$ differs significantly from zero, LD is said to exist. The degree of LD between two loci is dependent on both the recombination fraction, $\theta$, and time in generations, $t$ (e.g., since the origin of a new disease-causing mutation at time 0): $D_t = D_0(1 - \theta)^t$. Thus, $D$ will tend to be smaller when two loci are located further apart, and $D$ will decrease through time as a result of recombination. $D$ provides a simple indication of the frequency of recombination and, hence, the physical distance between two loci. Alternatively, if $\theta$ can be estimated, LD can be used to infer the age of a disease-causing mutation.

$D$ is dependent on allele frequencies in the population: its maximum value is given by $D_{max} = \min(p_1q_2, p_2q_1)$, whereas its minimum value, $D_{min}$, is given by

max($-p_1 q_1$, $-p_2 q_2$). $D$ can be scaled as $D' = D/D_{max}$ (Lewontin 1964). Another common scaling of $D$ is to divide it by $\sqrt{(p_1 p_2 q_1 q_2)}$ (Hill and Robertson 1968). This quantity, commonly labeled R or $\Delta$, is equal to $\sqrt{\chi^2/N}$, where the $\chi^2$ statistic can be obtained from the $2 \times 2$ table of haplotype frequencies (i.e., $P_{11}$, $P_{12}$, $P_{21}$, and $P_{22}$), and $N$ is the total number of haplotypes in the sample. This provides a means of testing the statistical significance of $R$. Alternatively, significance can be evaluated using permutation-based methods (Zaykin et al. 1995). The latter approach is especially useful with multiallelic microsatellite loci that yield many possible two-locus genotypes. Although $R$ can vary from –1 to 1, it is limited by the actual values of the allele frequencies, and thus, like $D$, it is a frequency-dependent measure.

Another common two-locus disequilibrium statistic, commonly labeled $\delta$, is similar to an attributable risk measure and is given by $D/(q_1 P_{22})$ (Bengtsson and Thomson 1981), in which $q_1$ is the population frequency of a disease allele, $B_1$, and $P_{22}$ is the frequency of chromosomes that contain marker allele $A_2$ and the normal allele, $B_2$. Several comparative analyses of two-locus LD measures have shown that, in most circumstances, $\delta$ and $D'$ give more reliable estimates of physical distance than do $D$ and $R$ because of the latter's dependence on allele frequencies (Ajioka et al. 1997; Devlin and Risch 1995; Guo 1997). Devlin and Risch (1995) showed that $\delta$ is directly proportional to the recombination fraction and is thus a desirable measure of genetic distance. However, even $\delta$ can show allele-frequency dependence when multiple disease-causing mutations have occurred on different haplotype backgrounds (Guo 1997).

The traditional LD measures, $R$ and $D$, implicitly assume a constant population, as is implied by the well-known relationship $E(R^2) = 1/(1 + 4\ N_e\theta)$, where $N_e$ is effective population size. Under a model of recombination-drift equilibrium, $N_e$ is proportional to the time since a new mutation occurred, giving rise to new disequilibrium (Hill and Robertson 1968; Kaplan et al. 1995). However, most human populations have undergone rapid population growth. This led investigators to adapt the Luria-Delbrück model of bacterial mutation to the process of recombination in exponentially growing human populations (Hästbacka et al. 1992). The $P_{excess}$ statistic derived by Hästbacka et al. is identical to $\delta$ (Devlin and Risch 1995). Kaplan et al. (1995) devised a likelihood approach that simulates a growing population and gives more accurate confidence limits than does the method of Hästbacka et al. Rannala and Slatkin (1998) derived the sampling distribution for a diallelic marker closely linked to a low-frequency mutation that arose once in a nonstationary population. Graham and Thompson (1998) applied coalescence theory to the problem of nonstationary populations and formulated a method that can account for any pattern of population growth.

## Controlling for Stratification: The Transmission Disequilibrium Test and Its Variants

LD testing is typically carried out as a case-control comparison in which marker frequencies are compared in samples of affected individuals and unaffected controls. Case-control studies can be confounded by population stratification, and there are several empirical examples of this problem (Knowler et al. 1988; Reich et al. 1999). LD analysis is no exception to this difficulty, and various strategies have been devised to control for stratification. An early example is the haplotype relative risk (HRR) method (Falk and Rubinstein 1987), which tests for association by defining the haplotype transmitted by a parent to an affected offspring as the "case" haplotype and the untransmitted parental haplotype as the "control". This ensures that case and control haplotypes come from the same population, reducing (but not eliminating) the potential for stratification. This test is also known, in modified form, as the transmission disequilibrium test, or TDT (Spielman et al. 1993). The TDT tests for linkage in the presence of LD and eliminates stratification effects completely, but a disadvantage is that it uses only heterozygous parents. When there is no stratification, it is less powerful than the HRR, which uses both homozygous and heterozygous parents (Schaid 1998). When there is stratification, the HRR (or similar tests such as AFBAC [Thomson 1995]) is more likely to yield false positive results (Spielman and Ewens 1996).

Numerous variants of the TDT have been devised, including extensions for multiallelic markers (Sham and Curtis 1995), multiple marker loci (Wilson 1997), quantitative disease loci (Allison 1997; Xiong et al. 1998), extended pedigrees (George et al. 1999), and families in which only one parent (Sun et al. 1999; Weinberg 1999) or only siblings are available (Allison et al. 1999; Spielman and Ewens 1998; Teng and Risch 1999). The sib-TDT is especially useful for diseases of late adulthood, in which multiple generations may not be available for study. However, it is somewhat less powerful than the traditional TDT (Schaid 1998). Another variant of the sib-TDT examines allele-sharing patterns in sibs who are discordant for a trait (Boehnke and Langefeld 1998). Again, this approach will be especially useful for late-onset traits in which multiple affected sibs may be difficult to collect.

In many LD-mapping studies, family data are already available because of their collection for an initial linkage analysis. It is thus practical to use the TDT in such situations to avoid the problem of stratification. It has been argued, however, that traditional case-control studies may be preferable to the TDT if family data are

not already available (Morton and Collins 1998). This is because the TDT is statistically only half as efficient as a case-control design and thus requires much more effort (i.e., the ascertainment of twice as many subjects) to gain an equivalent amount of information. Morton and Collins argue that stratification, which reduces the accuracy and power of the case-control design, is a problem only under rare circumstances. Although analyses of DNA forensic databases offer support for this assertion (National Research Council 1996), the issue should be addressed with further empirical data. When the extent of stratification is unknown, a practical solution is to assess associations between multiple unlinked markers in case and control populations to test and correct for stratification effects (Pritchard and Rosenberg 1999).

### Admixture Disequilibrium Tests

The admixture of genetically distinct populations can generate LD throughout the genome (Nei and Li 1973) and is often considered a liability in disequilibrium mapping. However, admixture can potentially be turned to an advantage because, following an admixture event, disequilibrium will decay as a function of the distance between a disease-causing gene and marker loci (Chakraborty and Weiss 1988). Several approaches have been devised to assess admixture-generated disequilibrium (Kaplan et al. 1998; McKeigue 1997; McKeigue 1998). Because many human populations have undergone extensive recent admixture (e.g., Hispanics in the United States), there are a number of potential candidates for admixture disequilibrium mapping. Important requirements for successful application of this approach are that the parental populations should show relative genetic homogeneity and that their allele frequencies should differ substantially (Stephens et al. 1994). This is often not the case in major human populations (e.g., African-Americans in whom the parental African population is highly diverse) (Dean et al. 1994). Further complications arise from the fact that admixture in modern human populations is seldom limited to a specific point in history but instead is a continuing process.

**Table 1.** A Summary of the Advantages and Disadvantages of Various Linkage Disequilibrium Measures and Methods

| Method | Advantages | Disadvantages |
|---|---|---|
| Two-locus measures: D and R | Commonly used; easy to compute; theory and sampling distributions well understood | Both measures can be strongly influenced by allele frequencies |
| Two-locus measures: D' and $\delta$ (latter is equivalent to $\lambda$ and $P_{excess}$) | Also easy to compute; less dependent on allele frequencies than D and R | Can still be influenced by allele frequencies in some cases; like all simple LD measures, do not take advantage of multilocus marker data |
| Transmission disequilibrium test (TDT) | Commonly used measure; unaffected by population stratification (in contrast to other two-locus measures listed above) | Less efficient than standard case-control design; uses only heterozygous parental genotypes |
| Haplotype relative risk (HRR) | Estimates transmitted allele frequencies from both homozygous and heterozygous parental genotypes and is thus useful for pooled DNA samples | Can produce false positive LD in the presence of population stratification |
| Admixture disequilibrium mapping | Takes advantage of a history of admixture in human populations | Entails many restrictive requirements about population history and homogeneity (see text) |
| Multilocus composite likelihood approaches (see text) | Use of multiple markers can yield additional power to detect LD | Assumptions of some methods are limiting or unrealistic; not all methods take adequate account of covariance among markers (see text) |
| Least-squares multilocus method (see text) | Takes covariance structure of marker loci into account; accommodates multiple founder mutations | Does not explicitly allow for disease mutations at multiple sites |
| Haplotype segment sharing methods | Takes covariance structure of marker loci into account | Some methods are compromised by allelic heterogeneity |

## Multilocus Disequilibrium Methods

As in traditional linkage analysis, the incorporation of information from multiple loci can enhance the power and accuracy of LD mapping. An early and popular method for using multilocus data was devised by Terwilliger (1995). This method uses a LD measure, λ, that is very similar to the δ statistic discussed above (Devlin and Risch 1995). The λ values for each marker and the disease locus are used to form marginal log-likelihoods, which are then summed to yield a multipoint test for LD. Another multipoint method has been formulated by Devlin et al. (1996). This composite likelihood approach attempts to take evolutionary variance (i.e., drift effects) into account and is thus potentially more realistic than Terwilliger's method. Xiong and Guo (1997) propose an elegant multipoint method in which a Taylor series expansion is used to approximate the likelihood. The advantages of this method are that multiple mutations at marker and disease loci can be accommodated, as well as variable models of population growth. However, this method, like those of Terwilliger and Devlin et al., does not account for covariance among markers (McPeek and Strahs 1999), and the Taylor series approximation may sometimes be inadequate (Rannala and Slatkin 1998). Another composite likelihood method is based on the Malécot isolation-by-distance model (Collins and Morton 1998). This method, which is derived from well-developed theory, accommodates multiple founder mutations and easily allows the pooling of heterogeneous data from multiple studies. Application of this method to various data sets (Collins and Morton 1998; Lonjou et al. 1998) indicates a relatively high level of resolving power.

Lazzeroni (1998) has formulated a least-squares approach in which piecewise nonlinear regression is used to fit a curve to the pattern of δ values for diallelic polymorphisms in the region containing a disease-causing locus. This curve predicts the most likely location of the locus. A bootstrap approach is used to estimate the sampling distribution of δ so that the covariance structure of the markers is taken into account. This method can be extended to accommodate multiple founder mutations and locus heterogeneity.

A related form of multilocus disequilibrium mapping involves the statistical analysis of haplotype regions shared in affected cases. Because these approaches are based on haplotypes rather than single marker loci, the relationships among groups of markers are necessarily taken into account. Service et al. (1999) propose a likelihood method that compares the distribution of haplotypes among cases with the distribution expected if all affected individuals are descended from a common ancestor who carried a disease-causing mutation. Computer simulations indicate that this method has greater power to detect disease-causing mutations under conditions of moderate heterogeneity than do methods such as those of Terwilliger; thus, this approach may be more useful in detecting loci underlying complex disease susceptibility. Another haplotype-sharing approach (McPeek and Strahs 1999) uses a coalescent model to account for the effects of population structure on covariance between marker loci. The ancestral haplotype that contains a disease-causing mutation is inferred, and a likelihood curve provides the estimated location of the disease locus and associated confidence limits. Yet another method (Lam et al. 2000) uses maximum parsimony to build an evolutionary phylogeny of disease haplotypes. Mutation and recombination probabilities are incorporated into the model, and various possible locations of a disease-causing mutation are evaluated by comparing likelihoods of evolutionary trees.

## Strategies for Mapping Complex Disease Genes

Having reviewed many of the statistical techniques that can be used to estimate LD, we turn now to several issues relevant to the design and execution of LD-based mapping studies. Because of the current focus on mapping complex disease genes, special attention is focused on this area.

### Statistical Power and Efficiency

The power and efficiency of LD statistics are affected by the methods used, the number of available samples, mode of inheritance, the patterns of recombination and mutation in a region, the age of the mutation(s), the degree of locus and allelic heterogeneity, the type of markers assayed, and many aspects of population history (Chapman and Wijsman 1998; Kaplan et al. 1997; Long and Langley 1999; Morton and Collins 1998; Page and Amos 1999; Risch and Merikangas 1996; Schaid 1998; Teng and Risch 1999; Xiong and Guo 1998; Xiong and Jin 1999; Zöllner and von Haeseler 2000). Because of this complexity, it is unlikely that a single technique or approach will provide optimal power under all circumstances. Power to detect LD tends to be greatest when a single disease-causing mutation that accounts for a large proportion of the phenotypic variance of a trait has arisen recently on a relatively uncommon haplotype background. These conditions promote large differences in marker allele frequencies in mutation carriers versus noncarriers. Locus and allelic heterogeneity, which are common in complex diseases, can produce dramatic decreases in power (Xiong and Guo 1998). Mutations responsible for complex diseases will often persist for long periods of time because they are typically less subject to the effects of natural selection than are mutations responsible for Mendelian diseases (Terwilliger and Weiss

**Table 2.**  A Compilation of Some Readily Available Software for Linkage Disequilibrium Analysis

| Program name | Description | Web Address | Reference |
|---|---|---|---|
| ALLASS | Estimates composite linkage disequilibrium for multilocus data using the Malécot isolation by distance equation | http://cedar.genetics.soton.ac.uk/pub/PROGRAM/ALLASS | (Collins and Morton 1998) |
| ARLEQUIN | Population genetic analysis package that includes haplotype estimation by the EM algorithm and LD analysis for locus pairs; significance tested by permutation method | http://anthro.unige.ch/arlequin/ | (Schneider et al. 2000; Slatkin and Excoffier 1996) |
| DISEQ | Multilocus disequilibrium estimation program | ftp://linkage.cpmc.columbia.edu/software/diseq | (Terwilliger 1995) |
| DMAP | Composite likelihood estimation for multilocus data | http://lib.stat.cmu.edu/~bdevlin/ | (Devlin et al. 1996) |
| ETDT | Uses logistic regression approach to perform TDT for multiallelic markers | http://www.gene.ucl.ac.uk/dcurtis/software.html | (Sham and Curtis 1995) |
| FINEMAP | Estimates evolutionary trees for multilocus disease and normal haplotypes to infer disease gene's location | http://www.stat.cmu.edu/cmu-stats/ | (Lam et al. 2000) |
| GASSOC | Performs various association tests, including TDT for multiple markers | http://www.mayo.edu/statgen | (Schaid 1996) |
| GDA | Population genetic analysis package that includes estimation of LD for pairs of loci; significance tested by permutation method | http://alleyn.eeb.uconn.edu/gda/ | (Weir 1996) |
| QTDT | Performs association tests and TDT for quantitative traits using a variance components approach | http://www.well.ox.ac.uk/asthma/QTDT | (Abecasis et al. 2000) |
| TDT/S-TDT | Performs TDT and sib-TDT | http://spielman07.med.upenn.edu/TDT.htm | (Spielman and Ewens 1996, 1998) |
| TRIMHAP | Shared haplotype analysis for estimation of disease gene location | http://www.vipbg.vcu.edu/trimhap | (MacLean et al. 2000) |

1998). This can further diminish LD and the resultant power to detect it.

Often, judicious study design will have a greater effect on statistical power than will the choice of analytic technique (Terwilliger and Göring 2000). Power may be increased substantially, for example, by minimizing environmental variance and by attempting to maximize a genetic signal through the selection of extreme phenotypes (Long and Langley 1999). As with traditional linkage analysis, the effects of locus heterogeneity can be decreased through careful definition of the disease phenotype.

As marker density continues to increase, and as genotyping costs decrease, whole genome scans for allelic associations are becoming feasible (Kruglyak 1997; Kruglyak 1999; Risch and Merikangas 1996). However, the volume of genotyping in such studies can be enormous, particularly if the number of cases is large. Pooling the DNA of cases and of controls and then estimating marker allele frequencies in each of the pooled

samples can cut costs considerably (Arnheim et al. 1985). This approach will allow the reliable detection of allele frequency differences of ~5% (Shaw et al. 1998), but, because heterozygous genotypes cannot be identified, the TDT cannot be used directly with pooled DNA samples. Instead, case-control approaches or the haplotype relative risk method may be used, provided that ethnic stratification is not a factor (Risch and Teng 1998; Shaw et al. 1998).

## Marker Characteristics

Most often, LD studies are carried out using microsatellite polymorphisms and/or SNPs. Because of its multiple alleles, one microsatellite usually provides more information for linkage analysis than does one SNP, but the situation is more complex for LD. Here, factors such as the age of the disease-causing mutation(s), mutation rate of the marker, mode of inheritance of the disease, and recombination distance between marker and disease loci will all influence power to detect LD. For example, the higher mutation rate of microsatellites will generally cause LD to decrease more rapidly; this effect will become more significant at small genetic distances, when mutation rates and recombination rates are similar in magnitude (Xiong and Jin 1999). Thus, the elevated microsatellite mutation rate will often result in a decrease in statistical power, unless a disease-causing mutation has arisen on a chromosome that contains a newly created microsatellite allele (producing a strong association and an increase in power). As a result of differences in models and assumptions, comparisons of the two types of markers have arrived at somewhat differing conclusions (Chapman and Wijsman 1998; Xiong and Jin 1999). It is clear, in any case, that the use of haplotypes containing multiple SNPs will increase the power to detect LD (Ott and Rabinowitz 1997; Zöllner and von Haeseler 2000). However, less information, and thus less power, will be contributed if the SNPs are in strong LD with one another.

Currently, many SNPs are being ascertained on the basis of relatively high heterozygosity in a multiethnic panel of individuals (Collins et al. 1998). Although high heterozygosity will generally increase the power to detect LD, SNPs that are highly polymorphic in most major human populations are likely to be ancient. Consequently, many generations will have elapsed during which LD between these SNPs and nearby disease-causing mutations can dissipate, reducing the power to detect LD. In at least some situations, younger SNPs with lower overall heterozygosity or a more restricted distribution in populations may therefore be preferable (Collins et al. 1999).

The density of markers required for effective association-based mapping is the subject of some controversy. A recent simulation study indicated that LD between mutations underlying complex diseases and surrounding SNPs may become nonsignificant after only 3 kb (Kruglyak 1999). However, this simulation did not consider the effects of natural selection, which will limit the persistence of disease-causing mutations in populations and decrease the length of time during which LD can dissipate. In addition, realistic demographic scenarios, such as recurrent population expansions, can produce much larger regions of disequilibrium (Collins et al. 1999; Thompson and Neel 1997). Importantly, a number of empirical studies of the extent of LD have now been performed on various outbred human populations, and most of these reveal significant disequilibrium for some locus pairs separated by at least 30–50 kb and often for loci separated by considerably greater distances (Collins et al. 1999; Goddard et al. 2000; Huttley et al. 1999; Jorde et al. 1994; Jorde et al. 2000; Kidd et al. 1998; Peterson et al. 1995; Watkins et al. 1994).

These empirical studies, as well as a recent study of a 9.7kb region in the LPL gene (Nickerson et al. 1998), show that levels of LD are quite variable in small genomic regions, ranging from highly significant to nonsignificant. In this context, it is important to keep in mind that the sampling variance of LD statistics becomes large in such regions (Golding 1984; Hudson 1985). In addition, these statistics can be strongly affected by evolutionary variance (i.e., the effects of stochastic factors such as genetic drift). On this basis alone, it is expected that closely linked sites (e.g., those <30–50 kb apart) will often fail to demonstrate significant LD and that LD is unlikely to provide an accurate prediction of the physical distance between closely linked sites (Jorde 1995; Jorde et al. 1994).

## Choice of Populations

Because LD reflects the history of recombination, populations with different demographic histories will often display different LD patterns. Although population comparisons of LD patterns are still relatively few, some generalizations can be made. In particular, most studies demonstrate higher levels of LD in recently founded populations than in "older" populations such as those in Africa (Jorde et al. 2000; Kidd et al. 1998; Kunst et al. 1996; Lonjou et al. 1999; Purandare et al. 1996; Tishkoff et al. 1996; Tishkoff et al. 1998). In recently founded groups, such as the Finnish (Peltonen 2000) or Mennonite populations (Puffenberger et al. 1994), LD may be seen for loci separated by several cM or more. These patterns have led to the suggestion that younger populations may be most useful for the initial detection of a disease locus via LD at large distances. Subsequently, older populations, in which more recombinants have accumulated, may be more useful for the fine-scale LD mapping of the disease locus. This approach assumes that the ages of disease-causing mu-

tations are correlated with the age of a population (i.e., that most major mutations arose near the time of the founding of the population). In addition, for complex diseases, it assumes that the relative effect of each susceptibility locus will be roughly similar in diverse populations, allowing one to extrapolate from an initial result in a young population to a fine-scale mapping effort in an older population.

Encouraged by the singular successes of LD-based mapping of Mendelian disorders in isolated populations (de la Chapelle and Wright 1998), many investigators are now turning to these populations in the search for loci underlying complex diseases (Peltonen 2000; Sheffield et al. 1998; Wright et al. 1999). The reasoning is simple: isolated populations typically have a simpler population history, with fewer founders and less population admixture. In effect, the ideal isolated population is a large pedigree with many, many generations. Therefore, it is expected that allelic and locus heterogeneity should be more limited, permitting easier detection of allelic associations.

This paradigm is not without its critics, however (Lonjou et al. 1999; Terwilliger and Weiss 1998). While allelic heterogeneity is often reduced for rare Mendelian diseases in isolated populations, it is unknown whether a similar reduction will be seen for loci underlying oligogenic disorders. For a relatively common disease that requires the contribution of, say, 3–5 predisposing loci, the total frequency of disease-causing variants in each oligogene would be relatively high. Only the most severe bottleneck, with a reduction to perhaps 10–100 unrelated individuals, would substantially reduce the number of disease-causing alleles at such a locus (Kruglyak 1999). Other difficulties with isolated populations are: (1) a potentially high level of background inbreeding, which decreases heterozygosity and hence the efficiency of the TDT (Morton and Collins 1998); (2) limitations on the sample size of unrelated (or, more precisely, distantly related) mutation-containing chromosomes (Bonné-Tamir et al. 1997); and (3) often, a relatively short population history, which will tend to increase the distance at which LD can be found but decrease the level of resolution of LD mapping (i.e., the length of shared haplotype segments in mutation-containing chromosomes tends to be large). Further discussion of these issues can be found elsewhere (Jorde et al. 2000). On the other hand, isolates are less likely to have experienced repeated admixture events, which can obscure disequilibrium signals. Also, they are more likely to have a relatively homogeneous environment, which should help to decrease the undesirable effects of factors such as phenocopies and reduced penetrance.

Many isolated populations have experienced rapid population growth in recent times. Indeed, genetic signatures of major population expansions are detectable in most human populations (Harpending et al. 1998). Rapid growth limits LD because genetic drift is minimal in such populations. This suggests that small, isolated populations of constant size, in which genetic drift can produce substantial LD (Slatkin 1994), may be especially effective in detecting complex disease loci (Laan and Pääbo 1997; Terwilliger et al. 1998; Zöllner and von Haeseler 2000). It remains to be seen whether many human populations meet the criteria of isolation, sufficient constancy of population size, and a sufficient number of unrelated disease haplotypes.

The ultimate test of the usefulness of isolates for mapping complex diseases will come from empirical data on the extent of locus and allelic heterogeneity in these populations. To date, such data are rare. One recent genome-wide scan for schizophrenia genes in an isolated Finnish subpopulation found evidence of multiple loci but no evidence, thus far, of LD (Hovatta et al. 1999). A population comparison of allelic diversity in the LPL locus, which can be considered a gene underlying susceptibility to a common disease, showed that allelic diversity is nearly as high in a Finnish population as in an outbred U.S. population (Nickerson et al. 1998). Clearly, additional data are needed on allelic and locus heterogeneity in isolated human populations and on patterns of LD in a variety of human populations.

## Conclusion

The past decade has witnessed a burgeoning development of methods for the analysis of LD. Investigators are no longer limited to a few frequency-dependent measures of simple LD. Several approaches now deal with nettlesome problems like population stratification. Techniques have also been developed to exploit specific attributes of some populations, such as admixture. Importantly, several new multilocus methods have been formulated, and these will become increasingly useful as more polymorphic markers accumulate.

As human geneticists move from the mapping of relatively tractable Mendelian conditions to the identification of loci underlying complex diseases, the usefulness of LD approaches remains an open question. Further methodological developments will be needed to deal effectively with the complexity underlying common disease. However, effective experimental design will be at least as important in determining the success of LD mapping for these diseases. Researchers must learn enough about the demographic history of a population to determine its usefulness for LD mapping, and they must design case-control or family studies to achieve maximum power and efficiency. The optimal choice of populations will depend on the distribution of genetic variation in populations, which in turn affects allelic and locus heterogeneity. Our understanding of this variation, and the factors that influ-

ence it, is still rudimentary and will require a thorough sampling of human genetic diversity. This information, combined with further methodological developments and well-informed study design, offer hope for the success of LD approaches in the search for complex disease genes.

## ACKNOWLEDGMENTS

## REFERENCES

Abecasis, G.R., Cardon, L.R., and Cookson, W.O. 2000. A general test of association for quantitative traits in nuclear families. *Am. J. Hum. Genet.* **66:** 279–292.

Ajioka, R.S., Jorde, L.B., Gruen, J.R., Yu, P., Dimitrova, D., Barrow, J., Radisky, E., Edwards, C.Q., Griffen, L.M., and Kushner, J.P. 1997. Haplotype analysis of hemochromatosis: Evolution of different linkage disequilibrium approaches and evolution of disease chromosomes. *Am. J. Hum. Genet.* **60:** 1439–1447.

Allison, D.B. 1997. Transmission-disequilibrium tests for quantitative traits. *Am. J. Hum. Genet.* **60:** 676–690.

Allison, D.B., Heo, M., Kaplan, N., and Martin, E.R. 1999. Sibling-based tests of linkage and association for quantitative traits. *Am. J. Hum. Genet.* **64:** 1754–1763.

Arnheim, N., Strange, C., and Erlich, H. 1985. Use of pooled DNA samples to detect linkage disequilibrium of polymorphic restriction fragments and human disease: studies of the HLA class II loci. *Proc. Natl. Acad. Sci.* **82:** 6970–6974.

Bengtsson, B.O., and Thomson, G. 1981. Measuring the strength of associations between HLA antigens and diseases. *Tissue Antigens* **18:** 356–363.

Boehnke, M., and Langefeld, C.D. 1998. Genetic association mapping based on discordant sib pairs: The discordant-alleles test. *Am. J. Hum. Genet.* **62:** 950–961.

Bonné-Tamir, B., Nystuen, A., Seroussi, E., Kalinsky, H., Kwitek-Black, A.E., Korostishevsky, M., Adato, A., and Sheffield, V.C. 1997. Usher syndrome in the Samaritans: Strengths and limitations of using inbred isolated populations to identify genes causing recessive disorders. *Am. J. Phys. Anthropol.* **104:** 193–200.

Chakraborty, R., and Weiss, K.M. 1988. Admixture as a tool for finding linked genes and detecting that difference from allelic association between loci. *Proc. Natl. Acad. Sci.* **85:** 9119–9123.

Chapman, N.H., and Wijsman, E.M. 1998. Genome screens using linkage disequilibrium tests: Optimal marker characteristics and feasibility. *Am. J. Hum. Genet.* **63:** 1872–1885.

Collins, A., Lonjou, C., and Morton, N.E. 1999. Genetic epidemiology of single-nucleotide polymorphisms. *Proc. Natl. Acad. Sci.* **96:** 15173–15177.

Collins, A., and Morton, N.E. 1998. Mapping a disease locus by allelic association. *Proc. Natl. Acad. Sci.* **95:** 1741–1745.

Collins, F.S., Brooks, L.D., and Chakravarti, A. 1998. A DNA polymorphism discovery resource for research on human genetic variation. *Genome Res.* **8:** 1229–1231.

de la Chapelle, A., and Wright, F.A. 1998. Linkage disequilibrium mapping in isolated populations: The example of Finland revisited. *Proc. Natl. Acad. Sci.* **95:** 12416–12423.

Dean, M., Stephens, J.C., Winkler, C., Lomb, D.A., Ramsburg, M., Boaze, R., Stewart, C., Charbonneau, L., Goldman, D., Albaugh, B.J., et al. 1994. Polymorphic admixture typing in human ethnic populations. *Am. J. Hum. Genet.* **55:** 788–808.

Devlin, B., and Risch, N. 1995. A comparison of linkage disequilibrium measures for fine-scale mapping. *Genomics* **29:** 311–322.

Devlin, B., Risch, N., and Roeder, K. 1996. Disequilibrium mapping: Composite likelihood for pairwise disequilibrium. *Genomics*

**36:** 1–16.

Falk, C.T., and Rubinstein, P. 1987. Haplotype relative risks: An easy reliable way to construct a proper control sample for risk calculations. *Am. J. Hum. Genet.* **51:** 227–233.

Feder, J.N., Gnirke, A., Thomas, W., Tsuchihashi, Z., Ruddy, D.A., Basava, A., Dormishian, F., Domingo Jr., R., Ellis, M.C., Fullan, A., et al. 1996. A novel MHC class I-like gene is mutated in patients with hereditary haemochromatosis. *Nat. Genet.* **13:** 399–408.

George, V., Tiwari, H.K., Zhu, X., and Elston, R.C. 1999. A test of transmission/disequilibrium for quantitative traits in pedigree data, by multiple regression. *Am. J. Hum. Genet.* **65:** 236–245.

Goddard, K.A., Hopkins, P.J., Hall, J.M., and Witte, J.S. 2000. Linkage disequilibrium and allele-frequency distributions for 114 single-nucleotide polymorphisms in five populations. *Am. J. Hum. Genet.* **66:** 216–234.

Golding, G.B. 1984. The sampling distribution of linkage disequilibrium. *Genetics* **108:** 257–274.

Graham, J., and Thompson, E.A. 1998. Disequilibrium likelihoods for fine-scale mapping of a rare allele. *Am. J. Hum. Genet.* **63:** 1517–1530.

Guo, S.W. 1997. Linkage disequilibrium measures for fine-scale mapping: A comparison. *Hum. Hered.* **47:** 301–314.

Harpending, H.C., Batzer, M.A., Gurven, M., Jorde, L.B., Rogers, A.R., and Sherry, S.T. 1998. Genetic traces of ancient demography. *Proc. Natl. Acad. Sci.* **95:** 1961–1967.

Hästbacka, J., de la Chapelle, A., Kaitila, I., Sistonen, P., Weaver, A., and Lander, E. 1992. Linkage disequilibrium mapping in isolated founder populations: Diastrophic dysplasia in Finland. *Nat. Genet.* **2:** 204–211.

Hästbacka, J., de la Chapelle, A., Mahtani, M.M., Clines, G., Reeve-Daly, M.P., Daly, M., Hamilton, B.A., Kusumi, K., Trivedi, B., Weaver, A., et al. 1994. The diastrophic dysplasia gene encodes a novel sulfate transporter: Positional cloning by fine-structure linkage disequilibrium mapping. *Cell* **78:** 1073–1087.

Hill, W.G., and Robertson, A. 1968. Linkage disequilibrium in finite populations. *Theor. Appl. Genet.* **38:** 226–231.

Hovatta, I., Varilo, T., Suvisaari, J., Terwilliger, J.D., Ollikainen, V., Arajärvi, R., Juvonen, H., Kokko-Sahin, M.L., Vaisanen, L., Mannila, H., et al. 1999. A genomewide screen for schizophrenia genes in an isolated Finnish subpopulation, suggesting multiple susceptibility loci. *Am. J. Hum. Genet.* **65:** 1114–1124.

Hudson, R.R. 1985. The sampling distribution of linkage disequilibrium under an infinite allele model without selection. *Genetics* **109:** 611–631.

Huttley, G.A., Smith, M.W., Carrington, M., and O'Brien, S.J. 1999. A scan for linkage disequilibrium across the human genome. *Genetics* **152:** 1711–1722.

Jennings, H.S. 1917. The numerical results of diverse systems of breeding, with respect to two pairs of characters, linked or independent, with special relation to the effects of linkage. *Genetics* **2:** 97–154.

Jorde, L.B. 1995. Linkage disequilibrium as a gene mapping tool. *Am. J. Hum. Genet.* **56:** 11–14.

Jorde, L.B., Watkins, W.S., Carlson, M., Groden, J., Albertsen, H., Thliveris, A., and Leppert, M. 1994. Linkage disequilibrium predicts physical distance in the adenomatous polyposis coli region. *Am. J. Hum. Genet.* **54:** 884–898.

Jorde, L.B., Watkins, W.S., Kere, J., Nyman, D., and Eriksson, A.W. 2000. Gene mapping in isolated populations: New roles for old friends? *Hum. Hered.* **50:** 57–65.

Kaplan, N.L., Hill, W.G., and Weir, B.S. 1995. Likelihood methods for locating disease genes in nonequilibrium populations. *Am. J. Hum. Genet.* **56:** 18–32.

Kaplan, N.L., Martin, E.R., Morris, R.W., and Weir, B.S. 1998. Marker selection for the transmission/disequilibrium test, in recently admixed populations. *Am. J. Hum. Genet.* **62:** 703–712.

Kaplan, N.L., Martin, E.R., and Weir, B.S. 1997. Power studies for the transmission/disequilibrium tests with multiple alleles. *Am. J. Hum. Genet.* **60:** 691–702.

Kerem, B., Rommens, J.M., Buchanan, J.A., Markiewicz, D., Cox, T.K., Chakravarti, A., Buchwald, M., and Tsui, L. 1989. Identification of the cystic fibrosis gene: Genetic analysis. *Science* **245:** 1073–1080.

Kidd, K.K., Morar, B., Castigione, C.M., Zhao, H., Pakstis, A.J., Speed, W.C., Bonné-Tamir, B., Lu, R.B., Goldman, D., Lee, C., et al. 1998. A global survey of haplotype frequencies and linkage disequilibrium at the DRD2 locus. *Hum. Genet.* **103:** 211–227.

Knowler, W.C., Williams, R.C., Pettitt, D., and Steinberg, A.G. 1988. Gm3;5,13,14 and type 2 diabetes mellitus: An association in American Indians with genetic admixture. *Am. J. Hum. Genet.* **43:** 520–526.

Kruglyak, L. 1997. What is significant in whole-genome linkage disequilibrium studies? *Am. J. Hum. Genet.* **61:** 810–812.

———. 1999. Prospects for whole-genome linkage disequilibrium mapping of common disease genes. *Nat. Genet.* **22:** 139–144.

Kunst, C.G., Zerylnick, C., Karickhoff, L., Eichler, E., Bullard, J., Chalifoux, M., Holden, J.J., Torroni, A., Nelson, D.L., and Warren, S.T. 1996. FMR1 in global populations. *Am. J. Hum. Genet.* **58:** 513–522.

Laan, M., and Pääbo, S. 1997. Demographic history and linkage disequilibrium in human populations. *Nat. Genet.* **17:** 435–438.

Lam, J.C., Roeder, K., and Devlin, B. 2000. Haplotype fine mapping by evolutionary trees. *Am. J. Hum. Genet.* **66:** 659–673.

Lazzeroni, L.C. 1998. Linkage disequilibrium and gene mapping: An empirical least-squares approach. *Am. J. Hum. Genet.* **62:** 159–170.

Lewontin, R.C. 1964. The interaction of selection and linkage. I. General considerations; heterotic models. *Genetics* **49:** 49–67.

Long, A.D., and Langley, C.H. 1999. The power of association studies to detect the contribution of candidate genetic loci to variation in complex traits. *Genome Res.* **9:** 720–731.

Lonjou, C., Collins, A., Ajioka, R.S., Jorde, L.B., Kushner, J.P., and Morton, N.E. 1998. Allelic association under map error and recombinational heterogeneity: A tale of two sites. *Proc. Natl. Acad. Sci.* **95:** 11366–11370.

Lonjou, C., Collins, A., and Morton, N.E. 1999. Allelic association between marker loci. *Proc. Natl. Acad. Sci.* **96:** 1621–1626.

MacLean, C.J., Martin, R.B., Sham, P.C., Wang, H., Straub, R.E., and Kendler, K.S. 2000. The trimmed-haplotype test for linkage disequilibrium. *Am. J. Hum. Genet.* **66:** 1062–1075.

McKeigue, P.M. 1997. Mapping genes underlying ethnic differences in disease risk by linkage disequilibrium in recently admixed populations. *Am. J. Hum. Genet.* **60:** 188–196.

———. 1998. Mapping genes that underlie ethnic differences in disease risk: methods for detecting linkage in admixed populations, by conditioning on parental admixture. *Am. J. Hum. Genet.* **63:** 241–251.

McPeek, M.S., and Strahs, A. 1999. Assessment of linkage disequilibrium by the decay of haplotype sharing, with application to fine-scale genetic mapping. *Am. J. Hum. Genet.* **65:** 858–875.

Morton, N.E., and Collins, A. 1998. Tests and estimates of allelic association in complex inheritance. *Proc. Natl. Acad. Sci.* **95:** 11389–11393.

National Research Council. 1996. The Evaluation of Forensic DNA Evidence. National Academy Press, Washington, DC.

Nei, M., and Li, W.H. 1973. Linkage disequilibrium in subdivided populations. *Genetics* **75:** 213–219.

Nickerson, D.A., Taylor, S.L., Weiss, K.M., Clark, A.G., Hutchinson, R.G., Stengard, J., Salomaa, V., Vartiainen, E., Boerwinkle, E., and Sing, C.F. 1998. DNA sequence diversity in a 9.7-kb region of the human lipoprotein lipase gene. *Nat. Genet.* **19:** 233–240.

Ott, J., and Rabinowitz, D. 1997. The effect of marker heterozygosity on the power to detect linkage disequilibrium. *Genetics* **147:** 927–930.

Page, G.P., and Amos, C.I. 1999. Comparison of linkage-disequilibrium methods for localization of genes influencing quantitative traits in humans. *Am. J. Hum. Genet.* **64:** 1194–1205.

Peltonen, L. 2000. Positional cloning of disease genes: Advantages of genetic isolates. *Hum. Hered.* **50:** 66–75.

Peterson, A.C., Di Rienzo, A., Lehesjoki, A.E., de la Chapelle, A., Slatkin, M., and Freimer, N.B. 1995. The distribution of linkage disequilibrium over anonymous genome regions. *Hum. Mol. Genet.* **4:** 887–894.

Pritchard, J.K., and Rosenberg, N.A. 1999. Use of unlinked genetic markers to detect population stratification in association studies. *Am. J. Hum. Genet.* **65:** 220–228.

Puffenberger, E.G., Kauffman, E.R., Bolk, S., Matise, T.C., Washington, S.S., Angrist, M., Weissenbach, J., et al. 1994. Identity-by-descent and association mapping of a recessive gene for Hirschsprung disease on human chromosome 13q22. *Hum. Mol. Genet.* **8:** 1217–1225.

Purandare, S.M., Cawthon, R., Nelson, L.M., Sawada, S., Watkins, W.S., Ward, K., Jorde, L.B., and Viskochil, D.H. 1996. Genotyping of PCR-based polymorphisms and linkage disequilibrium analysis at the NF1 locus. *Am. J. Hum. Genet.* **59:** 159–166.

Rannala, B., and Slatkin, M. 1998. Likelihood analysis of disequilibrium mapping, and related problems. *Am. J. Hum. Genet.* **62:** 459–473.

Reich, T., Hinrichs, A., Culverhouse, R., and Bierut, L. 1999. Genetic studies of alcoholism and substance dependence. *Am. J. Hum. Genet.* **65:** 599–605.

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

Risch, N., and Teng, J. 1998. The relative power of family-based and case-control designs for linkage disequilibrium studies of complex human diseases I. DNA pooling. *Genome Res.* **8:** 1273–1288.

Schaid, D.J. 1996. General score tests for associations of genetic markers with disease using cases and their parents. *Genet. Epidemiol.* **13:** 423–449.

———. 1998. Transmission disequilibrium, family controls, and great expectations. *Am. J. Hum. Genet.* **63:** 935–941.

Schneider, S., Kueffer, J.M., Roesslie, D., and Excoffier, L. 2000. Arlequin ver. 2.0: A software for population genetic data analysis. Genetics and Biometry Laboratory, University of Geneva, Geneva.

Service, S.K., Lang, D.W., Freimer, N.B., and Sandkuijl, L.A. 1999. Linkage-disequilibrium mapping of disease genes by reconstruction of ancestral haplotypes in founder populations. *Am. J. Hum. Genet.* **64:** 1728–1738.

Sham, P.C., and Curtis, D. 1995. An extended transmission/disequilibrium test (TDT) for multi-allele marker loci. *Ann. Hum. Genet.* **59:** 323–336.

Shaw, S.H., Carrasquillo, M.M., Kashuk, C., Puffenberger, E.G., and Chakravarti, A. 1998. Allele frequency distributions in pooled DNA samples: Applications to mapping complex disease genes. *Genome Res.* **8:** 111–123.

Sheffield, V.C., Stone, E.M., and Carmi, R. 1998. Use of isolated inbred human populations for identification of disease genes. *Trends Genet.* **14:** 391–396.

Slatkin, M. 1994. Linkage disequilibrium in growing and stable populations. *Genetics* **137:** 331–336.

Slatkin, M., and Excoffier, L. 1996. Testing for linkage disequilibrium in genotypic data using the expectation-maximization algorithm. *Heredity* **76:** 377–383.

Spielman, R.S., and Ewens, W.J. 1996. The TDT and other family-based tests for linkage disequilibrium and association. *Am. J. Hum. Genet.* **59:** 983–989.

———. 1998. A sibship test for linkage in the presence of association: The sib transmission/disequilibrium test. *Am. J. Hum. Genet.* **62:** 450–458.

Spielman, R.S., McGinnis, R.E., and Ewens, W.J. 1993. Transmission test for linkage disequilibrium: The insulin gene region and insulin-dependent diabetes mellitus (IDDM). *Am. J. Hum. Genet.* **52:** 506–516.

Stephens, J.C., Briscoe, D., and O'Brien, S.J. 1994. Mapping by admixture linkage disequilibrium in human populations: Limits and guidelines. *Am. J. Hum. Genet.* **55:** 809–824.

Sun, F., Flanders, W.D., Yang, Q., and Khoury, M.J. 1999. Transmission disequilibrium test (TDT) when only one parent is available: The 1-TDT. *Am. J. Epidemiol.* **150:** 97–104.

Teng, J., and Risch, N. 1999. The relative power of family-based and case-control designs for linkage disequilibrium studies of complex human diseases. II. Individual genotyping. *Genome Res.* **9:** 234–241.

Terwilliger, J.D. 1995. A powerful likelihood method for the analysis of linkage disequilibrium between trait loci and one or more polymorphic marker loci. *Am. J. Hum. Genet.* **56:** 777–787.

Terwilliger, J.D., and Göring, H.H. 2000. Gene mapping in the 20th and 21st centuries: Statistical methods, data analysis, and experimental design. *Hum. Biol.* **72:** 63–132.

Terwilliger, J.D., and Weiss, K.M. 1998. Linkage disequilibrium mapping of complex disease: fantasy or reality? *Curr. Opin. Biotechnol.* **9:** 578–594.

Terwilliger, J.D., Zöllner, S., Laan, M., and Pääbo, S. 1998. Mapping genes through the use of linkage disequilibrium generated by genetic drift: "Drift mapping" in small populations with no demographic expansion. *Hum. Hered.* **48:** 138–154.

Thompson, E.A., and Neel, J.V. 1997. Allelic disequilibrium and allele frequency distribution as a function of social and demographic history. *Am. J. Hum. Genet.* **60:** 197–204.

Thomson, G. 1995. Mapping disease genes: Family-based association studies. *Am. J. Hum. Genet.* **57:** 487–498.

Tishkoff, S.A., Dietzsch, E., Speed, W., Pakstis, A.J., Kidd, J.R., Cheung, K., Bonné-Tamir, B., Santachiara-Benerecetti, A.S., Moral, P., Krings, M., et al. 1996. Global patterns of linkage disequilibrium at the CD4 locus and modern human origins. *Science* **271:** 1380–1387.

Tishkoff, S.A., Goldman, A., Calafell, F., Speed, W.C., Deinard, A.S., Bonné-Tamir, B., Kidd, J.R., Pakstis, A.J., Jenkins, T., and Kidd, K.K. 1998. A global haplotype analysis of the myotonic dystrophy locus: implications for the evolution of modern humans and for the origin of myotonic dystrophy mutations. *Am. J. Hum. Genet.* **62:** 1389–1402.

Watkins, W.S., Zenger, R., O'Brien, E., Nyman, D., Eriksson, E., Renlund, M., and Jorde, L.B. 1994. Linkage disequilibrium patterns vary with chromosomal location: A case study from the von Willebrand factor region. *Am. J. Hum. Genet.* **55:** 348–355.

Weinberg, C.R. 1999. Allowing for missing parents in genetic studies of case-parent triads. *Am. J. Hum. Genet.* **64:** 1186–1193.

Weir, B.S. 1996. Genetic data analysis II. Sinauer, Sunderland, Mass.

Wilson, S.R. 1997. On extending the transmission/disequilibrium test (TDT). *Ann. Hum. Genet.* **61:** 151–161.

Wright, A.F., Carothers, A.D., and Pirastu, M. 1999. Population choice in mapping genes for complex diseases. *Nat. Genet.* **23:** 397–404.

Xiong, M., and Guo, S.W. 1997. Fine-scale genetic mapping based on linkage disequilibrium: theory and applications. *Am. J. Hum. Genet.* **60:** 1513–1531.

———. 1998. The power of linkage detection by the transmission/disequilibrium tests. *Hum. Hered.* **48:** 295–312.

Xiong, M., and Jin, L. 1999. Comparison of the power and accuracy of biallelic and microsatellite markers in population-based gene-mapping methods. *Am. J. Hum. Genet.* **64:** 629–640.

Xiong, M.M., Krushkal, J., and Boerwinkle, E. 1998. TDT statistics for mapping quantitative trait loci. *Ann. Hum. Genet.* **62:** 431–452.

Zaykin, D., Zhivotovsky, L., and Weir, B.S. 1995. Exact tests for association between alleles at arbitrary numbers of loci. *Genetica* **96:** 169–178.

Zöllner, S., and von Haeseler, A. 2000. A coalescent approach to study linkage disequilibrium between single-nucleotide polymorphisms. *Am. J. Hum. Genet.* **66:** 615–628.