Am. J. Hum. Genet. 68:1447–1456, 2001

The Effect That Genotyping Errors Have on the Robustness of Common Linkage-Disequilibrium Measures

Joshua M. Akey,[1,*] Kun Zhang,[1,*] Momiao Xiong,[1] Peter Doris,[2] and Li Jin[1]

[1]Human Genetics Center, School of Public Health, and [2]Institute of Molecular Medicine, University of Texas–Houston, Houston

The rapid development of a dense single-nucleotide–polymorphism marker map has stimulated numerous studies attempting to characterize the magnitude and distribution of background linkage disequilibrium (LD) within and between human populations. Although genotyping errors are an inherent problem in all LD studies, there have been few systematic investigations documenting their consequences on estimates of background LD. Therefore, we derived simple deterministic formulas to investigate the effect that genotyping errors have on four commonly used LD measures—*D′*, *r*, *Q*, and *d*—in studies of background LD. We have found that genotyping error rates as small as 3% can have serious affects on these LD measures, depending on the allele frequencies and the assumed error model. Furthermore, we compared the robustness of *D′*, *r*, *Q*, and *d*, in the presence of genotyping errors. In general, *Q* and *d* are more robust than *D′* and *r*, although exceptions do exist. Finally, through stochastic simulations, we illustrate how genotyping errors can lead to erroneous inferences when measures of LD between two samples are compared.

Introduction

Linkage disequilibrium (LD) is becoming an important tool in genetic studies because it is applicable to a wide variety of topics, including disease-gene mapping (Collins et al. 1997; Akey et al. 2001*a*), delineation of the demographic history of populations (Laan and Paabo 1997), and testing of hypotheses of human evolution (Tishkoff et al. 1996). However, the full utility of LD-based applications is currently limited, because relatively little is known about this complex population genetic phenomenon. To this end, a number of recent studies have attempted to characterize the magnitude and distribution of "background" LD (i.e., LD between anonymous genetic markers) throughout the genome both within and between human populations (Goddard et al. 2000; Gordon et al. 2000; Jorde et al. 2000; Kidd et al. 2000; Moffatt et al. 2000; Taillon-Miller et al. 2000; Zavattari et al. 2000; Abecasis et al. 2001). In the context of LD mapping, these studies are very important, because they provide the framework to address questions such as which populations are most suitable for LD mapping and what marker density will be required to map genes underlying complex diseases.

Received February 19, 2001; accepted for publication April 5, 2001; electronically published May 16, 2001.

Address for correspondence and reprints: Mr. Joshua Akey, Graduate School of Public Health, University of Texas–Houston, 1200 Herman Pressler, Houston, TX 77030. E-mail: jakey@gsbs3.gs.uth.tmc.edu

* The first two authors contributed equally to this work.

0002-9297/2001/6806-0016$02.00

Ideally, a measure of LD between marker loci is only a function of factors such as population history, age of the variants, and local rates of recombination. However, other nonbiological forces may have an impact on estimates of LD; for example, genotyping errors can occur in every LD study. Although the ramifications of genotyping errors have been investigated in the context of disease-gene mapping (Gordon et al. 1999; Goring and Terwilliger 2000), there have been no systematic studies documenting its consequences on estimates of background LD. In fact, it is generally not understood for which error rates, if any, various LD measures are robust to genotyping errors. Therefore, the purpose of the present article is to (1) demonstrate how genotyping errors affect four commonly used LD measures and (2) compare the robustness of these measures in the presence of genotyping errors.

LD Measures

In the present study, we restrict our analysis to LD between two single-nucleotide polymorphisms (SNPs), denoted as "locus 1" and "locus 2," which can be arranged into a 2 × 2 table, as shown in table 1. In table 1, loci 1 and 2 each have two alleles, denoted as "A" and "a" and "B" and "b," respectively. The frequencies of alleles A, a, B, and b are given by P_A, P_a, P_B, and P_b, and the haplotype frequencies of AB, Ab, aB, and ab gametes are given by P_{AB}, P_{Ab}, P_{aB}, and P_{ab}, respectively.

Numerous statistics have been proposed to measure the degree of LD between two biallelic markers (Hedrick 1987; Devlin and Risch 1995; Xiong and Guo

1997). A measure that is fundamental to many LD measures is the coefficient of LD, or "D" (Lewontin and Kojima 1960): $D = P_{AB} - P_A P_B = P_{AB}P_{ab} - P_{Ab}P_{aB}$. We consider four LD measures that are commonly used to estimate background LD: D' (Lewontin 1964), r (Hill and Robertson 1968), d (Nei and Li 1980), and Q (Yule 1900). The formulas for these measures are $D' = D/D_{max}$, $r = D/\sqrt{P_A P_a P_B P_b}$, $Q = D/(P_{AB}P_{ab} + P_{Ab}P_{aB})$, $d = D/P_B P_b$, where D_{max} is defined as $\min\{P_A P_b, P_a P_B\}$ if $D > 0$, and $D_{max} = \min\{P_a P_b, P_A P_B\}$ if $D < 0$. Note that the numerators of these measures are all D, and the formulas differ only in their denominators. For further information about the relationship between D', r, d, and Q and for alternative formulations, see the work of Devlin and Risch (1995).

Genotyping-Error Model

There are several potential sources of error in experimental studies that estimate LD between markers, including genotyping errors, haplotyping errors, and human errors (e.g., entering the wrong allele into a database). In the present study, we focused only on genotyping errors. We consider two models for genotyping errors: a stochastic-error model (SEM) and a directed-error model (DEM). As figure 1 illustrates, the SEM postulates that there is an equal probability for alleles at a locus to be erroneously genotyped, whereas the DEM postulates that there is a greater probability for one allele to be consistently misgenotyped. Our motivation for contemplating these two models is that different SNP-genotyping methods may be better characterized by the SEM or the DEM. For example, classic PCR-RFLP genotyping is prone to partial digestion (Wu et al. 2000), in which one allele is systematically misgenotyped; hence, the DEM may be more appropriate. Alternatively, the SEM may better describe errors that occur in genotyping methods that rely on hybridization for discriminating SNP alleles, such as *Taq*man (Livak et al. 1995) and oligonucleotide arrays (Halushka et al. 1999). In the results presented in the section "LD Measures in the Presence of Genotyping Errors," we assume, without loss of generality, that the genotyping-error rates at loci 1 and 2 (μ and ν, respectively) are equal (i.e., $\mu = \nu$).

Table 1

Layout and Notation for Haplotype Frequencies of SNP Loci 1 and 2

	Locus 2		
Locus 1	B	b	Overall
A	P_{AB}	P_{Ab}	P_A
a	P_{aB}	P_{ab}	P_a
Overall	P_B	P_b	1

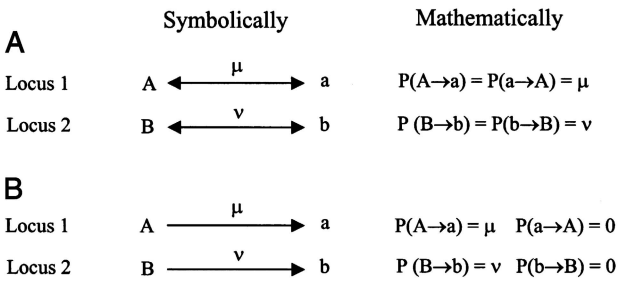

Figure 1 Symbolic and mathematical representations of error models used in analyses. The genotyping-error rates at loci 1 and 2 are denoted as μ and ν, respectively. Panels A and B correspond to the SEM and the DEM, respectively.

LD Measures in the Presence of Genotyping Errors

Deterministic Calculations

In this section, we derive simple, deterministic formulas to describe how genotyping errors affect LD measures. To accomplish this goal, it is necessary to describe how haplotype frequencies change in the presence of errors. For example, the change in haplotype frequency of P_{AB} is: $\Delta P_{AB} = P'_{AB} - P_{AB}$, where P'_{AB} and P_{AB} are the haplotype frequencies of AB gametes in the presence and the absence of errors, respectively. Table 2 provides the formulas for the change in haplotype frequencies for both the SEM and the DEM. Note that the change in haplotype frequency is a simple function both of the genotyping-error rates at loci 1 and 2 and of the "true" haplotype frequencies, in the absence of any errors. Using the formulas for the change in haplotype frequencies provided in table 2 and differentiating the formulas for the four LD measures described in the section "LD Measures" provides the changes in LD measures in the presence of genotyping errors:

$$\Delta r = r\left[\frac{\Delta D}{D} - \frac{1}{2}\left(\frac{\Delta P_A}{P_A} + \frac{\Delta P_B}{P_B} + \frac{\Delta P_a}{P_a} + \frac{\Delta P_b}{P_b}\right)\right],$$

$$\Delta Q = \frac{1}{(P_{AB}P_{ab} + P_{Ab}P_{aB})^2}[\Delta D\,(P_{AB}P_{ab} + P_{Ab}P_{aB}) - D\,(\Delta P_{AB}P_{ab} + P_{AB}\Delta P_{ab} + \Delta P_{Ab}P_{aB} + P_{Ab}\Delta P_{aB})],$$

and

$$\Delta d = \frac{1}{P_B^2 P_b^2}[\Delta D P_B P_b - D\,(\Delta P_B P_b + P_B \Delta P_b)],$$

Table 2

Change in Haplotype Frequencies in the Presence of Genotyping Errors, for Both the SEM and the DEM

Haplotype	SEM	DEM
ΔP_{AB}	$-(\mu+\nu)P_{AB}+\mu P_{aB}+\nu P_{Ab}$	$-(\mu+\nu)P_{AB}$
ΔP_{Ab}	$-(\mu+\nu)P_{Ab}+\mu P_{ab}+\nu P_{AB}$	$-\mu P_{Ab}+\nu P_{AB}$
ΔP_{aB}	$-(\mu+\nu)P_{aB}+\mu P_{AB}+\nu P_{ab}$	$\mu P_{AB}-\nu P_{aB}$
ΔP_{ab}	$-(\mu+\nu)P_{ab}+\mu P_{Ab}+\nu P_{aB}$	$\mu P_{Ab}+\nu P_{aB}$

where $\Delta D = \Delta P_{AB} - \Delta P_A P_B - P_A \Delta P_B$, $\Delta P_A = \Delta P_{AB} + \Delta P_{Ab}$, $\Delta P_a = -\Delta P_A$, $\Delta P_B = \Delta P_{AB} + \Delta P_{aB}$, and $\Delta P_b = -\Delta P_B$. Since the formula of D' involves the minimum of a certain quantity, a general solution for $\Delta D'$ is not readily obtainable. Therefore, we derived $\Delta D'$ in a piecewise fashion, as

$$\Delta D' = \begin{cases} \dfrac{1}{P_A^2 P_b^2}[\Delta D P_A P_b - D(\Delta P_A P_b + P_A \Delta P_b)] \\ \quad \text{if } D > 0 \text{ and } D_{max} = P_A P_b \\ \dfrac{1}{P_a^2 P_B^2}[\Delta D P_a P_B - D(\Delta P_a P_B + P_a \Delta P_B)] \\ \quad \text{if } D > 0 \text{ and } D_{max} = P_a P_B \\ \dfrac{1}{P_A^2 P_B^2}[\Delta D P_A P_B - D(\Delta P_A P_B + P_A \Delta P_B)] \\ \quad \text{if } D > 0 \text{ and } D_{max} = P_A P_B \\ \dfrac{1}{P_a^2 P_b^2}[\Delta D P_a P_b - D(\Delta P_a P_b + P_a \Delta P_b)] \\ \quad \text{if } D > 0 \text{ and } D_{max} = P_a P_b \end{cases}.$$

When $D = 0$, D is not differentiable.

The Effect That Genotyping Errors Have on LD Measures

Using the formulas given in the subsection "Deterministic Calculations," we have extensively explored how genotyping errors impact estimates of LD. Tables 3 and 4, for the SEM and the DEM, respectively, present values of D', r, Q, and d, in the presence and the absence of genotyping errors over a broad range of allele frequencies. Under the SEM in table 3, it is obvious that even small genotyping-error rates can have profound consequences on LD measures, particularly when the frequency of the minor SNP allele is low. For example, when $P_A = P_B = .9$ and there is complete LD ($D' = 1$) in the absence of genotyping errors, a 3% error rate reduces D', r, and d to .67, .67, and .65, respectively, whereas Q is unaffected. In fact, in the special case that $P_A = P_B$ and the true value of $Q = \pm 1$, we can show that Q is independent of genotyping errors. Moreover, as the minor-allele frequencies at SNP loci 1 and 2 increase, LD measures tend to become increasingly robust to genotyping errors.

Overall, these general observations are qualitatively similar, if we assume that genotyping errors follow the DEM (see table 4). A notable difference between these two error models is that genotyping errors tend to be less severe for a fixed error rate under the DEM compared to that under the SEM. When the minor-allele frequencies are low, the differences between the DEM and the SEM are negligible. For instance, when $P_A = P_B = .9$, and the genotyping-error rate is 3%, and there is complete LD ($D' = 1$) in the absence of genotyping errors, D' is reduced to .67 under the SEM (table 3) and .70 under the DEM (table 4). However, the differences between the SEM and the DEM become more pronounced when the minor-allele frequencies increase. Finally, it is interesting to note that, for a given set of haplotype frequencies, the values of D', r, Q, and d, in the absence of genotyping errors, can vary substantially (see table 3). For a review of conditions under which LD measures are or are not correlated, see the work of Hedrick (1987).

Comparing the Robustness of LD Measures in the Presence of Genotyping Errors

Although tables 3 and 4 are useful for providing a general overview of how genotyping errors affect the values of D', r, Q, and d, they do not allow for a direct comparison of which measure is more robust to errors. Therefore, to facilitate comparisons of how genotyping errors affect these four LD measures, the fractional error (FE) for each LD measure was calculated as $(\lambda_T - \lambda_E)/\lambda_T$, where λ_T denotes the true value of the LD measure in the absence of errors and where λ_E denotes the value of the LD measure in the presence of genotyping errors. For example, the FE value of D' is $(D'_T - D'_E)/D'_T$. For presentation purposes, it is useful to consider the fractional true (FT) value, which is simply 1−FE. Figure 2 plots the FT values for D', Q, r, and d, as a function of the genotyping-error rate for both the SEM and the DEM.

Several interesting observations emerge from figure 2. First, for both the SEM and the DEM, genotyping errors have a substantial impact on estimates of LD, and, as expected, the higher the error rate, the smaller the FT value. Moreover, it is evident that, as described in the subsection "The Effect That Genotyping Errors Have on LD Measures," genotyping errors that follow the DEM tend to be less severe than errors that follow the SEM. Second, under certain circumstances, D', r, Q, and d do differ in their robustness to genotyping errors. For low error rates (<2%), these four LD measures do not substantially differ, regardless of the assumed error model or the allele frequencies. However, as the error rate in-

Table 3

The Effect That Genotyping Errors Have on LD Measures, under the SEM, as a Function of Allele Frequency

	LD When Genotyping-Error Rate Is															
	3.00%								5.00%							
P_B	D'_T	D'_E	r_T	r_E	Q_T	Q_E	d_T	d_E	D'_T	D'_E	r_T	r_E	Q_T	Q_E	d_T	d_E
$P_A = .90$:																
.90	1.00	.67	1.00	.67	1.00	1.00	1.00	.65	1.00	.44	1.00	.44	1.00	1.00	1.00	.42
	.50	.33	.50	.33	.92	.83	.50	.32	.50	.22	.50	.22	.92	.76	.50	.21
.50	1.00	.64	.33	.26	1.00	.71	.20	.18	1.00	.40	.33	.21	1.00	.52	.20	.16
	.50	.32	.17	.13	.54	.37	.10	.09	.50	.20	.17	.10	.54	.25	.10	.08
.10	−1.00	−.67	−1.00	−.67	−1.00	−1.00	−1.00	−.91	−1.00	−.44	−1.00	−.44	−1.00	−1.00	−1.00	−.86
	−.50	−.33	−.50	−.33	−.92	−.83	−.50	−.46	−.50	−.22	−.50	−.22	−.92	−.76	−.50	−.43
$P_A = .70$:																
.90	1.00	.66	.51	.39	1.00	.86	.78	.50	1.00	.43	.51	.31	1.00	.77	.78	.32
	.50	.33	.25	.19	.68	.52	.39	.25	.50	.21	.25	.15	.68	.42	.39	.16
.50	1.00	.84	.65	.57	1.00	.94	.60	.53	1.00	.73	.65	.51	1.00	.89	.60	.48
	.50	.42	.33	.28	.65	.57	.30	.26	.50	.37	.33	.26	.65	.52	.30	.24
.10	−1.00	−.66	−.51	−.39	−1.00	−.86	−.78	−.71	−1.00	−.43	−.51	−.31	−1.00	−.77	−.78	−.67
	−.50	−.33	−.25	−.19	−.68	−.52	−.39	−.36	−.50	−.21	−.25	−.15	−.68	−.42	−.39	−.34
$P_A = .50$:																
.90	1.00	.64	.33	.26	1.00	.71	.56	.36	1.00	.40	.33	.21	1.00	.52	.56	.23
	.50	.32	.17	.13	.54	.37	.28	.18	.50	.20	.17	.10	.54	.25	.28	.12
.50	1.00	.88	1.00	.88	1.00	1.00	1.00	.88	1.00	.80	1.00	.80	1.00	1.00	1.00	.80
	.50	.44	.50	.44	.80	.74	.50	.44	.50	.40	.50	.40	.80	.70	.50	.40
.10	−1.00	−.64	−.33	−.26	−1.00	−.71	−.56	−.51	−1.00	−.40	−.33	−.21	−1.00	−.52	−.56	−.48
	−.50	−.32	−.17	−.13	−.54	−.37	−.28	−.25	−.50	−.20	−.17	−.10	−.54	−.25	−.28	−.24

NOTE.—LD values were generated by defining P_A, P_B, and D', from which haplotype frequencies were calculated via the formulas $D = D_{max}D'$, $P_{AB} = D + P_AP_B$, $P_{Ab} = P_A - P_{AB}$, $P_{aB} = P_B - P_{AB}$, and $P_{ab} = 1 - P_{AB} - P_{aB} - P_{Ab}$. Q_T, r_T, and d_T were then calculated from these haplotype frequencies.

creases, the differences between D', r, Q, and d become more pronounced. Third, whether one measure is "superior" strongly depends on the underlying allele and haplotype frequencies and on the error model. For example, in figure 2*A*, with the same set of allele and haplotype frequencies assumed, d is more robust for the SEM, whereas Q is more robust for the DEM.

Even if one assumes the same error model, these four LD measures demonstrate different patterns of robustness, depending on the allele and haplotype frequencies. For instance, notice that under the DEM in figure 2*A* (where $P_A = .6$ and $P_B = .4$) the order of FT values is $Q > r > D' > d$, whereas in figure 2*B* (where $P_A = .5$ and $P_B = .4$) the order changes to $D' > Q > r > d$. Interestingly, for this set of haplotype frequencies, D' is independent of genotyping errors for the DEM. In fact, if

$$\Delta D' = \frac{1}{P_A^2 P_B^2}[\Delta D P_A P_B - D(\Delta P_A P_B + P_A \Delta P_B)] ,$$

D' is independent of genotyping errors when $P_AP_B = D(\Delta P_{AB}/\Delta D - 1)$, which is satisfied only if a DEM is assumed.

To more systematically investigate for which parameters an LD measure is more robust than the other measures, we computed the ratio of FT values. For example, to compare D' to r, we compute the quantity $FT_{D'}/FT_r$. Thus, if $FT_{D'}/FT_r > 1$, then r is more affected by genotyping errors; conversely, if $FT_{D'}/FT_r < 1$, then D' is more affected by genotyping errors. In other words, in the presence of genotyping errors, if the ratio of FT values is <1, then the LD measure in the denominator is more robust, whereas, if the ratio of FT values is >1, then the LD measure in the numerator is more robust. Obviously, if the ratio of FT values equals 1, then the two measures are equally affected by genotyping errors.

Figure 3 plots the ratio of FT values for all six pairwise comparisons of LD measures, as a function of P_{AB}, for both the SEM and the DEM. Obviously, there is a complex relationship between the ratio of FT values and the underlying haplotype and allele frequencies. Moreover, there are notable differences between the robustness of measures, depending on the error model that is assumed. Compared to the SEM, the DEM shows a narrower range of ratios, which also tend to remain more constant over wider ranges of allele and haplotype frequencies. Under the SEM, there are also many ranges of allele and haplotype frequencies in which the ratio of FT values between two measures is ~1. For example, in figure 3*A*, under the SEM $FT_r/FT_{D'} = 1$, for all values of P_{AB} (and, as a consequence, Q/D' and r/Q are mirror images of one another). However, for both the SEM and the DEM,

Table 4

The Effect That Genotyping Errors Have on LD Measures under the DEM, as a Function of Allele Frequency

	LD When Genotyping-Error Rate Is															
	3.00%								5.00%							
P_B	D'_T	D'_E	r_T	r_E	Q_T	Q_E	d_T	d_E	D'_T	D'_E	r_T	r_E	Q_T	Q_E	d_T	d_E
$P_A = .90$:																
.90	1.00	.70	1.00	.70	1.00	1.00	1.00	.68	1.00	.50	1.00	.50	1.00	1.00	1.00	.47
	.50	.35	.50	.35	.92	.83	.50	.34	.50	.25	.50	.25	.92	.77	.50	.24
.50	1.00	.70	.33	.27	1.00	.76	.20	.18	1.00	.50	.33	.23	1.00	.60	.20	.17
	.50	.35	.17	.14	.54	.39	.10	.09	.50	.25	.17	.12	.54	.29	.10	.09
.10	−1.00	−.67	−1.00	−.83	−1.00	−1.00	−1.00	−.94	−1.00	−.44	−1.00	−.72	−1.00	−1.00	−1.00	−.89
	−.50	−.33	−.50	−.42	−.92	−.87	−.50	−.47	−.50	−.22	−.50	−.36	−.92	−.84	−.50	−.45
$P_A = .70$:																
.90	1.00	.70	.51	.41	1.00	.88	.78	.53	1.00	.50	.51	.34	1.00	.80	.78	.37
	.50	.35	.25	.20	.68	.54	.39	.26	.50	.25	.25	.17	.68	.44	.39	.18
.50	1.00	.90	.65	.60	1.00	.96	.60	.55	1.00	.83	.65	.57	1.00	.93	.60	.51
	.50	.45	.33	.30	.65	.60	.30	.27	.50	.42	.33	.28	.65	.56	.30	.26
.10	−1.00	−1.00	−.51	−.48	−1.00	−1.00	−.78	−.73	−1.00	−1.00	−.51	−.45	−1.00	−1.00	−.78	−.70
	−.50	−.50	−.25	−.24	−.68	−.66	−.39	−.36	−.50	−.50	−.25	−.23	−.68	−.65	−.39	−.35
$P_A = .50$:																
.90	1.00	.70	.33	.27	1.00	.76	.56	.38	1.00	.50	.33	.23	1.00	.60	.56	.26
	.50	.35	.17	.14	.54	.39	.28	.19	.50	.25	.17	.12	.54	.29	.28	.13
.50	1.00	.94	1.00	.94	1.00	1.00	1.00	.91	1.00	.90	1.00	.90	1.00	1.00	1.00	.85
	.50	.47	.50	.47	.80	.77	.50	.46	.50	.45	.50	.45	.80	.75	.50	.43
.10	−1.00	−1.00	−.33	−.32	−1.00	−1.00	−.56	−.52	−1.00	−1.00	−.33	−.31	−1.00	−1.00	−.56	−.50
	−.50	−.50	−.17	−.16	−.54	−.53	−.28	−.26	−.50	−.50	−.17	−.15	−.54	−.52	−.28	−.25

Note.—LD values were generated as described in table 3.

it is not uncommon to observe ratios of FT values >2.0 (or <0.5), indicating that the two LD measures' FT values differ twofold. Overall, Q and d appear to outperform r and D', in the presence of genotyping errors, although exceptions certainly do exist, as is evident in figures 2 and 3.

The Impact That Genotyping Errors Have on Comparison of LD-Measure Difference between Samples

Thus far, we have considered only how genotyping errors affect LD estimates from a single sample. However, the goal in many studies of background LD is to compare the extent of LD between two samples. Therefore, we have also investigated how genotyping errors affect the comparison of LD between two samples. Intuitively, one may conjecture that, if the same loci were genotyped with both the same genotyping method and the same genotyping-error rate, then the comparison of LD between the two samples compared would not be compromised by errors. However, as we will show, this is not necessarily the case.

We performed extensive stochastic simulations to determine if genotyping errors can lead to erroneous conclusions, when comparing LD between two samples. The simulation approach differs somewhat from the deterministic formulas given in the previous sections in that it introduces a stochastic element in sampling the observed gamete that undergoes a genotyping error. As described above, we considered both a SEM and a DEM for genotyping errors. We simulated genotyping errors in two samples, each consisting of 100 individuals (200 gametes). The simulations were iterated 100 times, and, with each iteration, D', r, Q, and d were calculated. For presentation purposes, the iterations will be referred to as "replicates." Across all replicates, the average values of D', r, Q, and d were within 3% of that calculated on the basis of the deterministic formulas previously described for both the SEM and the DEM (data not shown), implying that our simulation results were accurate.

Table 5 provides, for the SEM and the DEM, the average and the maximum absolute value in the LD-measure differences between the two simulated samples, which both have the same haplotype-frequency distributions. Thus, in the absence of genotyping errors, the LD-measure difference equals 0. As the error rate increases, the average absolute value in the LD-measure difference between samples increases moderately, to a maximum of .0722, .0718, .0703, and .0427 for r, D', d, and Q, respectively. However, the maximum absolute value in the LD-measure differences increases more dramatically as the genotyping-error rate increases. In other words, although the average LD-measure difference between the two samples does not appear to compromise

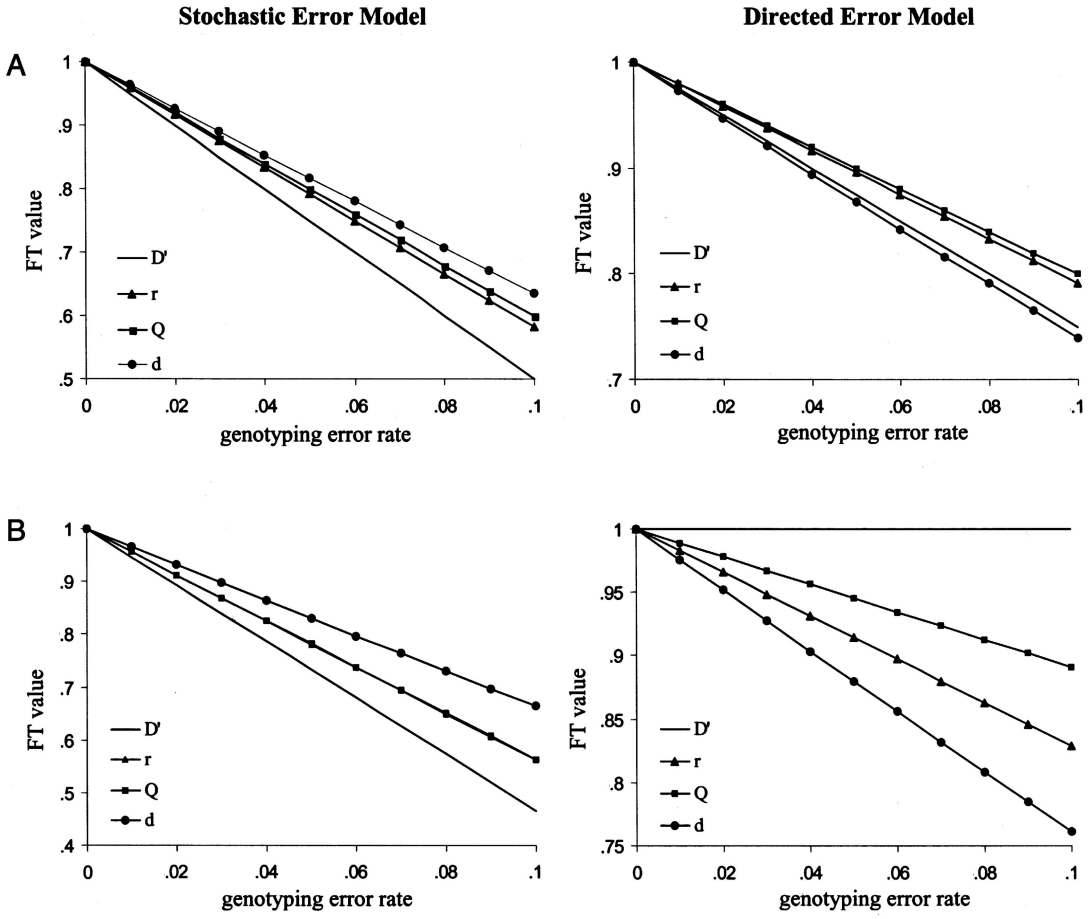

Figure 2 The effect that genotyping errors have on FT values, as a function of genotyping-error rate. In panel *A*, the haplotype frequencies are $P_{AB} = .3$, $P_{Ab} = .3$, $P_{aB} = .1$, and $P_{ab} = .3$; in panel *B*, the haplotype frequencies are $P_{AB} = .1$, $P_{Ab} = .4$, $P_{aB} = .3$, and $P_{ab} = .2$. Note that, in panel *B*, the lines corresponding to *r* and *Q* for the SEM are nearly superimposable.

the integrity of comparison of LD measures across samples, a comparison from any single experiment, in the presence of genotyping errors, may lead to erroneous inferences regarding the LD-measure difference between the two samples.

This idea is formalized in figure 4, which shows how the average absolute value in the D' difference between samples is distributed across replicates, under an SEM. In figure 4*A*, samples 1 and 2 have the same distribution of allele frequency ($P_A = P_B = .5$) and of haplotype frequency ($P_{AB} = P_{ab} = .5$, $P_{Ab} = P_{aB} = 0$), and, therefore, $|D'_1 - D'_2|$, in the absence of genotyping errors, is 0. If the genotyping-error rate is 1%, then ~80% of all replicates yield a D' difference of 0–.03. As the genotyping-error rate increases, the distribution gradually shifts away from the true difference of 0. For example, if the genotyping-error rate is 10%, then 22% of all replicates show a D' difference between samples that is >.09, and 6% of all replicates show a difference that is >.15. Figure 4*B* demonstrates how this problem is exacerbated when the allele frequencies become more extreme. Again, samples 1 and 2 have the same distribution of allele frequency ($P_A = P_B = .8$) and of haplotype frequency ($P_{AB} = .8$, $P_{ab} = .2$, $P_{Ab} = P_{aB} = 0$). With the more extreme allele frequencies, even an error rate of 1% causes a noticeable shift in the distribution of $|D'_1 - D'_2|$, away from the true value of 0, and an error rate of 10% leads to 58% of all replicates showing a D' difference between samples that is >.09. Although these examples are simplified, because they do not take into account the sampling variation in D', they illustrate how genotyping errors can make comparison of measures of LD between samples problematic.

Discussion

The development of the third-generation genetic map composed of SNPs (The International SNP Map Working Group 2001) has enabled LD to assume a prominent role in contemporary genetics research. Many studies on

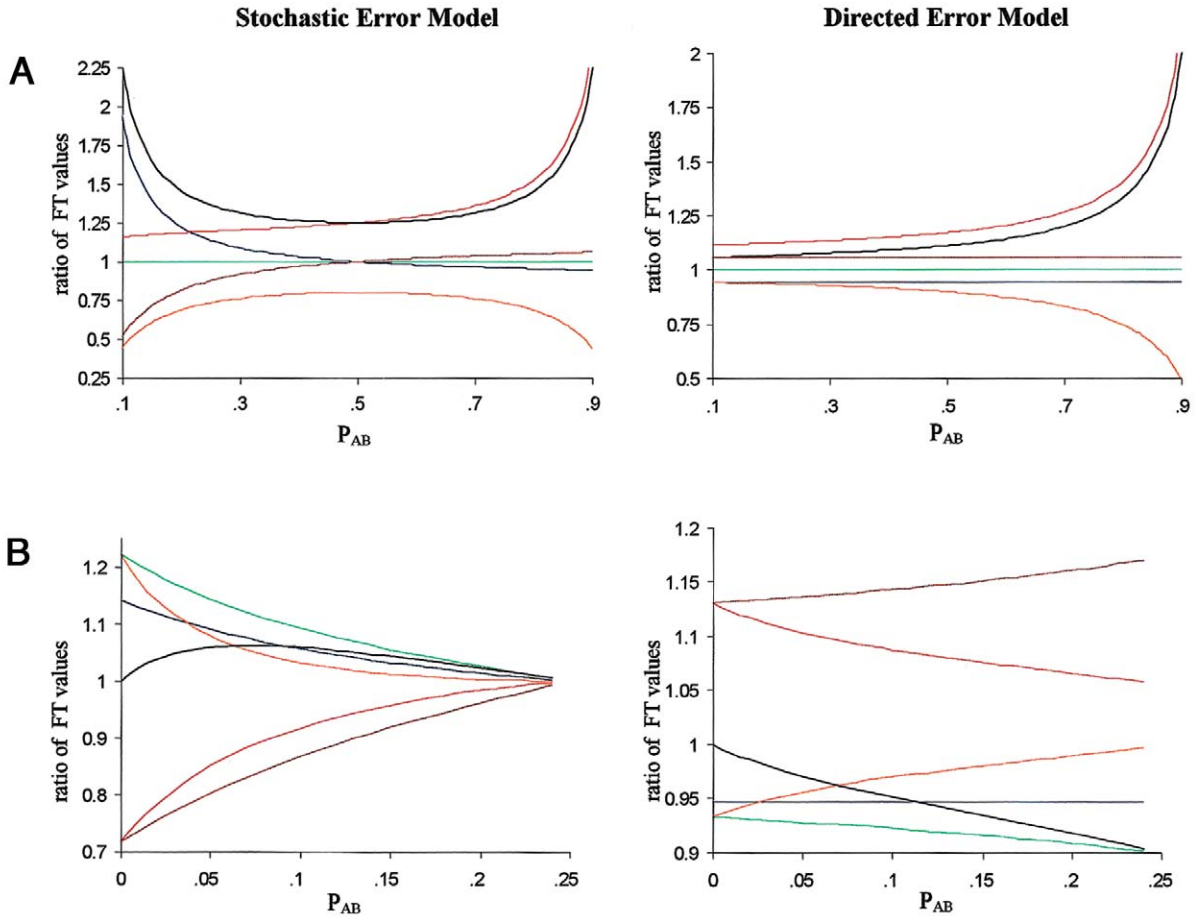

Figure 3 Comparison of the robustness of D', r, Q, and d. The ratio of FT values is plotted versus P_{AB}. The curves correspond to the ratios d/r (*blue*), r/D' (*green*), r/Q (*orange*), Q/d (*red*), Q/D' (*black*), and D'/d (*brown*). In panel *A*, the haplotype frequencies are $P_{Ab} = P_{aB} = 0$ and $P_{ab} = 1 - P_{AB}$. In panel *B*, the haplotype frequencies are $P_{Ab} = P_{aB} = .25$ and $P_{ab} = .5 - P_{AB}$. The genotyping-error rate was set at 5%.

background LD have been and will be conducted to better delimit the magnitude and distribution of LD, within and between human populations. In the planning and interpretation of LD studies, it is important to keep in mind the effect that genotyping errors have on LD measures. Through the use of deterministic formulas and stochastic simulations, we have demonstrated both that genotyping errors can have serious consequences with regard to estimates of LD and that LD measures show varying degrees of robustness in the presence of errors. It is important to note that, in our study, we have assumed that haplotypes were known without errors. Therefore, our results are likely to be optimistic, since haplotyping errors undoubtedly further impede accurate estimation of LD (Tishkoff et al. 2000).

Genotyping errors have long been recognized as problematic in genetic studies. In fact, the literature is rich in studies that have investigated the effect that genotyping errors have on linkage analysis. The general conclusions of these studies are that genotyping errors increase estimated values of the recombination fraction (Terwilliger et al. 1990), complicate the construction of high-resolution linkage maps (Buetow 1991; Shields et al. 1991; Lincoln and Lander 1992), and decrease the power to detect a disease locus (Terwilliger et al. 1990; Gordon et al. 1999; Goring and Terwilliger 2000). More specifically, Gordon et al. (1999) employed a simulation approach to study how genotyping errors influence the power of the transmission/disequilibrium test, using microsatellite markers, and concluded that error rates should be kept <5%. Moreover, Gordon and Ott (2001) investigated the consequences of SNP genotyping errors on the power of case-control studies, using Pearson's χ^2 as the test statistic, and proposed a novel reduced-penetrance model to increase the power to detect a disease locus. Although our study is complementary to the aforementioned investigations, it addresses several unique points. For example, we have provided

Table 5

Results of Stochastic Simulations for LD Measures, as a Function of the Genotyping-Error Rate and Model

LD-Measure Difference[a]	Value When Genotyping-Error Rate Is							
	1.00%		3.00%		5.00%		10.00%	
	SEM	DEM	SEM	DEM	SEM	DEM	SEM	DEM
$\lvert D'_1 - D'_2 \rvert$:								
Average	.0210	.0113	.0320	.0290	.0533	.0402	.0718	.0505
Maximum	.0614	.0417	.1210	.0808	.1670	.1250	.2257	.1406
$\lvert r_1 - r_2 \rvert$:								
Average	.0219	.0143	.0330	.0246	.0484	.0324	.0722	.0515
Maximum	.0602	.0398	.1300	.1093	.1704	.0992	.1792	.1204
$\lvert Q_1 - Q_2 \rvert$:								
Average	.0008	.0003	.0041	.0014	.0102	.0036	.0427	.0133
Maximum	.0042	.0013	.0272	.0075	.0467	.0104	.1052	.0300
$\lvert d_1 - d_2 \rvert$:								
Average	.0237	.0158	.0304	.0244	.0415	.0364	.0703	.0518
Maximum	.0797	.0400	.1499	.0900	.1530	.0900	.1784	.1200

Note.—The haplotype frequencies for both samples in the absence of errors were set at $P_{AB} = P_{ab} = .5$, $P_{Ab} = P_{aB} = 0$.

[a] Between two simulated populations.

the first detailed analysis of how genotyping errors affect the robustness of four commonly used LD measures. In addition, whereas those other studies have focused on disease-gene mapping, we provide practical information on how genotyping errors complicate estimates of background LD.

In our analyses, we considered only SNPs, although microsatellite markers are also commonly used in LD studies (Peterson et al. 1995; Eaves et al. 2000). Our results are not directly applicable to microsatellites, since genotyping-error models for multiallelic markers are much more complicated than either the SEM or the DEM used in this study. Although we can not make quantitative statements regarding the impact that genotyping errors have on the robustness of LD measures for microsatellites, our results can be approximately generalized to microsatellites if the alleles are grouped into two classes. In fact, there are several different strategies for the grouping of microsatellite alleles (see Akey et al. [2001*a*] for discussion) and it would be interesting to explore the possibility that the method of grouping influences the robustness of LD measures to genotyping errors.

Furthermore, we have assumed that genotyping errors follow either an SEM or a DEM. In reality, genotyping errors may also follow a hybrid SEM-DEM. Therefore, in the future, it may be worthwhile to investigate the impact that genotyping errors that follow more-complex error models have on estimates of LD, although the present study has captured the major features of this problem. In fact, the SEM and the DEM represent the two extreme cases of how genotyping errors occur, with the former leading to more-serious affects on LD measures than does the latter. Hence, for a fixed genotyping-error rate, the FE for an LD measure under a hybrid SEM-DEM is expected to be between the FE for an LD measure under a strict SEM and that for a strict DEM (i.e., $FE_{DEM} \leqslant FE_{SEM\text{-}DEM} \leqslant FE_{SEM}$). Computer simulations of D' have confirmed this expectation (data not shown). Moreover, comparing the values in tables 4 and 5 demonstrates that, for a broad range of allele frequencies, the differences between the SEM and the DEM are small—and thus the differences between a hybrid SEM-DEM and either a SEM or a DEM are likely even smaller.

An important question in LD studies is, What measure of disequilibrium should be used? D' and r (or r^2) are likely the two most commonly used measures. Although no existing LD measure is independent of allele frequencies (Lewontin 1988), the range of D' is independent of allele frequencies, making it attractive for comparisons between samples. However, on the basis of our data, Q and d are generally more robust to genotyping errors over a wide range of haplotype frequencies, than are D' and r, although exceptions certainly do exist (see figs. 2 and 3). Thus, the choice of LD measures is not straightforward. Because of the computational ease afforded by D', r, Q, and d, it is feasible to calculate all four measures and then examine them for any inconsistencies that may suggest the presence of genotyping errors.

Finally, it is interesting to consider published genotyping-error rates in the context of the present study. Reported error rates vary from ~1% (Pastinen et al. 2000; Akey et al. 2001*b*; Prince et al. 2001) to 30% (Wang et al. 1998; Cho et al. 1999; Hacia et al. 1999).

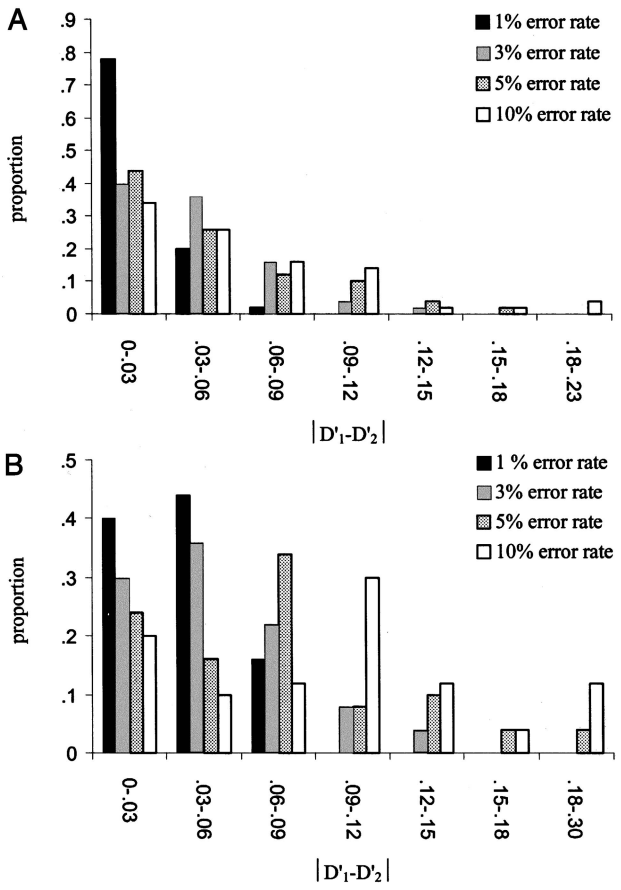

Figure 4 Distribution of the difference, in absolute value of D', between two simulated samples, in the presence of genotyping errors. In panel *A*, the allele frequencies are $P_A = P_B = .5$; in panel *B*, the allele frequencies are $P_A = P_B = .8$. For both panels, the value of $|D'_1 - D'_2|$ in the absence of genotyping errors is 0. Note that as the error rates increase the distribution begins to shift away from the expected difference of 0.

Obviously, a 30% error rate is unacceptable. However, in light of our data, even error rates as low as 3% can have important ramifications for LD measures. Thus, we conclude that, to extract meaningful information from LD studies, it is critical to minimize, if not eliminate, the extent of genotyping errors.

Acknowledgments

We would like to thank Michael Akey and Anthony Berella for helpful discussions related to this work. M.X. was supported by National Institutes of Health grants GM56515 and E509912.

References

Abecasis GR, Noguchi E, Heinzmann A, Traherne JA, Bhattacharyya S, Leaves NI, Anderson GG, Zhang Y, Lench NJ, Carey A, Cardon LR, Moffatt MF, Cookson WO (2001) Extent and distribution of linkage disequilibrium in three genomic regions. Am J Hum Genet 68:191–197

Akey J, Jin L, Xiong M (2001*a*) Haplotypes vs. single marker linkage disequilibrium tests: what do we gain? Eur J Hum Genet 9:291–300

Akey JM, Sosnoski D, Parra E, Dios S, Hiester K, Su B, Bonilla C, Jin L, Shriver MD (2001*b*) Melting curve analysis of SNPs (McSNP): a gel-free and inexpensive approach for SNP genotyping. Biotechniques 30:358–367

Buetow KH (1991) Influence of aberrant observations on high-resolution linkage analysis outcomes. Am J Hum Genet 49: 985–994

Cho RJ, Mindrinos M, Richards DR, Sapolsky RJ, Anderson M, Drenkard E, Dewdney J, Reuber TL, Stammers M, Federspiel N (1999) Genome-wide mapping with biallelic markers in *Arabidopsis thaliana*. Nat Genet 23:203–207

Collins FS, Guyer MS, Charkravarti A (1997) Variations on a theme: cataloging human DNA sequence variation. Science 278:1580–1581

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

Eaves IA, Merriman TR, Barber RA, Nutland S, Tuomilehto-Wolf E, Tuomilehto J, Cucca F, Todd JA (2000) The genetically isolated populations of Finland and Sardinia may not be a panacea for linkage disequilibrium mapping of common disease genes. Nat Genet 25:320–323

Goddard KA, Hopkins PJ, Hall JM, Witte JS (2000) Linkage disequilibrium and allele-frequency distributions for 114 single-nucleotide polymorphisms in five populations. Am J Hum Genet 66:216–234

Gordon D, Matise TC, Heath SC, Ott J (1999) Power loss for multiallelic transmission/disequilibrium test when errors introduced: GAW11 simulated data. Genet Epidemiol Suppl 17:S587–S592

Gordon D, Ott J (2001) Assessment and management of single nucleotide polymorphism genotype errors in genetic association analysis. Pac Symp Biocomput 18–29

Gordon D, Simonic I, Ott J (2000) Significant evidence for linkage disequilibrium over a 5-cM region among Afrikaners. Genomics 66:87–92

Goring HH, Terwilliger JD (2000) Linkage analysis in the presence of errors II: marker-locus genotyping errors modeled with hypercomplex recombination fractions. Am J Hum Genet 66:1107–1118

Hacia JG, Fan J-B, Ryder O, Jin L, Edgemon K, Ghandour G, Mayer RA, Sun B, Hsie L, Robbins CM (1999) Determination of ancestral alleles for human single nucleotide polymorphisms using high-density oligonucleotide arrays. Nat Genet 22:164–167

Halushka MK, Fan JB, Bentley K, Hsie L, Shen N, Weder A, Cooper R, Lipshutz R, Chakravarti A (1999) Patterns of single-nucleotide polymorphisms in candidate genes for blood-pressure homeostasis. Nat Genet 22:239–247

Hedrick JP (1987) Gametic disequilibrium measures: proceed with caution. Genetics 117:331–341

Hill WG, Robertson A (1968) Linkage disequilibrium in finite populations. Theor Appl Genet 38:226–231

International SNP Map Working Group, The (2001) A map

of human genome sequence variation containing 1.42 million single nucleotide polymorphisms. Nature 409:928–933

Jorde LB, Watkins WS, Kere J, Nyman D, Eriksson AW (2000) Gene mapping in isolated populations: new roles for old friends? Hum Hered 50:57–65

Kidd JR, Pakstis AJ, Zhao H, Lu RB, Okonofua FE, Odunsi A, Grigorenko E, Tamir BB, Friedlaender J, Schulz LO, Parnas J, Kidd KK (2000) Haplotypes and linkage disequilibrium at the phenylalanine hydroxylase locus, PAH, in a global representation of populations. Am J Hum Genet 66: 1882–1899

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

Lewontin RC (1964) The interaction of selection and linkage. I. General considerations: heterotic models. Genetics 49: 49–67

——— (1988) On measures of gametic disequilibrium. Genetics 120:849–852

Lewontin RC, Kojima K (1960) The evolutionary dynamics of complex polymorphisms. Evolution 14:458–472

Lincoln SE, Lander ES (1992) Systematic detection of errors in genetic linkage data. Genomics 14:604–610

Livak KJ, Flood SJ, Marmaro J, Giusti W, Deetz K (1995) Oligonucleotides with fluorescent dyes at opposite ends provide a quenched probe system useful for detecting PCR product and nucleic acid hybridization. PCR Methods Appl 4: 357–362

Moffatt MF, Traherne JA, Abecasis GR, Cookson WO (2000) Single nucleotide polymorphism and linkage disequilibrium within the TCR α/δ locus. Hum Mol Genet 9:1011–1019

Nei M, Li W-H (1980) Non-random association between electromorphs and inversion chromosomes in finite populations. Genet Res 35:65–83

Pastinen T, Raitio M, Lindroos K, Tainola P, Peltonen L, Syvanen AC (2000) A system for specific, high-throughput genotyping by allele-specific primer extension on microarrays. Genome Res 10:1031–1042

Peterson AC, Di Rienzo A, Lehesjoki AE, de la Chapelle A, Slatkin M, Freimer NB (1995) The distribution of linkage disequilibrium over anonymous genome regions. Hum Mol Genet 4:887–894

Prince JA, Feuk L, Howell WM, Jobs M, Emahazion T, Blennow K, Brookes AJ (2001) Robust and accurate single nucleotide polymorphism genotyping by dynamic allele-specific hybridization (DASH): design criteria and assay validation. Genome Res 11:152–162

Shields DC, Collins A, Buetow KH, Morton NE (1991) Error filtration, interference, and the human linkage map. Proc Natl Acad Sci USA 88:6501–6505

Taillon-Miller P, Bauer-Sardina I, Saccone NL, Putzel J, Laitinen T, Cao A, Kere J, Pilia G, Rice JP, Kwok PY (2000) Juxtaposed regions of extensive and minimal linkage disequilibrium in human Xq25 and Xq28. Nat Genet 25: 324–328

Terwilliger JD, Weeks DE, Ott J (1990) Laboratory errors in the reading of marker alleles cause massive reductions in LOD score and lead to gross overestimation of the recombination fraction. Am J Hum Genet Suppl 47:A201

Tishkoff SA, Dietzsch E, Speed W, Pakstis AJ, Kidd JR, Cheung K, Bonne-Tamir B, Santachiara-Benerecetti AS, Moral P, Krings M (1996) Global patterns of linkage disequilibrium at the CD4 locus and modern human origins. Science 271: 1380–1387

Tishkoff SA, Pakstis AJ, Ruano G, Kidd KK (2000) The accuracy of statistical methods for estimation of haplotype frequencies: an example from the CD4 locus. Am J Hum Genet 67:518–522

Wang DG, Fan JB, Siao CJ, Berno A, Young P, Sapolsky R, Ghandour G, et al. (1998) Large-scale identification, mapping, and genotyping of single-nucleotide polymorphisms in the human genome. Science 280:1077–1082

Wu YY, Delgado R, Costello R, Sunderland T, Dukoff R, Csako G (2000) Quantitative assessment of apolipoprotein E genotypes by image analysis of PCR-RFLP fragments. Clin Chim Acta 293:213–221

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

Yule GU (1900) On the association of attributes in statistics. Philos Trans R Soc Lond A 194:257–319

Zavattari P, Deidda E, Whalen M, Lampis R, Mulargia A, Loddo M, Eaves I, Mastio G, Todd JA, Cucca F (2000) Major factors influencing linkage disequilibrium by analysis of different chromosome regions in distinct populations: demography, chromosome recombination frequency and selection. Hum Mol Genet 12:2947–2957