Genetic Epidemiology 32 : 546–552 (2008)

# Does Strong Linkage Disequilibrium Guarantee Redundant Association Results?

**Dahlia M. Nielsen,[1,2]* Sunil Suchindran[2] and Christopher P. Smith[2]**

[1]*Program in Statistical Genetics, Department of Genetics, North Carolina State University, Raleigh, North Carolina*
[2]*Bioinformatics Research Center, North Carolina State University, Raleigh, North Carolina*

A substantial amount of effort has been expended recently towards the identification and evaluation of tag single nucleotide polymorphisms; markers that, due to linkage disequilibrium (LD) patterns in the genome, are able to act as ''proxies'' for other polymorphic sites. As such, these tag markers are assumed to capture, on their own, a large proportion of the genetic variation contributed by a much greater number of polymorphic sites. One important consequence of this is the potential ability to reduce the cost of genotyping in an association study without a corresponding loss of power. This application carries an implicit assumption that strong LD between markers implies high correlation between the accompanying association test results, so that once a tag marker is evaluated for association, its outcome will be representative of all the other markers for which it serves as proxy. We examined this assumption directly. We find that in the null hypothesis situation, where there is no association between the markers and the phenotype, the relationship between LD and the correlation between association test outcomes is clear, though it is not always ideal. In the alternative case, when genetic association does exist in the region, the relationship becomes much more complex. Here, reasonably high LD between markers does not necessarily imply that the association test result of one marker is a direct substitute for that of the other. In these cases, eliminating one of these markers from the set to be genotyped in an association study will lead to a reduction in overall power. *Genet. Epidemiol*. 32:546–552, 2008. © 2008 Wiley-Liss, Inc.

**Key words: linkage disequilibrium; association mapping; negative correlation; case control; transmission disequilibrium**

Contract grant sponsor: National Institutes of Health Grant; Contract grant number: GM 45344.
*Correspondence to: Dahlia M. Nielsen, Program in Statistical Genetics, Department of Genetics, North Carolina State University, Raleigh, NC 27695-7566, USA. E-mail: dahlia@statgen.ncsu.edu
Received 20 November 2007; Accepted 15 February 2008
Published online 7 April 2008 in Wiley InterScience (www.interscience.wiley.com).
DOI: 10.1002/gepi.20328

## INTRODUCTION

One of the goals of the HapMap project has been to provide insight into the patterns of linkage disequilibrium (LD) in the human genome; information that can be used to determine where redundancies between marker loci exist, and to subsequently reduce the cost of genetic association studies [The International HapMap Consortium, 2003]. This extensively genotyped reference panel achieves this goal by providing a means of identifying ''tag'' single nucleotide polymorphisms (SNPs); markers that provide the same, or close to the same, information as the entire set that was discovered in the panel. Cost reductions are possible in individual association studies, as the samples collected for these studies are genotyped only for the tag markers and not for the full set of sites identified in the reference panel. The assumption is that these tag SNPs are able to serve as proxies for the rest of the known polymorphisms.

Many selection strategies have been proposed to identify tag SNPs. While the underlying assumptions and the technical details of these methods differ, each relies on the underlying LD structure of the region to provide information regarding redundancies between polymorphic sites. Prevailing methods rely on LD estimates based on the $r^2$ measure [Carlson et al., 2004]. This is, in part, because these methods are not subject to definitions of physical boundaries along the genome and, importantly, appear to be reasonably robust to population choice [de Bakker et al., 2006; Montpetit et al., 2006; Service et al., 2007; Gu et al., 2007]. Additionally, a relationship has been shown to exist connecting the power of a marker to detect association and $r^2$ between this marker and an unobserved causal site [Pritchard and Przeworski, 2001; see also Terwilliger, 2006; Moskvina and O'Donovan, 2007].

It can only be assumed that tag SNPs indeed act as proxies of other sites if these other sites (along with the tag SNPs) were genotyped in the reference panel. LD between sites not observed in this panel and any genotyped markers remains unknown, and thus the relationship between the observed and unobserved sites is also unknown. Therefore, if the causal sites relevant for a given association study were not examined (or not polymorphic) in the reference panel, the tag SNPs may not serve as direct proxies for these causal sites. Tag SNPs, of course, do still serve a purpose in this situation. While they may not serve as direct proxies for the causal sites of interest, they do serve as proxies for many other known markers. The result remains that the cost of performing an

© 2008 Wiley-Liss, Inc.

association study is reduced, as redundant markers are not genotyped in that study. However, in this case, we must rely on the assumption that if one non-causal SNP serves as a proxy for another non-causal SNP (LD between these SNPs is high), the information gained by an association test using the tag SNP is highly redundant of the information that would be obtained by the untyped SNP. In other words, if association can be detected with some marker identified in the panel, it is expected that a tag SNP that serves as proxy for this marker is also highly likely to detect this association.

The assumption underlying this is that if the correlation between alleles (LD) is high, then the association test results will also be highly correlated. Intuitively, this assumption appears reasonable. For instance, if $r^2 = 1$ between two SNPs, alleles at one SNP completely predict the alleles at the other. In this case, the test statistics derived using data from these SNPs will be identical. An $r^2$ value of 1 between two SNPs indicates that the test statistics will be 100% correlated. At the other extreme, if $r^2 = 0$ between two SNPs, then the SNP genotypes are uncorrelated, and it seems reasonable to assume that the test statistics derived from these SNPs will also be uncorrelated. The question is what happens in between these extremes. Test statistics are functions of the data, and it is not clear that the correlation patterns between the test statistics should be equivalent to the correlation patterns of the original data.

We were interested in examining this assumption directly to determine how well LD between markers predicts redundancies between the association test results for these markers. This was done using various different types of association testing strategies as well as considering various genetic models.

## MATERIALS AND METHODS

We were interested in examining the relationship between association tests at different markers, compared with the amount of LD between these markers. To do this in a general setting, it is necessary to consider three loci at a time: the two markers to be tested and a functional site that can provide the framework for the marker-trait association. In this general setting, the two observed SNPs can exhibit any pattern of LD (including no LD) with the functional SNP and with each other. This also allows for the particular situation where one of the observed SNPs is functional itself, as this is equivalent to modeling an observed SNP that is in perfect LD ($r^2 = 1$) with the functional site.

In order to create LD patterns that avoided modeling assumptions and best reflected values found in true human populations, haplotypes were sampled following frequencies found in a densely genotyped human population. For this, we used the phased HapMap Japanese and Chinese samples from chromosome 22 [The International HapMap Consortium, 2003]. Only diallelic markers with minor allele frequencies of at least 3% were considered. Groups of three SNPs in nearby proximity were selected at random. If the haplotype frequencies for a selected SNP trio were sufficiently similar to the frequencies of a trio already sampled, one of the trios (selected at random) was dropped. Using this strategy, over 200,000 trios of SNPs were considered. These represented a very broad range of possible LD patterns and allele frequencies seen in real human populations. For each of the SNP trios used, the locus with the smallest minor allele frequency was determined to be the functional site. The minor allele was chosen to be the one associated with higher susceptibility.

Once appropriate haplotype patterns were established, 10,000 independent samples of individuals were generated for each SNP trio. Association tests were performed for both of the observed markers within an SNP trio using these samples. The resulting 10,000 replicate observations of paired association tests were used to estimate the correlation between the test statistics for that SNP trio. This procedure was repeated for the $>200{,}000$ SNP trios, each with different LD patterns and allele frequencies. Correlation between the test statistics could then be compared with LD between the observed markers.

In the null hypothesis case where the phenotype and the markers are not associated, data could be simulated by considering only markers; no functional locus is necessary. For consistency, the same three locus haplotypes were used under both the null and the alternate hypotheses. In the null case, the marker determined to be the functional site for a given trio of SNPs was effectively ignored.

We considered several types of association test procedures, including both the standard allele-based and genotype-based case-control tests, the single degree of freedom transmission disequilibrium test [TDT; Spielman et al., 1993], and a two degree of freedom TDT [Weinberg et al., 1998]. Simulations were performed using several different genetic models. Unless otherwise stated, 500 trios were simulated per TDT sample, and 500 cases and 500 controls were simulated for each case-control sample.

The choice of genetic models under the alternative hypothesis was driven by various factors. Any model can be described in terms of additive and non-additive effects (an additive model is one in which the non-additive components are equal to zero). Single degree of freedom tests tend to be sensitive to additive genetic effects only, so that for these tests, the non-additive components are generally unimportant. Two degrees of freedom tests are sensitive to both additive and non-additive terms [Nielsen and Weir, 2001]. To provide a range of reasonable terms, we chose an additive-only model and a recessive model for the genetic risk conferred by the functional site. In the recessive case, even though the functional locus follows this model, an associated marker will only do so if there is perfect LD between it and the functional site. Otherwise, the markers will display a range of both additive and non-additive effects, depending on the degree of LD between the marker and the functional site [Nielsen and Weir, 2001]. As the association tests are performed on the markers and not the functional site itself, it is the indirect effects seen for these sites that are relevant.

To assign penetrances to the functional genotypes for both additive and recessive models, we followed the method described by de Bakker et al. [2006]. With this strategy, the overall population prevalence of the phenotype is fixed for each locus (in our case at 5%), and the penetrances assigned so that, if the causative locus itself were tested, the expected power to detect association would be about 90%. This means that causal loci with rarer minor alleles have higher penetrances assigned to their risk genotypes than do loci with more common minor alleles. Maintaining the power of the causal locus at a fixed level provides, to some degree, a more stable comparison

of power levels at the associated markers. To verify that this strategy did not affect our results, we also performed the simulations using fixed values of the penetrances assigned to all functional loci, and found no differences in the conclusions.

# RESULTS

The question we are chiefly concerned with examining is ''what does LD between markers tell us about the relationship between association tests results?'' Specifically, we wanted to know how well LD between SNPs predicts the relationship between their respective association test outcomes. To address this, we examined pairs of SNPs that exhibited various degrees of LD with one another, and tested each of these SNPs for association with a trait. Correlations between the test statistics corresponding to these two SNPs were then estimated and compared with LD between the markers.

## NULL HYPOTHESIS: NO ASSOCIATION

In the case where no association exists between the trait and the markers under examination, population frequencies were used to sample marker haplotypes for both affected and unaffected individuals. In this case, the only factor creating dependencies between the test statistics for the two markers is LD between the markers. The results for the TDT are shown in Figure 1A. Along the horizontal axis is $r^2$ between each pair of observed markers, and along the vertical axis is the estimated correlation between the test statistics. Each point on this plot represents one pair of markers, both tested for 10,000 simulated samples. The results from this simulation were somewhat surprising in that the relationship here is linear, and it appears to be completely predictive. The linearity is interesting, as the *X*-axis of this plot represents squared correlations (LD represented as $r^2$), while the *Y*-axis is the unsquared correlation between test statistics ($\rho$). Results shown in Figure 1A are for the single degree of freedom TDT, but the results for the allele-based case-control test were identical (not shown). These results are illuminating in that they appear to verify the intuition that LD between SNPs (captured by $r^2$) is predictive of the relationship between the association results. For example, in this situation, a value of $r^2 = 0.8$ between SNPs implies that the association test result for one SNP will be 80% correlated with the test statistic of the other SNP. We have also shown this result analytically [Sunil Suchindran, in preparation].

An important consideration, however, is that an association test does not generally end with the calculation of the test statistic. Instead, a *P*-value is derived and statistical significance is determined based on an appropriate threshold. Therefore, the real issue might not be the level of correlation of the test statistics themselves, but the relationship between the conclusions of these tests. We examined this relationship as well, considering the correlation between the outcomes of ''detected association'' (rejected the null hypothesis) versus ''did not detect association'' for both SNPs. The results of this comparison are shown in Figure 1B. We now find that the relationship between LD and correlation of these conclusions is not linear, but instead, the points fall below the 45° line. For example, correlations between conclusions (whether or not SNPs are associated) are less than 60% when LD between SNPs is $r^2 = 0.8$. We have previously exploited this relationship to derive a multiple testing adjustment strategy that takes the amount of LD between markers into consideration when calculating appropriate association test thresholds [Nielsen et al., 2004; at that time the quantity we refer to as correlation between test statistics is actually the squared correlation].

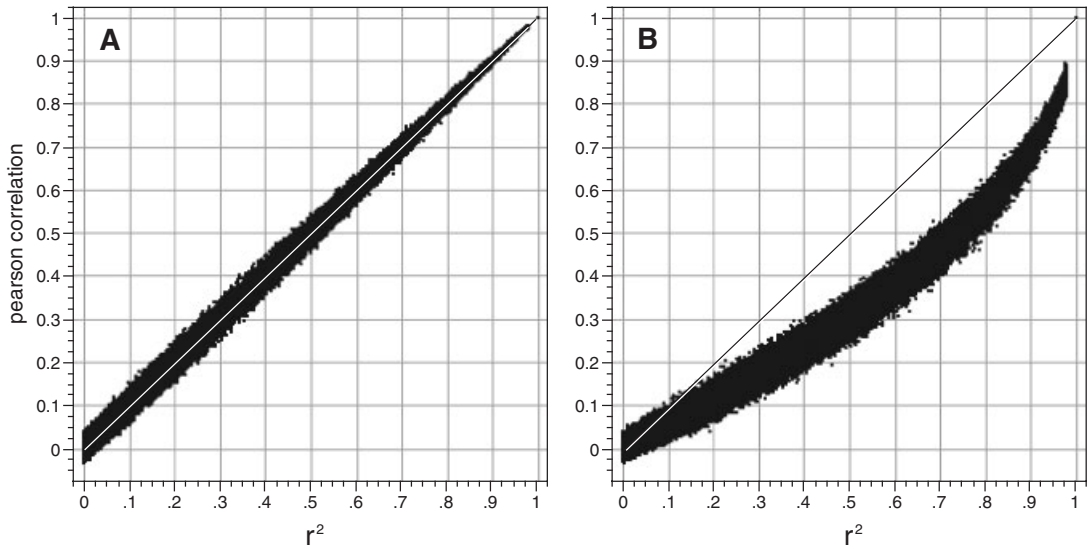


**Fig. 1. Results for the TDT under the null hypothesis (no association between SNPs and the phenotype). Horizontal axis is LD presented as $r^2$: (A) Correlation of the test statistics, which, in this situation, lies along the 45° line and (B) correlation between test conclusions (reject or do not reject the null hypothesis of no association). When this binary outcome, rather than the test statistic itself, is considered, the relationship is no longer linear. Instead, the correlation between conclusions is smaller than LD ($r^2$). LD, linkage disequilibrium.**

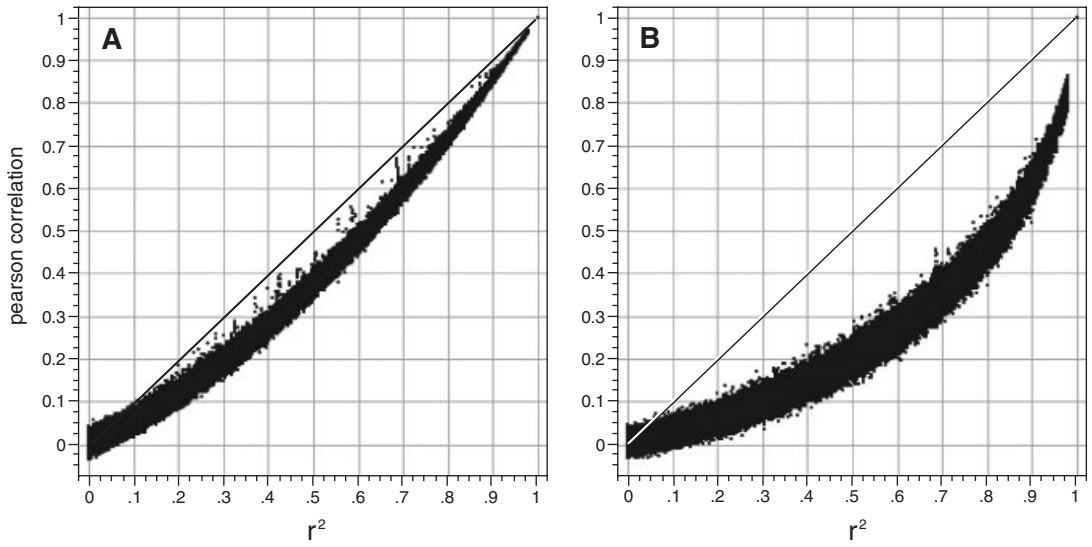


**Fig. 2. Results for the genotype-based case-control test under the null hypothesis. Horizontal axis is $r^2$ (LD): (A) Correlation of the test statistics and (B) correlation between binary test outcomes (reject or do not reject the null hypothesis of no association). In both these cases, correlation of the test results is lower than $r^2$ (LD). LD, linkage disequilibrium.**

The results described so far are for single degree of freedom tests (the allele-based case-control test and the TDT). These tests are sensitive to the additive effects of the traits only [Nielsen and Weir, 2001]. We also examined two degree of freedom tests, which are sensitive to both additive and dominance effects [Nielsen and Weir, 2001]. The results for the genotype-based case-control test are shown in Figure 2A. As before, the *X*-axis here is LD between SNPs ($r^2$) and the *Y*-axis is correlation between test statistics. As can be seen, this relationship is again not linear; most of the points fall below the 45° line. Comparing the outcomes from these tests (associated versus not associated) provides results further from unity (Fig. 2B). In this situation, LD of $r^2 = 0.8$ predicts a correlation of less than 50% for the outcomes of the association tests, possibly a much lower level of prediction than one might hope.

All the results described so far represent observations when there was no association between the SNPs and the trait (the null hypothesis). In this case, the observations essentially address the issue of correlation between false-positive results. This is indeed of interest when a researcher detects a positive association signal between multiple adjacent SNPs, and wants to evaluate whether this is strong evidence of true association in the region or merely correlated false positives.

## ALTERNATIVE HYPOTHESIS: ASSOCIATION EXISTS

A possibly more interesting situation to consider is that in which there is true association in the region. To perform these simulations, we consider a general scenario in which we model two observed SNPs and a third, unobserved functional site. When association does exist, the relationship between association test results depends on a wider set of factors than just LD between the observed SNPs. Contributing here are LD values between each of these SNPs and the functional site, multi-SNP LD (that which exists when considering all three sites jointly), and the manner in which the functional site affects the trait itself (the genetic model). Of course, of these factors, only LD between the genotyped SNPs is observable. It is this relationship that we rely on to provide us with information regarding tag SNP selection.

We again use LD patterns determined from real data for the simulations under the alternative hypothesis case. Here, as mentioned above, these patterns involve three sites: two observed markers and a functional site that is not tested directly in the simulation procedure. We consider an additive genetic model and a purely recessive model. The results for the TDT under the additive model are shown in Figure 3. The axes are as before, with LD ($r^2$) on the horizontal axis and correlation between test statistics on the vertical. Again, each point represents a correlation estimate derived from 10,000 replicated samples; the graph contains >200,000 such points. Not surprisingly, it is apparent that LD between SNPs is no longer completely predictive of the correlation between association test statistics, as now there are more unseen elements involved. More interestingly, we find a large proportion of situations in which non-zero LD between SNPs is seen together with negative correlation between the test statistics. This is true even in cases where LD between SNPs is in the vicinity of $r^2 = 0.8$. This implies that when the test statistic for one SNP is high, the result for the other SNP tends to be low, and vice versa. In these cases, selecting only one of these SNPs for genotyping implies relying on luck as to whether a good choice was made.

In order to understand what might cause this type of result to occur, we examined a number of these points in detail. Variations of the scenario depicted in Figure 4A appear to stand out. For ease of explanation, here the alleles of both observed SNP markers are labeled "C" and "T." In this specific scenario, three haplotypes appear in the population. One of the haplotypes, carrying a "C" allele at each marker (denoted as the "C-C" haplotype), is also the only one to carry a risk allele at the functional site. From this, we expect that "C" alleles for both SNPs will tend to aggregate in (or be transmitted to) affected

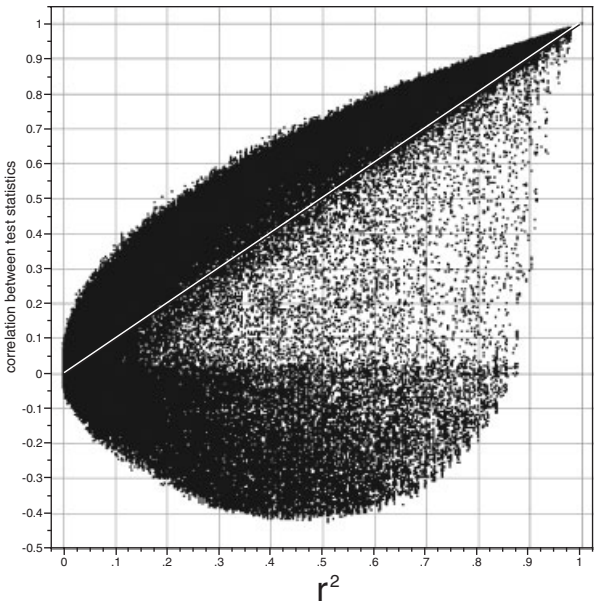


**Fig. 3. Results for the TDT when association exists. Horizontal axis is LD between pairs of SNPs ($r^2$) and vertical axis is correlation between association test statistics for each SNP in the pair. While a large proportion of the points are above the 45° line, indicating that the correlation between the test statistics is greater than LD between SNPs, some points are actually below the zero line. For these points, the association test statistics are negatively correlated, even in the presence of strong LD between SNPs. SNP, single nucleotide polymorphism; LD, linkage disequilibrium.**

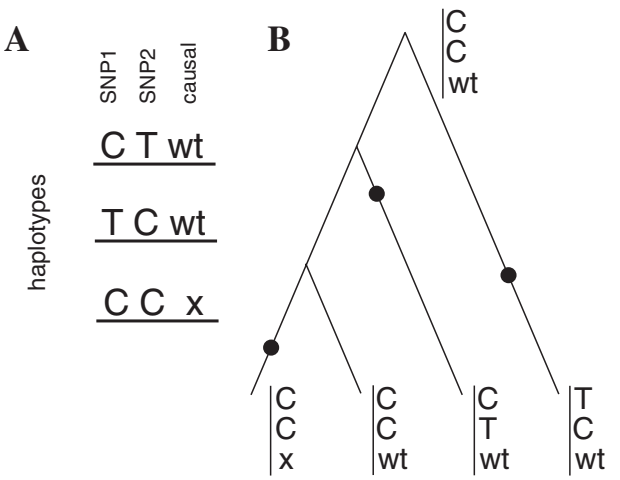


**Fig. 4. (A) Haplotype scenario leading to negative correlations between test statistics in the presence of strong LD between markers and (B) genealogy leading to such a population of haplotypes (the extra haplotype arising here does not alter the outcome). LD, linkage disequilibrium.**

individuals, and association evidence will be at its strongest when we detect this effect at the same time as we detect ''T'' alleles among unaffected individuals (non-transmitted chromosomes; for the following, we restrict our explanation to the case-control design). On average, sample proportions of marker alleles among cases and controls are expected to equal their theoretical frequencies, which are dependent on the multilocus genotype frequencies in the population and on the genetic model. In any given sample, however, these alleles will appear at rates that vary around their expected values. If one allele (''C'' or ''T'') at a given SNP happens to be sampled at a higher proportion than expected, the other allele for this SNP is necessarily sampled at a lower frequency for a given sample size. This variation will cause some samples to provide better power than others. In this example, the best results will be provided by samples with an overrepresentation of ''C'' alleles among cases and an overrepresentation of ''T'' alleles among controls. Because of the haplotype structure in the overall population, ''T'' alleles at one SNP appear only on haplotypes with a ''C'' allele at the other SNP. Therefore, the samples in which ''T'' alleles are overrepresented in controls for one SNP are the same samples in which ''T'' alleles are underrepresented in controls for the other SNP. Hence, samples that provide stronger evidence for association at one SNP are those that provide weaker evidence for association at the other SNP.

This specific situation involving three haplotypes is useful for illustration of the phenomenon, however, more complicated scenarios involving a larger number of haplotypes are also possible. In all cases we examined in which negative correlations occurred, a variation on the basic theme was evident. While this does not exclude the possibility that other causes for negative correlation may also exist, these causes were not readily apparent.

One question that might arise from this discussion is whether a scenario such as this is particularly plausible in real data. In addition to having modeled our haplotype frequencies on values estimated from real data, a simple genealogy can be created representing the type of haplotype structure described above (Fig. 4B). In this genealogy, a fourth haplotype exists, composed of the ''C-C'' SNP alleles together with a non-risk allele at the functional site. The presence of this haplotype in the population merely reduces the overall power to detect association (as it increases the observations of ''C'' alleles among unaffected individuals). It will not change the phenomenon of negative correlation between association tests.

As the effect of negative correlation between test statistics arises because of sampling variation, it was interesting to see how an increase in sample size would influence the results. We repeated the simulation procedure for the TDT and the genotype-based case-control tests under each of the genetic models, but increasing the sample size collected for each run by a factor of 10. The results from these simulations for the TDT under the additive model (using sample sizes of 5,000 trios per run) are seen in Figure 5. We find that the correlations became more extreme with larger samples, in both the negative and positive directions. This migration to more extreme values occurs because of the increase in the magnitude of the test statistics; as these values move further away from zero, the correlation patterns between values become more apparent.

**The effect of allele frequency on correlation.** We examined whether allele frequency had an effect on the

correlation patterns that we detected, restricting our focus to SNP combinations where both observed markers had 20% or greater power to detect association (there were over 150,000 such SNP trios). We find that when the functional allele had a frequency larger than around 20%, non-zero correlations between test statistics were much more likely to be positive than negative. In this case, it is likely that the functional mutation appears on more than one marker haplotype background, creating haplotype patterns that are not consistent with those that cause negative correlations between test statistics.

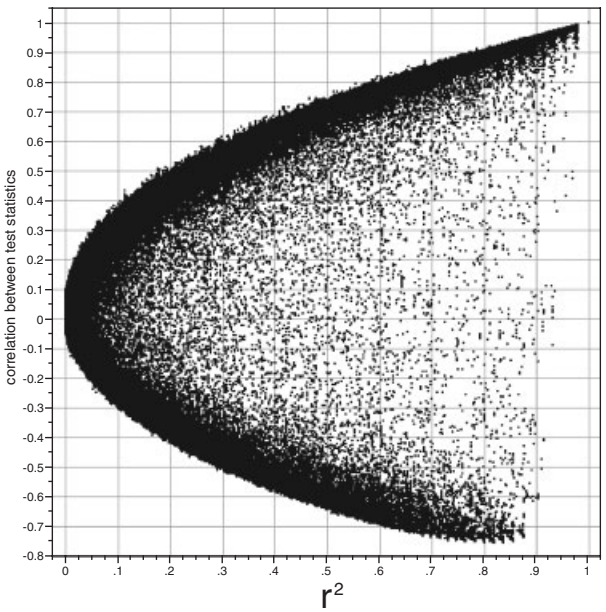


**Fig. 5. The results from the single degree of freedom TDT with sample sizes of 5,000 trios per run. The trend is the same as that seen with smaller sample sizes (Fig. 3), though the values for the correlation estimates are more extreme. This appears to occur because as sample sizes increase, the $\chi^2$ values move further away from the null hypothesis situation. This allows correlation estimates to better reflect their "true" values under the alternative hypothesis. TDT, transmission disequilibrium test.**

For each of the three-SNP combinations considered above, we also examined the allele frequency of the marker that had higher power of the pair of observed markers, by at least 2%. When the test statistics were positively correlated, 45% of the time the power results of the two markers were within 2% of one another (thus, they were considered to be ties). When one marker was observed to have come out ahead, about 25% of the time it was the SNP with the largest minor allele frequency, and 30% of the time it was the one with the rarer minor allele. For the SNP combinations in which negative correlations between test statistics were seen, nearly 80% of the time the SNP with the higher minor allele frequency displayed higher power. Thirteen percent of the time the other SNPs came out ahead; the remaining 7% went to ties.

## CORRELATION OF ASSOCIATION CONCLUSIONS

For each of the figures shown here illustrating our results under the alternative hypothesis (association between the markers and the phenotype does exist), we find the optimistic result that many of the points lie above the 45° line. This implies that, at least to some extent, LD between the markers underestimates the correlation between test statistics. The consequence of this is that if two SNPs are in high LD, unless luck is against you, the association test statistic of one SNP is highly predictive of the result of the other SNP. This, however, is the situation when one is considering the test statistics themselves, rather than the overall conclusions. If, instead, we consider

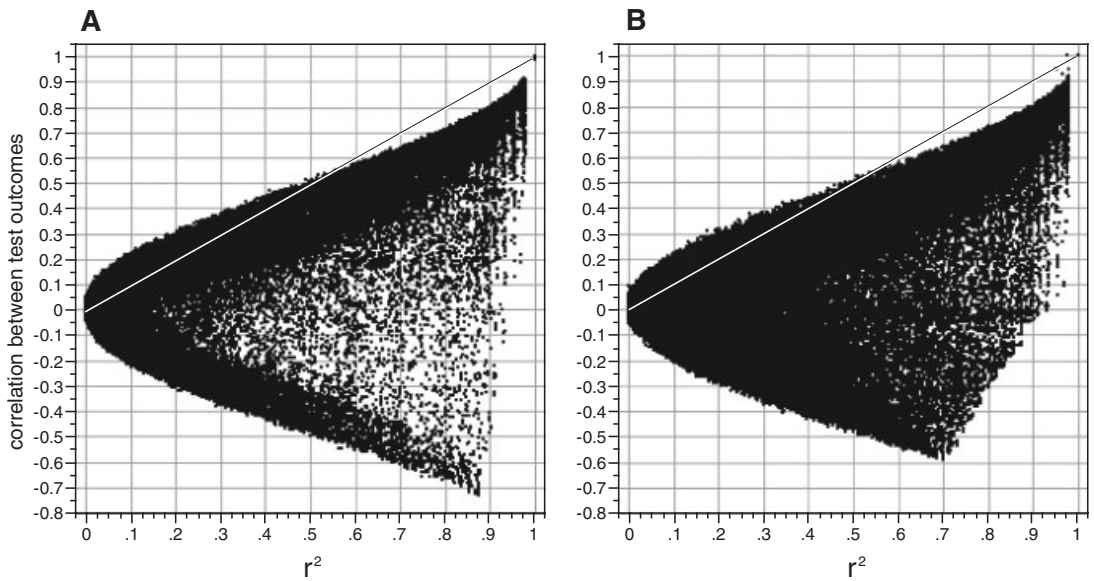


**Fig. 6. Correlation between test conclusions (reject or do not reject the null hypothesis of no association) under the alternative hypothesis: (A) Using a significance threshold of $\alpha = 0.05$ and (B) $\alpha = 0.0005$.**

the correlation between outcomes ''SNP is associated'' versus ''don't detect association,'' the picture changes somewhat. Figure 6 depicts this relationship for the TDT under the additive model, using a significance threshold of 0.05 (A) and 0.0005 (B). In both cases, SNP combinations for which one or both of the markers had power rates of more than 99% were eliminated, as correlation estimates were not reliable in this situation (as the value of the indicator variable is consistently, or nearly consistently, equal to one). As can be seen in the figure, LD between SNPs is generally larger than the correlation between association results. Here, LD of $r^2$ about 0.8 appears to indicate a correlation of around 60% between association test outcomes.

## DISCUSSION

As technology advances and costs continue to drop, it will eventually become feasible to perform association studies using study samples that have been fully sequenced. At this point, all available polymorphisms in the sample will be available for analysis, presumably including the functional sites themselves. Until this time, it is necessary to direct genotyping efforts to a reduced number of known polymorphisms. In this respect, implementing a well-thought-out tag selection strategy will be a necessary and unavoidable component of association mapping, at least for whole-genome studies.

This approach, however, is not free of consequences. As we see here, SNPs can be correlated in terms of their LD in the population, and yet might not provide similarly correlated association test results. This implies that non-tag SNPs may not be as uninformative as is hoped. In some cases, it might come down to sheer luck as to whether a good choice was made regarding the selection of SNPs to be genotyped. One trend we did observe in our results was that, in the cases where test results were negatively correlated, the best power was largely connected with the marker with the higher minor allele frequency. This consideration may aid in the tag SNP selection process.

An intermediate step toward overcoming this problem will be the generation of more complete information in the reference panels. As more and more polymorphisms are identified in these samples, the likelihood of capturing causal sites among them increases. As this happens, it will be more likely that the selected tag markers will act as effective proxies for these sites. This ability is, of course, also limited by the number of individuals included in the reference panels. It is reasonable to assume that samples collected for association studies will be enriched for alleles at functional sites. If these alleles are rare in the general population, the probability of capturing them in a limited reference sample is low. In this case, even with maximal density screening of the reference panel, the problem of imperfect proxies remains.

In the end, the use of tag SNPs selected based on high values of $r^2$ appears to be a reasonable approach. Multi-marker-based techniques [de Bakker et al., 2005], in particular, appear promising. Caution regarding the overall conclusions one makes based on null findings is, however, well warranted.

## ACKNOWLEDGMENTS

This work was supported in part by National Institutes of Health Grant GM 45344.

## REFERENCES

Carlson CS, Eberle MA, Rieder MJ, Yi Q, Kruglyak L, Nickerson DA. 2004. Selecting a maximally informative set of single-nucleotide polymorphisms for association analyses using linkage disequilibrium. Am J Hum Genet 74:106–120.

de Bakker PIW, Yelensky R, Pe'er I, Gabriel SB, Daly MJ, Altshuler D. 2005. Efficiency and power in genetic association studies. Nat Genet 37:1217–1223.

de Bakker PIW, Burtt NP, Graham RR, Guiducci C, Yelensky R, Drake JA, Bersaglieri T, Penney KL, Butler J, Young S, Onofrio RC, Lyon HN, Stram DO, Haiman CA, Freedman ML, Zhu X, Cooper R, Groop L, Kolonel LN, Henderson BE, Daly MJ, Hirschhorn JN, Altshuler D. 2006. Transferability of tag SNPs in genetic association studies in multiple populations. Nat Genet 38:1298–1303.

Gu S, Pakstis AJ, Li H, Speed WC, Kidd JR, Kidd KK. 2007. Significant variation in haplotype block structure but conservation in tagSNP patterns among global populations. Eur J Hum Genet 15:302–312.

The International HapMap Consortium. 2003. The International HapMap project. Nature 426:789–796.

Montpetit A, Nelis M, Laflamme P, Magi R, Ke XY, Remm M, Cardon L, Hudson TJ, Metspalu A. 2006. An evaluation of the performance of tag SNPs derived from HapMap in a Caucasian population. PLOS Genet 2:282–290.

Moskvina V, O'Donovan MC. 2007. Detailed analysis of the relative power of direct and indirect association studies and the implications for their interpretation. Hum Hered 64:63–73.

Nielsen DM, Weir BS. 2001. Association studies under general disease models. Theor Popul Biol 60:253–263.

Nielsen DM, Ehm MG, Zaykin DV, Weir BS. 2004. Effect of two and three locus linkage disequilibrium on the power to detect marker/phenotype associations. Genetics 168:1029–1040.

Pritchard JK, Przeworski M. 2001. Linkage disequilibrium in humans: models and data. Am J Hum Genet 69:1–14.

Service S, Sabatti C, Freimer N. 2007. Tag SNPs chosen from HapMap perform well in several population isolates. Genet Epidemiol 31:189–194.

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

Terwilliger JD. 2006. An utter refutation of the 'Fundamental Theorem of the HapMap'. Eur J Hum Genet 14:426–437.

Weinberg CR, Wilcox AJ, Lie RT. 1998. A log-linear approach to case-parent-triad data: assessing effects of disease genes that act either directly or through maternal effects and that may be subject to parental imprinting. Am J Hum Genet 62:969–978.