Statistical Applications in Genetics and Molecular Biology

Volume 9, *Issue* 1 2010 *Article* 35

Assessment of LD Matrix Measures for the Analysis of Biological Pathway Association

David R. Crosslin* Xuejun Qin†

Elizabeth R. Hauser‡

*University of Washington, davidcr@u.washington.edu

†Duke University Medical Center, xjq@chg.duhs.duke.edu

‡Duke University Medical Center, hause006@chg.duhs.duke.edu

Assessment of LD Matrix Measures for the Analysis of Biological Pathway Association

David R. Crosslin, Xuejun Qin, and Elizabeth R. Hauser

Abstract

Complex diseases will have multiple functional sites, and it will be invaluable to understand the cross-locus interaction in terms of linkage disequilibrium (LD) between those sites (epistasis) in addition to the haplotype-LD effects. We investigated the statistical properties of a class of matrix-based statistics to assess this epistasis. These statistical methods include two LD contrast tests (Zaykin et al., 2006) and partial least squares regression (Wang et al., 2008). To estimate Type 1 error rates and power, we simulated multiple two-variant disease models using the SIMLA software package. SIMLA allows for the joint action of up to two disease genes in the simulated data with all possible multiplicative interaction effects between them. Our goal was to detect an interaction between multiple disease-causing variants by means of their linkage disequilibrium (LD) patterns with other markers. We measured the effects of marginal disease effect size, haplotype LD, disease prevalence and minor allele frequency have on cross-locus interaction (epistasis).

In the setting of strong allele effects and strong interaction, the correlation between the two disease genes was weak ($r = 0.2$). In a complex system with multiple correlations (both marginal and interaction), it was difficult to determine the source of a significant result. Despite these complications, the partial least squares and modified LD contrast methods maintained adequate power to detect the epistatic effects; however, for many of the analyses we often could not separate interaction from a strong marginal effect. While we did not exhaust the entire parameter space of possible models, we do provide guidance on the effects that population parameters have on cross-locus interaction.

KEYWORDS: epistasis, linkage disequilibrium, complex disease, cardiovascular disease

1 Introduction

Genotyping technologies are providing unprecedented SNP coverage of the human genome, but the inference from those analyses remains difficult. While SNPs have different functional profiles, they are most often used to mark genetic variation within a gene or across the genome. Understanding which if any of the many SNPs represent a contribution to disease risk can be difficult. Furthermore, the concept of defining a gene measure based on multiple SNPs is not well-developed. The probability of declaring SNPs to be significantly associated to a phenotype if they are truly associated while keeping the probability of making false declarations of association low remains a difficult task (Lee and Whitmore, 2002). A large sample size is needed to detect the smaller gene effect sizes, which are normally overlooked when controlling for multiple testing (Loos et al., 2008). These may play an important role in understanding complex pathways leading to disease phenotypes. These effects may also hinder candidate gene replication.

One extension of correlating single-marker genotypes with phenotypes (allelic association) is the characterization of multiple SNP patterns known as haplotypes. Haplotypes are specific combinations of alleles aligned (in phase) along a chromosomal region. Haplotypes may provide more information on the complex relationship between DNA or whole-gene variation and disease-related phenotypes compared to a single SNP (Schaid et al., 2002, Stephens et al., 2001). Specific allele combinations along a chromosome for a gene may provide more information than a single SNP. Haplotype analyses localize a susceptibility gene via linkage disequilibrium (LD) with adjacent genetic markers or on the influence that the entire haplotype has on the trait (Schaid et al., 2002). LD is the non-random association of two alleles at different loci on the same gamete.

When selecting which SNPs to represent a haplotype, one must consider many factors. The two main approaches to haplotype construction are either choosing SNPs within (intra) or across (inter) LD bins, thus producing a higher-order haplotype. The intra-LD bin haplotypes are a set of alleles together on the same chromosome surrounded by recombination hotspots defined by a correlation threshold. Higher-order haplotypes use tagging SNPs (tSNP) to capture all variation at a given LD level. The crucial subset of markers to type would be those that distinguish one haplotype from another (Zondervan and Cardon, 2004). Also, many methods use windowing strategies which test sets of perhaps contiguous SNPs. Obstacles with this method include non-contiguous LD patterns and unevenly spaced SNPs throughout the genome. Understanding the benefits and shortfalls with each method are important when considering analysis strategies.

Another extension of single SNP methods is SNP tagging. For a given sample of a population, tagging SNPs are used to represent blocks of LD or haplotypes

by a threshold of correlation. The idea is that a few SNPs can represent larger regions of a chromosome or gene of interest, thus reducing the number needed to genotype. SNP tagging is common practice with gene fine mapping and many GWAs panels are based on this format. Understanding what effect gene-tagging and the resulting LD pattern have on detecting gene-gene interactions will be beneficial. It is also important to note that the widely used strategies for tagging SNP (tSNP) selection are based on a single-disease-gene model, which may not apply to complex disease (Zhao et al., 2006). This approach assumes a common variant that is associated with a common disease with no interaction effects. The ability to detect interaction effects (gene-gene) may be attenuated with tagging strategies.

In addition to the one-to-many relationship of multiple SNPs per gene, there is an added layer of complexity when considering multiple genes in a metabolic pathway. An example of a complex metabolic pathway is the leukotriene biosynthesis pathway in coronary artery disease (CAD) (Crosslin et al., 2009). Leukotrienes are arachidonic acid derivatives long known for their inflammatory properties and their involvement with a number of complex human diseases, most particularly asthma. Several genetic linkage and associations studies as well as gene expression studies have shown an association of the leukotriene biosynthesis pathway to CAD (Dwyer et al., 2004, Helgadottir et al., 2004, 2005b,a, Lohmussaar et al., 2005, Shah et al., 2008). Helgadottir et al. (2004) identified a four-SNP haplotype (HapA) spanning the *ALOX5AP* gene that conferred a nearly two times greater risk of MI and stroke in a case-control cohort. The *ALOX5AP* gene is mapped to a locus on chromosome 13q12.3 and encodes a protein that, with 5-lipoxygenase (*ALOX5*), is required for leukotriene synthesis. A second haplotype, HapB, was also identified in *ALOX5AP*, although the risk conferred by this haplotype was considerably smaller and not replicated. In addition, a ten-SNP haplotype (HapK) spanning the *LTA4H* gene encoding leukotriene A4 hydrolase (*LTA4H*), a protein in the same biological pathway as *ALOX5AP*, was shown to confer a modest risk of MI in the Icelandic cohort (Helgadottir et al., 2005b). The multiple haplotypes suggest two independent disease-causing mutations, but also may indicate gene involvement within a pathway context. While haplotyping does characterize a single gene's association with disease using multiple SNPs, no gene-gene interaction is considered. Like single-SNP effects, smaller effects from multiple haplotypes for a given pathway may play an important role in more complex pathway analyses. Harnessing the additive effects and gaining power from considering multiple SNPs in multiple genes in a biological pathway could aid in the understanding of complex diseases like CAD.

DOI: 10.2202/1544-6115.1561

1.1 Tests for interaction generalizing LD measures

More generalized correlation between SNPs can be due to interactions and functional relationships related to the disease phenotype, or to other factors such as population genetic history and sampling on case status. Joint frequencies of alleles at different loci can be measured using a specific type of correlation called linkage disequilibrium (LD). LD is the non-random association of two alleles at different loci. As a generalized measure of disequilibrium or alternatively correlation at two loci, LD patterns may be useful for assessing a measure of association (Zaykin et al., 2006, Wang et al., 2007, Zhao et al., 2006).

Zaykin et al. (2006) presented two methods for contrasting LD patterns. The first method is a permutation-based statistic (Z_1) that measures the difference between two spaces (e.g. the case LD space and the control LD space) using spectral decomposition of the pairwise LD matrix. The second method is a sum-of-squared-differences statistic (Z_2) that measures the overall difference in the corresponding pairwise LD. Simulations presented by Zaykin et al. (2006) showed that the sum-of-squared-differences statistic (Z_2) was the more powerful statistic for a single gene disease model.

Zaykin et al. (2006) evaluated performance of the whole-matrix LD statistic by comparing methods designed to detect either single-SNP effects or SNP interactions when the effects are associated with entire haplotypes, but these tests were based on single variant disease models. There is strong evidence that several mutations within a single gene can interact to create a "super allele" that has a large effect on the observed phenotype (Schaid et al., 2002). In a separate study, Zhao et al. (2006) presented a similar test of measuring LD differences to test the interaction effects among cases and controls. Interestingly, Zhao et al. (2006) showed that interaction between two unlinked loci can be represented by LD.

Complex diseases will have multiple functional sites, and it will be invaluable to understand the cross-locus interaction in terms of correlation between those sites in addition to the within-gene (or within locus) LD effect. Zhao et al. (2006) indicated that the method proposed by Zaykin et al. (2006) is similar to their method, but that while the LD contrast test was originally designed to test association between a single gene and disease, it has not been extended to testing gene-gene interaction. The interaction will indicate the joint action of two genes in the development of disease (Zhao et al., 2006). Modeling a trait as an additive combination of single-locus and interaction terms is likely to limit the power to detect interaction and a combined measure may be more powerful (Zhao et al., 2006). Because the criteria for tSNP selection are based on only one pairwise LD measure between the marker and disease locus, the LD between tSNPs and loci may not be strong enough to ensure that indirect interaction between two loci will be detected. Thus,

if the interacting loci are not selected as tSNPs, many loci with interactions will be missed (Zhao et al., 2006).

Wang et al. (2008) presented a partial least-square (PLS) approach for modeling gene-gene as well as gene-environment interactions with multiple markers. In this study they compared the PLS approach to other methods to jointly test the interaction (termed as cross-locus gametic disequilibrium) and main effects. More importantly, they described advantages and disadvantages compared to other methods including Tukey's one-degree-of-freedom model, logistic regression with factors generated from principal component analysis (PCA) and baseline logistic regression both with and without interaction terms. The arguments outlined by Wang et al. (2008) present a case for developing alternative methods for modeling gene-gene interaction using multiple SNPs. For instance, using logistic regression to exhaustively model all pair-wise interactions between two genes could present a large number of degrees of freedom with a large number of parameters (Wang et al., 2008). While PCA is a commonly used tool for data reduction of predictors, no correlations of SNPs with the trait are characterized. So modeling a trait using logistic regression with the first few components may not perform well in certain LD structures (Wang et al., 2008). These first few components may be driven by correlation (LD) not associated with the trait. Tukey's approach is based on the assumption that a SNP's interaction effect on the trait is approximately proportional to it's marginal effects on the trait. The model is not optimal in power when there exists no, or small, marginal effects (Wang et al., 2008). The methods presented are tested on simulated data sets with the assumption that the causal variants are genotyped, and then tested with the assumption that the variants are not genotyped. Expanding on this approach will help converge on a model for complex diseases whose etiology is most likely derived from the interaction of multiple genes and environmental risk factors that ultimately affect metabolic states.

We performed a comprehensive simulation study to evaluate LD matrix-based measures in case-control data under a two-locus genetic model. We assessed how the matrix measures performed in measuring cross-locus interaction in the presence of varying LD (including equilibrium), MAF and marginal relative risk (*RR*). We hypothesized that we could detect an interaction between two unobserved disease-causing variants by means of their LD patterns with other markers. This is a subtle but important difference from measuring different LD patterns with the two disease-causing variants included in the model. This could obviously produce significant results, depending on the magnitude of the difference. In effect, our goal was to determine if the interaction in terms of *RR* can significantly change the LD patterns across SNP markers composing different haplotypes, such as those seen in the analysis of CAD association with *ALOX5AP* and *LTA4H* (Helgadottir et al., 2004, 2005b, Crosslin et al., 2009).

DOI: 10.2202/1544-6115.1561

In summary, we evaluated these matrix-based methods in the setting of multiple risk variants in a simulation study. With the knowledge gained from the simulations, we explored using the matrix-based measures to determine if there was an interaction between haplotypes in the leukotriene biosynthesis pathway. Having the prior knowledge of haplotype and single SNP association results for the leukotriene pathway provided an excellent application for the matrix-based methods results. Our results suggest that matrix methods are powerful methods to detect complex haplotype interactions among multiple unmeasured disease-causing variants.

2 Methods

We investigated the statistical properties of a class of matrix-based statistics (Section 1.1) to detect interaction between two loci (epistasis) in simulated data. We define epistasis as the departure from multiplicativity on the penetrance scale (Cordell, 2002). Our goal was to detect an interaction between multiple disease-causing variants by means of their linkage disequilibrium (LD) patterns with other haplotype markers. We considered two statistics based on contrasting pairwise matrices of LD patterns between cases and controls. The statistics include Zaykin's LD permutation-based test (Z_1) and the LD contrast test (Z_2) (Zaykin et al., 2006). Next, we considered Wang's interaction test using partial least squares (PLS) (Wang et al., 2008).

Our goal was to approximate models of genetic interaction of two loci involving multiple markers composing two separate haplotypes. We simulated multiple two-variant disease models with haplotypes to gain an understanding of pathway interactions in terms of correlation patterns. Our first goal was to model a complex disease, with the potential for multiple risk variants in the same gene or in different genes which may or may not interact to increase disease risk, all in the presence of LD. For a complex trait system independent of epistasis, it is useful to consider four framework parameters (Zondervan and Cardon, 2004):

1. The effect size of a disease locus
   - Relative Risk (RR) = $\frac{Pr(\mathrm{D}|\text{risk allele})}{Pr(\mathrm{D}|\text{normalallele})}$
2. The frequency of disease allele(s)
3. The frequency of marker allele(s)
4. The extent of linkage disequilibrium (LD) between the marker and disease locus
   - Correlation coefficient (r)

These parameters were varied to create different null and genetic interaction models.

2.1 Sample simulations

SIMLA can simulate two bi-allelic disease loci spread over three chromosomes, with the third chromosome providing the option to analyze markers completely unlinked to disease loci. Blocks of LD can be simulated by selecting a subset of markers (SNPs) to be in LD with a disease locus, while other markers are in linkage equilibrium with the disease locus (Schmidt et al., 2005). The disease risk model we simulated included two disease loci, each of which was included in distinct three-SNP haplotypes. Figure 1a illustrates a two-variant (V_1 & V_2) disease model in a single disease gene, for example HapA and HapB in *ALOX5AP*. Each of the two variants was simulated at independent levels of *RR*, and was in LD with two independent sets of markers forming an independent haplotype. There may also be interaction (epistasis) between the variants in terms of *RR*. This can also be illustrated in the context of pathway interaction (Figure 1b). Each variant is hypothetically located in separate disease genes with equal marginal effects (*RR*) and an interaction (epistasis) effect. We varied the values of these parameters for the simulation as described in Table 1.

2.1.1 Null and alternative disease models

Null models were simulated to evaluate Type I error rates. Based on a nominal significance level of $\alpha = 0.05$, we calculated the proportion of rejections to evaluate whether the size/level of a test achieves the nominal α when the null hypothesis is true (Davidian, Spring 2005). To evaluate power, we generated alternative models. We calculated the proportion of rejections based on the nominal significance level of $\alpha = 0.05$ when in fact the alternative hypothesis is true.

Each disease locus was generated with varying LD with genotypes assigned according to the specific LD pattern (see Table 1). Because in real data analysis the haplotype phase cannot be directly observed, the composite disequilibrium $\widehat{D}_{AB}$ and correlation $\widehat{r}$ were calculated using dilocus counts and sample allele frequencies (Weir and Cockerham, 1979, Weir, 1996, Zaykin, 2004, Zaykin et al., 2006). Zaykin et al. (2006) emphasizes that with the EM algorithm, the likelihood is constructed assuming HWE on the level of haplotypes that cannot be inferred from single SNP HWE assessment. The composite disequilibrium approach provides results similar to those of the EM-based method under HWE, is computationally simpler, and avoids the assumption of haplotypic HWE (Zaykin et al., 2006):

DOI: 10.2202/1544-6115.1561

$$\widehat{D}_{AB} = \frac{1}{n}(2n_{AABB} + n_{AABb} + n_{AaBB} + \frac{1}{2}n_{AaBb}) - 2\tilde{p}_A\tilde{p}_B \quad (1)$$

In this equation $\tilde{p}_A$ and $\tilde{p}_B$ are sample allele frequencies and n_{AABB} is the dilocus sample genotype count. The composite correlation is given as

$$\widehat{r} = \frac{D_{AB}}{\sqrt{(P_A \cdot P_a + D_A)(P_B \cdot P_b + D_B)}} \quad (2)$$

where $D_A = P_{AA} - p_A^2$ is the Hardy-Weinberg disequilibrium coefficient.

Pairwise composite correlation $r_{composite}$ matrices were generated for input data. In the statistical analyses, we assumed that the disease variants V_1 and V_2 were not genotyped; however, the genotypes were generated to measure the interaction in terms of $r_{composite}$ and to validate the parameter settings for SIMLA.

Null models Null models were used to investigate the statistical properties of the tests under the null hypothesis of no interaction. The central null condition is the absence of any interaction ($Interaction_{RR} = 1.0$) between the two variants with and without marginal effects. There is also a null model in the presence of LD, but no marginal effects. The null models can be summarized into three categories (Table 1): 1. Null models with no marginal effects - V_1 with disease $RR = 1.0$, V_2 with disease $RR = 1.0$ and no LD with other markers (model 1); 2. Null models with marginal effects V_1 and V_2 $RR > 1.0$ and no LD with other markers (models 2 & 3); and 3. Null models in the presence of LD with other markers and no marginal effects, $RR = 1.0$ (model 4). As shown in Table 1, all null models have interaction *RR* set to 1.0. For several of the null models, there are genetic effects, but there is no interaction between the variants.

Alternative models Alternative models were used to evaluate power of the matrix measures and to identify marginal and interaction effects in complex disease models. The alternative models can be summarized into three categories (Table 1): 1. Alternative models in the presence of LD, marginal effects and no interaction (models 5 & 6); 2. Alternative models in the presence of LD, interaction effects and no marginal effects (model 7); and 3. Alternative models in the presence of LD, marginal effects and interaction (models 8 & 9). For models 7, 8 and 9, we generated a *RR* interaction of 3.0 and 10.0. We also considered the effect of varying allele frequencies.

LD and effect simulations The *haplotype_select* feature in SIMLA was used to simulate two three-SNP haplotypes with two markers (M_1 and M_2), and one disease variant (V) in LD, thus generating a six SNP system. We used the multiplicative model where risk of developing disease is increased by a factor for each allele carried. If the homozygote susceptibility $k = RR(V_iV_i) > 1.0$, then $ln(RR(V_iV_j)) = 0.5 \times ln(k)$ for the heterozygote risk. Given an already assigned V_i (susceptibility) or V_j (non-susceptibility) allele at the disease variant locus, a marker haplotype (set of alleles) for all individual founder chromosomes were randomly generated based on the conditional haplotype probabilities. For each haplotype, marker M_1 is in LD with disease variant V, in addition to LD with M_2 (see Figure 1a). LD cannot be controlled between M_2 and the disease variant V (see figure 1a) given SIMLA's algorithm for generating haplotypes.

In Equation 3, we set $p(M_{1i}, M_{2i}, V_i)$ for alleles i and subsequently j to maximize the three locus correlation r^2 (Equation 2). The D symbol is the composite disequilibrium (Equation 1).

$$r^2(M_{1i}, M_{2i}, V_i) = \frac{D^2(M_{1i}, M_{2i}, V_i)}{var(D(M_{1i}, M_{2i}, V_i))} \tag{3}$$

where

$$\begin{aligned} D(M_{1i}, M_{2i}, V_i) = p(M_{1i}, M_{2i}, V_i) - p(M_{2i})D(M_{1i}, V_i) - \\ p(V_i)D(M_{1i}, M_{2i}) - p(M_{1i})D(M_{2i}, V_i) - p(M_{1i})p(M_{2i})p(V_i) \end{aligned} \tag{4}$$

Interaction was simulated in terms of the *RR* in addition to marginal effect *RR*. The susceptibility variants V_1 and V_2 were generated, but were not used in the analysis; LD was measured between the disease variants to validate the simulation parameters. Chromosome-specific genetic maps with inter-marker distances were specified in Morgans using the Haldane map function. While maintaining an interaction *RR* of 1.0 between variant 1 (V_1) and variant 2 (V_2), we simulated minor allele frequencies (MAF) of $0.05, 0.15, 0.35$ and 0.50 for the variants and the haplotype markers associated with each (see Table 1). Allele frequencies for V_1,V_2 and all haplotype markers were kept equivalent. LD between the haplotype-specific markers was generated at $r^2 = 0.4, 0.7$ and 0.9 with equal values for both haplotypes. We then considered the effect of increasing *RR* for the two variants. We generated marginal *RR* of $1.0, 1.5$ and 3.0 effect sizes for our models. We generated a prevalence frequency of 0.10 for all models. Finally, we generated a *RR* interaction between both variants of $1.0, 3.0$ and 10.0.

DOI: 10.2202/1544-6115.1561

For each parameter set, we generated samples of 500 cases and 500 controls. We simulated 1000 replicates for each model listed in Table 1. We compared mean values and standard deviations to assess the stability of the measures. Type I error rates and empirical power were estimated using the proportion of achieved significance levels $<= 0.05$ based on 1000 simulations of the null and alternative models respectively.

2.2 LD contrast test statistics

2.2.1 LD-contrast and eigenvector permutation-based tests

As square matrices, composite-LD matrices have spectral decompositions which can be reduced to canonical form, represented by eigenvalues and eigenvectors. Following Zaykin et al. (2006), we considered two statistics based on the eigenvectors and eigenvalues for the composite LD matrix. The Z_1 statistic measures the difference between two spaces defined by the first k eigenvectors with the sum of squared cosines of the angles (θ) between the eigenvectors. The matrix E_i is formed by the first k eigenvectors. This statistic and the subsequent permutation-based test were originally described by Krzanowski (Zaykin et al., 2006, Krzanowski, 1979, 1993):

$$Z_1 = \sum_{i=1}^{k} \lambda_i = traceTT' = \sum_{i=1}^{k}\sum_{j=1}^{k} cos^2\theta_{ij} \tag{5}$$

In this equation the value $T = E_{case}E'_{control}$. The values i and j represent the principal components of two separate groups. Krzanowski described this sum as a measure of similarity between two spaces. The value can lie between k (coincident spaces) and 0 (orthogonal spaces) Krzanowski (1979). This can be interpreted as a geometric correlation. Because our simulated matrix consists of four SNPs, we generated vectors or matrices consisting of one, two and three eigenvectors. The tests were implemented in SAS/IML. The sum-of-squared-differences statistic (Z_2) utilizes the overall sample mean to center the genotypes and measures the overall difference in the corresponding pairwise LD where r is the correlation matrix:

$$Z_2 = trace[(\hat{r}_{case} - \hat{r}_{control})^T(\hat{r}_{case} - \hat{r}_{control})] \tag{6}$$

2.2.2 Partial least squares

Finally, we considered the partial least squares (PLS) approach presented by Wang et al. (2008) for modeling gene-gene interactions with multiple markers. PLS accounts for spectral decomposition of response values in addition to the correlation patterns of the predictor variables. For binary response variables (case-control status), the PLS is a discriminant analysis and identifies the linear combination of the predictors that best explains the difference between case and control status. Wang et al. (2008) suggest detecting an association by a likelihood ratio test based on a logistic regression model:

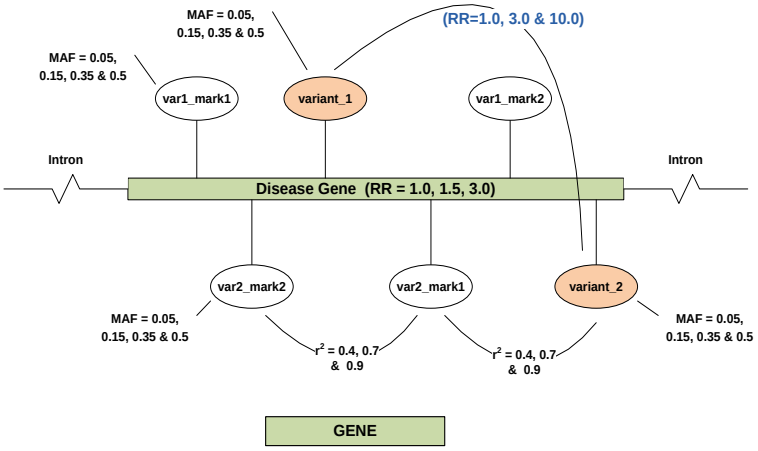
$$logit(Pr(D)) = \beta_0 + \sum_{i=1}^{k_1} \beta_{1i} s_{1i} + \beta_{g1g2} T_1^{1PLS} T_2^{1PLS} \tag{7}$$

or

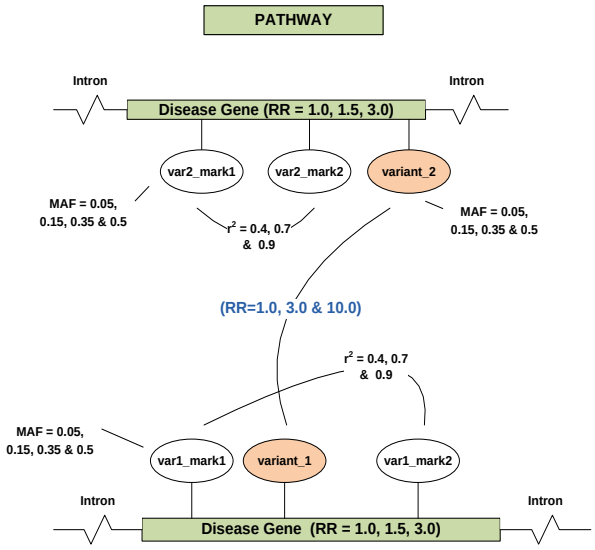
$$logit(Pr(D)) = \beta_0 + \sum_{j=1}^{k_2} \beta_{2j} s_{2j} + \beta_{g1g2} T_1^{1PLS} T_2^{1PLS} \tag{8}$$

In this equation $k =$ the number of SNPs for a given gene 1 or 2, or in our example disease causing haplotypes 1 or 2. This two-variant disease model is illustrated as a single gene in Figure 1a, and as separate genes in Figure 1b. The $\beta_{g1g2} T_1^{1PLS} T_2^{1PLS}$ term is the interaction between the first factor for each gene / haplotype 1 and 2. We reported both the likelihood ratio-based results testing the global null hypothesis ($\beta = 0$) as well as the single effect significance for the haplotype interaction term.

DOI: 10.2202/1544-6115.1561



(a) Gene annotation for interaction simulations using the SIMLA package.



(b) Pathway annotation for interaction simulations using the SIMLA package.

Figure 1: Gene and pathway annotation for interaction simulations using the SIMLA package.

Table 1: Descriptive summary of the nine models for simulation using the package SIMLA Schmidt et al. (2005). Each model was simulated with MAF = 0.05, 0.15, 0.35 and 0.50 at a prevalence of 0.10.

Model	$LD(r^2)$ V_1M_1	$LD(r^2)$ V_1M_2	$LD(r^2)$ V_2M_1	$LD(r^2)$ V_2M_2	RR V_1	RR V_2	RR $V_1 \times V_2$	MAF
Null models with no LD, marginal or interaction effects								
1	0	0	0	0	1.0	1.0	1.0	0.05, 0.15 0.35, 0.50
Null models with marginal effects, but no LD or interaction effects								
2	0	0	0	0	1.5	1.5	1.0	0.05, 0.15 0.35, 0.50
3	0	0	0	0	3.0	3.0	1.0	0.05, 0.15 0.35, 0.50
Null models in the presence of LD, but no marginal or interaction effects								
4	0.4 0.7 0.9	0.4 0.7 0.9	0.4 0.7 0.9	0.4 0.7 0.9	1.0	1.0	1.0	0.05, 0.15 0.35, 0.50
Alternative models in the presence of LD, marginal effects and no interaction								
5	0.4 0.7 0.9	0.4 0.7 0.9	0.4 0.7 0.9	0.4 0.7 0.9	1.5	1.5	1.0	0.05, 0.15 0.35, 0.50
6	0.4 0.7 0.9	0.4 0.7 0.9	0.4 0.7 0.9	0.4 0.7 0.9	3.0	3.0	1.0	0.05, 0, 15 0.35, 0.50
Alternative models in the presence of LD, interaction and no marginal effects								
7	0.4 0.7 0.9	0.4 0.7 0.9	0.4 0.7 0.9	0.4 0.7 0.9	1.0	1.0	3.0 10.0	0.05, 0.15 0.35, 0.50
Alternative models in the presence of LD, marginal effects and interaction								
8	0.4 0.7 0.9	0.4 0.7 0.9	0.4 0.7 0.9	0.4 0.7 0.9	1.5	1.5	3.0 10.0	0.05, 0.15 0.35, 0.50
9	0.4 0.7 0.9	0.4 0.7 0.9	0.4 0.7 0.9	0.4 0.7 0.9	3.0	3.0	3.0 10.0	0.05, 0.15 0.35, 0.50

DOI: 10.2202/1544-6115.1561

3 Results

3.1 Evaluation of simulations

We evaluated the realized correlation matrices for our simulated data sets. While we did not exhaust the model space, we did observe trends in correlation patterns between the two disease variants V_1 and V_2. The results for realized correlation patterns are presented in Section 3.1.1. We then evaluated the matrix measures statistics for testing marginal and interaction effects in these complex models.

3.1.1 Composite correlation for the two variant disease model

Tables 2, 3 and 4 contain the composite correlation between variants V_1 and V_2 for both cases and controls. While the variants were not included in the correlation matrices assessed, we wanted to understand what effects the disease effect size (relative risk *RR*), haplotype LD, disease prevalence and MAF have on cross-locus interaction (Figures 1a & 1b). The table is divided into three sections for marginal disease *RR* of $1.0, 1.5$ and 3.0 for each variant. Within each section there are four columns representing MAF of $0.5, 0.15, 0.35$ and 0.50 for both markers and variants. Each row represents *LD* in terms of correlation for the haplotypes with values $0.0, 0.4, 0.7$ and 0.9. Table 2 contains the correlations for the null models. These models can be summarized into three categories (Table 1): 1. Null models with no marginal effects and no LD (model 1); 2. Null models with marginal effects and no LD (model 2 & 3); and 3. Null models in the presence of LD and no marginal effects (model 4). Table 2 also contain the correlations for alternative models 5 and 6 (Table 1). These alternative models are in the presence of LD, marginal effects but no interaction.

Tables 3 and 4 contain the composite correlations for the interaction *RR* of 3.0 and 10.0 respectively. Because these are alternate models with interaction $RR > 1.0$, we did not consider $r^2 = 0.0$ for the haplotypes. These alternate models can be summarized into two categories (Table 1): 1. Alternate models in the presence of LD, marginal effects and no interaction; and 2. Alternate models in the presence of LD, marginal effects and interaction.

Figure 2 illustrates the MAF ($x - axis$) by case composite correlations ($y - axis$) for the interaction *RR* of $1.0, 3.0$ and 10.0 represented as a solid, dashed and dotted line respectively. Each trellis window represents a different combination of marginal disease *RR* of $1.0, 1.5$ and 3.0 (moving up the trellis), and LD in terms of r^2 for each haplotype (moving across the trellis) of $0.4, 0.7$ and 0.9. At an interaction *RR* of 1.0 (solid line), we detected a composite correlation close to 0.0 for almost every combination of marginal relative risk and haplotype LD as evident throughout

the trellis plot in Figure 2. However, there is a slight dip below 0.0 in correlation that is illustrated in the top row as a marginal *RR* of 3.0 as MAF increases from 0.05 to 0.5. This is most likely due to the case selection bias with a marginal *RR* of 3.0. At a marginal $RR \geq 1.0$, a single disease variant is sufficient to cause disease and thus to be sampled as a case.

For an interaction *RR* of 3.0 (dashed line) in Figure 2, we observed different interaction effects (r^2) based on different combinations of parameters (MAF, LD and *RR*). As we increase the marginal *RR* from 1.0 to 1.5 (bottom and middle row of trellis), there are similar effects on the interaction correlation with the maximum of approximately 0.10 at MAF 0.35 and 0.5. At *RR* of 3.0, the correlation is lower (approximately 0.05) because of the case selection bias as described above.

For an interaction *RR* of 10.0 (dotted line) in Figure 2, we also observed similar effects as we increased the marginal *RR* from 1.0 to 1.5 (bottom and middle row of trellis). This is also where we observed the maximum composite correlation of approximately 0.20 at MAF = 0.35. In all of the trellis windows (every combination of LD and marginal *RR*) we observed a decrease in interaction correlation as we increased MAF from 0.35 to 0.5. This is likely due to the fact that we have equal representation of each allele in the population and the chance that a case subject is selected based on the presence of variants at either V_1 or V_2 is greater in addition to the joint effects (interaction) disease. Cases could have disease due to variant 1, variant 2 or both.

3.1.2 Matrix measures results

Z_1 permutation test Supplementary Table 8 contains the Type I error rates for the Z_1 permutation-based test using the null models with no LD (models 1, 2 & 3 in Table 1). Supplementary Table 9 contains the Type I error rates for the Z_1 permutation-based test using the null models in the presence of LD and no marginal or interaction effects (model 4 in Table 1). The Type I error rate remained close to the nominal significance level of 0.05 for all simulation conditions.

Figure 3a illustrates empirical power for the Z_1 permutation-based test where LD is present and no interaction *RR*, but the marginal $RR > 1.0$ (models 5 and 6 in Table 1). This model is testing the correlation differences between cases and controls due to the correlation (LD) between haplotype markers and the ungenotyped disease variant. For the four-snp matrix, we considered featurevector of $k = 1,2$ and 3 eigenvalues, but we only present $k = 1$ and 2. Figure 3b illustrates $k = 1$ or one eigenvector.

The respective correlation between the two variants can be referenced in Figure 2. Like the interaction LD plot (Figure 2), each window in the trellis plot

DOI: 10.2202/1544-6115.1561

is represented by a marginal *RR* (moving up the trellis) and haplotype LD (moving across the trellis) in terms of r^2. The maximum power for the Z_1 test was 0.11, thus the strong marginal effects ($RR = 3.0$) and high disease marker-variant LD $r^2 = 0.9$ appears to have little effect on the power of the Z_1 test statistic.

Figure 3b illustrates empirical power for the Z_1 permutation-based test where interaction $RR > 1.0$ (models 7, 8 and 9 in Table 1). The interaction *RR* of 3.0 and 10.0 are represented by a solid and dashed line respectively. For the interaction *RR* of 3.0 (solid line) in Figure 3b, we observed an increase in power to detect interaction between the variants. Zaykin et al. (2006) indicated that both the featurevector-permutation (Z_1) and LD-contrast (Z_2) have limitations when allele frequencies are low; lower MAF can give rise to spurious correlation structures. Our results illustrated in Figure 3b at MAF = 0.05, and for the most part 0.15, support those findings. At low haplotype LD ($r^2 = 0.4$; left column of trellis), the power to detect an interaction is low (< 0.20). As haplotype LD is increased (LD = 0.9, *RR* = 3.0) in Figure 3b, the power to detect an interaction at MAF = 0.5 is 0.70, which is the maximum power across all simulation conditions for *RR* = 3.0.

At an interaction *RR* of 10.0 (dashed line) in Figure 3b, we observed an increase in power throughout all combinations of parameter settings. As MAF increases, we observed an increase in power. There is also an increase in power as LD is increased. The increase of marginal *RR* (moving up the trellis) has less of an effect on power as compared to the effects of LD of marker-variant and MAF. However, the power for the Z_1 permutation-based test to detect an interaction between V_1 and V_2 increases as both parameters are increased. At higher LD = 0.7 and 0.9, we observe power > 0.80 for MAF 0.35 and 0.5 when marginal $RR > 1.0$. The empirical power plots for the Z_1 permutation-based test mimic the interaction LD patterns illustrated in Figure 2 suggesting that this measure is appropriately sensitive to the extent of correlation between the disease variants.

Figure 4 illustrates $k = 2$ eigenvectors. There is a total loss of power to detect interaction between the disease variants V_1 and V_2 when $k = 2$. As Zaykin et al. (2006) indicated, Krzanowski suggested using the value of k that is the largest integer smaller than $\frac{L}{2}$ where in our case L is the total number of markers. Values $k \geq L/2$ will cause the subspaces defined by the two sets of subspaces to intersect in at least one dimension (Zaykin et al., 2006, Krzanowski, 1979). Our results in Figure 4 support this fact. We also weighted the eigenvectors by their corresponding eigenvalues (Figures 5a and 5b) to assess a benefit of creating a $k = 2$ matrix by the level of significance. The eigenvectors accounting for more of the variance as defined by the eigenvalues are given greater weight.

This indicates that the second eigenvector contributes minimal information with respect to the first, which seems reasonable for the small matrices here. The empirical power where LD is present and no interaction *RR*, but the marginal

$RR > 1.0$ (models 5 and 6 1) is different than the unweighted $k = 1$. For the weighted $k = 2$, Z_1 method, power approaches 0.10 for $RR = 1.5$. For $RR = 3.0$, the power increases as MAF increases, but there is a decrease of power as LD increases (moving across the trellis). For the the interaction $RR > 1.0$ (models 7, 8 & 9 in Table 1), there is improved power compared to the $k = 1$ approach. Adding an additional eigenvector and weighting by significance (eigenvalue) increased power under this disease model.

Z_2 LD contrast test Supplementary Table 8 contains the Type I error rates for the Z_2 permutation-based test using the null models with no LD (models 1, 2 & 3 in Table 1). The Type I error remained close to the nominal significance level of 0.05. Supplementary Table 9 contains the Type I error rates for the Z_1 permutation-based test using the null models in the presence of LD and no marginal or interaction effects (model 4 in Table 1). Like the Z_1, the Type I error remains close to the nominal significance level of 0.05.

Figure 6a illustrates empirical power for the Z_2 permutation-based test where LD is present and no interaction *RR*, but the marginal $RR > 1.0$ (models 5 and 6 in Table 1). At LD 0.4 and MAF of 0.5., we observed an increase in power approaching 0.23. All other combination of parameters had very low power (0.05).

Figure 6b illustrates empirical power for the Z_2 LD contrast test where interaction $RR > 1.0$ (models 7, 8 and 9 in Table 1). For an interaction *RR* of 3.0 (solid line), we observed mixed results. The empirical power does not exceed 0.40 for any combination of marginal *RR* (moving up the trellis) and haplotype LD (moving across the trellis) except for the top left panel (LD = 0.4, *RR* = 3.0). At MAF = 0.35 and 0.5, the observed power is approximately 0.5 and 0.5. In the presence of low LD (r^2), high marginal *RR*(3.0) and high MAF (> 0.35) this test seems to perform well. For an interaction *RR* of 10.0 (dotted line), we observed higher power. As Zaykin et al. (2006) suggest, the LD contrast test is optimal at higher MAF which we observed. At higher LD (middle and right column) of $r^2 = 0.7$ and 0.9, the power results strongly mimic the interaction correlation pattern we observed in Figure 2. Interestingly, at this level of LD, we observed diminishing effectiveness as the marginal *RR* of 1.0 (power is approximately 0.8) is increased to 3.0 (power is approximately 0.6). At a lower level of LD ($r^2 = 0.4$), we observe the opposite effect (power increases from ~ 0.6 to ~ 1.0). The higher marginal effects in the presence of higher haplotype LD masked the interaction correlation, which we attribute to the high relative risk conferred by a single variant.

Partial least squares approach Supplementary Table 8 contains the Type I error rates for the PLS approach using the null models with no LD (models 1, 2

DOI: 10.2202/1544-6115.1561

& 3 in Table 1). We reported the P-value for the likelihood ratio test based on a logistic regression model including marginal and interaction effects as suggested by Wang et al. (2008). The likelihood ratio test is based on a regression model to test $H_0 : \beta_{g1g2}$ and $\beta_{1i} = 0$ where $i = 1$ to the total number of SNPs for a given gene/haplotype. We also reported the single effects of the interaction term ($\beta_{g1g2} T_1^{1PLS} T_2^{1PLS}$) based on the logistic regression model in equation 7 (Supplementary Figures 8 and 9). Supplementary Table 9 contains the Type I error rates for the PLS likelihood test and interaction term using the null models in the presence of LD but no *RR* (models 4 Table 1). The Type I error rates remained close to the nominal significance level of 0.05 for both measures.

Figure 7a and Supplementary 8 illustrate empirical power for the PLS likelihood ratio test and interaction term where LD is present and no interaction *RR*, but the marginal $RR > 1.0$ (models 5 and 6 in Table 1). The PLS likelihood ratio test had greater power for both marginal $RR = 1.5$ (approaching 0.40) and 3.0 (approaching > 0.90). There is a slight increase in power as LD increases (moving across trellis). As expected, the PLS interaction term had low power (0.05).

Figures 7b and Supplementary 9 illustrate the empirical power for the PLS likelihood ratio test and interaction term where interaction $RR > 1.0$ (models 7, 8 and 9 in Table 1). Overall, there was an increase in power as the MAF and marginal *RR* (moving up the trellis) were increased. We only observed a slight increase in power as the LD was increased (moving across the trellis).

While powerful, this test may not be appropriate to separate marginal ($RR = 3.0$) from interaction effects which is illustrated in the top row of Figure 7b. For the likelihood ratio test, at least one β term must be significant. A significant result could be reporting on the marginal effect β's. Reporting on the interaction term β helped control the strong influence of the marginal *RR* which is illustrated in Supplementary Figure 9. For an interaction *RR* of 3.0 (solid line) and 10.0 (dashed line), we observed an increase in power as LD is increased (moving across the trellis). This observation may be due to the emphasis on predicting the responses. Our results in Supplementary Figure 9 (bottom and middle row of trellis) with power approaching 0.80 at LD = 0.7 and 0.9 support this.

3.2 Results of leukotriene pathway using matrix measures

With the knowledge gained from our simulations, we applied these methods to SNPs in the leukotriene biosynthesis pathway. Crosslin et al. (2009) assessed the role of the leukotriene pathway in CVD pathogenesis with association studies of *ALOX5AP* and *LTA4H* in a non-familial data set of EOCAD. LD plots for *ALOX5AP* and *LTA4H* are included in Supplementary Figures 10 and 11. As evident in the LD

plots for *ALOX5AP* (Supplementary Figure 10), there is minimal correlation for both Caucasians and African Americans. There is a difference in the LD block consisting of SNPs rs4769874, rs9551963 and rs9315050 for the African Americans. There is an increase in correlation for *LTA4H* (Figure 11) as compared to *ALOX5AP*.

3.2.1 Matrix measure results

Table 5 depicts the matrix measure results for the leukotriene biosynthesis pathway for both race groups. For the LD permutation Z_1, LD contrast Z_2 and background-corrected contrast test, we analyzed the SNPs composing the following: 1. HapA (*ALOX5AP*); 2. HapB (*ALOX5AP*); 3. HapK (*LTA4H*); 4. HapA $\times$ HapB; and HapA $\times$ HapK. For the PLS interaction method, we considered only HapA $\times$ HapB and HapA $\times$ HapK as a test of interaction. We reported both the likelihood ratio-based results testing the global null hypothesis (all $\beta = 0$) as well as the single effect significance for the haplotype interaction term (probability of observing a Wald $\chi^2 \geq$ to observed). The P-values presented for Z_1 and Z_2 are the empirical P-values achieved from permuting the phenotype labels of EOCAD. For reference, we added the EOCAD haplotype association results generated using the haplo.stats R package (Schaid et al., 2002) (last column in Table 5) from Crosslin et al. (2009). Haplo.stats expands on the likelihood approach to account for phase ambiguity in case-control studies by using a generalized linear model to test for haplotype association which allows for adjustment of non-genetic covariables (Schaid et al., 2002).

Tables 6 and 7 present the race stratified results for African Americans and Caucasians respectively. The race stratified LD plots for *ALOX5AP* and *LTA4H* comparing cases and controls can be found in Supplementary Figures 10 and 11 respectively. We added the EOCAD haplotype association results from Crosslin et al. (2009).

For both race groups combined and stratified, generally we observed Z_1 and Z_2 associations results in agreement to what Crosslin et al. observed in the haplotype analyses (Table 5). For instance, the Z_1 results ($P = 0.284$) are less supportive of LD differences between cases and controls than the Z_2 results ($P = 0.066$) compared to the haplotype result for HapA ($P = 0.02$) in both race groups (Table 5). However, the Z_1 results ($P = 0.064$) is more supportive than the Z_2 ($P = 0.115$) compared to the haplotype result for HapK ($P = 0.04$). For HapA in Caucasians, we could not reproduce the haplotype results ($P = 0.01$) using the Haplo.stats package through R Statistical Computing. Both the permutation Z_1 ($P = 0.178$) and LD

DOI: 10.2202/1544-6115.1561

contrast Z_2 ($P = 0.216$) tests were not significant. We observed similar results with HapK (Table 7).

We were most intrigued by the HapK for both tests, most notably in the African Americans (Table 6). Helgadottir et al. reported a significant effect ($RR = 6.50$) in the Philadelphia African American sample (Helgadottir et al., 2005b). This haplotype was not significant in the CATHGEN African American sample ($P = 0.27$), but generally we observe a qualitative difference in the LD patterns as illustrated in Supplementary Figure 11. We believe power is lost beyond a seven-SNP haplotype because the correct estimation of phase becomes an issue, especially with a smaller n. The risk allele-pattern in *LTA4H* may not be driven from a single disease-associated haplotype, but rather a global haplotype pattern driven by LD for a subset of SNPs.

We did not observe a significant interaction between HapA and HapB for any method. This result could be explained by the fact that the marker shared by both HapA and HapB (rs10507391) has different alleles included in the HapA and HapB risk haplotypes, suggesting they are independent susceptibility alleles. We did observe interesting results when assessing the HapA and HapK interaction, using both the permutation Z_1 ($P = 0.087$) and LD contrast Z_2 tests ($P = 0.139$), although as we saw in the simulation study, this suggestive interaction may be due to the marginal effects of HapK (Z_1 $P = 0.121, 0.018$ $(k = 2)$; Z_2 $P = 0.023$).

The PLS method did not show evidence with either the global test or the interaction test. We did not expect to see significant results with PLS using SNPs from *ALOX5AP* and *LTA4H* in a multivariate logistic regression model. Our lack of significant (*ALOX5AP*) and (*LTA4H*) univariate SNP association results for EO-CAD do not support using this method (Crosslin et al., 2009).

Table 2: Composite correlation between V_1 and V_2 for cases (top) and controls (bottom) for the null models 1, 2, 3 and 4, and for the alternate models 5 and 6 (Table 1). Numbers in the table are the mean for 1000 simulations. Prevalence = 0.10 and interaction relative risk = 1.0.

Interaction Relative Risk (*RR*) = 1.0												
RR for V_1, V_2	1.0 (null 1,2, 3 & 4)				1.5 (alternate 5)				3.0 (alternate 6)			
MAF	0.05	0.15	0.35	0.50	0.05	0.15	0.35	0.50	0.05	0.15	0.35	0.50
$r^2_{Hap_{1,2}} = 0.0$	-0.003	0.002	-0.002	-0.001	0.000	0.002	0.000	-0.002	-0.006	-0.014	-0.016	-0.014
	0.000	-0.002	0.000	-0.001	0.000	0.000	-0.003	-0.002	-0.003	-0.008	-0.013	-0.014
$r^2_{Hap_{1,2}} = 0.4$	0.001	0.001	-0.001	0.001	-0.001	-0.001	-0.002	-0.002	-0.009	-0.012	-0.018	-0.012
	0.000	0.000	0.001	0.000	0.001	0.002	-0.002	-0.004	-0.003	-0.010	-0.011	-0.010
$r^2_{Hap_{1,2}} = 0.7$	0.000	-0.000	0.001	0.001	0.001	0.001	-0.003	-0.002	-0.008	-0.010	-0.017	-0.014
	0.003	0.002	0.002	0.002	-0.001	-0.003	-0.002	0.000	-0.003	-0.010	-0.013	-0.012
$r^2_{Hap_{1,2}} = 0.9$	0.001	0.001	0.001	-0.001	-0.001	0.001	0.000	-0.002	-0.007	-0.011	-0.018	-0.014
	-0.003	-0.001	0.000	0.001	-0.001	-0.001	0.001	-0.005	-0.004	-0.007	-0.013	-0.013

Table 3: Composite correlation between V_1 and V_2 for cases (top) and controls (bottom) for the alternative models 7, 8 and 9 (Table 1). Numbers in the table are the mean for 1000 simulations. Prevalence = 0.10 and interaction relative risk = 3.0.

Interaction Relative Risk (*RR*) = 3.0												
RR for V_1, V_2	1.0				1.5				3.0			
MAF	0.05	0.15	0.35	0.50	0.05	0.15	0.35	0.50	0.05	0.15	0.35	0.50
$r^2_{Hap_{1,2}} = 0.4$	0.027	0.069	0.112	0.116	0.024	0.072	0.107	0.103	0.023	0.052	0.065	0.065
	-0.001	-0.009	-0.018	-0.017	-0.006	-0.013	-0.021	-0.026	-0.010	-0.026	-0.038	-0.040
$r^2_{Hap_{1,2}} = 0.7$	0.025	0.067	0.114	0.117	0.027	0.069	0.105	0.101	0.025	0.053	0.070	0.065
	-0.004	-0.007	-0.017	-0.014	-0.003	-0.012	-0.025	-0.026	-0.011	-0.024	-0.041	-0.039
$r^2_{Hap_{1,2}} = 0.9$	0.026	0.069	0.114	0.117	0.026	0.068	0.104	0.103	0.024	0.049	0.071	0.067
	-0.002	-0.008	-0.015	-0.018	-0.005	-0.010	-0.024	-0.026	-0.010	-0.027	-0.042	-0.040

Table 4: Composite correlation between V_1 and V_2 for cases (top) and controls (bottom) for the alternate models 7, 8 and 9 (Table 1. Numbers in the table are the mean for 1000 simulations. Prevalence = 0.10 and interaction relative risk = 10.0.

Interaction Relative Risk (*RR*) = 10.0												
RR for V_1, V_2	1.0				1.5				3.0			
MAF	0.05	0.15	0.35	0.50	0.05	0.15	0.35	0.50	0.05	0.15	0.35	0.50
$r^2_{Hap_{1,2}} = 0.4$	0.061	0.151	0.220	0.207	0.063	0.146	0.192	0.171	0.058	0.114	0.119	0.107
	-0.005	-0.020	-0.041	-0.044	-0.009	-0.028	-0.049	-0.054	-0.019	-0.046	-0.071	-0.072
$r^2_{Hap_{1,2}} = 0.7$	0.062	0.151	0.219	0.206	0.059	0.146	0.192	0.170	0.056	0.117	0.117	0.108
	-0.008	-0.022	-0.040	-0.043	-0.010	-0.030	-0.052	-0.056	-0.017	-0.045	-0.072	-0.073
$r^2_{Hap_{1,2}} = 0.9$	0.065	0.152	0.218	0.205	0.064	0.146	0.191	0.173	0.059	0.116	0.118	0.108
	-0.010	-0.022	-0.039	-0.043	-0.010	-0.028	-0.052	-0.052	-0.017	-0.044	-.069	-0.073

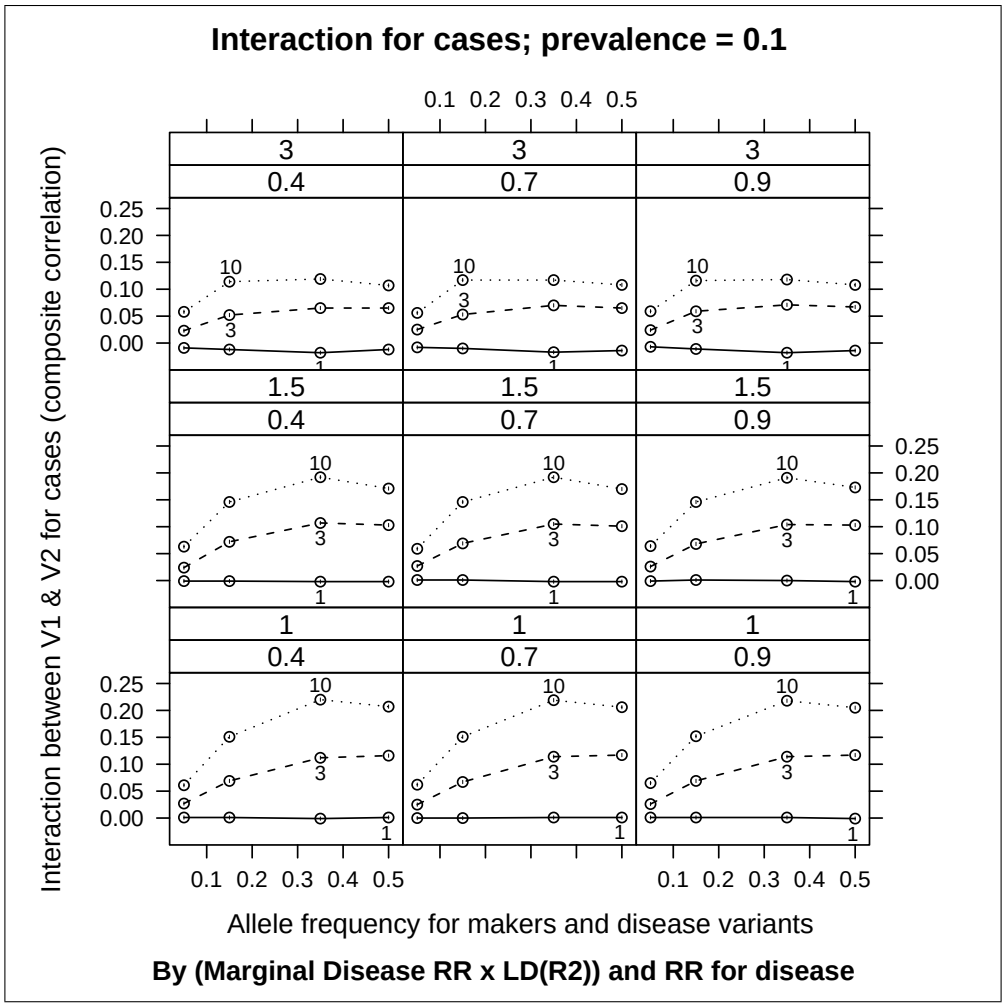
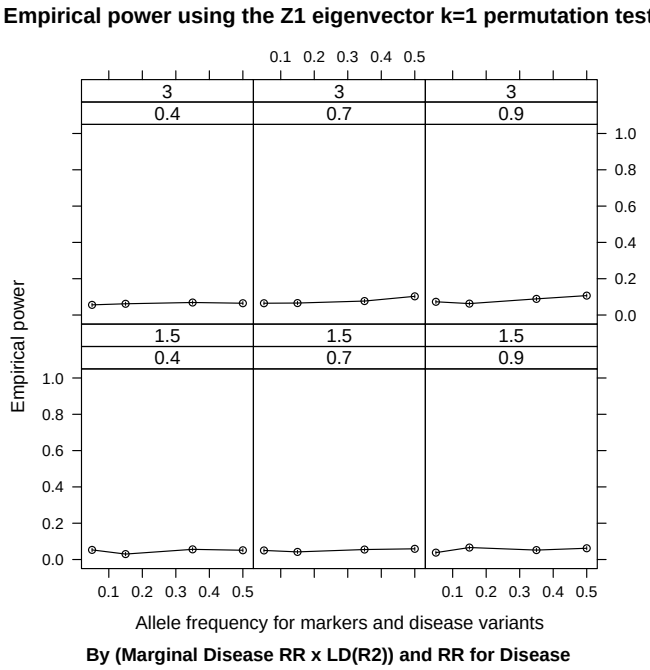


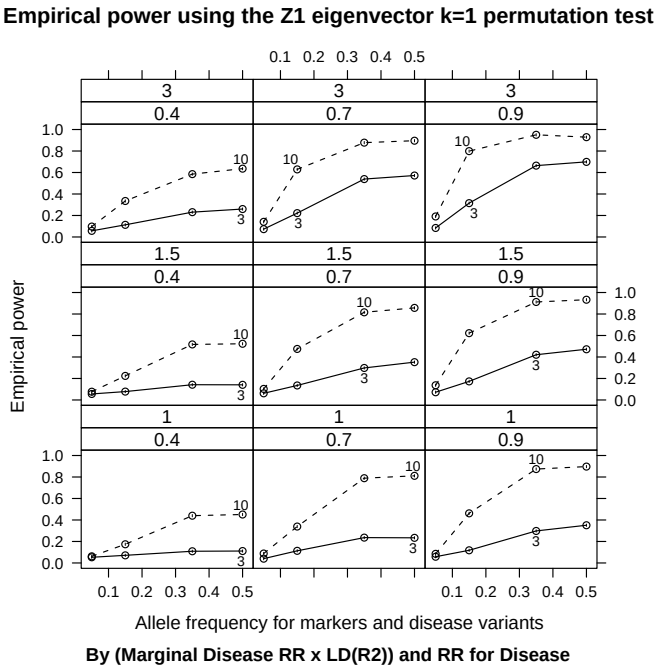
Figure 2: Composite correlation coefficients between V_1 and V_2 in cases for disease model simulations using the SIMLA package.

.

DOI: 10.2202/1544-6115.1561



(a) Empirical power for alternative models 5 and 6 with interaction $RR = 1.0$.



(b) Empirical power for alternative models 7, 8 and 9 with interaction $RR = 3.0$ and 10.0.

Figure 3: Empirical power using the Z_1 LD eigenvector ($k = 1$) permutation test for alternative models $5 \rightarrow 9$ (Table 1). Interaction $RR = 1.0, 3.0$ and 10.0.

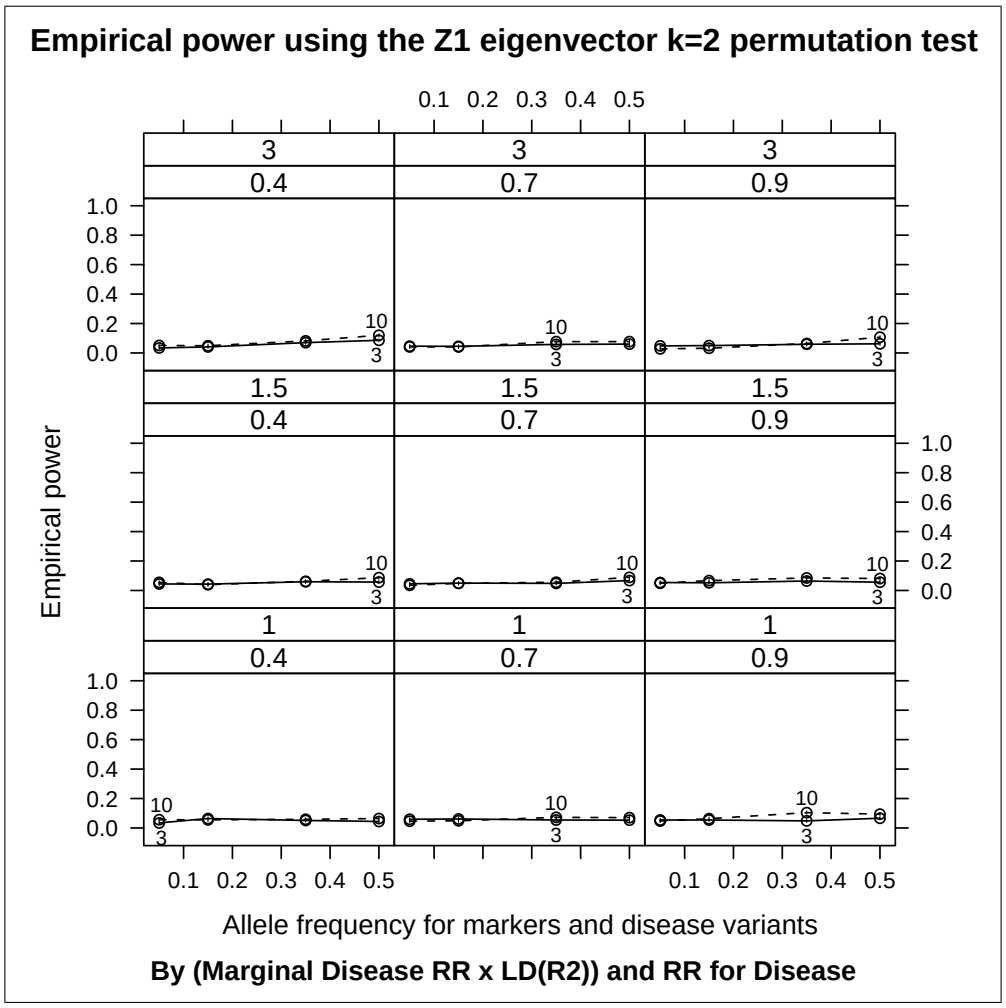
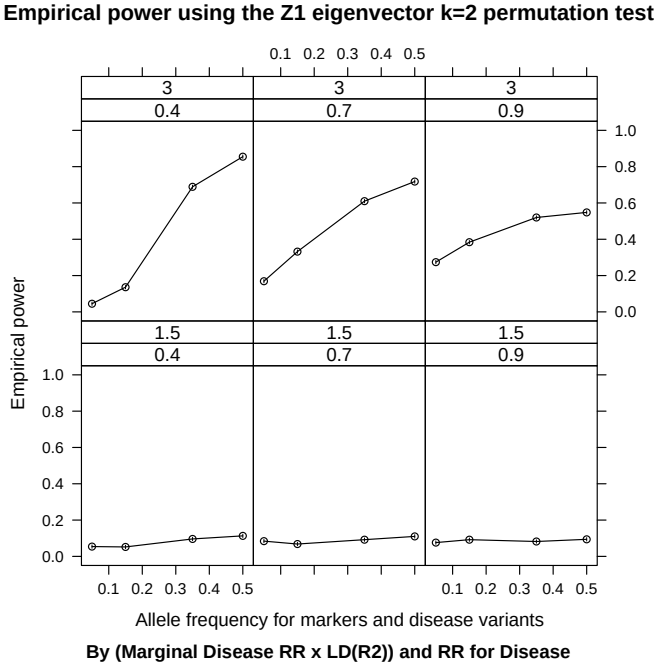
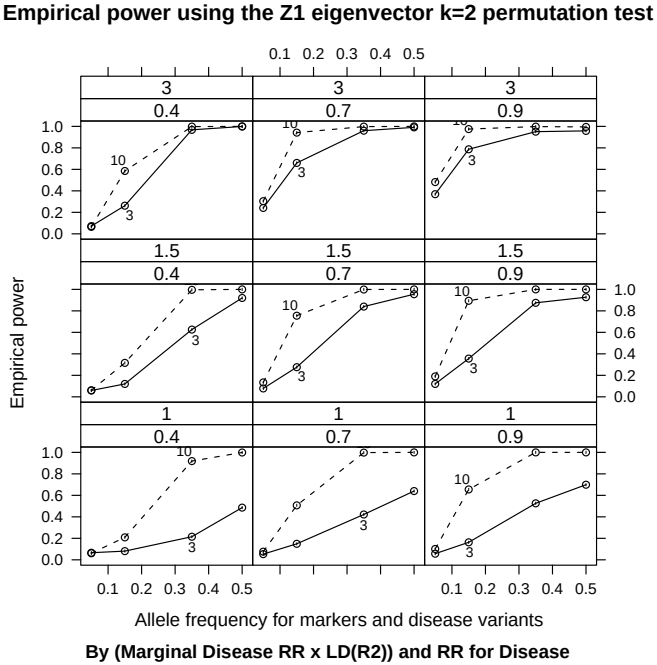


Figure 4: Empirical power using the Z_1 LD eigenvector ($k = 2$) permutation test for alternative models 7, 8 and 9. Interaction RR = 3.0 and 10.0.

DOI: 10.2202/1544-6115.1561

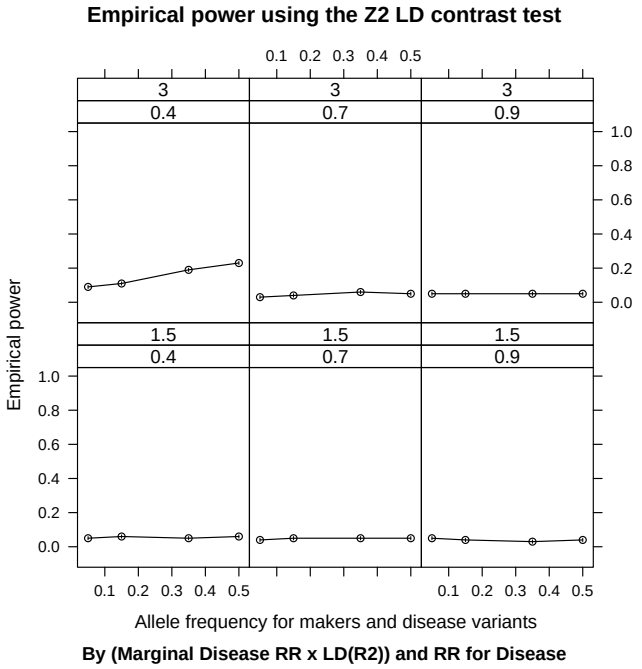


(a) Empirical power for alternative models 5 and 6 with interaction $RR = 1\ 0.\ .$

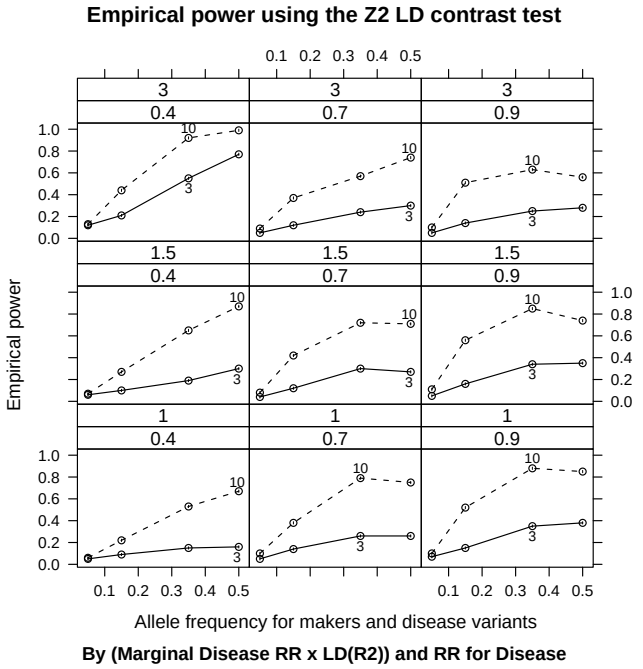


(b) Empirical power for alternative models 7, 8 and 9 with interaction $RR = 3.0$ and 10.0.

Figure 5: Empirical power using the Z_1 LD eigenvector ($k = 2$) weighted by the eigenvalue permutation test for alternative models $5 \rightarrow 9$ (Table 1). Interaction $RR = 1.0, 3.0$ and 10.0.

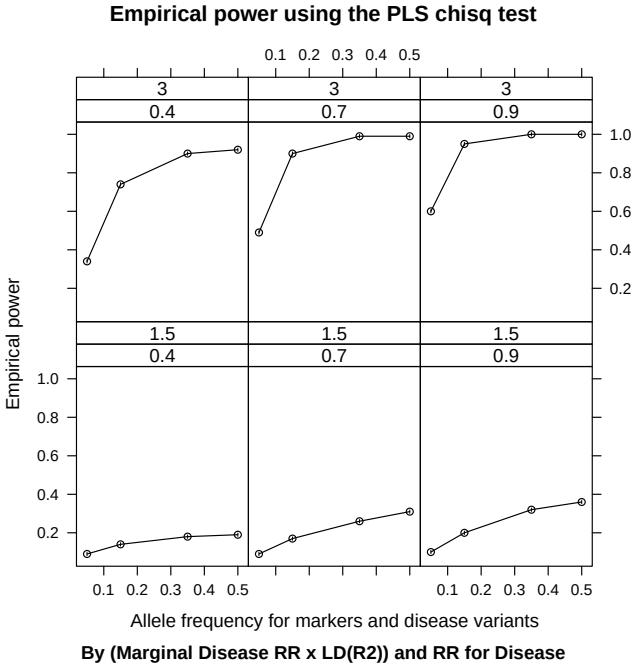


(a) Empirical power for alternative models 5 and 6 with interaction $RR = 1.0$.

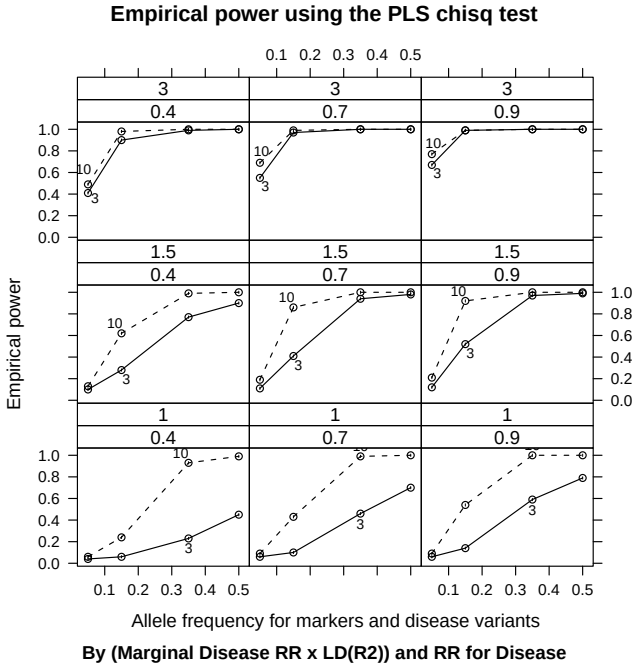


(b) Empirical power for alternative models 7, 8 and 9 with interaction $RR = 3.0$ and 10.0.

Figure 6: Empirical power using the Z_2 LD contrast test for alternative models $5 \rightarrow 9$ (Table 1). Interaction $RR = 1.0, 3.0$ and 10.0.



(a) Empirical power for alternative models 5 and 6 with interaction $RR \neq 0$.



(b) Empirical power for alternative models 7, 8 and 9 with interaction $RR = 3.0$ and 10.0.

Figure 7: Empirical power using the partial least square χ^2 test for alternative models $5 \rightarrow 9$ (Table 1). Interaction $RR = 1.0, 3.0$ and 10.0.

Table 5: Results using matrix measures to assess genetic effects in the leukotriene biosynthesis pathway in the CATHGEN sample $n = 656(\text{case})/406(\text{control})$. For the PLS test, the first P-value is for the likelihood test and the second is for the interaction term.

	Matrix Measure				Haplotype
	LD Permutation Z_1 $(k=1)$	LD Contrast Z_2	PLS χ^2	PLS Inter.	Results (Crosslin et al., 2009)
ALOX5AP HapA	$P = 0.284$	$P = 0.066$	N/A	N/A	$P = 0.02$
ALOX5AP HapB	$P = 0.063$	$P = 0.396$	N/A	N/A	$P = 0.39$
LTA4H HapK	$P = 0.064, 0.040$ $(k=2)$	$P = 0.115$	N/A	N/A	$P = 0.04$
HapA $\times$ HapB	$P = 0.122$	$P = 0.464$	$P = 0.528$	$P = 0.231$	N/A
HapA $\times$ HapK	$P = 0.089$	$P = 0.276$	$P = 0.699$	$P = 0.566$	N/A

Table 6: Results using matrix measures to assess genetic effects in the leukotriene biosynthesis pathway for African Americans in the CATHGEN sample $n = 143(\text{case})/77(\text{control})$.

	Matrix measure				Haplotype
	LD Permutation Z_1 $(k=1)$	LD Contrast Z_2	PLS χ^2	PLS Inter.	Results (Crosslin et al., 2009)
ALOX5AP HapA	$P = 0.784$	$P = 0.697$	N/A	N/A	$P = 0.67$
ALOX5AP HapB	$P = 0.016$	$P = 0.090$	N/A	N/A	$P = 0.37$
LTA4H HapK	$P = 0.121, 0.018$ $(k=2)$	$P = 0.023$	N/A	N/A	$P = 0.27$
HapA $\times$ HapB	$P = 0.515$	$P = 0.230$	$P = 0.715$	$P = 0.629$	N/A
HapA $\times$ HapK	$P = 0.087$	$P = 0.139$	$P = 0.750$	$P = 0.952$	N/A

Table 7: Results using matrix measures to assess genetic effects in the leukotriene biosynthesis pathway for Caucasians in the CATHGEN sample $n = 465(\text{case})/289(\text{control})$.

	Matrix measure				Haplotype
	LD Permutation Z_1 $k=1$	LD Contrast Z_2	PLS χ^2	PLS Inter.	Results (Crosslin et al., 2009)
ALOX5AP HapA	$P = 0.178$	$P = 0.216$	N/A	N/A	$P = 0.01$
ALOX5AP HapB	$P = 0.375$	$P = 0.203$	N/A	N/A	$P = 0.44$
LTA4H HapK	$P = 0.178, 0.105$ $(k=2)$	$P = 0.309$	N/A	N/A	$P = 0.03$
HapA $\times$ HapB	$P = 0.233$	$P = 0.602$	$P = 0.119$	0.667	N/A
HapA $\times$ HapK	$P = 0.247$	$P = 0.090$	$P = 0.127$	0.916	N/A

4 Discussion

Our applied work in genetic association of cardiovascular disease suggests that models will be complex, including genetic and allelic heterogeneity along with genetic interactions. Our motivation for studying these methods was to find alternate robust measures for whole gene(s) and interaction(s) (epistasis) associations using multiple SNPs. In the context of whole-genome scans, measures incorporating multiple SNPs at the gene-level will allow for gene set (pathway) analysis. To evaluate LD contrast methods (Zaykin et al., 2006) we performed a comprehensive evaluation by simulating two separate disease variants with marginal effects (*RR*) in the presence of LD, with and without joint effects (*RR*). We also expanded on previous versions of the test methods. We introduced the concept of weighting the eigenvector by the corresponding eigenvalue for the permutation method (Z_1). This approach may allow the detection of marginal effects in the presence of LD (see Figure 5a). We compared the PLS approach introduced by Wang et al. (2008) to the LD contrast methods. Finally, we utilized these measures to contrast CVD-phenotype case and control subjects using SNPs found in a candidate pathway. Based on correlation patterns observed, we were able to detect contrasts between cases and controls when stratifying by race groups.

4.1 The simulated disease model

One purpose in simulating the two-variant disease models with haplotypes was to gain an understanding of pathway interactions in terms of correlation patterns. Our simulations better capture the range of allele frequencies that were observed in the leukotriene pathway example. One of the biggest hurdles with this model was developing a reasonable two-locus genetic model and specifying the parameters for input into the SIMLA software package. Our goal was to understand what effect haplotype LD, MAF and marginal *RR* have on the interaction between disease variant 1 (V_1) and variant 2 (V_2) in terms of composite correlation. We simulated a *RR* interaction between both variants of 1.0, 3.0 and 10.0 (the limit for SIMLA). By no means did we exhaust the entire parameter space of possible models (Marchini et al., 2005, Evans et al., 2006). We simulated the data using a multiplicative disease model based on *RR*. Other models to be addressed in future research include dominant, codominant and recessive models as these may introduce different correlation patterns.

While the models we simulated seem quite extreme (i.e. interaction $RR = 10.0$), the purpose was to observe a trend in correlation patterns based on different combinations of model parameters including joint *RR*. The rationale behind this

DOI: 10.2202/1544-6115.1561

approach is that the joint effect of two variants would generate different LD patterns in cases and controls (Wang et al., 2008). The effect of MAF and *RR* are strong influences on correlation, but we also observed the influence of case selection bias based on marginal effects in the presence of interaction.

We simulated a range of minor allele frequencies(see Table 1). Zaykin et al. (2006) indicated that their tests have limitations when allele frequencies are low because the correlation structures are difficult to estimate. Equally important, LD between the haplotype-specific markers was simulated at $r^2 = 0.4, 0.7$ and 0.9. For instance, criteria for tSNP selection may affect LD patterns when analyzing candidate pathways. Finally, under the two-disease variant model, understanding the effect of increasing *RR* for the two variants was crucial. We generated *RR* of $1.0, 1.5$ and 3.0 marginal effect sizes to cover a range of power for our models. We generated a prevalence frequency of 0.10 for all models. We do note that certain age-dependent diseases like CVD can approach frequencies of 0.25 at older ages. The impact of age-dependent prevalence on application of these methods merits further investigation.

4.2 Matrix test statistics

We evaluated the realized correlation for complex models as described in Section 2. We investigated the statistical properties of the matrix-based statistics to measure epistasis under various model conditions. Zaykin et al. (2006) presented the LD matrix measures to provide statistical support in quantifying difference in a graphical display of LD of cases and controls. This was our original motivation for exploring these measures. One major drawback with these matrix methods is that they do not produce effect sizes. The models were difficult to assess statistically; in a complex system such as this with multiple correlations (both marginal and interaction), it was difficult to determine the source of a significant result. There are very few data driven models that produced high correlation. The underlying correlation is low even in strong interaction models.

Allele frequency differences between cases and controls due to systematic ancestry differences, can cause spurious associations in disease studies (Price et al., 2008). Because the effects of stratification vary in proportion to the number of samples, stratification is an increasing problem in large scale association studies (Price et al., 2008); however, as more markers are added to a model, subtle background population differences can be observed. Future work will determine if these measures are robust to admixture.

4.2.1 Z_1 LD permutation test

We did not observe Type I error rate inflation for any of the null models using the Z_1 method. For alternate models 5 and 6 (Table 1), we observed low power. These models are the paradigm for common-disease/common-variant genetic association studies using SNPs. The LD permutation test gained power with the synergistic epistatic model (model 7 in Table 1) where each variant independently has no risk, although interaction $RR = 3.0$ only reaches 0.40 under high LD (0.9). The $RR = 10.0$ performed better, but may be unrealistic for common diseases. There was additional gain in power for the full models of LD, interaction and marginal effects (Table 5b).

One drawback with this measure is deciding *a priori* how many eigenvectors to use in the featurevector (matrix of eigenvectors). Krazanowski's suggestion of using the first integer below 50% of the original number of variables is a valid argument and supported by our Z_1 power results when $k = 2$ (Figure 4). We introduced weighting the eigenvector by its corresponding eigenvalue as an alternate approach. We observed an increase in power for alternate models 5 and 6 (table 1) at marginal $RR = 3.0$. There was a gain in adding the additional weighted eigenvector to detect the indirect risk via LD. There was additional gain in power for the full models of LD, interaction and marginal effects (Figure 5a). Our results suggest that a four-SNP matrix will most likely have a single important component, but variance explained by additional eigenvalues/eigenvectors is important. Depending on the amount of variance explained by each eigenvalue-eigenvector, the value of k can conceivably exceed 50% of the original number of variables for the weighted vector approach. Further exploration for the precision of k with respect to variance explained (weighted) remains to be studied. In addition, this approach on higher order multivariate data will need to be explored.

4.2.2 Z_2 LD contrast test

We again did not observe Type I error rate inflation for any of the null models using the Z_2 method. We observed low power for alternate models 5 and 6 (table 1), but there was a slight increase (0.20) at low LD (0.4) and high marginal risk (3.0) (Figure 6a). We observed an increase in power for the synergistic epistatic model, but again only reaching 0.40 under high LD (0.9) for $RR = 3.0$. There was additional gain in power for the full models of LD, interaction and marginal effects illustrated in Figure 6b. We will need to explore the increase in power at low LD in the alternate models in the presence of LD and marginal effects. This method in combination with the permutation test would be a viable approach to assess correlation differences.

DOI: 10.2202/1544-6115.1561

4.2.3 Partial least squares

As evident in the power plots for the PLS χ^2 method (Figures 7a and 7b), this is a powerful test to detect the marginal effects as well as effects due to interaction. If the purpose of the analysis is to specifically test for interaction in the presence of overall genetic effects, this test was not the most powerful in our simulations. Reporting the interaction term P-value will assist in identifying the interaction significance while controlling for the main effects. Mathematically, it is logical to choose factors that best predict outcome, but adding SNPs from one gene in a multivariate model may not be the best approach. More SNPs in LD may provide a more significant interaction via PLS.

4.3 Future directions

4.3.1 LD background-corrected

Wang et al. (2007) described a method that is similar to the sum of squared differences statistic (Z_2), but the pair-wise elements are the sum of subject-specific mean-corrected cross-products, by use of the best linear unbiased predictor (BLUP) of the means. In a mixed model, linear combinations of the fixed and random effects can be formed from linear combinations of the conditional means (Littell et al., 2006). We attempted to expand the LD background-corrected contrast method introduced by Wang et al. (2007) to include multiple SNPs from two disease causing variants (data not shown). We used the linear mixed-effects model (lme) function in R to estimate the best linear unbiased predictor (BLUP) (Dr. Tao Wang personal communication). Instead of using the overall sample mean to center the genotypes, which is equivalent to the composite correlation-based LD contrast test statistic, Wang et al. (2007) proposed centering the genotype by use of the individual (i) specific means which absorb background LD. Under certain conditions the model failed to converge. As such, our inability to derive nominal Type 1 significance prevented us from calculating empirical power. However, we still view this approach as potentially useful and would like to make further comparisons to the methods we evaluated in this work.

4.3.2 Adjusting for covariates and matrix-based contrast methods

In most complex diseases like CAD it is important to adjust for covariates. The PLS approach introduced by Wang et al. (2008) is based in a regression framework,

thus allowing for the adjustment of covariates. In the context of accounting for covariates with the matrix-based contrast methods introduced by Zaykin et al. (2006), there are multiple methods to explore in the future. Krzanowski (1979) originally presented the permutation-based method (Z_1) to allow for comparisons across multiple categories, not just dichotomous variables. This stratification approach will assist in identifying the effects of covariates in the presence of LD patterns on the phenotype. Another approach includes weighting the correlations (LD) patterns based on covariate values. This method can be extended to both the permutation-based method (Z_1) and the LD contrast method (Z_2). Finally, additional variables (covariates) can be incorporated to the matrices as correlation patterns with SNPs and additional covariates. This can be applied to both methods.

4.3.3 Extension of matrix-based contrast methods to quantitative traits

We would like to extend the matrix-based contrast methods to quantitative traits. As referenced, Krzanowski (1979) presented the permutation-based method (Z_1) to allow for comparisons across multiple categories. Breaking a continuous variable into intervals (quantiles) would allow a multiple comparison. These same intervals could be used for the LD contrast method (Z_2), in the context of observing a trend (or lack of) of significance. The PLS method is built in a regression framework and can be extended to ordinary least squares.

4.3.4 Genome-wide analysis and Gene Set Enrichment Analysis (GSEA)

Genome-wide analysis It does not seem likely that these approaches could be applied on a genome-wide scale, rather our goal was to evaluate matrix methods of summarizing genes, regions and pathways. Evans et al. (2006) investigated epistasis in genome-wide association using different research strategies, including a single locus-search, an exhaustive two-locus search (pairwise), and two (two-stage) procedures in which a subset of loci initially identified with single locus tests are analyzed using a full two-locus model. We view the genome-wide methods as a means to reduce the set of SNPs to include in the matrix. Then the matrix-based measures will provide an additional tool to assess epistasis and may increase power by harnessing data from more SNPs compared to the current two-locus approach.

Another future goal is to understand what level of pathway correlation patterns are observed in the presence of intra-gene LD. Using pathway data from KEGG, we will derive gene/SNP sets to analyze for association. Given prior knowledge of a biological pathway, creating a matrix measure of correlation patterns of

DOI: 10.2202/1544-6115.1561

SNPs for genes constituting that pathway will provide insight into SNP-SNP and gene-gene interaction and the association to disease.

GSEA The LD-based statistics we generate can be used in pathway ascertainment tools that require scores on a gene-level. We could compute pathway-level association measures by deriving an enrichment scores from GSEA. Providing a single statistical measure per gene using multiple SNPs will also enable the combination of multiple types of genomic data at a gene-level, including expression data.

4.4 Conclusion

In summary, our model design for simulations provides insight into the effect of case status selection on correlation patterns. This insight is not only relevant for the methods we assessed, but also for standard genetic association methods. We have demonstrated that LD matrix statistics are feasible and powerful in detecting moderate to strong interactions and may be useful in understanding contributions of multiple haplotypes to disease susceptibility. Zaykin et al. (2006) suggested that the LD contrast test (Z_2) had the highest power based on the models they simulated. The permutation-based method (Z_1) had similar or better power that the Z_2 based on the alternate models we simulated. Neither method without modification was able to detect an association in terms of correlation with the disease variants when the marginal $RR = 1.5$ and 3.0. Weighting the eigenvectors by their respective level of significance (eigenvalue) increased the power without inflating the Type 1 error rate. Additional simulations with more SNP markers are required. Nevertheless, the weighted Z_1 test holds promise for detecting complex interactions and may allow differentiation of marginal from joint effects. The magnitude of weighting would depend on the SNP loadings on those components. By examining the loading structure on each component, an investigator can identify a correlation pattern that represents a joint effect versus a marginal effect, and can properly weight as appropriate. The PLS method of isolating variance that correlates with outcome does demonstrate power in assessing interactions.

We were encouraged to see these methods recapitulate what we observed in the initial leukotriene biosynthesis pathway analyses (Crosslin et al., 2009). Our applied results suggest greater complexity in the relationship between the HapK haplotype in *LTA4H* with genetic variation at other loci in the leukotriene pathway. The ability to take these complex relationships into account in future study design may cause variability in association results.

A Supplementary

Table 8: Type 1 error rates for null models 1, 2 and 3 (Table 1) with no LD for the haplotypes. Interaction relative risk = 1.0.

$r^2_{Hap_1} = r^2_{Hap_2} = 0.0$												
RR for V_1, V_2	1.0				1.5				3.0			
MAF	**0.05**	**0.15**	**0.35**	**0.50**	**0.05**	**0.15**	**0.35**	**0.50**	**0.05**	**0.15**	**0.35**	**0.50**
Permutation Z_1	0.04	0.04	0.03	0.06	0.05	0.05	0.06	0.05	0.05	0.04	0.05	0.04
Contrast Z_2	0.04	0.04	0.04	0.06	0.05	0.05	0.04	0.04	0.04	0.04	0.04	0.06
PLS χ^2	0.04	0.05	0.05	0.04	0.03	0.05	0.05	0.05	0.04	0.04	0.04	0.04
PLS Interaction	0.03	0.04	0.05	0.04	0.04	0.04	0.06	0.03	0.02	0.04	0.04	0.05

Table 9: Type 1 error rates for null model 4 (Table 1). Interactive relative risk = 1.0.

RR for V_1, V_2	1.0			
MAF	0.05	0.15	0.35	0.50
LD Permutation Z_1 $k = 1$				
$r^2_{Hap_1}, r^2_{Hap_2} = 0.4$	0.05	0.04	0.05	0.06
$r^2_{Hap_1}, r^2_{Hap_2} = 0.7$	0.06	0.05	0.06	0.05
$r^2_{Hap_1}, r^2_{Hap_2} = 0.9$	0.05	0.05	0.05	0.06
LD Permutation Z_1 $k = 2$ **weighted**				
$r^2_{Hap_1}, r^2_{Hap_2} = 0.4$	0.05	0.06	0.06	0.06
$r^2_{Hap_1}, r^2_{Hap_2} = 0.7$	0.06	0.06	0.06	0.05
$r^2_{Hap_1}, r^2_{Hap_2} = 0.9$	0.05	0.05	0.05	0.06
LD Contrast Z_2				
$r^2_{Hap_1}, r^2_{Hap_2} = 0.4$	0.05	0.05	0.05	0.04
$r^2_{Hap_1}, r^2_{Hap_2} = 0.7$	0.05	0.05	0.05	0.06
$r^2_{Hap_1}, r^2_{Hap_2} = 0.9$	0.06	0.04	0.05	0.04
PLS χ^2				
$r^2_{Hap_1}, r^2_{Hap_2} = 0.4$	0.06	0.04	0.05	0.05
$r^2_{Hap_1}, r^2_{Hap_2} = 0.7$	0.07	0.05	0.07	0.06
$r^2_{Hap_1}, r^2_{Hap_2} = 0.9$	0.06	0.04	0.06	0.06
PLS Interaction				
$r^2_{Hap_1}, r^2_{Hap_2} = 0.4$	0.03	0.04	0.05	0.04
$r^2_{Hap_1}, r^2_{Hap_2} = 0.7$	0.04	0.04	0.06	0.04
$r^2_{Hap_1}, r^2_{Hap_2} = 0.9$	0.03	0.05	0.06	0.05

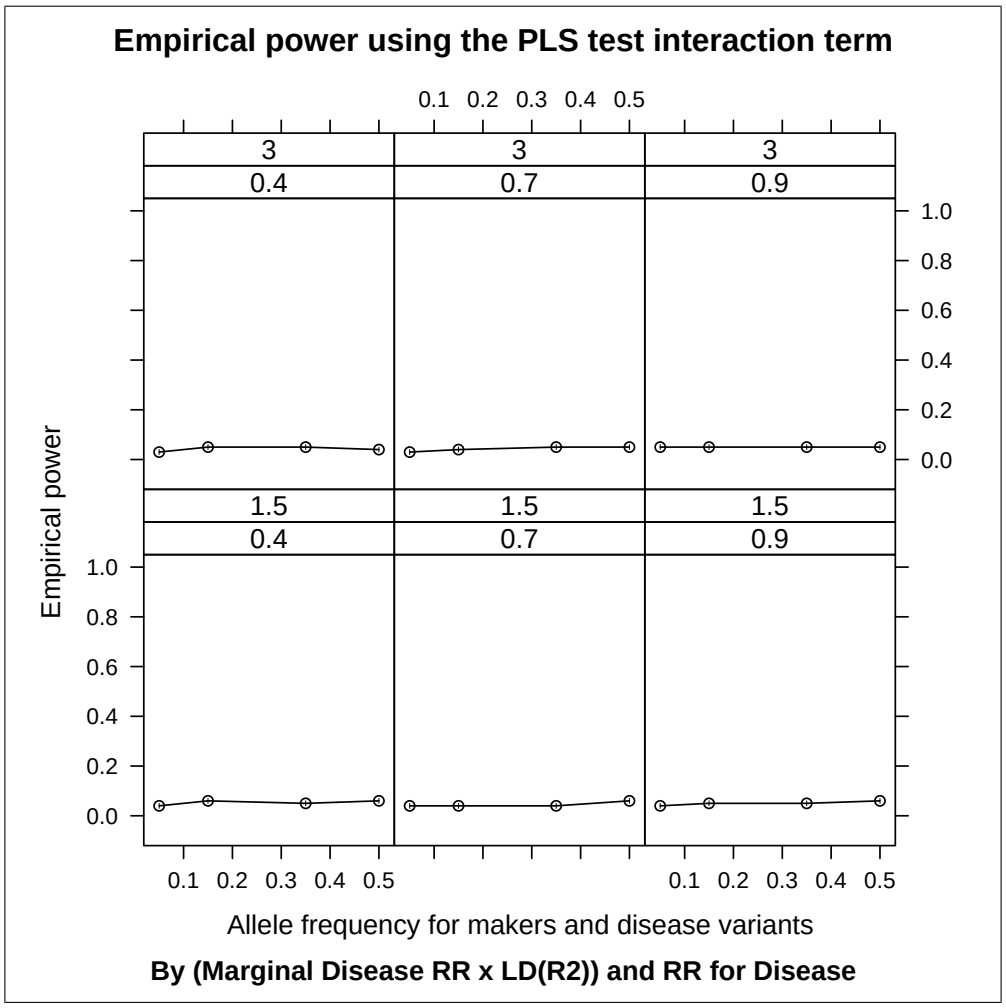


Figure 8: Empirical power using the partial least square test interaction term for alternative models 5 and 6 (Table 1). Interaction $RR = 1.0$.

DOI: 10.2202/1544-6115.1561

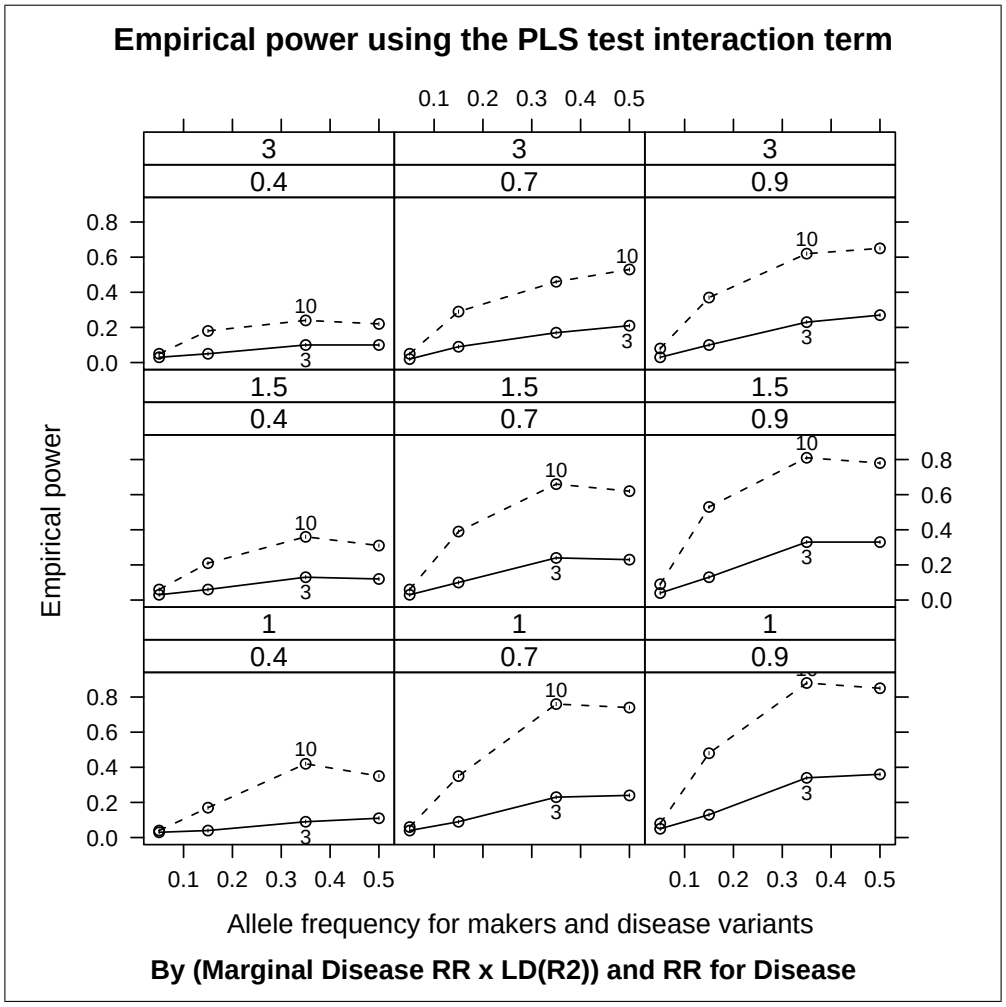


Figure 9: Empirical power using the partial least square test interaction term for alternative models 7, 8 and 9 (Table 1). Interaction *RR* = 3.0 and 10.0.

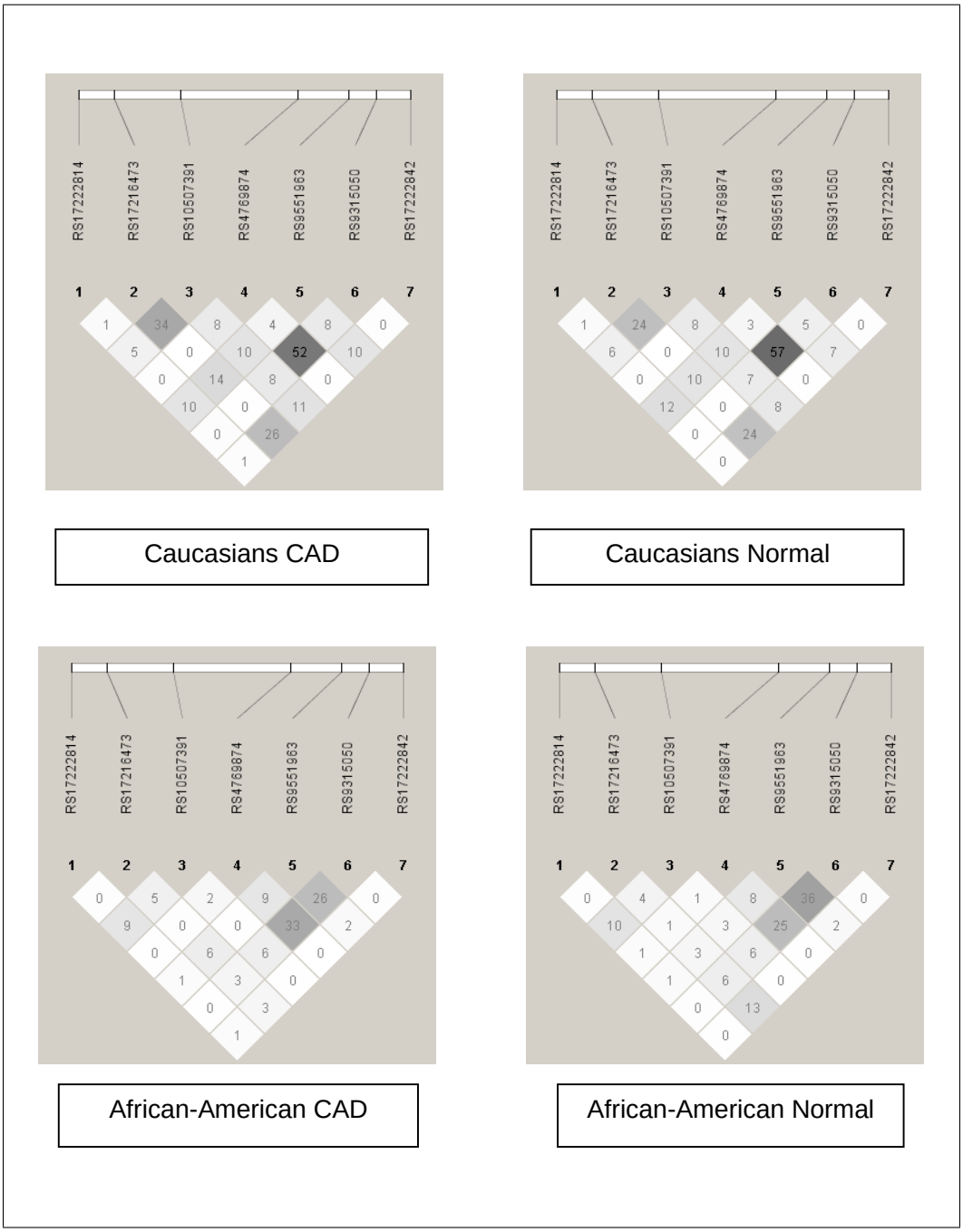


Figure 10: *ALOX5AP* (including SNPs for HapA and HapB) linkage disequilibrium figures for CAD and controls by ethnicity in the CATHGEN sample.

DOI: 10.2202/1544-6115.1561

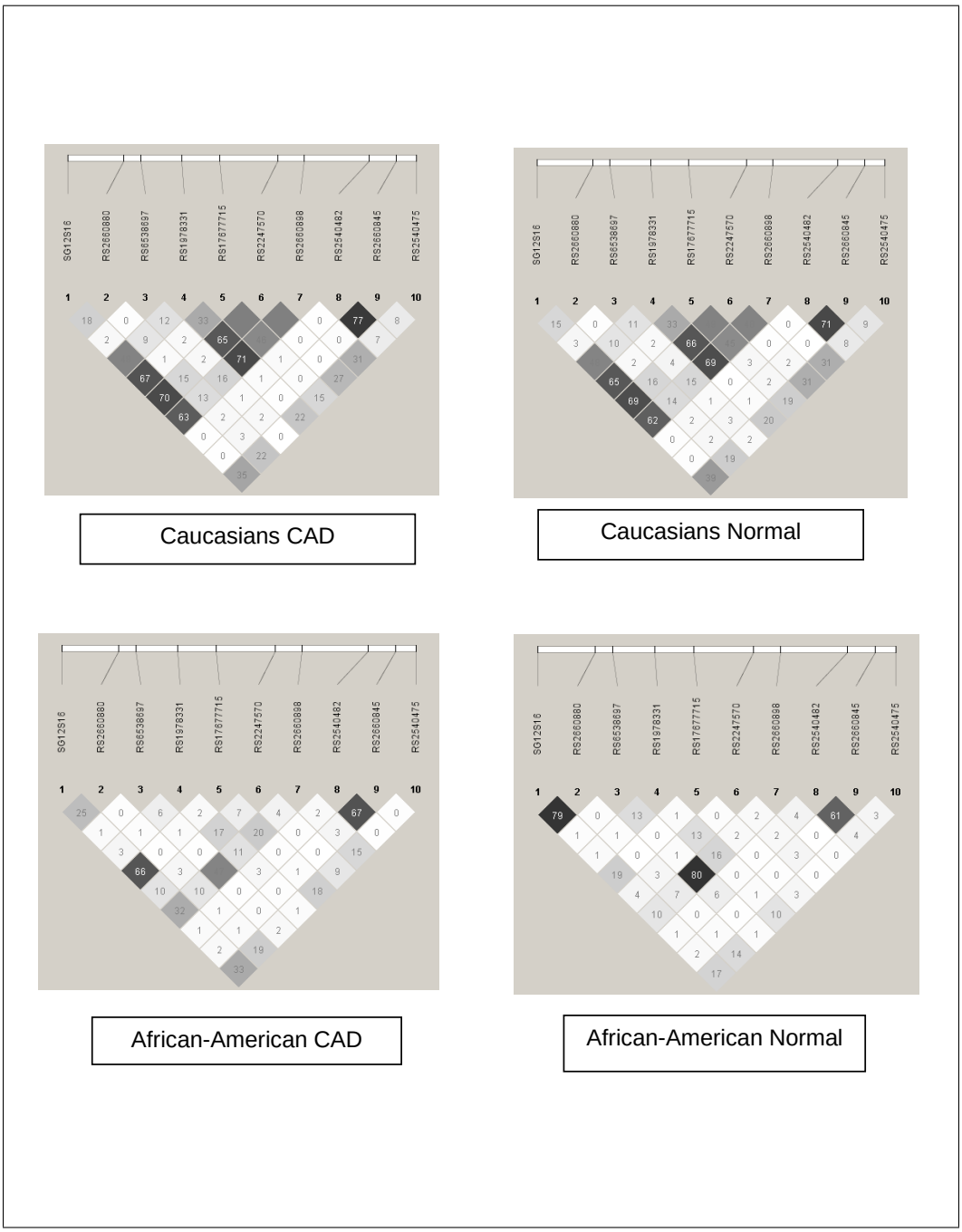


Figure 11: *LTA4H* (including SNPs for HapK) linkage disequilibrium figures for CAD and controls by ethnicity in the CATHGEN sample.

References

Cordell, H. (2002): "Epistasis: what it means, what it doesnt mean,and statistical methods to detect it in humans," *Human Molecular Genetics*, 11, 24632468.

Crosslin, D., S. Shah, S. Nelson, C. Haynes, J. Connelly, S. Gadson, P. Goldschmidt-Clermont, J. Vance, J. Rose, C. Granger, D. Seo, S. Gregory, W. Kraus, and E. Hauser (2009): "Genetic effects in the leukotriene biosynthesis pathway and association with atherosclerosis," *Human Genetics*, 125, 217–229.

Davidian, M. (Spring 2005): *Simulation Studies in Statistics; ST 810A*.

Dwyer, J., H. Allayee, K. Dwyer, J. Fan, H. Wu, R. Mar, A. Lusis, and M. Mehrabian (2004): "Arachidonate 5-lipoxygenase promoter genotype, dietary arachidonic acid, and atherosclerosis," *N. Engl. J. Med.*, 350, 29–37.

Evans, D., J. Marchini, A. Morris, and C. L.R. (2006): "Two-stage two-locus models in genome-wide association," *PLoS Genetics*, 2, e157.

Helgadottir, A., S. Gretarsdottir, C. D. St, A. Manolescu, J. Cheung, G. Thorleifsson, A. Pasdar, S. Grant, L. Whalley, H. Hakonarson, U. Thorsteinsdottir, A. Kong, J. Gulcher, K. Stefansson, and M. MacLeod (2005a): "Association between the gene encoding 5-lipoxygenase-activating protein and stroke replicated in a scottish population," *Nature Genetics*, 76, 505–509.

Helgadottir, A., A. Manolescu, A. Helgason, G. Thorleifsson, U. Thorsteinsdottir, D. Gudbjartsson, S. Gretarsdottir, K. Magnusson, G. Gudmundsson, A. Hicks, T. Jonsson, S. Grant, J. Sainz, S. O'brien, S. S. E. Valdimarsson, S. Matthiasson, A. Levey, J. Abramson, M. Reilly, V. Vaccarino, M. Wolfe, V. Gudnason, A. Quyyumi, E. Topol, D. Rader, G. Thorgeirsson, J. Gulcher, H. Hakonarson, A. Kong, and K. Stefansson (2005b): "A variant of the gene encoding leukotriene a4 hydrolase confers ethnicity-specific risk of myocardial infarction," *Nature Genetics*, 1–7.

Helgadottir, A., A. Manolescu, G. Thorleifsson, S. Gretarsdottir, H. Jonsdottir, U. Thorsteinsdottir, N. Samani, G. Gudmundsson, S. Grant, G. Thorgeirsson, S. Sveinbjornsdottir, E. Valdimarsson, S. Matthiasson, H. Johannsson, O. Gudmundsdottir, M. Gurney, J. Sainz, M. Thorhallsdottir, M. Andresdottir, M. Frigge, E. Topol, A. Kong, V. Gudnason, H. Hakonarson, J. Gulcher, and K. Stefansson (2004): "The gene encoding 5-lipoxygenase activating protein confers risk of myocardial infarction and stroke," *Nature Genetics*, 36, 233–239.

Krzanowski, W. (1979): "Between-groups comparison of principal components," *Diabetes*, 74, 703–707.

Krzanowski, W. (1993): "Permutational tests for correlation matrices," *Statistics and Computing*, 3, 37–44.

Lee, M. and G. Whitmore (2002): "Power and sample size for dna microarray studies," *Stat. Med.*, 21, 3543–3570.

DOI: 10.2202/1544-6115.1561

Littell, R., G. Milliken, W. Stroup, R. Wolfinger, and O. Schabenberger (2006): *SAS for Mixed Models*, Cary, NC: SAS Institute Inc., 2nd edition.

Lohmussaar, E., A. Gschwendtner, J. Mueller, T. Org, E. Wichmann, G. Hamann, T. Meitinger, and M. D. M (2005): "Alox5ap gene and the pde4d gene in a central european population of stroke patients," *Stroke*, 36, 731–736.

Loos, R., C. Lindgren, S. Li, E. Wheeler, J. Zhao, I. Prokopenko, M. Inouye, R. Freathy, A. Attwood, J. Beckmann, S. Berndt, K. Jacobs, S. Chanock, R. Hayes, S. Bergmann, A. Bennett, S. Bingham, M. Bochud, M. Brown, S. Cauchi, J. Connell, C. Cooper, G. Smith, I. Day, C. Dina, S. De, E. Dermitzakis, A. Doney, K. Elliott, P. Elliott, D. Evans, I. S. Farooqi, P. Froguel, J. Ghori, C. Groves, R. G. D. Hadley, A. Hall, A. Hattersley, J. Hebebrand, I. Heid, C. L. C. Gieger, T. Illig, T. Meitinger, H. Wichmann, B. Herrera, A. Hinney, S. Hunt, M. Jarvelin, T. Johnson, J. Jolley, F. Karpe, A. Keniry, K. Khaw, R. Luben, M. Mangino, J. Marchini, W. McArdle, R. McGinnis, D. Meyre, P. Munroe, A. M. A. Ness, M. N. A. Nica, K. Ong, S. O'Rahilly, K. Owen, C. Palmer, K. Papadakis, S. Potter, A. Pouta, L. Qi, J. Randall, N. Rayner, S. Ring, M. Sandhu, A. S. M. Sims, K. Song, N. Soranzo, E. Speliotes, H. Syddall, S. Teichmann, N. Timpson, J. Tobias, M. Uda, C. Vogel, C. Wallace, D. Waterworth, M. Weedon, C. Willer, Wraight, X. Yuan, E. Zeggini, J. Hirschhorn, D. Strachan, W. Ouwehand, M. Caulfield, N. Samani, T. Frayling, P. Vollenweider, G. Waeber, V. Mooser, P. Deloukas, M. McCarthy, N. Wareham, I. Barroso, K. Jacobs, S. Chanock, R. Hayes, C. Lamina, C. Gieger, T. I. T. Meitinger, H. Wichmann, P. Kraft, S. Hankinson, D. Hunter, F. Hu, H. Lyon, B. Voight, M. Ridderstrale, L. Groop, P. Scheet, S. Sanna, G. A. G., Albai, R. N. D. Schlessinger, A. Jackson, J. Tuomilehto, and F. C. M. B. K. Mohlke (2008): "Common variants near mc4r are associated with fat mass, weight and risk of obesity," *Nature Genetics*, 40, 768–775.

Marchini, J., P. D. P., and L. Cardon (2005): "Genome-wide strategies for detecting multiple loci that influence complex diseases," *Nat.Genet.*, 37, 413–417.

Price, A., N. Patterson, R. Plenge, M. Weinblatt, N. Shadick, and D. Reich (2008): "Principal components analysis corrects for stratification in genome-wide association studies," *Nat. Genet.*, 38, 904–909.

Schaid, D., C. Rowland, D. Tines, R. Jacobson, and G. Poland (2002): "Score tests for association between traits and haplotypes when linkage phase is ambiguous," *Am. J. Hum. Genet.*, 70, 425–434.

Schmidt, M., E. Hauser, E. Martin, and S. Schmidt (2005): "Extension of the simla package for generating pedigrees with complex inheritance patterns: Environmental covariates, gene-gene and gene-environment interaction," *Stat. Appl. Genet. Mol. Biol.*, 4, Article15.

Shah, S., E. Hauser, D. Crosslin, L. Wang, C. Haynes, J. Connelly, S. Nelson,

Shah, S., E. Hauser, D. Crosslin, L. Wang, C. Haynes, J. Connelly, S. Nelson, J. Johnson, S. Gadson, C. Nelson, D. Seo, S. Gregory, W. Kraus, C. Granger, P. Goldschmidt-Clermont, and L. Newby (2008): "Alox5ap variants are associated with in-stent restenosis after percutaneous coronary intervention," *Atherosclerosis*, 201, 148–154.

Stephens, M., N. Smith, and P. Donnelly (2001): "A new statistical method for haplotype reconstruction from population data," *Am. J. Hum. Genet.*, 68, 978–989.

Wang, T., G. Ho, K. Ye, H. Strickler, and R. Elston (2008): "A partial least-square approach for modeling gene-gene and gene-environment interactions when multiple markers are genotyped," *Genet. Epidemiol.*

Wang, T., X. Zhu, and R. Elston (2007): "Improving power in contrasting linkage-disequilibrium patterns between cases and controls," *Am. J. Hum. Genet.*, 80, 911–920.

Weir, B. (1996): *Genetic Data Analysis II*, Sunderland , MA: Sinauer Associates, Inc., 2nd edition.

Weir, B. and C. Cockerham (1979): "Estimation of linkage disequilibrium in randomly mating populations," *Heredity*, 4, 105–111.

Zaykin, D. (2004): "Bounds and normalization of the composite linkage disequilibrium coefficient," *Genet. Epidemiol.*, 27, 252–257.

Zaykin, D., Z. Meng, and M. Ehm (2006): "Contrasting linkage-disequilibrium patterns between cases and controls as a novel association-mapping method," *Am. J. Hum. Genet.*, 78, 737–746.

Zhao, J., L. Jin, and M. Xiong (2006): "Test for interaction between two unlinked loci," *Am. J. Hum. Genet.*, 79, 831–845.

Zondervan, K. and L. Cardon (2004): "The complex interplay among factors that influence allelic association," *Nature Reviews Genetics*, 5, 89–100.

DOI: 10.2202/1544-6115.1561