A Testing Framework for Identifying Susceptibility Genes in the Presence of Epistasis

Joshua Millstein,[1] David V. Conti,[2] Frank D. Gilliland,[2] and W. James Gauderman[2]

[1]National Oceanic and Atmospheric Administration/National Marine Fisheries Service, Alaska Fisheries Science Center, Seattle; and [2]Department of Preventive Medicine, Division of Biostatistics, University of Southern California, Los Angeles

An efficient testing strategy called the "focused interaction testing framework" (FITF) was developed to identify susceptibility genes involved in epistatic interactions for case-control studies of candidate genes. In the FITF approach, likelihood-ratio tests are performed in stages that increase in the order of interaction considered. Joint tests of main effects and interactions are performed conditional on significant lower-order effects. A reduction in the number of tests performed is achieved by prescreening gene combinations with a goodness-of-fit χ^2 statistic that depends on association among candidate genes in the pooled case-control group. Multiple testing is accounted for by controlling false-discovery rates. Simulation analysis demonstrated that the FITF approach is more powerful than marginal tests of candidate genes. FITF also outperformed multifactor dimensionality reduction when interactions involved additive, dominant, or recessive genes. In an application to asthma case-control data from the Children's Health Study, FITF identified a significant multilocus effect between the nicotinamide adenine dinucleotide (phosphate) reduced:quinone oxidoreductase gene (*NQO1*), myeloperoxidase gene (*MPO*), and catalase gene (*CAT*) (unadjusted $P = .00026$), three genes that are involved in the oxidative stress pathway. In an independent data set consisting primarily of African American and Asian American children, these three genes also showed a significant association with asthma status ($P = .0008$).

The importance of accounting for gene-gene interactions in the search for susceptibility genes for complex diseases has been widely suggested to explain difficulties in replicating significant findings. Recent human and animal studies of complex diseases have identified susceptibility genes that marginally contribute to a common trait, to a minor extent only or not at all, but that interact significantly in combined analyses (Kuida and Beier 2000; Naber et al. 2000; Williams et al. 2000; Hsueh et al. 2001; Kim et al. 2001; Tripodis et al. 2001; Ukkola et al. 2001; Barlassina et al. 2002; De Miglio et al. 2004; Yanchina et al. 2004; Yang et al. 2004; Aston et al. 2005; Dong et al. 2005; Roldan et al. 2005). Several investigators have found alleles that have opposite effects depending on the genetic background (Balmain and Harris 2000; Staessen et al. 2001), which further raises the likelihood of overlooking epistatic susceptibility genes in single-gene analyses (Culverhouse 2002).

Accounting for interactions is not a trivial task, because of the serious multiple-testing problem created by the large number of possible interactions for even a relatively small set of candidate genes. For example, in the Children's Health Study (CHS), a prospective study of children's respiratory health, we are studying ~20 candidate genes related to oxidative stress and inflammatory pathways (Gilliland et al. 1999). These 20 genes yield 190 possible two-gene interactions and 1,140 possible three-gene interactions. If the multiple testing problem is ignored, type I error rates will be greatly inflated, leading to false conclusions and to studies that are difficult to replicate.

Foulkes et al. (2005) applied a combined dimension-reduction and mixed-modeling approach to four SNPs in three lipase genes to assess risk of cardiovascular disease. Although their approach accounts for possible interactions and allows controlling for possible confounders, it is unclear what the performance or proper implementation would be for a larger set of candidates. Devlin et al. (2003) showed that type I error rates were extremely inflated for model-selection methods such as the Lasso (Tibshirani 1996). Another multilocus approach is the set-association approach (Hoh et al. 2001), which uses sums of statistics based on locus-specific association and Hardy-Weinberg disequilibrium to test a global null hypothesis. This approach may be powerful for finding many small effects that combine to have an important effect on the phenotype but does not explicitly account for possible epistatic interactions.

Several data-mining approaches have been developed to address the problem of identifying susceptibility genes involved in epistatic interactions (Ritchie et al. 2001, 2003*b*; Moore and Hahn 2002; Bastone et al. 2004; Cook et al. 2004; Culverhouse et al. 2004; Foulkes et al. 2004); however, their performance may be limited

Received July 6, 2005; accepted for publication October 5, 2005; electronically published November 11, 2005.

Address for correspondence and reprints: Dr. Joshua Millstein, National Oceanic and Atmospheric Administration/National Marine Fisheries Service, Alaska Fisheries Science Center, 7600 Sand Point Way NE, Building 4, 1133B, Seattle, WA 98115. E-mail: josh.millstein@noaa.gov

Am. J. Hum. Genet. 2006;78:15–27. 0002-9297/2006/7801-0004$15.00

in the presence of main effects or genetic heterogeneity. Also, properties related to type I error rates and power have not been thoroughly compared with more traditional approaches. A data-mining approach that has generated some recent interest is multifactor dimensionality reduction (MDR) (Ritchie et al. 2001, 2003*a*; Hahn et al. 2003; Bastone et al. 2004; Cho et al. 2004; Coffey et al. 2004; Hahn and Moore 2004; Moore 2004; Tsai et al. 2004; Williams et al. 2004; Qin et al. 2005; Soares et al. 2005). MDR is a nonparametric method designed to detect genes involved in high-order interactions in case-control studies (Ritchie et al. 2001). To implement this method, the investigator must first specify the number of interacting genes, k, to consider (throughout this article, we will consider three genes). The data are divided into 10 equal parts, and the phenotypes of subjects in each 1/10 of the data are predicted by the MDR model derived from the remaining 9/10 of the data. For each 9/10 of the data, several steps are performed. For every set of k genes, MDR classifies each multilocus genotype as "high risk" or "low risk," depending on the ratio of cases to controls. The subjects in the "high risk" groups are then pooled. The k-gene set that maximizes the cases: controls ratio in the pooled "high risk" group is selected as the "best" gene set. Disease status for subjects in the remaining 1/10 of the data is then predicted on the basis of genotype risk for the "best" gene set. The overall "best" gene set is determined by the data split with the lowest prediction error. Prediction error is averaged over the 10 data splits and is used as a measure of predictive power. Another useful measure, termed "consistency," is the number of data splits with the same "best" set of factors.

We developed a new search strategy designed to identify susceptibility genes among a group of candidate genes in the presence of gene-gene interactions. The candidate genes may be selected for their role in a specific biochemical pathway or from a prior genome scan for linkage. A powerful testing framework based on likelihood-ratio tests (LRTs) is presented here that simultaneously tests multilocus effects across various orders of interaction. Our search strategy also employs a screening statistic to reduce the total number of gene sets that are tested for multilocus effects. We present an assessment of power and type I error from simulation analysis and compare the method's performance with that of MDR. We then apply both our method and MDR to a case-control data set from the CHS that includes 12 candidate loci measured in asthmatic and nonasthmatic subjects.

Methods

Consider a disease phenotype, D, and a sample of cases ($D = 1$) and controls ($D = 0$) selected from some population. We assume that genotypes are obtained for each subject for a set of diallelic, autosomal candidate loci. For each candidate locus, indexed by $i, j, k,\ldots$, we define a covariate, G, with possible values 0, 1, or 2, corresponding to genotypes aa, Aa, and AA, respectively. This defines a log-additive coding scheme, a robust approach when the specific genetic model is unknown (Schaid 1996). We note, however, that the methods presented here are readily adaptable to alternative risk models (e.g., dominant, recessive, or codominant). We adopt a logistic model to relate genes to D. For example, the fully saturated model for a set of three candidate genes has the form

$$\begin{aligned}\text{logit}[P(D = 1)] = {} & \beta_0 + \beta_i G_i + \beta_j G_j + \beta_k G_k \\ & + \beta_{ij} G_i G_j + \beta_{ik} G_i G_k + \beta_{jk} G_j G_k \\ & + \beta_{ijk} G_i G_j G_k \ . \end{aligned} \tag{1}$$

The model contains three main effects, three two-way interactions, and one three-way interaction. An analogous saturated model for two genes would be

$$\text{logit}[P(D = 1)] = \beta_0 + \beta_i G_i + \beta_j G_j + \beta_{ij} G_i G_j \ , \tag{2}$$

whereas a model for a single gene would be

$$\text{logit}[P(D = 1)] = \beta_0 + \beta_i G_i \ . \tag{3}$$

LRTs can be used to identify susceptibility genes by testing the parameters in the above models. An LRT statistic is computed as $\chi^2 = 2(L_{\text{full}} - L_{\text{reduced}})$, where L_{full} is the log-likelihood of the data computed under a fully specified model and L_{reduced} is the log-likelihood computed under the constraint that one or more parameters equal zero. Under the null hypothesis, this statistic has a χ^2 distribution with df equal to the difference in the number of unconstrained parameters between the full and reduced models. Three LRT testing strategies for identification of genes will be considered.

Marginal Effects

The simple model in equation (3) is used to test the null hypothesis $\beta_i = 0$ for each candidate gene. We refer to this test as the marginal test of G_i, since the estimated effect from this model, β_i, represents an average of the main effect of G_i and any interactive effects with other loci. With a total of K candidate genes, there are K marginal tests. The threshold for significance is adjusted for multiple testing by controlling false-discovery rates (FDRs) (Benjamini and Hochberg 1995), although other approaches (e.g., Bonferroni adjustment) could be adopted. In brief, Benjamini and Hochberg (1995) defined FDR as the ratio of the number of falsely rejected null hypotheses to the total number of rejected null hypotheses. They showed that the expected FDR can be controlled by a procedure that applies a cutoff to the unadjusted ordered P values, $P_{(1)}, P_{(2)},\ldots,P_{(i)},\ldots,P_{(m)}$. All null hypotheses with P values at or below cutoff t are rejected; specifically,

$$t = \max\left\{P_{(i)}: P_{(i)} \leq \frac{i\alpha}{m}\right\} \ .$$

Interaction Testing Framework (ITF)

In this strategy, tests are performed in a series of stages, with an incremental increase in the highest-order interaction parameter considered at each subsequent stage. The first stage tests the main effect of each gene, the second stage tests all possible two-way interactions, the third stage tests all three-way interactions, and so forth. To avoid retesting the same effects, a test in a higher stage (e.g., test of a specific two-way interaction in stage 2) is conditioned on any component factors (e.g., either of the two genes involved in that two-way interaction) that were already declared significant in a lower stage (e.g., stage 1). Gene sets are tested for multilocus effects, whether or not marginal effects were found. Type I error is controlled by dividing the overall α level by the number of stages and allocating this adjusted α level, α^*, to each stage. Within each stage, the threshold for significance is adjusted by controlling FDR. The specific stages are as follows.

1. *First stage.* Perform marginal LRTs of β_i for each of the K candidate genes. Declare a test significant if $P_i < \alpha^*_1$, where P_i is the P value that corresponds to the ith LRT and α^*_1 denotes the significance threshold for first-stage tests corrected to control FDR. A total of K tests are conducted in this stage.
2. *Second stage.* For all possible two-gene sets ($K(K-1)/2$), the full model (eq. [2]) is tested against the reduced model,

$$\text{logit}[P(D=1)] = \beta_0 + \beta_i G_i I() + \beta_j G_j I() \ ,$$

   where $I()$ is an indicator function that assumes the value 1 if the corresponding term was statistically significant in a first-stage test and 0 otherwise. Thus, if both β_i and β_j were statistically significant in the first stage, the reduced model would be $\beta_0 + \beta_i G_i + \beta_j G_j$, and the interaction between G_i and G_j would be tested in a 1-df test in this second stage. On the other hand, if neither β_i nor β_j was statistically significant in the first stage, then a 3-df test of β_i, β_j, and β_{ij} would be conducted in the second stage. This selective conditioning is done to avoid retesting effects that have already been declared significant. Significance is declared if $P_{ij} < \alpha^*_2$, where P_{ij} is the P value that corresponds to the ijth LRT and α^*_2 denotes the significance threshold for second-stage tests corrected to control FDR.
3. *Third stage.* All three-gene sets are tested (the number of tests is $K(K-1)(K-2)/6$) in a fashion similar to the method in stage 2. The saturated model (eq. [1]) is tested against the reduced model,

$$\begin{aligned}\text{logit}[P(D=1)] = {} & \beta_0 + \beta_i G_i I() + \beta_j G_j I() \\ & + \beta_k G_k I() + \beta_{ij} G_i G_j I() \\ & + \beta_{ik} G_i G_k I() + \beta_{jk} G_j G_k I() \ ,\end{aligned}$$

   where, again, the indicator function $I()$ assumes the value 1 if the term was in a model that achieved statistical significance in a previous stage and 0 otherwise. It should be stated explicitly that a model that includes higher-order terms would always include the component lower-order terms.

The ITF approach described thus far can be directly generalized to multilocus effects involving four or more genes.

Focused ITF (FITF)

It is clear that the number of tests conducted in the ITF method can be quite large when K is large. Adjusting the type I error for so many tests may cause an unacceptable loss in power. We developed a method for prescreening all possible gene sets, to focus attention on those that are most likely to be informative in the ITF.

Let G_{ijk} denote a multilocus genotype over a set of three candidate genes i, j, and k. Then, by the Bayes theorem, the probability that a case possesses the particular genotype G_{ijk} is

$$P(G_{ijk} \mid D=1) = \frac{P(D=1 \mid G_{ijk}) P(G_{ijk})}{P(D=1)} \ .$$

The factor $P(G_{ijk})$ describes the population distribution of G_{ijk}, which, under our assumption of locus independence, is simply a product of the corresponding genotype frequencies. If the three loci combine to affect disease risk, $P(G_{ijk}|D=1)$ will differ from $P(G_{ijk})$ by an amount that depends on the magnitude of risk that G_{ijk} confers. One might compute a measure of difference between the observed distribution of G_{ijk} in cases and that expected on the basis of the product of genotype frequencies and then focus the third stage of the ITF on only those sets with a difference that exceeds some threshold. However, the use of only cases in this screening step will induce a bias into the ITF because of the explicit use of disease status. Rather, we propose to compute this difference measured with the pooled sample of cases and controls, to avoid this bias. A deviation from the expected prevalence of G_{ijk} in the entire case-control sample could be the result of a deviation from the expected prevalence of G_{ijk} in cases and could thus indicate association with disease.

The measure of difference we propose to use is a χ^2 goodness-of-fit statistic that compares the observed with the expected distribution of G_{ijk} in the combined case-control sample. The χ^2 statistic is then used as the criterion by which to choose gene combinations for inclusion in ITF—that is, only gene sets with a calculated χ^2 statistic above a selected cutoff value are analyzed. The form of the χ^2 statistic should match the underlying assumptions of risk—in other words, for the risk model in equation (1), the genotype groups would be chosen to match risk levels associated with each interaction term. For instance, there would be four genotype groups for two-gene sets, corresponding to $G_i \times G_j = 0$, 1, 2, or 4, and five genotype groups for three-gene sets, corresponding to $G_i \times G_j \times G_k = 0$, 1, 2, 4, or 8. The χ^2 statistic, henceforth referred to as the "CSS" (chi-squared subset) statistic, would then take the form

$$\text{CSS} = \sum_{i=1}^{r} \frac{[n_i - E(n_i)]}{E(n_i)} \ .$$

Table 1

Sample Size and Relative Efficiency of LRTs for Two Genes Involved in a Gene-Gene Interaction

Test	logit[$P(D = 1)$]		df	N[a]	RE[b]
	Saturated	Null			
1	$\beta_0 + \beta_1 G_1$	β_0	1	130	1
2	$\beta_0 + \beta_1 G_1 + \beta_2 G_2$	β_0	2	90	1.44
3	$\beta_0 + \beta_1 G_1 + \beta_2 G_2 + \beta_{12} G_1 G_2$	β_0	3	73	1.78
4	$\beta_0 + \beta_1 G_1 + \beta_2 G_2 + \beta_{12} G_1 G_2$	$\beta_0 + \beta_1 G_1 + \beta_2 G_2$	1	192	.68

NOTE.—The prevalence of D was set to 0.1, and both allele frequencies were set to 0.3. Genes G_1 and G_2 were assumed to be log-additive with no main effects ($\beta_1 = 0$; $\beta_2 = 0$) but with an interaction OR of 2 ($\beta_{12} = \log(2)$). The df value for a test is the difference in the number of parameters between the null and saturated models.

[a] N denotes the number of case-control pairs required to achieve 80% power with an α level of 0.05 for an LRT comparing the saturated model with the null model.

[b] RE = relative efficiency = N(test 1)/N(test i).

Here, n_i is the observed number of subjects, irrespective of case status, in the ith genotype group, and r is the total number of genotype groups. The expected n_i, $E(n_i)$, is estimated on the basis of the sample marginal genotype frequencies of each gene. For example, let n_4 equal the observed number of subjects with $G_i \times G_j = 4$—in other words, genotype AA at locus i and BB at locus j—then, for two-gene sets, $E(n_4) = (n_{AA} n_{BB})/N^2$, where N denotes the total sample size.

We emphasize the point that use of the CSS statistic to limit the number of gene sets considered does not bias subsequent tests. Under the global null hypothesis of independence between genotype **G** and phenotype D, any variable that is strictly a function of **G** will also be independent of D. Specifically, the reduced set of gene combinations ($\mathbf{G}^*$) that results from screening with the CSS statistic is strictly a function of **G**, since case-control status is not used in computing CSS. Therefore, the reduced set is also statistically independent of D under the global null.

Simulations and Results

Proof of Concept

As an initial proof of concept, we first provide evidence to show that accounting for interactions leads to increased efficiency in tests of candidate genes. We assume a model with no main effects and a two-gene interaction with an odds ratio (OR) of 2.0—that is, under equation (2), β_{ij} is set to log(2), and β_i and β_j are set to zero. Phenotype prevalence was set to 10%, allele frequencies were set to 0.3, power was set to 80%, and the significance level was assumed to be 0.05 with a two-sided alternative hypothesis. Conditional on all of these parameter settings, the method of Longmate (2001) was used to estimate required sample sizes for a variety of LRTs derived from equation (2). Test 1 (see table 1) shows the sample size required ($N = 130$) to detect G_1 (or G_2) by use of a standard marginal test. A 2-df test of β_1 and β_2 (test 2) requires only $N = 90$, a 44% increase in efficiency. A 3-df test of the saturated model with both main effects and their interaction (test 3) further increases power, providing a 78% improvement over test 1. Interestingly, test 4, a 1-df test of the interaction parameter β_{12} alone, is less powerful than any of the other tests. In general, this example demonstrates the potential power of multi-df tests that include both main and interaction effects, and these findings are the impetus for our investigations of the ITF strategy.

Type I Error

To evaluate type I error, populations were simulated as having 20 independent diallelic candidate genes, $G_1,\ldots,G_{20}$, with allele frequencies between 0.1 and 0.33. A binary disease phenotype with a population prevalence of 10% was randomly assigned to individual subjects, independent of genotype. A case-control data set with 200 cases and 200 controls was then sampled from the simulated population for analysis. This process was repeated 1,000 times, and the number of replicates with one or more false-positive tests was recorded. For the ITF and FITF strategies, the experimentwise significance level was set to 0.05. For these and most subsequent analyses in this study, the CSS cutoff values for the FITF method were set to 3 for the second stage and 6 for the third stage. The MDR method was also used to analyze all 1,000 null data sets. For each data set, MDR was run on the entire group of 20 candidate genes, with the instruction to consider sets of 3 candidate genes only. Although it is possible to instruct MDR to consider gene sets of sizes other than three, we chose this size in an effort to optimize the potential power of MDR in our comparisons, since it corresponded to the number of genes we simulated to affect disease risk. The 5th and 95th percentiles of prediction error and consistency were

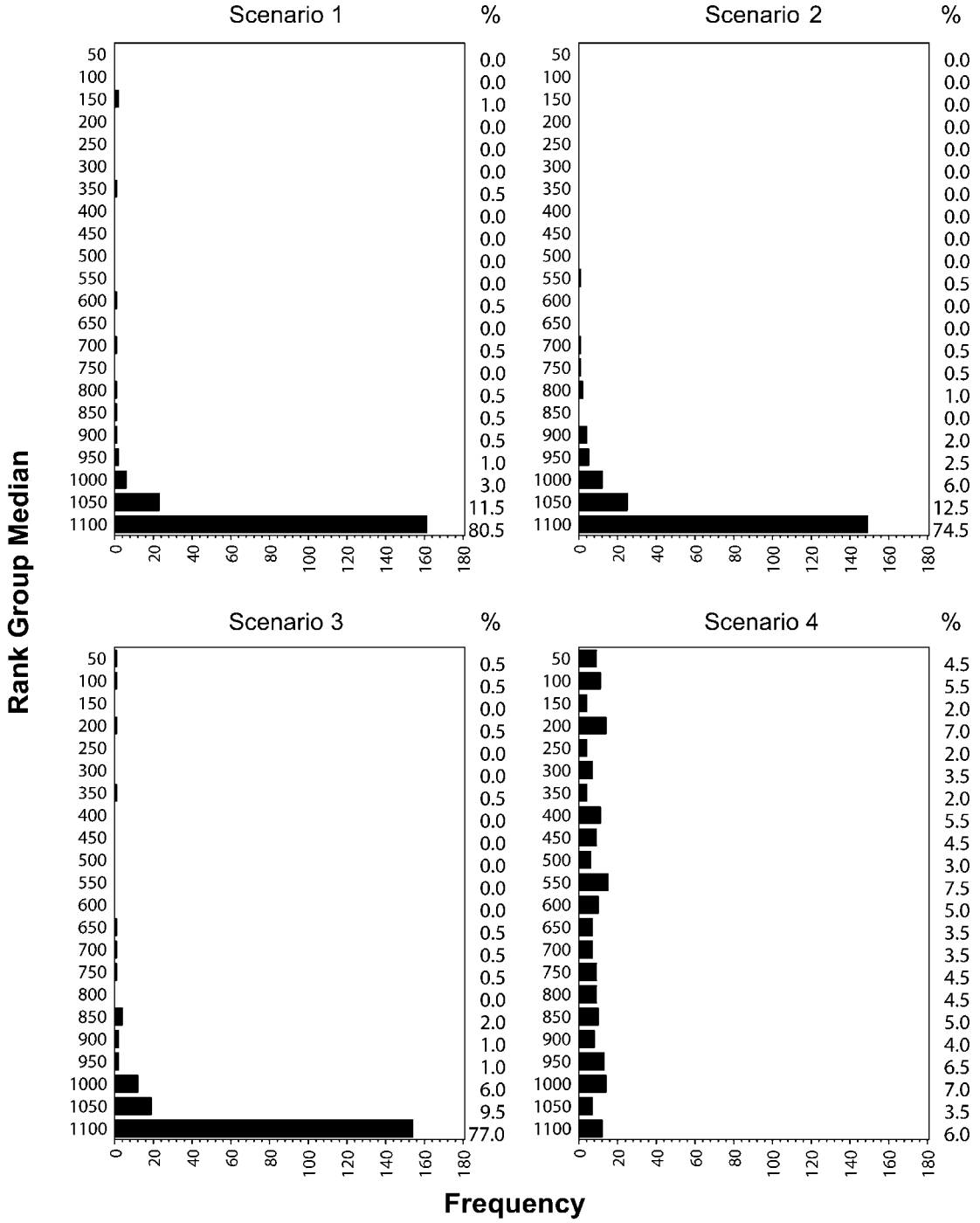

Figure 1 Histograms for scenarios 1–4. The rank of the observed CSS statistic for the true set of susceptibility genes, gene set $\{G_1,G_2,G_3\}$, was recorded for each of the 200 simulations per risk scenario. With 20 candidate genes, there are 1,140 three-gene combinations; thus, there were 1,140 possible ranks. Histograms were constructed for each set of the 200 ranks of gene set $\{G_1,G_2,G_3\}$ for each risk scenario. The four risk scenarios (following the form of eq. [1]) were as follows: scenario 1, $\beta_{123} = \log(3)$; scenario 2, $\beta_1 = \log(1.5)$ and $\beta_{123} = \log(3)$; scenario 3, $\beta_1 = \log(1.5)$, $\beta_2 = \log(.67)$, and $\beta_{123} = \log(3)$; and scenario 4, $\beta_1 = \log(1.5)$, $\beta_2 = \log(1.5)$, and $\beta_3 = \log(1.5)$. All other nonintercept parameters were set to zero.

Table 2

Frequencies of Susceptibility Genes Identified from Simulated Data Sets by Use of Five Testing Strategies

Risk Scenario (ORs[a]) and Testing Strategy	No. of Susceptibility Genes Identified			
	Marginal	2nd Stage[b]	3rd Stage[b]	Total (%)
Scenario 1 (3, 1, 1, 1):				
Marginal	148	…	…	148 (24.7)
ITF	97	79	43	219 (36.5)
FITF	97	90	85	272 (45.3)
MDR(e)	…	…	…	41 (6.8)
MDR(c)	…	…	…	20 (3.3)
Scenario 2 (3, 1.5, 1, 1):				
Marginal	294	…	…	294 (49.0)
ITF	242	40	16	298 (49.7)
FITF	242	62	34	338 (56.3)
MDR(e)	…	…	…	131 (21.2)
MDR(c)	…	…	…	31 (5.2)
Scenario 3 (3, 1.5, .67, 1):				
Marginal	198	…	…	198 (33.0)
ITF	160	49	19	228 (38.0)
FITF	160	60	73	293 (48.8)
MDR(e)	…	…	…	102 (17.0)
MDR(c)	…	…	…	38 (6.3)
Scenario 4 (1, 1.5, 1.5, 1.5):				
Marginal	174	…	…	174 (29.0)
ITF	109	17	6	132 (22.0)
FITF	109	17	16	142 (23.7)
MDR(e)	…	…	…	53 (8.8)
MDR(c)	…	…	…	30 (5.0)

Note.—Tests were performed on 1,000 simulated case-control data sets with 20 diallelic candidate genes and a binary phenotype. Each data set consisted of 200 cases and 200 controls. Allele frequencies were set between 0.1 and 0.33, and disease prevalence was set to 0.1. For FITF, CSS cutoffs were set to 3 for second-stage tests and 6 for third-stage tests. MDR(e) denotes a test using the average prediction error for the evaluating data sets obtained from MDR software with a cutoff of 42.185%. MDR(c) denotes a test using "consistency" of MDR results with a cutoff of 8—that is, a consistency of 9 or 10 results in a rejected test. OR_1, OR_2, and OR_3 denote the ORs for the main effects of the three susceptibility genes—for example, $OR_1 = \exp(\beta_1)$. OR_{123} denotes the OR for the interaction parameter—that is, $OR_{123} = \exp(\beta_{123})$. The number of susceptibility genes identified was based on the analysis of 200 simulated data sets for each risk scenario, each with three susceptibility genes (genes simulated to confer risk in accordance with one of four risk scenarios) out of a total of 20 candidate genes—that is, the total number of susceptibility genes was $200 \times 3 = 600$ (the total number of candidate genes was $200 \times 20 = 4{,}000$).

[a] (OR_{123}, OR_1, OR_2, OR_3).

[b] Susceptibility genes identified by LRTs that were not identified in previous stages.

used as significance thresholds for the tests denoted by MDR(e) and MDR(c), respectively. These two tests, by construction, have approximately correct type I error rates under these conditions. Significance thresholds for MDR computed under the null hypothesis will be fixed in subsequent simulations, to evaluate power under various alternative hypotheses.

Estimates of type I error rates for the marginal, ITF, and FITF strategies were close to the nominal 5% level (6.4, 3.5, and 4.7, respectively). For MDR, the 5th percentile of prediction error was equal to 42.19, and declaring the MDR(e) test significant when prediction error was equal to or below this value resulted in false-positive tests in exactly 5.0% of the replicates. Consistency is an integer that ranges from 0 to 10. Setting the significance level of the MDR(c) test such that significance was declared when consistency was equal to 9 or 10 produced false-positive tests in 4.7% of the replicates.

Power

To compare the power of the five testing strategies, populations were simulated with 20 independent diallelic candidate genes, $G_1,\ldots,G_{20}$, each with allele frequencies randomly sampled between 0.1 and 0.33. A binary disease phenotype, with overall population prev-

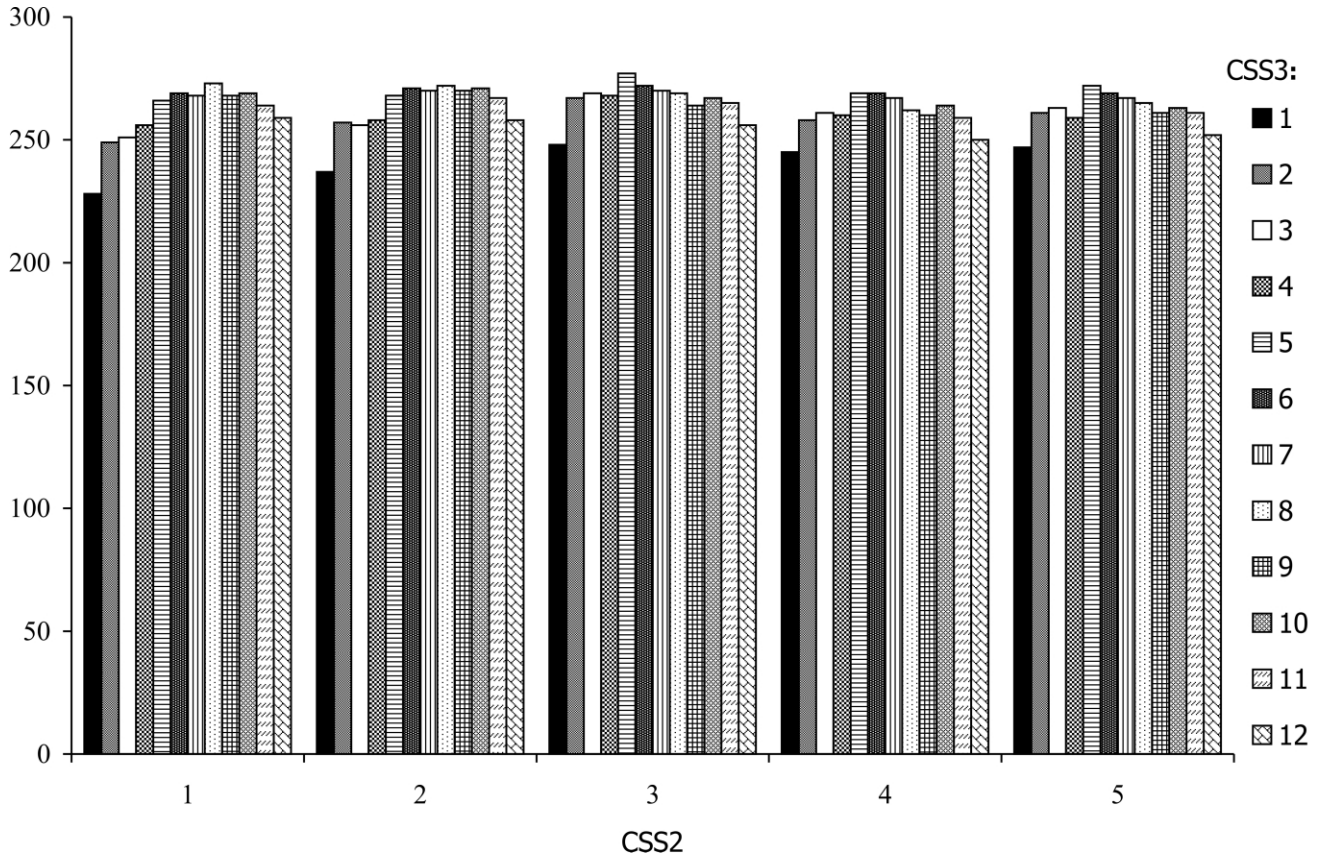

Figure 2 Genes identified using CSS cutoffs. Bar height represents the total number of true susceptibility genes identified by the FITF method with the use of a range of CSS cutoff values: CSS2 for the second stage and CSS3 for the third stage.

alence 10%, was assigned in accordance with one of four risk scenarios. The four scenarios, each of which involved the three genes G_1, G_2, and G_3, were as follows: scenario 1, $\beta_{123} = \log(3)$; scenario 2, $\beta_1 = \log(1.5)$ and $\beta_{123} = \log(3)$; scenario 3, $\beta_1 = \log(1.5)$, $\beta_2 = \log(.67)$, and $\beta_{123} = \log(3)$; and scenario 4, $\beta_1 = \log(1.5)$, $\beta_2 = \log(1.5)$, and $\beta_3 = \log(1.5)$. The remaining parameters in equation (1) (excluding the intercept) were set to zero in each scenario, and none of the genes $G_4,\ldots,G_{20}$ were simulated to influence disease risk. A case-control data set with 200 cases and 200 controls was sampled from each simulated population for analysis. This process was repeated 200 times for each of the four risk scenarios.

The distribution of the 200 ranks of the CSS statistics for the true set of susceptibility genes, $\{G_1,G_2,G_3\}$, was examined for each scenario (fig. 1). The rank of this statistic should be uniformly distributed across replicates if it is not capturing interaction information and should be centered over the higher ranks if it is. In the absence of interaction (scenario 4), the ranks were uniformly distributed over the entire range, as expected. In contrast, in the presence of interaction ($\beta_{123} = \log(3)$ under scenarios 1–3), the CSS statistic for gene set $\{G_1,G_2,G_3\}$ was among the highest ranked sets in most replicates. These results imply that, under these conditions, one can ignore low-ranking gene sets in the search for gene-gene interactions without an appreciable probability of ignoring gene sets with true interaction effects.

Table 2 presents the number of susceptibility genes found by each of the five testing strategies applied to the 200 simulated case-control data sets for each of the four risk scenarios. For each data set, MDR was run on the entire group of 20 candidate genes, with the instruction to consider sets of 3 candidate genes only. A susceptibility gene was considered identified by MDR when the gene was in the final set of the three "best" genes and the prediction error was below the significance threshold. Prediction error and consistency were considered separately. With a log-additive interaction between genes G_1, G_2, and G_3 but no main effects (scenario 1), FITF was clearly the most powerful method, identifying almost 24% more susceptibility genes than the ITF method did and almost 84% more than the marginal method did. FITF also performed much better than MDR, detecting >6 times the number of susceptibility genes found by MDR(e) and >13 times the number found by

Table 3

Frequencies of Susceptibility Genes Identified from Data Simulated in Accordance with Logical Epistasis Rules

Risk Scenario and Testing Strategy	No. of Susceptibility Genes Identified			
	Marginal	2nd Stage[a]	3rd Stage[a]	Total (%)
Scenario 5[b]:				
Marginal	457	...	...	457 (76.2)
ITF	429	167	4	600 (100.0)
FITF	429	170	1	600 (100.0)
MDR(e)	...	...	...	600 (100.0)
MDR(c)	...	...	...	600 (100.0)
Scenario 6[c]:				
Marginal	16	...	...	16 (2.7)
ITF	11	133	267	411 (68.5)
FITF	11	181	182	374 (62.3)
MDR(e)	...	...	...	600 (100.0)
MDR(c)	...	...	...	600 (100.0)

Note.—Tests were performed on 1,000 simulated case-control data sets with 20 diallelic candidate genes and a binary phenotype. Each data set consisted of 200 cases and 200 controls. Allele frequencies were set between 0.1 and 0.33, and disease prevalence was set to 0.1. For FITF, CSS cutoffs were set to 3 for second-stage tests and 6 for third-stage tests. MDR(e) denotes a test that uses the average prediction error for the evaluated data sets obtained from MDR software with a cutoff of 42.185%. MDR(c) denotes a test that uses "consistency" of MDR results with a cutoff of 8—that is, a consistency of 9 or 10 results in a rejected test. For risk scenario 6, the susceptibility genes are also genes 1, 2, and 3, and the susceptible homozygous genotypes may be wild type or variant. The number of susceptibility genes identified was based on the analysis of 200 simulated data sets for each risk scenario, each with 3 susceptibility genes out of a total of 20 candidate genes—that is, the total number of susceptibility genes was 200 × 3 = 600 (the total number of candidate genes was 200 × 20 = 4,000).

[a] Susceptibility genes identified by LRTs that were not identified in previous stages.

[b] Epistasis rules: If $G_1 + G_2 + G_3 = 2$, then penetrance = 0.2; otherwise, penetrance = 0.

[c] Epistatis rules: If any two genes are homozygous and the third is heterozygous, then penetrance = 0.1; otherwise, penetrance = 0.

MDR(c). In the presence of main effects and interactions (scenarios 2 and 3), FITF again outperformed the other methods in identifying susceptibility genes. With main effects but no interaction (scenario 4), FITF detected 18% fewer susceptibility genes than the marginal method did, whereas the ITF, MDR(e), and MDR(c) detected 24%, 70%, and 83% fewer genes, respectively, than the marginal method did. For all four risk scenarios, the marginal, ITF, and FITF methods clearly outperformed MDR. For ITF and FITF, most of the identified susceptibility genes were detected in the first stage, with diminishing but nontrivial numbers first identified in the second and third stages.

Our choice of the CSS cutoff values in the FITF method (3 for stage 2 and 6 for stage 3) may seem somewhat arbitrary. To address the sensitivity of results to this choice, we reanalyzed the simulated data in scenario 1, using a range of alternative CSS values. The results, shown in figure 2, demonstrate that nearly equal power was obtained across a range of stage 2 and stage 3 CSS cutoff values. This indicates that the FITF method is reasonably insensitive to the specific cutoff values used.

Two additional risk scenarios (scenarios 5 and 6) were simulated to assess the comparison of the marginal, ITF, FITF, and MDR methods. These scenarios were equivalent to two special types of interactions that were previously used to test the performance of MDR (Ritchie et al. 2001, 2003*a*; Moore 2004). In scenario 5, disease risk was set to 0.2 if $G_1 + G_2 + G_3 = 2$ and to 0.0 otherwise. Under this scenario, for example, the multilocus genotypes {AA,aa,aa} and {aa,aA,aA} each produce risk of 0.2. In scenario 6, when any two of the three susceptibility genes (G_1, G_2, and G_3) were homozygous and the third was heterozygous, disease risk was set to 0.1; otherwise, risk was set to 0.0. For example, the multilocus genotypes {aa,aa,aA} and {AA,aA,aa} would each produce risk of 0.1. As in the work by Ritchie et al. (2001, 2003*a*, 2003*b*), our simulations demonstrated that MDR was able to identify 100% of the susceptibility genes in both scenarios 5 and 6 (table 3). In scenario 5, ITF and FITF also detected 100% of the true genes, whereas the marginal method only detected 76%. In scenario 6, ITF identified almost 68% of the susceptibility genes, whereas FITF identified 62%, and the marginal method found only 2.7%. In retrospect, these results are not too surprising, given that scenarios 5 and 6 demonstrate interactions that deviate strongly from our underlying logistic model of additivity of effects (eq. [1]).

Another type of nonadditive interaction to consider is the joint effect of genes that act in a dominant or recessive pattern. Four scenarios (scenarios 7–10) were simulated for interactions of this type. Again, susceptibility at genes G_1, G_2, and G_3 was simulated to increase disease risk ($\beta_{123} = \log(8.0)$), where now each gene was coded as either dominant ($G = 1$ if g = AA or Aa; $G = 0$ otherwise) or recessive ($G = 1$ if g = aa; $G = 0$ otherwise). Note that the susceptible genotype for the recessive pattern was defined to be the common homozygous type, which was required to have adequate numbers of susceptible individuals in the data. For this set of simulations, minor-allele frequencies were based

Table 4

Asthma-Related Genes Included in Case-Control Analysis of CHS Data

The table is available in its entirety in the online edition of *The American Journal of Human Genetics*.

Table 5

Frequencies of Susceptibility Genes Identified from Simulated Data Sets in Accordance with Dominant and Recessive Epistatic Risk Patterns

Risk Scenario (Gene Pattern[a]) and Testing Strategy	No. of Susceptibility Genes Identified			
	Marginal	2nd Stage[b]	3rd Stage[b]	Total (%)
Scenario 7 (dom, dom, dom):				
Marginal	337	...	...	337 (56.2)
ITF	292	45	26	363 (60.5)
FITF	292	46	47	385 (64.2)
MDR(e)	...	...	...	311 (51.8)
MDR(c)	...	...	...	186 (31.0)
Scenario 8 (dom, dom, rec):				
Marginal	334	...	...	334 (55.2)
ITF	287	43	25	355 (59.2)
FITF	287	55	55	397 (66.2)
MDR(e)	...	...	...	310 (51.7)
MDR(c)	...	...	...	161 (26.8)
Scenario 9 (dom, rec, rec):				
Marginal	329	...	...	329 (54.8)
ITF	287	48	22	357 (59.5)
FITF	287	48	64	399 (66.5)
MDR(e)	...	...	...	296 (49.3)
MDR(c)	...	...	...	135 (22.5)
Scenario 10 (rec, rec, rec):				
Marginal	298	...	...	298 (49.7)
ITF	269	44	20	333 (55.5)
FITF	269	45	26	340 (56.7)
MDR(e)	...	...	...	264 (44.0)
MDR(c)	...	...	...	123 (20.5)

Note.—Tests were performed on 1,000 simulated case-control data sets with 20 diallelic candidate genes and a binary phenotype. Each data set consisted of 200 cases and 200 controls. Minor-allele frequencies were based on the minor-allele frequencies of 12 genes given in table 4 for whites and Hispanics from the CHS data. Disease prevalence was set to 0.1. For FITF, CSS cutoffs were set to 3 for second-stage tests and 6 for third-stage tests. MDR(e) denotes a test that uses the average prediction error for the evaluated data sets obtained from MDR software with a cutoff of 41.33%. MDR(c) denotes a test that uses "consistency" of MDR results with a cutoff of 8—that is, a consistency of 9 or 10 results in a rejected test. Subjects with susceptible genotypes at all three loci had a relative risk of 8. The number of susceptibility genes identified was based on the analysis of 200 simulated data sets for each risk scenario, each with 3 susceptibility genes out of a total of 20 candidate genes—that is, the total number of susceptibility genes was 200 × 3 = 600 (the total number of candidate genes was 200 × 20 = 4,000).

[a] Gene pattern for G_1, G_2, G_3 ($OR_i = 8$). dom = dominant; rec = recessive.

[b] Susceptibility genes identified by LRTs that were not identified in previous stages.

on the minor-allele frequencies of 12 genes in whites and Hispanics from the CHS cohort (table 4). Cutoff values for the MDR(e) and MDR(c) tests were reassessed under these allele frequencies, which resulted in a cutoff of 41.33 for MDR(e) and no change for MDR(c). Allowing the significance level of the MDR(c) test to remain, such that significance was affirmed when consistency was equal to 9 or 10, produced false-positive tests in 5.0% of 1,000 simulated replicate data sets under the null hypothesis.

In all four of the risk scenarios, FITF identified the most susceptibility genes (table 5). The remaining methods, in decreasing order of effectiveness, were the ITF, marginal method, MDR(e), and MDR(c). The increase in power of the FITF over the ITF is evident at both the second and third stages.

In each of the above simulations, genes $G_4,\ldots,G_{20}$ represented non–disease-related loci that potentially could have been identified (falsely) as susceptibility genes. When a multilocus test was declared significant but included non–disease-related loci, these loci were considered to be falsely identified as susceptibility genes. The percentage of these genes in statistically significant tests (the "null percentage") was substantially lower for the marginal, ITF, and FITF approaches than for MDR (table 6).

Table 6

Average Percentage of Non–Disease-Related Genes Found in Significant Tests

The table is available in its entirety in the online edition of *The American Journal of Human Genetics*.

Application to Asthma Case-Control Data Set

Oxidative stress, resulting from increased amounts of reactive oxygen species (ROS), such as superoxide radicals and hydrogen peroxide (H_2O_2), has been implicated in the pathogenesis of asthma, a chronic inflammatory airway disease (MacNee 2001; Mak et al. 2004). The marginal, ITF, FITF, and MDR methods were applied to a case-control data set from the CHS with candidate genes that have common functional polymorphisms and that are related to the oxidative stress and inflammatory pathways (table 4). Informed consent was obtained from CHS subjects or their parents. For each locus, *G* was coded as 0, 1, or 2 by the number of variant alleles present, with the exception of the glutathione S-transferase M1 gene (*GSTM1* [MIM 138350]) and glutathione S-transferase T1 gene (*GSTT1* [MIM 600436]), which were coded present (0) or null (1).

The analysis was restricted to 2,089 non-Hispanic and Hispanic white children with nonmissing genetic data from CHS cohorts. This group included 321 children with physician-diagnosed asthma at study entry. The sample was further restricted to create a balanced case-control data set by random selection of 321 children without physician-diagnosed asthma at study entry to be used as controls, creating a total sample size of 642 subjects. For the MDR analysis, an additional case-control pair was randomly excluded to produce a data set of 640 subjects, since MDR version 1.4.1 required a sample size divisible by 10 for the purpose of 10-fold cross-validation (in the most recent version, this is no longer a requirement).

Analysis of these data resulted in statistically significant effects, after adjustment for multiple comparisons, detected by both the marginal and the FITF methods. The marginal method identified a protective effect for the nicotinamide adenine dinucleotide (phosphate) reduced:quinone oxidoreductase gene (*NQO1* [MIM 125860]) (OR = 0.67; adjusted $P = .035$). The FITF method identified a three-gene effect involving *NQO1,* the myeloperoxidase gene (*MPO* [MIM 606989]), and the catalase gene (*CAT* [MIM 115500]) (unadjusted $P = .00026$; significance threshold 0.00052). In sensitivity analyses, this set of three genes maintained the same level of significance after adjustment for sex, study community, study cohort, maternal tobacco-smoke exposure during pregnancy, and parental physician-diagnosed asthma (data not shown). The MDR method also identified *NQO1, MPO,* and *CAT* as the overall "best" three-gene set. However, the average prediction error was high (53.54%) and did not meet the threshold for statistical significance (43.14% for 12 genes). Also, consistency was low (2 of 10) and did not meet the threshold for statistical significance (>8). The ITF method alone also did not find any statistically significant effects.

Additional information can be gained by examining the effect estimates for the significant model involving *NQO1, MPO,* and *CAT* (table 7). In general, the estimates imply protective effects for the variant alleles at these loci. However, the degree of protection depends on the combination of loci with variant alleles—for example, possessing variant alleles at *MPO* and *CAT* but not *NQO1* is more protective than possessing variant alleles at *NQO1* and either *MPO* or *CAT*.

Replicate Sample

Independent data for 445 CHS subjects from black, Asian, Indian, and other racial/ethnic groups were then analyzed in an attempt to replicate the significant findings from the FITF method. This sample included 72 children with physician-diagnosed asthma (i.e., 72 cases) and 373 unaffected controls. The fully saturated model in equation (1), with genes *NQO1, MPO,* and *CAT,* was applied to this sample and yielded a highly significant multi-df test ($P = .0008$; 7 df). This three-gene combination remained statistically significant in this sample after conditioning on the effect of racial/ethnic background ($P = .0033$). The effect estimates from this seven-parameter model were generally similar to the estimates for the Hispanic and non-Hispanic white group (table 8).

Discussion

It is clear from the simulation results that accounting for interactions, when they exist, by use of the ITF or FITF methods significantly increases the number of susceptibility genes that are above the detection limit, compared with results of testing genes by their marginal effects. It was also demonstrated that the cost of using this approach in the presence of main effects but no interactions is small. The CSS screening process increased the power of the method by integrating information that would otherwise have been discarded by the ITF approach, which uses equally weighted tests of all gene combinations. The distributions of ranks of the CSS statistic for the susceptibility gene set imply that this statistic could also be helpful in screening combinations of genes for formal interactions. There are factors other than gene-gene interactions that could cause association among candidate genes not in close proximity on a chromosome—for example, population stratification and dif-

Table 7

Asthma Risk Estimates from Significant FITF Model for Whites and Hispanics

Effect	OR (95% CI)
NQO1	.49 (.32–.72)
MPO	.75 (.49–1.13)
CAT	.88 (.56–1.40)
NQO1 × *MPO*	1.48 (.88–2.49)
NQO1 × CAT	1.39 (.77–2.50)
MPO × *CAT*	.51 (.25–.99)
NQO1 × *MPO* × *CAT*	1.14 (.51–2.51)

NOTE.—Effects were simultaneously estimated using logistic regression (eq. [1]). ORs for *NQO1, MPO,* and *CAT* indicate risk for a single-variant allele, with allelic risk assumed to be log-additive.

ferential survival between genotype groups for reasons other than the phenotype under investigation. Although these factors could decrease power by adding noise to the CSS statistic, they will not bias the LRTs if they are not associated with the phenotype under investigation. In this study, the overall α level was divided equally between the three stages, and the CSS cutoff values were set to 3 and 6; however, in practice, the investigator may choose different weights and CSS cutoff values to suit the focus of the study. For instance, if the focus is on marginal effects and there are a large number of candidate genes in the study, then the investigator may decide to allocate a larger proportion of the α level to the first stage and to increase the CSS cutoff values to control the number of gene sets that require testing in the second and third stages.

In the current state of technology, in which large amounts of data per study unit are easily available to geneticists, it seems increasingly important to create ways of reducing the total number of tests without inflating the number that are falsely rejected. Regardless of the initial number of gene sets that potentially have multilocus effects, a set that is manageable in size can always be determined using the FITF strategy by setting the CSS cutoffs suitably high.

The marginal, ITF, and FITF methods were substantially more powerful than both MDR(e) and MDR(c) when the interactions involved additive, dominant, or recessive genes. Most of these simulated scenarios involved genes with no main effects but some weak marginal effects. If lower- and higher-order effects and allele frequencies had been adjusted such that risk exactly canceled out to result in no marginal effects, it is unclear whether the performance of FITF, relative to that of MDR, would be changed. This comparison is warranted for future research. MDR(e) and MDR(c) performed better when the interactions did not involve additive, dominant, or recessive genes but rather conformed to certain logical rules that result in a set of high-risk multilocus genotypes for genes with little or no marginal effects. The differences between ITF and MDR are partially attributable to modeling assumptions and, in this way, reflect general differences between parametric and nonparametric methods. However, the basic ideas developed in the ITF and FITF methods do not depend on a particular susceptibility pattern.

As is often the case, it is difficult to directly compare two different methods. For example, a fundamental difference between MDR and FITF is the number of possible "significant" gene sets. A single MDR analysis produces exactly one "best" set of genes, with its associated consistency and prediction error. In contrast, in FITF analysis, the number of significant gene sets is a function of the data. Additionally, the significance threshold for MDR must be determined by simulation, whereas the threshold for FITF can be determined using a standard multiple-testing approach, such as FDR or Bonferroni adjustment. It is not clear how these operational issues contributed to our reported differences in power between the two approaches. However, we did demonstrate that both approaches controlled the experimentwise type I error rate under the null distribution.

In an application to real case-control data of asthmatic and nonasthmatic white and Hispanic children from the CHS, the marginal method identified a statistically significant protective effect of *NQO1,* and the FITF method identified a statistically significant three-gene model for *NQO1, MPO,* and *CAT.* These results were replicated in an independent data set that included children with other racial/ethnic backgrounds from the CHS. The enzymes encoded by these genes play an important role in the physiological response to oxidative stress. High levels of ROSs such as superoxide have been found in the airways of asthmatics and correlate with the severity of reactivity of the airways (Andreadis et al. 2003). NQO1

Table 8

Asthma Risk Estimates for Non-Hispanic, Nonwhite Subjects

Effect	OR (95% CI)
NQO1	.42 (.21–.77)
MPO	1.60 (.93–2.75)
CAT	.71 (.21–1.86)
NQO1 × *MPO*	1.29 (.62–2.57)
NQO1 × *CAT*	.76 (.096–3.90)
MPO × *CAT*	.28 (.035–1.45)
NQO1 × *MPO* × *CAT*	2.12 (.26–14.06)

NOTE.—Effects were simultaneously estimated using logistic regression (eq. [1]). ORs for *NQO1, MPO,* and *CAT* indicate risk for a single-variant allele, with allelic risk assumed to be log-additive. Of the 445 subjects, 72 reported doctor-diagnosed asthma.

has the ability to scavenge superoxide, resulting in the production of hydrogen peroxide (Siegel et al. 2004). Hydrogen peroxide is an ROS and, as such, can contribute to inflammatory changes in the asthmatic airway. Our results for *NQO1* are consistent with previous findings of a protective effect in childhood asthma for the *NQO1* Ser allele (David et al. 2003). The peroxidase MPO can cause amplification of the oxidizing potential of hydrogen peroxide, leading to increased oxidative injury (Andreadis et al. 2003). Myeloperoxidase converts hydrogen peroxide to highly reactive hypochlorous acid (Ambrosone et al. 2005), which has been found to cause tissue damage and vascular dysfunction in the lungs of sheep (Turan et al. 2000). Thus, the product of an NQO1 reaction, hydrogen peroxide, is a substrate for MPO. The variant *MPO* allele confers lower transcriptional activation and is presumed to be associated with lower ROS levels (Ambrosone et al. 2005). Catalase plays an important role in the antioxidant defense system by catalyzing hydrogen peroxide into water (Sindhu et al. 2005). In a study of Swedish adults, Forsberg et al. (2001) found catalase levels to be higher in carriers of the variant allele. Thus, the three genes compose a broad and interconnected part of the oxidative stress pathway, and it is conceivable that functional polymorphisms at these loci could influence asthma susceptibility.

In summary, we have presented a testing framework that accounts for possible interactions while preserving power and is parsimonious in the interpretation of effects. For commonly considered types of interactions, this method was more powerful than MDR or marginal tests of candidate genes. When applied to real data, the FITF method was able to detect effects that were undetectable marginally, even after controlling for multiple tests (software used to produce the results in this article is provided free by the authors at the FITF Software Web site). The argument has been presented here that it is both feasible and desirable to account for possible interactions in case-control studies of candidate genes.

Acknowledgments

We acknowledge Dr. John M. Peters, Dr. Yu-Fen Li, and many others who put an enormous amount of work into the Children's Health Study and the collection of the genetic data. This research was funded by National Institutes of Health grant ES10421.

Web Resources

The URLs for data presented herein are as follows:

FITF Software, http://hydra.usc.edu/fitf (for software used to produce the results in this article; provided free by the authors)

Online Mendelian Inheritance in Man (OMIM), http://www.ncbi.nlm.nih.gov/Omim/ (for *GSTM1, GSTT1, NQO1, MPO,* and *CAT*)

References

Ambrosone CB, Ahn J, Singh KK, Rezaishiraz H, Furberg H, Sweeney C, Coles B, Trovato A (2005) Polymorphisms in genes related to oxidative stress (*MPO, MnSOD, CAT*) and survival after treatment for breast cancer. Cancer Res 65:1105–1111

Andreadis AA, Hazen SL, Comhair SA, Erzurum SC (2003) Oxidative and nitrosative events in asthma. Free Radic Biol Med 35:213–225

Aston CE, Ralph DA, Lalo DP, Manjeshwar S, Gramling BA, Defreese DC, West AD, Branam DE, Thompson LF, Craft MA, Mitchell DS, Shimasaki CD, Mulvihill JJ, Jupe ER (2005) Oligogenic combinations associated with breast cancer risk in women under 53 years of age. Hum Genet 116:208–221

Balmain A, Harris CC (2000) Carcinogenesis in mouse and human cells: parallels and paradoxes. Carcinogenesis 21:371–377

Barlassina C, Lanzani C, Manunta P, Bianchi G (2002) Genetics of essential hypertension: from families to genes. J Am Soc Nephrol Suppl 3 13:S155–S164

Bastone L, Reilly M, Rader DJ, Foulkes AS (2004) MDR and PRP: a comparison of methods for high-order genotype-phenotype associations. Hum Hered 58:82–92

Benjamini Y, Hochberg Y (1995) Controlling the false discovery rate: a practical and powerful approach to multiple testing. J R Stat Soc Ser B 57:289–300

Cho YM, Ritchie MD, Moore JH, Park JY, Lee KU, Shin HD, Lee HK, Park KS (2004) Multifactor-dimensionality reduction shows a two-locus interaction associated with type 2 diabetes mellitus. Diabetologia 47:549–554

Coffey CS, Hebert PR, Ritchie MD, Krumholz HM, Gaziano JM, Ridker PM, Brown NJ, Vaughan DE, Moore JH (2004) An application of conditional logistic regression and multifactor dimensionality reduction for detecting gene-gene interactions on risk of myocardial infarction: the importance of model validation. BMC Bioinformatics 5:49

Cook NR, Zee RY, Ridker PM (2004) Tree and spline based association analysis of gene-gene interaction models for ischemic stroke. Stat Med 23:1439–1453

Culverhouse R (2002) A perspective on epistasis: limits of models displaying no main effects. Am J Hum Genet 70:461–471

Culverhouse R, Klein T, Shannon W (2004) Detecting epistatic interactions contributing to quantitative traits. Genet Epidemiol 27:141–152

David GL, Romieu I, Sienra-Monge JJ, Collins WJ, Ramirez-Aguilar M, del Rio-Navarro BE, Reyes-Ruiz NI, Morris RW, Marzec JM, London SJ (2003) Nicotinamide adenine dinucleotide (phosphate) reduced:quinone oxidoreductase and glutathione S-transferase M1 polymorphisms and childhood asthma. Am J Respir Crit Care Med 168:1199–1204

De Miglio MR, Pascale RM, Simile MM, Muroni MR, Virdis P, Kwong KM, Wong LK, Bosinco GM, Pulina FR, Calvisi DF, Frau M, Wood GA, Archer MC, Feo F (2004) Polygenic control of hepatocarcinogenesis in Copenhagen × F344 rats. Int J Cancer 111:9–16

Devlin B, Roeder K, Wasserman L (2003) Analysis of multilocus models of association. Genet Epidemiol 25:36–47

Dong C, Li WD, Li D, Price RA (2005) Interaction between obesity-susceptibility loci in chromosome regions 2p25-p24 and 13q13-q21. Eur J Hum Genet 13:102–108

Forsberg L, Lyrenas L, de Faire U, Morgenstern R (2001) A common functional C-T substitution polymorphism in the promoter region of the human catalase gene influences transcription factor binding, reporter gene transcription and is correlated to blood catalase levels. Free Radic Biol Med 30:500–505

Foulkes AS, De Gruttola V, Hertogs K (2004) Combining genotype groups and recursive partitioning: an application to human immunodeficiency virus type 1 genetics data. Appl Stat 53:311–323

Foulkes AS, Reilly M, Zhou L, Wolfe M, Rader DJ (2005) Mixed modelling to characterize genotype-phenotype associations. Stat Med 24:775–789

Gilliland FD, McConnell R, Peters J, Gong HJ (1999) A theoretical basis for investigating ambient air pollution and children's respiratory health. Environ Health Perspect 107:403–407

Hahn LW, Moore JH (2004) Ideal discrimination of discrete clinical endpoints using multilocus genotypes. In Silico Biol 4:183–194

Hahn LW, Ritchie MD, Moore JH (2003) Multifactor dimensionality reduction software for detecting gene-gene and gene-environment interactions. Bioinformatics 19:376–382

Hoh J, Wille A, Ott J (2001) Trimming, weighting, and grouping SNPs in human case-control association studies. Genome Res 11:2115–2119

Hsueh WC, Cole SA, Shuldiner AR, Beamer BA, Blangero J, Hixson JE, MacCluer JW, Mitchell BD (2001) Interactions between variants in the β_3-adrenergic receptor and peroxisome proliferator–activated receptor-γ2 genes and obesity. Diabetes Care 24:672–677

Kim JH, Sen S, Avery CS, Simpson E, Chandler P, Nishina PM, Churchill GA, Naggert JK (2001) Genetic analysis of a new mouse model for non-insulin-dependent diabetes. Genomics 74:273–286

Kuida S, Beier DR (2000) Genetic localization of interacting modifiers affecting severity in a murine model of polycystic kidney disease. Genome Res 10:49–54

Longmate JA (2001) Complexity and power in case-control association studies. Am J Hum Genet 68:1229–1237

MacNee W (2001) Oxidative stress and lung inflammation in airways disease. Eur J Pharmacol 429:195–207

Mak JC, Leung HC, Ho SP, Law BK, Lam WK, Tsang KW, Ip MS, Chan-Yeung M (2004) Systemic oxidative and antioxidative status in Chinese patients with asthma. J Allergy Clin Immunol 114:260–264

Moore JH (2004) Computational analysis of gene-gene interactions using multifactor dimensionality reduction. Expert Rev Mol Diagn 4:795–803

Moore JH, Hahn LW (2002) A cellular automata approach to detecting interactions among single-nucleotide polymorphisms in complex multifactorial diseases. Pac Symp Biocomput 53–64

Naber CK, Husing J, Wolfhard U, Erbel R, Siffert W (2000) Interaction of the *ACE D* allele and the *GNB3 825T* allele in myocardial infarction. Hypertension 36:986–989

Qin S, Zhao X, Pan Y, Liu J, Feng G, Fu J, Bao J, Zhang Z, He L (2005) An association study of the *N*-methyl-D-aspartate receptor NR1 subunit gene (*GRIN1*) and NR2B subunit gene (*GRIN2B*) in schizophrenia with universal DNA microarray. Eur J Hum Genet 13:807–814

Ritchie MD, Hahn LW, Moore JH (2003*a*) Power of multifactor dimensionality reduction for detecting gene-gene interactions in the presence of genotyping error, missing data, phenocopy, and genetic heterogeneity. Genet Epidemiol 24:150–157

Ritchie MD, Hahn LW, Roodi N, Bailey LR, Dupont WD, Parl FF, Moore JH (2001) Multifactor-dimensionality reduction reveals high-order interactions among estrogen-metabolism genes in sporadic breast cancer. Am J Hum Genet 69:138–147

Ritchie MD, White BC, Parker JS, Hahn LW, Moore JH (2003*b*) Optimization of neural network architecture using genetic programming improves detection and modeling of gene-gene interactions in studies of human diseases. BMC Bioinformatics 4:28

Roldan V, Gonzalez-Conejero R, Marin F, Pineda J, Vicente V, Corral J (2005) Five prothrombotic polymorphisms and the prevalence of premature myocardial infarction. Haematologica 90:421–423

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

Siegel D, Gustafson DL, Dehn DL, Han JY, Boonchoong P, Berliner LJ, Ross D (2004) NAD(P)H:quinone oxidoreductase 1: role as a superoxide scavenger. Mol Pharmacol 65:1238–1247

Sindhu RK, Ehdaie A, Farmand F, Dhaliwal KK, Nguyen T, Zhan CD, Roberts CK, Vaziri ND (2005) Expression of catalase and glutathione peroxidase in renal insufficiency. Biochim Biophys Acta 1743: 86–92

Soares ML, Coelho T, Sousa A, Batalov S, Conceicao I, Sales-Luis ML, Ritchie MD, Williams SM, Nievergelt CM, Schork NJ, Saraiva MJ, Buxbaum JN (2005) Susceptibility and modifier genes in Portuguese transthyretin V30M amyloid polyneuropathy: complexity in a single-gene disease. Hum Mol Genet 14:543–553

Staessen JA, Wang JG, Brand E, Barlassina C, Birkenhager WH, Herrmann SM, Fagard R, Tizzoni L, Bianchi G (2001) Effects of three candidate genes on prevalence and incidence of hypertension in a Caucasian population. J Hypertens 19:1349–1358

Tibshirani R (1996) Regression shrinkage and selection via the Lasso. J R Stat Soc B 58:267–288

Tripodis N, Hart AA, Fijneman RJ, Demant P (2001) Complexity of lung cancer modifiers: mapping of thirty genes and twenty-five interactions in half of the mouse genome. J Natl Cancer Inst 93:1484–1491

Tsai CT, Lai LP, Lin JL, Chiang FT, Hwang JJ, Ritchie MD, Moore JH, Hsu KL, Tseng CD, Liau CS, Tseng YZ (2004) Renin-angiotensin system gene polymorphisms and atrial fibrillation. Circulation 109:1640–1646

Turan NN, Demiryurek AT, Kanzik I (2000) Hypochlorous acid-induced responses in sheep isolated pulmonary artery rings. Pharmacol Res 41:589–596

Ukkola O, Perusse L, Chagnon YC, Despres JP, Bouchard C (2001) Interactions among the glucocorticoid receptor, lipoprotein lipase and adrenergic receptor genes and abdominal fat in the Quebec Family Study. Int J Obes Relat Metab Disord 25:1332–1339

Williams SM, Addy JH, Phillips JA 3rd, Dai M, Kpodonu J, Afful J, Jackson H, Joseph K, Eason F, Murray MM, Epperson P, Aduonum A, Wong LJ, Jose PA, Felder RA (2000) Combinations of variations in multiple genes are associated with hypertension. Hypertension 36:2–6

Williams SM, Ritchie MD, Phillips JA 3rd, Dawson E, Prince M, Dzhura E, Willis A, Semenya A, Summar M, White BC, Addy JH, Kpodonu J, Wong LJ, Felder RA, Jose PA, Moore JH (2004) Multilocus analysis of hypertension: a hierarchical approach. Hum Hered 57:28–38

Yanchina ED, Ivchik TV, Shvarts EI, Kokosov AN, Khodzhayantz NE (2004) Gene-gene interactions between glutathione-s transferase M1 and matrix metalloproteinase 9 in the formation of hereditary predisposition to chronic obstructive pulmonary disease. Bull Exp Biol Med 137:64–66

Yang P, Bamlet WR, Ebbert JO, Taylor WR, de Andrade M (2004) Glutathione pathway genes and lung cancer risk in young and old populations. Carcinogenesis 25:1935–1944