Evidence for Non-additive Influence of Single Nucleotide Polymorphisms within the Apolipoprotein E Gene

S. C. Hamon[1,]*, J. H. Stengard[1,2], A. G. Clark[3], V. Salomaa[2], E. Boerwinkle[4] and C. F. Sing[1]

[1]*Department of Human Genetics, University of Michigan, Ann Arbor, MI, USA*
[2]*Department of Epidemiology and Health Promotion, National Public Health Institute, Helsinki, FINLAND*
[3]*Department of Molecular Biology and Genetics, Cornell University, Ithaca, NY, USA*
[4]*Human Genetics Center, University of Texas Health Science Center, Houston, TX, USA*

Summary

We analyzed 13 single nucleotide polymorphisms (SNPs) within the apolipoprotein E (*APOE*) gene, to identify pairs of SNPs that interact in a non-additive manner to influence genotypic mean levels of the ApoE protein in blood. An overparameterized general linear model of two-SNP genotype means was applied to data from 456 female and 398 male unrelated European Americans from Rochester, MN, USA. We found statistically significant evidence for non-additivity between SNPs within the male sample, but not within the female sample. We observed nine pairs of SNPs with evidence of non-additivity at the $\alpha = 0.05$ level of statistical significance within the male sample, when approximately three were expected by chance. Five of the nine pairs involved three SNPs (560, 624 and 1163) that did not have a statistically significant influence when considered separately in a single-site analysis. Three of the nine pairs involving four SNPs (832, 1998, 3937 and 4951) showed significant evidence for non-additivity in at least one of two other male samples from Jackson, MS, USA and North Karelia, Finland. Although all four of these SNPs had a statistically significant influence in Rochester when considered separately, only SNP 3937 gave a significant result in the other male samples. The four SNPs are located in the promoter, intronic and exonic regions, and 3' to the polyadenylation signal in the *APOE* gene. Our study suggests that analyses that only consider SNPs located in exons and ignore contexts such as those indexed by gender and population, and disregard non-additivity of SNP effects, may inappropriately model the contribution of a gene to the genetic architecture of a trait that has a complex multifactorial etiology.

Introduction

A major impediment in studying the contribution of genetic variation to interindividual variation in quantitative levels of apolipoproteins, and other risk factors for developing coronary heart disease, is that the available statistical models are not realistic representations of the biological complexity of genotype-phenotype relationships (Page *et al*. 2003; Sing *et al*. 2003). Phe-

nomena such as allelic and genotypic heterogeneity, pleiotropy, gene by environment interaction and gene by gene interaction (epistasis) are realities that cannot easily be modelled, or estimated and tested, using population based samples of human data (Clark, 2000). In this study, we investigate the role of pairwise non-additivity of the effects of variable sites in the apolipoprotein E (*APOE)* gene (19q13.2) in predicting mean genotypic levels of the apolipoprotein E (ApoE) protein in the blood. We have applied an overparameterized general linear model of two-single nucleotide polymorphism (SNP) genotype means to data collected from unrelated European Americans from Rochester, MN, USA

*Correspondence to: Dr. C. F. Sing Department of Human Genetics University of Michigan 1241 E. Catherine Street 5928 Buhl Building Ann Arbor, MI 48109-0618 USA Tel: 734-764-5492 Fax: 734-763-5277 Email: csing@umich.edu

in order to estimate and statistically test for pairwise non-additivity.

Various models have been employed to evaluate epistasis. These models have been useful in modelling trait variation in experimental organisms, but have limitations when applied to observational data from human populations. W. Bateson first modelled epistasis as the pattern of segregation of discrete phenotypes that results when variation at one locus masks the phenotypic effects of another locus (Bateson, 1909). Experimental biologists have used this model of epistasis in studies of *Drosophila* and *Saccharomyces cerevisiae* to determine the position of a gene product relative to the position of a second gene product in the same biochemical pathway, by comparing single and double mutant organisms (Avery & Wasserman, 1992; Potter *et al.* 2001; Pouliot *et al.* 2001; Tong *et al.* 2001).

The first statistical model for estimating the contribution of epistasis to quantitative trait variation was presented in 1918 by R. A. Fisher. He modelled epistasis between pairs of loci as the deviations of average phenotypic values of the two-locus genotypes from those expected, by summing the effects due to additivity and dominance of alleles at separate loci. Cockerham (1954) and Kempthorne (1954) expanded on Fisher's model by partitioning the phenotypic variance attributable to epistatic effects determined by two loci into four orthogonally defined genotypic effects that are the products of the effects due to single-locus additivity and dominance of alleles, weighted by the relative allele frequencies at the two separate loci (i.e., additive × additive, dominance × additive, additive × dominance and dominance × dominance). This formulation assumes Hardy-Weinberg equilibrium of the relative genotype frequencies at each locus, and no correlation between loci of the single-locus genotype frequencies.

The Cockerham-Kempthorne model has been applied in experimental studies of a variety of organisms, such as barley, *Drosophila*, soybeans, maize and rice, to estimate the phenotypic variance attributable to epistatic effects of pairs of loci on quantitative traits such as yield (Brim & Cockerham, 1961; Russell & Eberhart, 1970; Stuber & Moll, 1971; Hallauer, 1981; Luo *et al.* 2001; Edwards & Lamkey, 2002). For example, in the progeny of crosses between four nearly isogenic lines of barley to create the nine two-locus genotypic classes, researchers observed that for seven yield component traits an average of 32% of the genotypic variance was attributable to epistasis. For one, four and six of the traits there was significant ($\alpha = 0.05$) evidence for additive × dominance, dominance × dominance and additive × additive epistatic effects, respectively (Fasoulas & Allard, 1962). In another example, Clark & Wang (1997) constructed eight crosses between P-element bearing lines of *Drosophila*, and then estimated the phenotypic variance attributable to single-locus additive and dominance effects and two-locus epistatic effects in the F_2 generation of each cross for 16 metabolic traits. They partitioned the phenotypic variance attributable to epistatic effects into the four orthogonal components defined by Cockerham (1954). The results showed statistically significant evidence for phenotypic variance attributable to additive × dominance and/or dominance × dominance epistatic effects for 27% of the 128 trait-locus tests. When the genotype frequencies at the loci involved are correlated, an alternative statistical model of epistasis suggested by Cheverud & Routman (1995) can be used. This model includes only the phenotypic averages of the two-locus genotypes and does not assume a panmictic population. With this model it is often possible to detect epistasis defined by comparing genotypic values that cannot be detected using the Cockerham method of modeling epistasis, which is a function of comparing both genotypes values and relative allele frequencies. One application, to an F_2 sample of mice, resulted in 9% of the possible pairs of 76 marker loci showing statistically significant evidence of additive × additive, additive × dominance, dominance × additive or dominance × dominance epistatic components of genotypic values for the trait body weight (Cheverud & Routman, 1995; Routman & Cheverud, 1997).

Using variance component linkage models, the component of phenotypic variance attributable to the additive × additive epistatic effects of unmeasured loci was shown to be statistically significant for human quantitative traits such as proneness to anxiety and plasma concentrations of the angiotensin-converting enzyme (Cloninger *et al.* 1998; Blangero *et al.* 2000; Zhu *et al.* 2001). In most observational studies of humans, non-independence of single-locus genotype frequencies or missing two-locus genotypic classes make it impossible to partition the phenotypic variance attributable to

epistatic effects into the four orthogonal components defined by Cockerham (1954) and Kempthorne (1954). Furthermore, missing genotypic classes will limit the number of epistatic components of genotypic values that can be estimated using the method suggested by Cheverud and Routman (1995). Given these considerations, we present as an alternative application of an overparameterized model (Searle, 1971) to test the null hypothesis that the separate influences of variable DNA sites within the *APOE* gene combine additively to influence levels of ApoE protein in Rochester. To determine which statistically significant estimates of non-additivity between pairs of SNPs in the Rochester sample are unlikely to be Type I errors, we evaluated whether such findings could be replicated in two other samples from Jackson, MS, USA and North Karelia, Finland.

Methods

Samples

This study considers 13 SNPs within the *APOE* gene (Figure 1) identified by completely resequencing 24 unrelated European Americans from Rochester for a 5.5 kb region, including 1059 bp of 5' flanking sequence, the entire coding sequence with four exons, and the intervening three introns of the *APOE* gene spanning 3586 bp, as well as 846 bp 3' to the polyadenylation signal. Seven additional SNPs that were identified by resequencing 48 individuals (24 from Jackson, MS, USA and 24 from North Karelia, Finland) were monomorphic in the sample from

Rochester (Nickerson *et al.* 2000). There may be additional SNPs within the *APOE* gene that were not identified by resequencing 144 chromosomes, but because of the rarity of allelic variations at these SNPs it would not be possible to estimate non-additivity in the available samples. Three of the SNPs genotyped Rochester are located 5' to the first exon, five are located within the first two introns, three are located in the third and fourth exons and two are located 3' of the polyadenylation signal. Genotyping in the larger epidemiological samples used in the analyses presented here was carried out using the OLA genotyping method (Fullerton *et al.* 2000; Nickerson *et al.* 2000). Additional information on these SNPs can be found at http://droog.gs.washington.edu/mdecode/data/apoe/.

A sample of 456 unrelated female and 398 unrelated male European Americans (ages 34-64) were selected from 281 multigeneration pedigrees ascertained without regard to health through elementary school children in Rochester (Turner *et al.* 1989). A detailed description of the African American Jackson sample (219 unrelated males ages 32-86) and the European North Karelia sample (124 unrelated males ages 45-65), which were used to identify which of the significant findings in the Rochester study replicated in a second sample, is given in Stengård *et al.* (2002). Levels of ApoE protein in the blood for the Jackson and Rochester samples were measured at the Mayo Clinic, Rochester, using published methods (National Institutes of Health, 1974; Barr *et al.* 1981; Kaprio *et al.* 1991). ApoE levels for the North Karelia sample were measured at the
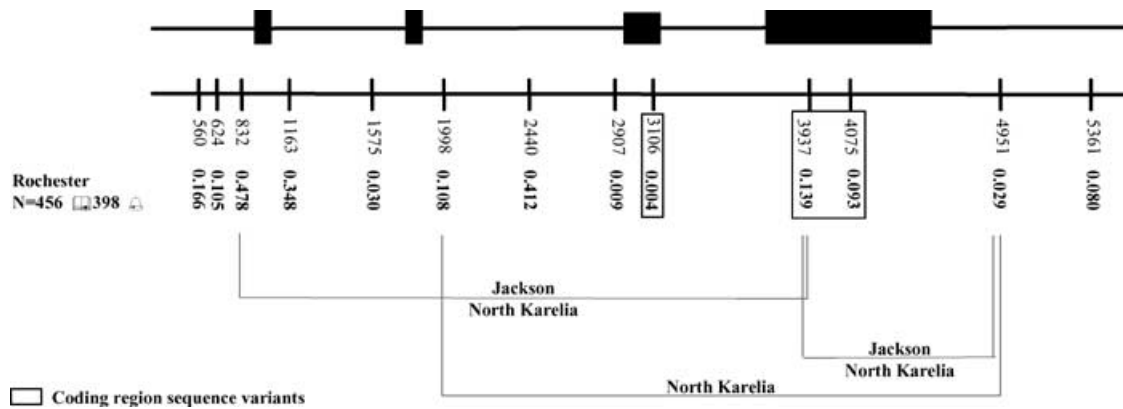

Figure 1 SNP locations and relative frequencies of the least common SNP allele in female and male samples from Rochester, and the location within the *APOE* gene of the three cases of statistically significant evidence for non-additivity that replicate across the male samples.

Department of Biochemistry, National Public Health Institute, Helsinki, using standard enzymatic assays (Salomaa *et al.* 1994; Schiele *et al.* 2000)

Statistical Methods

The distribution of ApoE levels was significantly positively skewed in both genders. The natural log (ln) transformation of ApoE reduced skewness to non-significant values (less than 0.05) in each sample. This transformed variable (lnApoE) was used in this study to accommodate statistical tests that assume normality. We carried out separate analyses of females and males because of the well-documented gender-specific differences in the natural history of the risk of developing cardiovascular disease (Barrett-Connor, 1997; Hayward *et al.* 2000; Matthews *et al.* 1989; Thomas & Braus, 1998; Xhignesse *et al.* 1991). To document such gender specificity in our study, we used Fisher's F-ratio to test whether there was a statistically significant difference in the phenotypic variance of age, height, weight, BMI or lnApoE between females and males (Sokal & Rohlf, 1995). Student's t-test was used to test the statistical significance of the difference between gender means when the F-ratio was not significant, and Satterthwaite's modification of the t-test (Sokal & Rohlf, 1995) was used when the F-ratio for inequality of the variances was significant.

Neither the approaches of Cockerham (1954) and Kempthorne (1954) nor Cheverud & Routman (1995) can be applied in this study because there are missing genotypic classes and the genotype frequencies of the two loci are correlated. Instead we used an overparameterized statistical model as an alternative to estimate and compare average phenotypic values of two-SNP genotypes, to determine if there is evidence that single-SNP genotypes interact in a non-additive manner to influence ApoE levels (Searle, 1971). This model can be represented as:

$$\gamma_{ijk} = \mu + \alpha_i + \beta_j + \phi_{ij} + \varepsilon_{ijk}$$

where γ_{ijk} is the phenotypic value of the kth individual with the ij^{th} two-SNP genotype ($i, j = 1, 2, 3$), μ is the population mean, α_i is the influence associated with the ith genotype of the first SNP, β_j is the influence associated with the jth genotype of the second SNP, ϕ_{ij} is the measure of the non-additive influ-

ence associated with the ij^{th} two-SNP genotype combination not attributable to the separate influences of each of the two SNPs, and ε_{ijk} represents the residual influence of unmeasured effects and measurement error. This model has the advantage that it does not require either α_i or β_j to be non-zero in order for ϕ_{ij} to be non-zero. The expectations of linear functions of the estimates of two-SNP genotype cell means of the form $(\hat{\mu}_{ij} - \hat{\mu}_{ij'}) - (\hat{\mu}_{i'j} - \hat{\mu}_{i'j'})$ are linear functions of the corresponding parameters of non-additivity, $(\phi_{ij} - \phi_{ij'}) - (\phi_{i'j} - \phi_{i'j'})$, where i and iN (i, iN = 1, 2, 3) are two possible genotypes for the first SNP, and j and jN (j, jN = 1, 2, 3) are two possible genotypes for the second SNP. When all nine two-SNP genotypes are observed there are four independent comparisons, each involving four two-SNP genotype means. They can be written as functions of the nine possible genotypes (ij, iNj, ijN and iNjN = 11, 12, 13, 21, 22, 23, 31, 32 and 33): $(\hat{\mu}_{11} - \hat{\mu}_{12}) - (\hat{\mu}_{21} - \hat{\mu}_{22})$, $(\hat{\mu}_{12} - \hat{\mu}_{13}) - (\hat{\mu}_{22} - \hat{\mu}_{23})$, $(\hat{\mu}_{21} - \hat{\mu}_{22}) - (\hat{\mu}_{31} - \hat{\mu}_{32})$ and $(\hat{\mu}_{22} - \hat{\mu}_{23}) - (\hat{\mu}_{32} - \hat{\mu}_{33})$. Figure 2 gives these four independent comparisons, and examples of four possible non-additive outcomes when three genotype levels (AA, AG and GG), at the first SNP and three genotype levels (CC, CT and TT), at the second SNP are considered. When fewer than nine two-SNP genotypes are observed, which happens to be the case in our study for most pairs of SNPs, all four independent comparisons are not estimable.

When there are missing two-locus genotype classes it is not possible to obtain tests of the single-locus genotype differences that are independent of the non-additive interactions. However, Searle (1971) presents an analytical strategy for testing the statistical significance of non-additivity for each pair of SNPs that is independent of the influence of single-locus genotype differences. The test compares the sum of squares associated with non-additive influences captured by the comparisons of two-SNP genotypes with the error sum of squares using an F–test (Searle, 1971). For a particular ij^{th} pair of SNPs the influence of non-additivity is computed as the difference between the model sum of squares over all observations obtained, fitting an overparameterized model that includes the non-additivity, $\gamma_{ijk} = \alpha_i + \beta_j + \phi_{ij} + \varepsilon_{ijk}$, and the model sum of squares over all observations, fitting a model that does
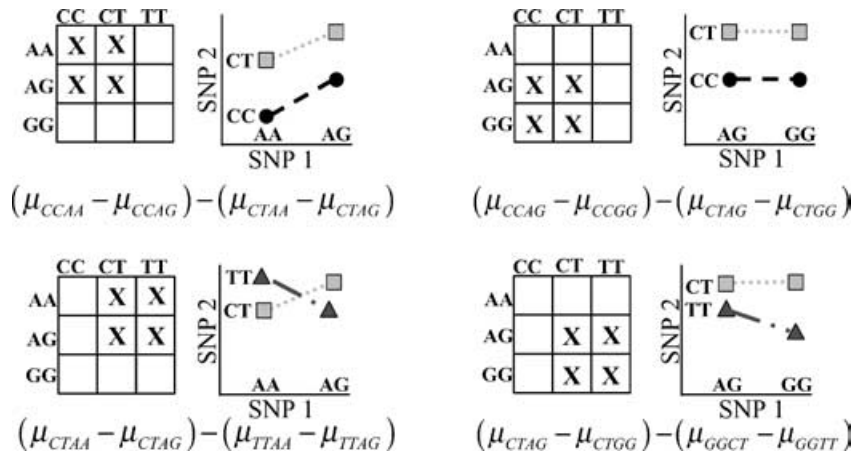
Figure 2 The four possible comparisons for two-SNP genotypes, corresponding equations, and example mean levels that illustrate types of non-additivity.

not include the non-additive influence of a two-SNP genotype, $\gamma_{ijk} = \alpha_i + \beta_j + \varepsilon_{ijk}$. The number of degrees-of-freedom associated with this difference equals the number of independent comparisons of non-additivity (presented in Figure 2) that are estimable from the data. The error sum of squares for testing for significant non-additivity is:

$$\sum_i \sum_j \sum_k \gamma_{ijk}^2 - \sum_i \sum_j n_{ij} \hat{\mu}_{ij}^2,$$

and the degrees of freedom associated with this quantity is:

$$\sum_i \sum_j n_{ij} - (a - 1) - (b - 1)$$

− number of independent comparisons - 1,

where a is the number of genotypes that are observed for the first SNP, b is the number of genotypes that are observed for the second SNP, and n_{ij} is the total number of individuals having ij^{th} genotype.

Resampling methods were used to create a test of the null hypothesis that the observed number of significant F-tests to detect pairwise non-additivity within each gender strata is equal to the expected number of significant F-tests. Within each gender strata the phenotypes were permuted among individuals keeping the two-SNP genotypic structure fixed (Good, 2000). The number of F-tests to detect pairwise non-additivity that were significant at the 5% level was determined for each of the 1000 permutations to create the null distribution (Churchill & Doerge, 1994). The observed number of F-tests significant at the 5% level of probability in the original data was then compared to the null distribution

of the number of F-tests significant at the 5% level, to evaluate the statistical significance of the observed number of significant F-tests.

With the overparameterized model described above it is possible to construct comparisons of two-SNP genotype means that are independent of genotypic frequencies. We take advantage of this independence in our study, to ask whether a particular statistically significant comparison of two-SNP genotypes, observed in the Rochester sample, replicates in independently collected samples from two different populations in which both SNPs segregate but relative genotype frequencies may differ.

Results

Summary of samples

A description of the anthropometric characteristics and ApoE concentrations in the female and male samples are summarized in Table 1. The average age and level of lnApoE did not differ significantly between females and males. On average, males are significantly taller and heavier and have a significantly greater BMI than females. Interindividual variability of age and BMI was significantly greater in females. The locations and the relative frequencies of the least frequent allele for each of the 13 SNPs in the combined female and male sample from Rochester are shown in Figure 1. Relative allele frequencies ranged from 0.004 to 0.412. For eight of the 13 SNPs the relative allele frequency of the least frequent allele was greater than 0.08. A summary of

	Females (N = 456)		Males (N = 398)		P-values t-test F-test	
Variable	Mean	Variance	Mean	Variance	Mean	Variance
Age	48.46	93.09	48.14	74.84	0.61	0.03
Height (cm)	163.78	33.02	177.44	38.33	<0.0001	0.12
Weight (kg)	69.2	189.12	86.84	188.65	<0.0001	0.98
BMI (kg/m2)	25.83	26.6	27.59	17.84	<0.0001	<0.0001
lnApoE (mg/dl)	1.62	0.1	1.63	0.11	0.47	0.43

Table 1 Distributions of anthropometric characteristics and measures of lipid metabolism in the Rochester samples.

D' and r^2 for the thirteen sites that were polymorphic in Rochester males and females is presented in Table 2. The D' and r^2 values ranged from 0.02-1 and 0.00-0.704 respectively, and 9, 1 and 42 of the 78 pairs of SNPs showed significant evidence for linkage disequilibrium at the 0.05, 0.01 and 0.001 levels of significance, respectively.

Characteristics of pairwise SNP combinations

Table 3 summarizes, separately for the Rochester female and male samples, the number of sites that are segregating, the number of possible pairwise tests of non-additivity if all two-SNP genotypes were observed, the number of these pairwise tests that are found to have 0-4 independent comparisons of two-SNP genotypes in the samples studied, the number of tests to detect non-additivity significant at the 5% level of probability and the experiment-wise probability, i.e., the probability of observing that many significant tests at the 5% level by chance alone. A statistically significant excess in the number of pairwise significant tests was observed for the male sample. From this point forward, we focus on the significant pairwise tests observed in this sample.

A summary of the nine pairs of SNPs giving significant evidence of non-additivity at the $\alpha = 0.05$ level is presented in Table 4. Five of these nine pairs involved one comparison of two-SNP genotypes, three involved two independent comparisons and one involved three independent comparisons. Five of the nine pairs involved three SNPs (560, 624 and 1163) that did not have a statistically significant influence on levels of the ApoE protein when considered separately in a single-site analysis (p > 0.05). Four of these five tests did not have a replicate comparison in either the Jackson or North Karelia male samples. The fifth pair (832-1163)

had one of the two comparisons replicated in Jackson and both replicated in North Karelia. Neither test was significant at the $\alpha = 0.05$ level of probability. Three of the remaining four pairs (832-3937, 1998-4951 and 3937-4951) had comparisons that when replicated in at least one of the two additional samples were statistically significant.

Discussion

We report here a systematic study of the role of non-additivity in determining the influences of variable sites within the *APOE* gene in humans. We established that there was significant non-additivity within the *APOE* gene in a sample of Rochester males, and use the Jackson and/or North Karelia male samples to confirm our results as discussed below. Our findings illustrate the importance of heterogeneity of relative genotype frequencies among samples from different populations in studying non-additivity, the fact that all areas within a gene may be involved in non-additive interactions, and that deviations due to non-additivity can make an important contribution to deviations of individual phenotypic values from the population mean.

Population of inference and estimation of non-additivity

In order to determine which pairs of SNPs in Rochester males showed significant evidence of non-additivity unlikely to be due to Type I errors, we asked whether such pairs also exhibited significant non-additivity in male samples from Jackson and/or North Karelia. Differences in genetic structure among samples determines whether replicated tests of non-additivity are possible. Only 63 of the 78 pairs of SNPs defined by the 13 SNPs in the Rochester male sample had one or more

Table 2 Measure of linkage disequilibrium r^2 (lower left hand corner) and D' (upper right hand corner) for 13 SNPs genotyped in Rochester males and females.

SNP	0560	0624	0832	1163	1575	1998	2440	2907	3106	3937	4075	4951	5361
0560		0.622	0.240	0.339	0.675	0.876	0.649	0.630	1	0.878	0.370	0.767	0.873
0624	0.010***		0.067	0.030	0.919	0.198	0.790	1	0.020	0.226	0.427	0.816	0.531
0832	0.012***	0.001		0.984	0.894	0.929	0.933	0.758	0.758	0.552	0.809	0.919	0.779
1163	0.044***	0.000	0.565***		0.828	0.985	0.984	0.615	0.644	0.988	0.983	1	0.959
1575	0.003*	0.236***	0.028***	0.041***		0.660	0.911	0.032	0.860	0.736	0.943	1	0.771
1998	0.019***	0.001	0.115***	0.065***	0.002		0.949	1	1	0.969	0.970	0.952	0.804
2440	0.058***	0.052***	0.560***	0.361***	0.019***	0.077***		0.541	1	0.970	1	1	0.952
2907	0.001	0.001	0.005**	0.006***	0.000	0.001	0.002			1	1	1	1
3106	0.001	0.000	0.003*	0.001	0.000	0.031***	0.003*	0.000		0.855		0.977	1
3937	0.025***	0.001	0.052***	0.085***	0.003*	0.704***	0.103***	0.001	0.023***			0.789	0.791
4075	0.070***	0.159***	0.062***	0.055***	0.003*	0.011***	0.066***	0.001	0.001	0.016***			1
4951	0.004*	0.002*	0.024***	0.016***	0.001	0.004*	0.018***	0.000	0.000	0.170***	0.002		0.502
5361	0.013***	0.003*	0.049***	0.042***	0.002	0.007***	0.113***	0.001	0.000	0.008***	0.009***	0.001	

* at $\alpha = 0.05$ ** at $\alpha = 0.01$ and *** at $\alpha = 0.001$ indicates a statistically significant evidence for linkage disequilibrium.

independent comparisons of two-SNP genotypes necessary to estimate non-additivity. An even smaller number of two-SNP combinations were also observed in the Jackson and/or North Karelia male samples, making confirmation of significant findings in a sample from a second population problematic. Of the 63 pairwise tests of SNPs observed in the Rochester sample, only 27 (42.9%), involving 41 comparisons, had at least one comparison that was observed in at least one of the other two samples. These 27 pairs did not include SNPs 2907, 3106 and 5361. Twenty-two of the 41 comparisons were observed in Jackson and North Karelia, 10 were observed only in Jackson, and 9 only in North Karelia.

Five of the nine pairs that showed significant evidence of non-additivity in the Rochester male sample had at least one independent comparison in the Jackson and/or North Karelia sample. For the 832–1163 pair of SNPs, where two comparisons were observed for Rochester, only one of the comparisons was observed in the Jackson sample while both were observed in the North Karelia sample. For the 832–3937 pair, where three of the comparisons were observed in Rochester, only two of the comparisons were observed in Jackson while all three were observed in North Karelia. The 832–4075 pair had two comparisons observed in Rochester, but only one of these comparisons was observed in Jackson and North Karelia. The one comparison observed for the 1998–4951 pair in the Rochester sample was only observed in the North Karelia sample. Finally, for the 3937–4951 pair, the same comparison was observed in the Rochester, Jackson and North Karelia samples. Our study makes clear that differences in the genetic structure of the samples considered can dramatically reduce the utility of replication to sort out those significant pairwise estimates of non-additivity that are Type I errors from those that are not. Increasing the sample size can often improve the likelihood that comparisons across population samples can be made, but real differences in the relative allele frequencies and linkage disequilibrium imply that the problem of cross-population replication is not always solvable by having a larger sample size.

Replication and Non-additivity

Differences in relative genotype frequencies among samples allowed us only to consider replication of pair-wise non-additivity in the Jackson and North Karelia samples

Table 3 Characteristics of pairwise SNP combinations and corresponding experimentwise p-values for the Rochester sample.

Gender	Number of Sites that Segregate	Possible Tests of Pairwise Additivity	Tests of Pairwise Additivity with 0-4 Comparisons					Tests of Pairwise Additivity with > 0 Comparisons	Significant Tests of Pairwise Additivity (α = 0.05)	Experimentwise P-value
			0	1	2	3	4			
Females (N = 456)	13	78	11	49	10	6	2	67	1 (1.5%)	0.931
Males (N = 398)	13	78	15	40	14	6	3	63	9 (14.3%)	0.039

First SNP	Second SNP	Number of Comparisons	P-value	Number of Comparisons that were Observed in	
				Jackson	North Karelia
560	2907**	1	0.0150	0	0
624	1998***	1	0.0171	0	0
624	3937***	2	0.0002	0	0
624	4951***	1	0.0397	0	0
832**	1163	2	0.0017	1	2
832**	3937***	3	0.0009	2†	3
832**	4075***	2	0.0366	1	1
1998***	4951***	1	0.0324	0	1†
3937***	4951***	1	0.0003	1†	1†

Table 4 P-values for each of the nine pairs of SNPs that showed significant evidence (p≤0.05) for non-additivity in the sample of Rochester males.

** at α = 0.01 and *** at α = 0.001 indicates a statistically significant influence on levels of the ApoE protein when considered separately in a single-site analysis.
†Indicates evidence for significant non-additivity at α = 0.05 level of probability in a male sample from a second population.

for five of the nine pairs (832-1163, 832-3937, 832-4075, 1998-4951 and 3937-4951) that were significant in the Rochester sample. Three of these five pairs (832-3937, 1998-4951 and 3937-4951) were observed to exhibit statistically significant non-additivity in either the Jackson and/or North Karelia male samples (Table 4).

Significant evidence for a non-additive influence of SNPs 1998 and 4951 (Figure 3A) in the Rochester male sample (p = 0.03) was replicated in the North Karelia male sample (p = 0.004). This comparison could not be estimated in the Jackson male sample because of missing genotype classes. The replicated comparison of non-additivity between SNPs 1998 and 4951 is summarized in Figure 3A as a greater difference between the average lnApoE levels of the AG and GG genotypes for SNP 1998, in the presence of the AC genotype at SNP 4951, compared to the difference between these genotypes in the presence of the AA genotype at SNP 4951. This comparison can also be summarized in terms of a greater difference between the average lnApoE levels of

the AA and AC genotypes for SNP 4951, in the presence of the AG genotype at SNP 1998, compared to the difference in the presence of the GG genotype at SNP 1998.

Significant evidence for a non-additive influence of the SNPs 3937 and 4951 (Figure 3B) in the Rochester male sample (p = 0.0003) was replicated in the Jackson male sample (p = 0.04) and the North Karelia male sample (p = 0.03). In the Rochester and North Karelia samples there was a greater difference between the average lnApoE levels of the CC and CT genotypes for SNP 3937, in the presence of the AC genotype at SNP 4951, than when in the presence of the AA genotype at SNP 4951. However, in the Jackson sample there was a greater difference between the average lnApoE levels of the CC and CT genotypes for SNP 3937, in the presence of the AA genotype at SNP 4951, than when in the presence of the AC genotype at SNP 4951. These different influences of the same comparison suggest a higher order interaction with unmeasured
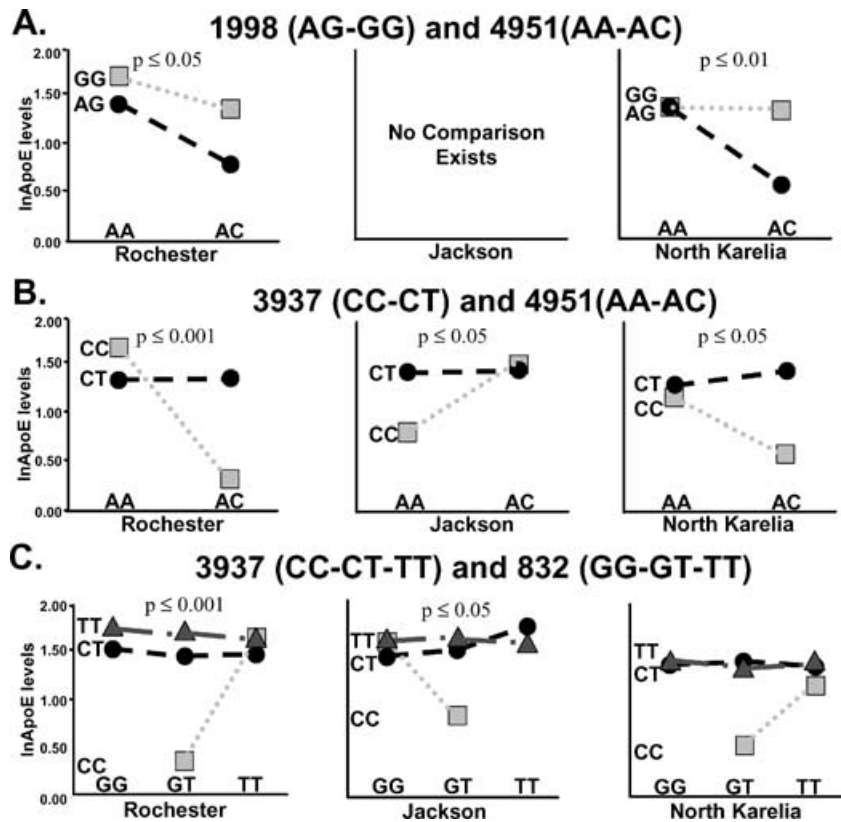
Figure 3 Levels of lnApoE plotted for genotypes that make up the comparisons for the three pairs (3A: 1998-4951, 3B: 3937-4951, 3C: 3937-832) of SNPs that showed statistically significant evidence for non-additivity in Rochester males, as well as Jackson and/or North Karelia males.

genetic and/or environmental factors that vary among populations.

Significant evidence for a non-additive influence of the SNPs 832 and 3937 (Figure 3C) in the Rochester male sample (p = 0.001) is replicated in the Jackson male sample (p = 0.03), and there is suggestive evidence of non-additivity in the North Karelia male sample (p = 0.13). The comparison that contributed to the statistical significance of non-additivity between SNPs 832 and 3937 in the Rochester sample (p = 0.0004) was observed only in the North Karelia (p = 0.036) sample (Figure 4A). There was a greater difference between the average phenotypic levels of the CT and CC genotypes for SNP 3937, in the presence of the GT genotype at SNP 832, compared to the difference in the presence of the TT genotype at SNP 832. The other two replicated comparisons (Figure 4B-C) did not contribute to the statistical significance of non-additivity in any of the three samples (p > 0.05). A fourth comparison, that

contributed to significant non-additivity between the SNPs in the Jackson sample (p = 0.017), was not observed in either the North Karelia or Rochester samples (Figure 4D). In this case, there was a greater difference between the average lnApoE levels of the CT and CC genotypes for SNP 3937, in the presence of the GG genotype at SNP 832, compared to the difference in the presence of the GT genotype at SNP 832. Overall, these results serve to illustrate that the replication of additivity and the replication of non-additivity can only be expected for a very small fraction of intragenic pairs of SNPs, because of the large role heterogeneity in genetic structure among samples plays in determining the number of replicated comparisons. Context dependency, and the embedding of these pairwise tests in a higher-order interacting system of unmeasured genetic and environmental effects, are possible explanations for the non-replication of pairwise tests of non-additivity. Observation of a significant interaction will
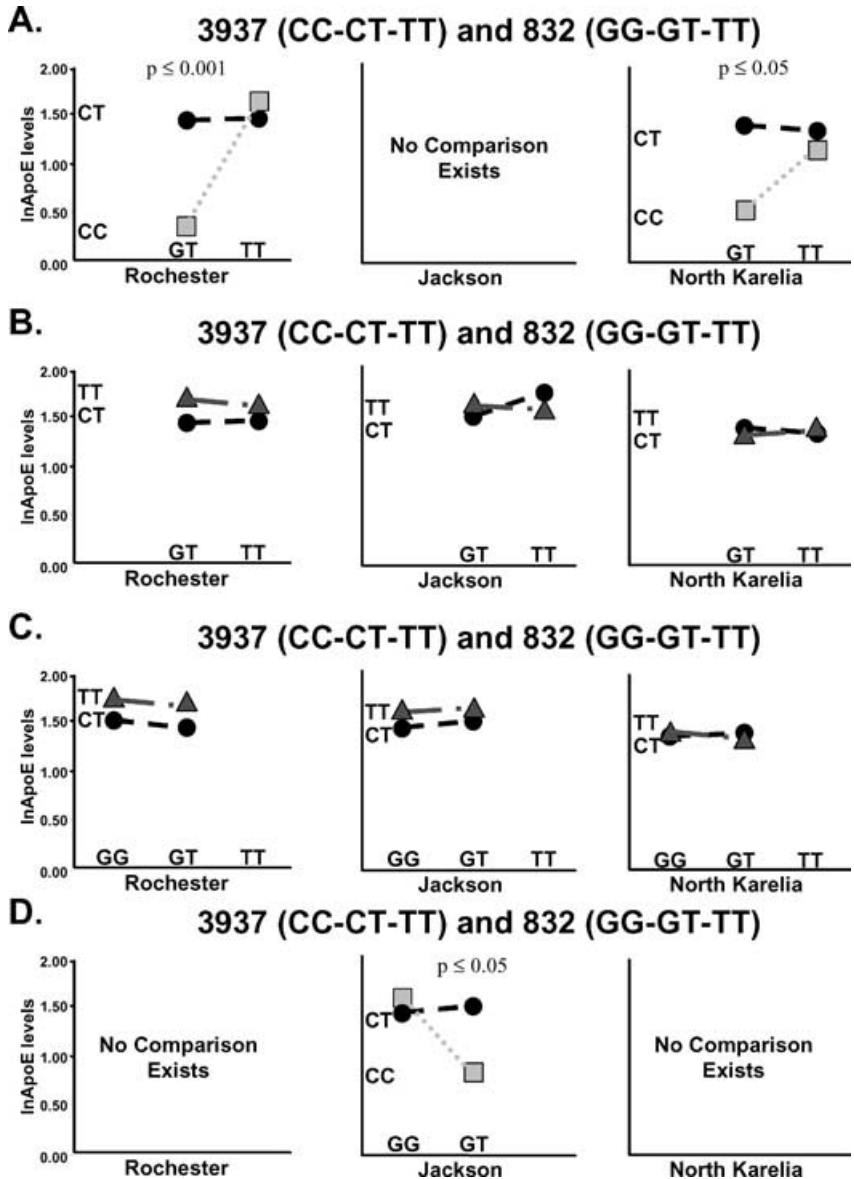
Figure 4 Levels of lnApoE plotted for genotypes that make up the comparisons for the pair 3937-832 (Figure 3C), plotted so that each comparison can be viewed individually.

seldom imply that SNPs directly cause an interaction, because these pairs of variables are embedded in a higher dimension system of interacting agents (Clark, 2000). Furthermore, given the lability of environmental effects in determining CVD risks, the expectation that these interactions will be context dependent borders on certainty.

Intragenic Non–additivity

Many researchers estimate interaction only between loci that have statistically significant marginal effects

(Fedorowicz *et al.* 1998; Blangero *et al.* 2000; Mackay, 2001). Recent theoretical work by Culverhouse *et al.* (2002) shows that it is possible to have a statistically significant component of genetic variance attributable to epistatic effects between variable loci that have no statistically significant marginal effects. Nelson *et al.* (2001) reported an example of joint effects of variable sites in the gene cluster *APOA1/C3/A4* (11q23-q24) and low density lipoprotein receptor (*LDLR)* (19p13.3) gene in determining interindividual variability in triglyceride levels, when genotypic variation in each variable site,

considered separately, was not statistically significant. In our study all four SNPs involved in the three tests to detect non-additivity, which replicated in at least one other sample of males, had a statistically significant influence on levels of the ApoE protein when considered separately in a single-site analysis in the Rochester sample of males. However, only SNP 3937 had a statistically significant influence on lnApoE levels when considered separately in a single-site analysis in the Jackson and North Karelia samples. Ignoring SNPs at positions 832, 1998 and 4951 because they are not making a separate contribution to trait variation would exclude important predictors of lnApoE levels in the Jackson and North Karelia samples. This result is some of the first evidence in humans that supports the findings in yeast that it is the interaction between genetic variations, not the genetic variations themselves, which are causes of phenotypic variability (Hartwell, 2004).

Other studies in humans have estimated the joint effects of variable sites located in different unlinked genes. Templeton (2000) reviewed published examples of human studies to evaluate the role of gene-gene interaction effects in determining variation in traits having a multifactorial etiology. Nelson *et al.* (2001) showed that effects of variable sites in the *APOA1/C3/A4* gene cluster combine non-additively with effects of variable sites in the *LDLR* gene to significantly influence interindividual variability in blood triglyceride levels. Small *et al.* (2002) found that a combination of two variants, one in the alpha-2C adrenergic receptor gene (*ADRA2C*Del322-325, 4p16) and the other in the beta-1 adrenergic receptor gene (*ADRB1*Arg389, 10q24-q26), interact to increase the risk of developing heart failure. Our study clearly establishes that SNPs within a gene can also interact non-additively to influence levels of a quantitative trait.

Context Dependency

Our study provides further documentation that the impact of variations in the *APOE* gene on measures of lipid metabolism are gender specific. This result is not unexpected because most genetic and environmental agents only have an influence in the presence of a particular environment or genetic background (Holdrege, 1996; Lewontin, 2000). Many studies have established that age,

gender, smoking and alcohol consumption modify the influence of variation in the *APOE* gene on interindividual variation in measures of lipid metabolism (Zerba *et al.* 1996; Jarvik *et al.* 1997; Davignon *et al.* 1999; Lussier-Cacan *et al.* 2002). The age specific distribution of plasma ApoE, and the association of plasma lipid and apolipoprotein traits with variation in the *APOE* gene have all been shown to be dependent upon gender (Reilly *et al.* 1991; 1992; 1994; Cobb *et al.* 1992; Jarvik *et al.* 1997). Likewise, our study provides further evidence that gender is an important index of unmeasured interacting factors that influence the quantitative levels of ApoE.

Location and Function of Interacting Variable Sites

Our analyses imply that the functional effects of gene variation may involve more than one region of a gene. The locations within the *APOE* gene of the pairs of SNPs that gave replicated evidence of non-additive influence of SNPs in the Rochester male sample and at least one other sample of males are shown in Figure 1. These three pairwise tests to detect non-additivity involve four SNPs located in four separate regions of the gene: in the promoter, intronic, exonic and 3' to the polyadenylation signal. The locations of SNPs 832, 1998, 3937 and 4951 serve as further evidence of the importance of considering non-exonic sites when measuring functional gene variation (Stengård *et al.* 2002). Laboratory studies have shown transcriptional regulation of the *APOE* gene is influenced by multiple cis-acting regulatory elements within the promoter region, marked by SNP 832. These include three upstream regulatory elements and two GC boxes (Taylor *et al.* 1987; Paik *et al.* 1988; Smith *et al.* 1988; Chang *et al.* 1990; Jo *et al.* 1995). Lower levels of the ApoE protein are associated with the T allele compared to the G allele, an effect that may be attributable to differential binding of nuclear proteins (Artiga *et al.* 1998). The SNP at site 1998 is located in the second intronic region of the gene, and SNP 4951 is located 3' to the polyadenylation signal. Our group is one of the first to study the impact of these two SNPs on interindividual variation in ApoE, and other measures of lipid metabolism in the population at large (Stengård *et al.* 2002). Possible roles for these

types of regions include regulation of translation initiation or translation efficiency (Chen *et al.* 2002). The SNP at site 3937 in combination with the SNP at site 4075 code for changes in residues 112 and 158 of the 299 amino acid ApoE protein. Lower circulating levels of the ApoE protein are associated with the C allele at SNP 3937 compared to the T allele (Davignon *et al.* 1999).

Our results are consistent with the expectation that the promoter variant at site 832 and the exonic variant at site 3937 may combine to regulate *APOE* levels in males. Interindividual variability in circulating levels of the ApoE protein has been associated with each of these variants (Artiga *et al.* 1998; Davignon *et al.* 1999). Within our own study, differences between those individuals who are heterozygous at SNP 832 versus those who are homozygous for either the C or T allele at SNP 832 only occurs in the presence of those individuals who are homozygous for the C allele at SNP 3937. Specifically, the GT 832 heterozygotes have lower mean ApoE levels than either of the SNP 832 homozygotes, in the presence of the CC genotype at SNP 3937. Statistical methods applied to population based data cannot distinguish which SNPs within a gene are responsible for the observed genotype-phenotype associations. Therefore, we are forced to rely on laboratory and clinical studies to determine the biological significance of these four sites. However, experimental and clinical studies may fall short because they cannot replicate the genetic background and environmental histories responsible for the observed genotype-phenotype relationships.

Contribution of Non-additivity to Phenotype

Non-additivity may be the largest contributor to an individual's phenotype (Culverhouse *et al.* 2002). We estimated the deviation from the grand mean due to the additivity and non-additivity for each two-locus genotype for the three pairs of SNPs (832-3937, 1998-4951 and 3937-4951) that exhibited statistically significant non-additivity in the Rochester male sample, and at least one of the other two male samples. The distribution of genotypic deviations due to additivity for these four SNPs (832, 1998, 3937 and 4951) in the three male samples ranged from 2.72 to 4.31 ApoE, with a mean contribution of 3.09 mg/dl. The combined distribution of deviations due to additivity for these three pairs of SNPs ranged from 2.74 to 9.97 ApoE, with a mean combined contribution of 4.05 mg/dl. The distribution of genotypic deviations due to non-additivity among genotypes defined by these three pairs of SNPs ranged from 2.72 to 17.99 ApoE, with a mean contribution of 3.39 mg/dl. These results suggest that ignoring the contribution of a SNP to non-additive interactions may underestimate the value of a SNP for predicting levels of ApoE protein.

Linkage Disequilibrium and Non-additivity

The average measures of non-additivity did not differ significantly between for those pairs of SNPs that showed significant evidence for non-additivity ($D' = 0.688$, $r^2 = 0.095$) and those that did not ($D' = 0.769$ $r^2 = 0.049$) (D' p $= 0.2575$, r^2p $= 0.2824$). Linkage disequilibrium may also be used to investigate whether pairs that replicated across samples are measuring the same non-additivity. The pairs of SNPs 832-3937 and 3937-4951, and the pairs of SNPs 1998-4951 and 3937-4951, each share one SNP in common. It is unlikely that the pairs of SNPs 832-3937 and 3937-4951 are measuring the same non-additive effects because the r^2 value for the 832-4951 pair is 0.024. On the other hand, pairs of SNPs 1998-4951 and 3937-4951 may be measuring the same non-additive effects, because the r^2 value for the 1998-3937 pair is 0.704.

Conclusion

Sing *et al.* (1996, 2003) and Wright *et al.* (2003) have suggested that the genetic architecture of a complex trait consists of many genes with common alleles (relative allele frequencies greater than 0.01) that have a small effect on a particular phenotype, and a few genes with rare alleles that will have relatively larger effects. An alternative hypothesis assumes that genetic architecture of a complex trait is defined by a few genes with common alleles that have large phenotypic effects (Lander, 1996). Both of these models assume that each genetic variation will make an independent contribution to variability in the phenotype of interest (Wright *et al.* 2003). Our study suggests this may be a biologically unrealistic assumption for particular combinations of variable sites within a gene.

Many associations between common genetic variants and measures of risk of disease have not been reproducible (Hirschhorn et al. 2002). Meta-analyses assume that inconsistencies between studies, or subgroups within studies, are due to chance rather than biological differences attributable to context dependency (Efron & Morris, 1977). Our study suggests that by considering the context in which a genetic variant may have an influence, it is possible to observe results that replicate across samples from different populations. It follows that studies that consider the level of an interacting genetic or environmental context should maximize our ability to detect a gene effect and increase the likelihood of replicating a finding (Wright, 1968). Our study demonstrates that heterogeneity of population structure, i.e., relative allele and genotype frequencies, can be a major factor in determining the comparisons that are available for estimating and testing the statistical significance of non-additive interactions using samples from multiple populations. Furthermore, our study illustrates that non-additivity can be an important consideration in selecting SNPs relevant for defining genotype-phenotype relationships, that all areas within a gene may be involved in non-additive interactions, and that deviations due to non-additivity can make an important contribution to deviations of individual phenotypic values from the population mean.

Acknowledgements

References

Artiga, M.J., Bullido, M.J., Sastre, I., Recuero, M., Garcia, M.A., Aldudo, J., Vazquez, J. & Valdivieso, F. (1998) Allelic polymorphisms in the transcriptional regulatory region of apolipoprotein E gene. *FEBS Lett*. **421**, 105–108.

Avery, L. & Wasserman, S. (1992) Ordering gene function: the interpretation of epistasis in regulatory hierarchies. *Trends Genet*. **8**, 312–316.

Barr, S.E., Kottke, B.A. & Map, S.J.T. (1981) Improved method for determination of triglyc-erides in plasma lipoproteins by an enzymic kit method. *Clin Chem*. **27**, 1142–1144.

Barrett-Connor E. (1997) Sex differences in coronary heart disease. Why are women so superior? The 1995 Ancel Keys Lecture. *Circulation*. **95**, 252–264.

Bateson, W. (1909) *Mendel's principles of heredity*. Cambridge, MA: Cambridge University Press.

Blangero, J., Williams, J.T. & Almasy, L. (2000) Quantitative trait locus mapping using human pedigrees. *Hum Biol*. **72**, 35–62.

Brim, C.A. & Cockerham, C.C. (1961) Inheritance of quantitative characters in soybeans. *Crop Sci*. **1**, 187–190.

Chang, D.J., Paik, Y.K., Leren, T.P., Walker, D.W., Howlett, G.J. & Taylor, J.M. (1990) Characterization of a human apolipoprotein E gene enhancer element and its associated protein factors. *J Biol Chem*. **265**, 9496–9504.

Chen, C., Gentles, A.J., Jurka, J. & Karlin, S. (2002) Genes, pseudogenes, and Alu sequence organization across human chromosomes 21 and 22. *Proc Natl Acad Sci*. *U.S.A*. **99**, 2930–2935.

Cheverud, J.M. & Routman, E.J. (1995) Epistasis and its contribution to genetic variance components. *Genetics*. **139**,1455–1461.

Churchill, G.A. & Doerge, R.W. (1994) Empirical threshold values for quantitative trait mapping. *Genetics*. **138**, 963–971.

Clark, A.G. (2000) Limits to prediction of phenotypes from knowledge of genotypes. In: *Evolutionary Biology*. **Vol 32** (eds.M.T. Clegg), pp.205–224. New York: Plenum Publishers.

Clark, A.G. & Wang, L. (1997) Epistasis in measured genotypes: Drosophila P-element insertions. *Genetics*. **147**, 157–163.

Cloninger, C.R., Van Eerdewegh P., Goate, A., Edenberg, H.J., Blangero, J., Hesselbrock, V., Reich, T., Nurnberger, J., Schuckit, M., Porjesz, B., Crowe, R., Rice, J.P., Foroud, T., Przybeck, T.R., Almasy, L., Bucholz, K., Wu, W., Shears, S., Carr, K., Crose, C., Willig, C., Zhao, J., Tischfield, J.A., Li, T.K. & Conneally, P.M., et al. (1998) Anxiety proneness linked to epistatic loci in genome scan of human personality traits. *Am J Med Genet*. **81**, 313–317.

Cobb, M., Teitlebaum, H., Risch, N., Jekel, J. & Ostfeld, A. (1992) Influence of dietary fat, apolipoprotein E phenotype, and sex on plasma lipoprotein levels. *Circulation*. **86**, 849–857.

Cockerham, C.C. (1954) An extension of the concept of partitioning hereditary variance for analysis of covariances among relatives when epistasis is present. *Genetics*. **39**, 51–54.

Culverhouse, R., Suarez, B.K., Lin, J. & Reich, T. (2002) A perspective on epistasis: limits of models displaying no main effect. *Am J Hum Genet*. **70**, 461–471.

Davignon, J., Cohn, J.S., Mabile, L. & Bernier, L. (1999) Apolipoprotein E and atherosclerosis: insight from animal and human studies. *Clin Chim Acta.* **286**, 115–143.

Edwards, J.W. & Lamkey, K.R. (2002) Quantitative Genetics of Inbreeding in a Synthetic Maize Population. *Crop Sci.* **42**, 1094–1104.

Efron, B. & Morris, C. (1977) Stein's paradox in statistics. *Sci Am.* **236**, 119–127.

Fasoulas, A.C. & Allard, R.W. (1962) Nonallelic gene interactions in the inheritance of quantitative characters in barley. *Genetics.* **47**, 899–907.

Fedorowicz, G.M., Fry, J.D., Anholt, R.R. & Mackay, T.F. (1998) Epistatic interactions between smell-impaired loci in Drosophila melanogaster. *Genetics.* **148**, 1885–1891.

Fisher, R.A. (1918) The correlation between relatives on the supposition of Mendelian inheritance. *Trans R Soc Edinb.* **52**, 399–433.

Fullerton, S.M., Clark, A.G., Weiss, K.M., Nickerson, D.A., Taylor, S.L., Stengård J.H.,

Salomaa, V., Vartiainen, E., Perola, M., Boerwinkle, E. & Sing, C.F. (2000) Apolipoprotein E variation at the sequence haplotype level: implications for the origin and maintenance of a major human polymorphism. *Am J Hum Genet.* **67**, 881–900.

Good, P.I. (2000) *Permutation tests: a practical guide to resampling methods for testing hypotheses.* New York: Springer-Verlag.

Hallauer, A.R. (1981) *Quantitative genetics in maize breeding.* Ames, IA: Iowa State University Press.

Hartwell, L. (2004) Robust Interactions. *Science.* **303**, 774–775.

Hayward, C.S., Kelly, R.P. & Collins, P. (2000) The roles of gender, the menopause and hormone replacement on cardiovascular function. *Cardiovasc Res.* **46**, 28–49.

Hirschhorn, J.N., Lohmueller, K., Byrne, E. & Hirschhorn, K. (2002) A comprehensive review of genetic association studies. *Genet Med.* **2**, 45–61.

Holdrege, C. (1996) *Genetics and the manipulation of life: the forgotten factor of context (renewal in science).* Hudson, NY: Lindisfarne Books.

Jarvik, G.P., Goode, E.L., Austin, M.A., Auwerx, J., Deeb, S., Schellenberg, G.D. & Reed, T. (1997) Evidence that the apolipoprotein E-genotype effects on lipid levels can change with age in males: a longitudinal analysis. *Am J Hum Genet.* **61**, 171–181.

Jo, D.W., Leren, T.P., Yang, Z.Y., Chung, Y.H., Taylor, J.M. & Paik, Y.K. (1995) Characterization of an upstream regulatory element of the human apolipoprotein E gene, and purification of its binding protein from the human placenta. *J Biochem.* **117**, 915–922.

Kaprio, J., Ferrell, R.E., Kottke, B.A., Kamboh, M.I. & Sing, C.F. (1991) Effects of polymorphisms in apolipoproteins E, A-IV, and H on quantitative traits related to risk

for cardiovascular disease. *Arterioscler Thromb.* **11**, 1330–1348.

Kempthorne, O. (1954) The correlation between relatives in a random mating population. *Proc R Soc Lond B Biol Sci.* **143**, 103–113.

Lander, E.S. (1996) The new genomics: global views of biology. *Science.* **274**, 536–539.

Lewontin, R.C. (2000) *The triple helix: gene, organism, and environment.* Cambridge, MA: Harvard University Press.

Luo, L.J., Li, Z.K., Mei, H.W., Shu, Q.Y., Tabien, R., Zhong, D.B., Ying, C.S., Stansel, J.W., Khush, G.S. & Paterson, A.H. (2001) Over dominant epistatic loci are the primary genetic basis of inbreeding depression and heterosis in rice. II. Grain yield components. *Genetics.* **158**, 1755–1771.

Lussier-Cacan S., Bolduc, A., Xhignesse, M., Niyonsenga, T. & Sing, C.F. (2002) Impact of alcohol intake on measures of lipid metabolism depends on context defined by gender, body mass index, cigarette smoking, and apolipoprotein E genotype. *Arterioscler Thromb Vasc Biol.* **22**, 824–831.

Mackay, T.F. (2001) The genetic architecture of quantitative traits. *Annu Rev Genet.* **35**, 303–339.

Matthews, K.A. (1989) Interactive effects of behavior and reproductive hormones on sex differences in risk for coronary heart disease. *Health Psychol.* **8**, 373–387.

National Institute of Health (1974) *Lipid research clinics program manual of laboratory operations.* Washington, DC, Department of Health Education and Welfare Publication No. 75–628.

Nelson, M.R., Kardia, S.L., Ferrell, R.E. & Sing, C.F. (2001) A combinatorial partitioning method to identify multilocus genotypic partitions that predict quantitative trait variation. *Genome Res.* **11**, 458–470.

Nickerson, D.A., Taylor, S.L., Fullerton, S.M., Weiss, K.M., Clark, A.G., Stengård J.T., Salomaa, V., Boerwinkle, E. & Sing, C.F. (2000) Sequence diversity and large-scale typing of SNPs in the human apolipoprotein E gene. *Genome Res.* **10**, 1532–1545.

Page, G.P., George, V., Go, R.C., Page, P.Z., Allison, D.B. (2003) "Are we there yet?": Deciding when one has demonstrated specific genetic causation in complex diseases and quantitative traits. *Am J Hum Genet.* **73**, 711–719.

Paik, Y.K., Chang, D.J., Reardon, C.A., Walker, M.D., Taxman, E. & Taylor, J.M. (1988) Identification and characterization of transcriptional regulatory regions associated with expression of the human apolipoprotein E gene. *J Biol Chem.* **263**, 13340–13349.

Potter, C.J., Huang, H. & Xu, T. (2001) Drosophila Tsc1 functions with Tsc2 to antagonize insulin signaling in regulating cell growth, cell proliferation, and organ size. *Cell.* **105**, 357–368.

Pouliot, J.J., Robertson, C.A. & Nash, H.A. (2001) Pathways for repair of topoisomerase I covalent complexes in Saccharomyces cerevisiae. *Genes Cells.* **6**, 677–687.

Reilly, S.L., Ferrell, R.E., Kottke, B.A., Kamboh, M.I. &

Sing, C.F. (1991) The gender-specific apolipoprotein E genotype influence on the distribution of lipids and apolipoproteins in the population of Rochester MN. I. Pleiotropic effects on means and variances. *Am J Hum Genet*. **49**, 1155–1166.

Reilly, S.L., Ferrell, R.E., Kottke, B.A. & Sing, C.F. (1992) The gender-specific apolipoprotein E genotype influence on the distribution of plasma lipids and apolipoproteins in the population of Rochester, Minnesota. II. Regression relationships with concomitants. *Am J Hum Genet*. **51**, 1311–1324.

Reilly, S.L., Ferrell, R.E. & Sing, C.F. (1994) The gender-specific apolipoprotein E genotype influence on the distribution of plasma lipids and apolipoproteins in the population of Rochester MN. III. Correlations and covariances. *Am J Hum Genet*. **55**, 1001–1018.

Routman, E.J. & Cheverud, J.M. (1997) Gene Effects on a quantitative trait: two locus epistasis effects measured at microsatellite markers and at estimated QTL. *Evolution*. **51**, 1654–1662.

Russell, W.A. & Eberhart, S.A. (1970) Effects of three gene loci in the inheritance of quantitative characters in maize. *Crop Sci*. **10**, 165–169.

Salomaa, V.V., Rasi, V.P., Vahtera, E.M., Pekkanen, J., Pursiainen, M., Jauhiainen, M., Vartiainen, E., Ehnholm, C.P. & Myllyla, G. (1994) Haemostatic factors and lipoprotein (a) in three geographical areas in Finland: the Finrisk Haemostasis Study. *J Cardiovasc Risk*. **1**, 241–248.

Schiele, F., De Bacquer D., Vincent-Viry M., Beisiegel, U., Ehnholm, C., Evans, A., Kafatos, A., Martins, M.C., Sans, S., Sass, C., Visvikis, S., De Backer G., Siest, G. & The ApoEurope Group (2000) Apolipoprotein E serum concentration and polymorphism in six European countries: The ApoEurope Project. *Atherosclerosis*. **152**, 475–488.

Searle, S.R. (1971) *Linear models*. New York: John Wiley & Sons.

Sing, C.F., Haviland, M.B. & Reilly, S.L. (1996) Genetic architecture of common multifactorial diseases. In: *Variation in the Human Genome (Ciba Found Symp197)* (eds.D.J. Chadwick & G. Cardew) pp 211–232. Chichester, UK: John Wiley & Sons.

Sing, C.F., Stengård J.H. & Kardia, S.L.R. (2003) Genes, environment and coronary artery disease. *Arterioscler Thromb Vasc Bio*. **23**, 1190–1196.

Small, K.M., Wagoner, L.E., Levin, A.M., Kardia, S.L.R. & Liggett, S.B. (2002) Synergistic polymorphisms of beta1- and alpha2C-adrenergic receptors and the risk of congestive heart failure. *N Engl J Med*. **347**, 1135–1142.

Smith, J.D., Melian, A., Leff, T. & Breslow, J.L. (1988) Expression of the human apolipoprotein E gene is regulated by multiple positive and negative elements. *J Biol Chem*. **263**, 8300–8308.

Sokal, R.R. & Rohlf, F.J. (1995) *Biometry: the principles and practice of statistics in biological research. 3rd ed*. New York: W. H. Freeman and Co.

Stengård J.H., Clark, A.G., Weiss, K.M., Kardia, S., Nickerson, D.A., Salomaa, V., Ehnholm, C., Boerwinkle, E. & Sing, C.F. (2002) Contributions of 18 additional DNA sequence variations in the gene encoding apolipoprotein E to explaining variation in quantitative measures of lipid metabolism. *Am J Hum Genet*. **71**, 501–517.

Stuber, C.W. & Moll, R.H. (1971) Epistasis in maize (zea mays l.). II. Comparison of selected with unselected populations. *Genetics*. **67**, 137–149.

Taylor, J.M., Lauer, S., Elshourbagy, N., Reardon, C., Taxman, E., Walker, D., Chang, D. & Paik, Y.K. (1987) Structure and evolution of human apolipoprotein genes: identification of regulatory elements of the human apolipoprotein E gene. *Ciba Found Symp*. **130**, 70–86.

Templeton, A.R. (2000) Epistasis and Complex Traits. In: *Epistasis and the Evolutionary Process* (eds.J.B. Wolf E.D. Brodie, M.J. Wade), pp. 41–57. New York: Oxford University Press.

Thomas, J.L. & Braus, P.A. (1998) Coronary artery disease in women. A historical perspective. *Arch Intern Med*. **158**, 333–337.

Tong, A.H., Evangelista, M., Parsons, A.B., Xu, H., Bader, G.D., Page, N., Robinson, M., Raghibizadeh, S., Hogue, C.W., Bussey, H., Andrews, B., Tyers, M. & Boone, C. (2001) Systematic genetic analysis with ordered arrays of yeast deletion mutants. *Science*. **294**, 2364–2368.

Turner, S., Weidman, W., Michels, V., Reed, T., Ormson, C., Fuller, T. & Sing, C.F. (1989) Distribution of sodium-lithium countertransport and blood pressure in Caucasians five to eighty–nine years of age. *Hypertension*. **13**, 378–391.

Wright, A., Charlesworth, B., Rudan, I., Carothers, A. & Campbell, H. (2003) A polygenic basis for late-onset disease. *Trends Genet*. **19**, 97–106.

Wright, S (1968) *Evolution and the genetics of populations; a treatise*. Chicago: University of Chicago Press.

Xhignesse, M., Lussier-Cacan S., Sing, C.F., Kessling, A.M. & Davignon, J. (1991) Influences of common variants of apolipoprotein E on measures of lipid metabolism in a sample selected for health. *Arterioscler Thromb*. **11**, 1100–1110.

Zerba, K.E., Ferrell, R.E. & Sing, C.F. (1996) Genotype-environment interaction: apolipoprotein E (ApoE) gene effects and age as an index of time and spatial context in the human. *Genetics*. **143**, 463–478.

Zhu, X., Bouzekri, N., Southam, L., Cooper, R.S., Adeyemo, A., McKenzie, C.A., Luke, A., Chen, G., Elston, R.C. & Ward, R. (2001) Linkage and association analysis of angiotensin I–converting enzyme (ACE)-gene polymorphisms with ACE concentration and blood pressure. *Am J Hum Genet*. **68**, 1139–1148.