Epistatic Interactions

Tyler J. VanderWeele[*]

[*]Harvard University, tvanderw@hsph.harvard.edu

Epistatic Interactions[*]

Tyler J. VanderWeele

Abstract

The term "epistasis" is sometimes used to describe some form of statistical interaction between genetic factors and is alternatively sometimes used to describe instances in which the effect of a particular genetic variant is masked by a variant at another locus. In general statistical tests for interaction are of limited use in detecting "epistasis" in the sense of masking. It is, however, shown that there are relations between empirical data patterns and epistasis that have not been previously noted. These relations can sometimes be exploited to empirically test for "epistatic interactions" in the sense of the masking of the effect of a particular genetic variant by a variant at another locus.

KEYWORDS: epistasis, gene-gene interaction, synergism

Introduction

Writing in 1909, Bateson used the term "epistasis" to describe instances in which the effect of a particular genetic variant was masked by a variant at another locus so that variation of phenotype with genotype at one locus was only apparent amongst those with certain genotypes at the second locus (Bateson, 1909). In recent papers, Cordell (2002, 2009) has argued that the statistical tests that are often used to assess interactions (Ritchie et al., 2001; Hahn et al., 2003; Moore, 2004; Chung et al., 2007; Purcell et al., 2007; Zhang and Liu, 2007; Ferreira et al., 2007; Gavan et al., 2008) are of limited use in elucidating the type of biologic interaction that Bateson had originally conceived. Recent developments have extended interaction tests for case-control design to settings of case-only designs (Piegorsch et al., 1994; Khoury and Flanders, 1996; Yang et al., 1999; Weinberg and Umback, 2000) and to family-based association studies (Cordell and Clayton, 2002; Cordell et al., 2004; Laird and Lange, 2006; Martin et al., 2006; Kotti et al., 2007; Lou et al., 2008; Hoffmann et al., 2009); however, these developments are arguably also subject to Cordell's critique (2002). In this paper, it is argued that there are relations between empirical data patterns and epistasis in the sense of masking that have not been previously noted and that can sometimes be exploited to empirically test for epistasis as originally conceived by Bateson.

Empirical Tests for Epistasis

Under Bateson's original conception, epistasis would be said to be present if variation of phenotype with genotype at one locus was only apparent amongst those with certain genotypes at the second locus. Those with other genotypes at the second locus would show no effect at the first. Consider first a setting in which genotypes at both loci can effectively be considered binary as in Table 1; below we will consider more general settings.

Table 1. Example of a table of phenotypes for a particular individual for the effects of different genotypes at two loci exhibiting epistasis under Bateson's (1909) original definition

	Genotype at locus B	
Genotype at locus A	*b/b* or *b/B*	*B/B*
a/a or *a/A*	0	0
A/A	0	1

Table 1 describes a potential phenotype pattern for a particular individual such that the effect of genotype at locus A is only present for the *B/B* variant; if the genotype at locus B is not *B/B* then the effect of variation at locus A is not apparent. The effect of genetic variation at locus A can be masked by that at locus B and we might say that that locus B is epistatic to locus A. By symmetry in this example, it is also the case the effect of genotype at locus B is only present for the *A/A* variant so the effect at locus B can be masked by that at locus A and we might thus also say that that locus A is epistatic to locus B (Cordell, 2002).

In this simple setting in which the genotype at these two loci can effectively be considered binary, one way to conceive of epistasis then is whether there are any individuals for whom the response pattern follows that in Table 1. In populations with heterogeneity and for complex traits with non-Mendelian inheritance, the response patterns may vary between individuals but we may be interested whether there are any individuals whose phenotype response patterns manifest such epistasis. Let X_1 be a binary indicator for genotype at locus A (in the example, $X_1=0$ for genotype *a/a* or *a/A* and $X_1=1$ for genotype *A/A*); let X_2 be a binary indicator for genotype at locus B (in the example, $X_2=0$ for genotype *b/b* or *b/B* and $X_2=1$ for genotype *B/B*). Let D be a binary indicator of phenotype, indicating the presence of some dichotomous trait. For each individual in the population let D_{ij} denote what the trait would have been if X_1 were i and if X_2 were j. For each individual we could thus consider D_{11}, D_{10}, D_{01} and D_{00} i.e. what would have happened to the individual under the presence or absence of each of the two factors. An epistatic interaction, in Bateson's original sense of masking, would be present if there were an individual for whom Table 1 describes the phenotype response pattern for that individual. In others words, an epistatic interaction would be present if there were an individual for whom:

$$D_{11}=1 \text{ but } D_{10}=0, D_{01}=0, D_{00}=0. \tag{1}$$

Another way of asking whether there are individuals for whom relation (1) is satisfied is to ask whether there are individuals for whom

$$D_{11} - D_{10} - D_{01} - D_{00} > 0. \tag{2}$$

In general, we may not be able to infer what all of D_{11}, D_{10}, D_{01}, D_{00} would be for a particular individual. However, in a genetic association study we might hope to be able to estimate the values of D_{11}, D_{10}, D_{01}, D_{00} on average for a population by careful control for confounding by stratification and admixture. If we let C denote a genetic marker for population substructure based on many loci (Pritchard and Rosenberg, 1999; Pritchard, 2000; Satten et al., 2001; Hoggart et al., 2003; Price et al., 2006) then if control for the marker suffices to control for

confounding then we can estimate the average likelihood of the outcome D when $X_1=i$ and $X_2=j$ for those with genetic marker C=c by:

$$P(D_{ij}=1|C=c) \approx P(D=1|X_1=i,X_2=j,C=c). \tag{3}$$

If marker C suffices to control for confounding, then the average effects of genetic factors X_1 and X_2 on D can be estimated with data and (3) will hold (Hernán, 2004). We could thus test whether there are any individuals with C=c whose response patterns satisfy (2) (i.e. for whom the response pattern is that in Table 1) by testing:

$$p_{11c} - p_{10c} - p_{01c} - p_{00c} > 0 \tag{4}$$

where $p_{ijc} = P(D=1|X_1=i,X_2=j,C=c)$. Provided the genetic marker C suffices to control for confounding by stratification and admixture so that relation (3) holds, then if for some value of the genetic marker we find that $p_{11c} - p_{10c} - p_{01c} - p_{00c} > 0$ then there must be some individuals with C=c for whom the response pattern is given by Table 1 i.e. for whom an epistatic interaction, in the sense of Bateson, is present. Condition (4) is not the usual statistical test for interaction but it can be tested empirically from data to draw conclusions about whether there are at least some individuals with an epistatic response pattern. We will discuss below the relationship between condition (4) and the usual statistical tests for interactions. Note that the implication described above is one-way; if condition (4) is satisfied then there are individuals for whom $D_{11}=1$ but $D_{10}=0$, $D_{01}=0$, $D_{00}=0$; however, if condition (4) is not satisfied we cannot necessarily conclude that there are no individuals for whom $D_{11}=1$ but $D_{10}=0$, $D_{01}=0$, $D_{00}=0$. Condition (4) is a sufficient condition for an epistatic interaction but not a necessary condition.

Using the same logic as that given above, we could empirically test for other epistatic response patterns. We could test whether there are individuals for whom $D_{00}=1$ but $D_{11}=D_{10}=D_{01}=0$ by testing $p_{00c} - p_{11c} - p_{10c} - p_{01c} > 0$; we could test whether there are individuals for whom $D_{10}=1$ but $D_{11}=D_{01}=D_{00}=0$ by testing $p_{10c} - p_{11c} - p_{01c} - p_{00c} > 0$; finally, we could test whether there are individuals for whom $D_{01}=1$ but $D_{11}=D_{10}=D_{00}= 0$ by testing $p_{01c} - p_{11c} - p_{10c} - p_{00c} > 0$.

Tests under Additional Assumptions

In some cases, we might be willing to assume that genotype $X_1=1$ (as compared with $X_1=0$) never prevents the outcome so that if $D_{00}=1$ then it is also the case that $D_{10}=1$ and if $D_{10}=0$ then it must also the case that $D_{00}=0$ and similarly, if $D_{01}=1$ then it is also the case that $D_{11}=1$ and if $D_{11}=0$ then it must also the case that $D_{01}=0$. In such cases, $X_1=1$ (as compared with $X_1=0$) is never preventive in that it

has either a neutral or causative effect on all individuals. Such cases of no preventive effects are sometimes referred to as monotonicity relationships. Stated more succinctly, we may say that X_1 has a monotonic effect on D if for all individuals in the population, $D_{1j} \geq D_{0j}$ for j=0,1. Similarly, we say that X_2 has a monotonic effect on D if for all individuals in the population, $D_{i1} \geq D_{i0}$ for i=0,1. Monotonicity is a strong assumption and will often not hold. For X_1 to have a monotonic effect on D, it must be the case that $X_1=1$ (as compared with $X_1=0$) is either neutral or causative of the outcome D=1 for all individuals in the population i.e. $X_1=1$ (as compared with $X_1=0$) never prevents the outcome. For X_2 to have a monotonic effect on D, it must be the case that $X_2=1$ (as compared with $X_2=0$) is either neutral or causative of the outcome D=1 for all individuals in the population i.e. $X_2=1$ never prevents the outcome. Whenever a particular genetic variant is such that it makes the outcome more likely in some populations but less likely in others, this monotonicity relation will not hold. In some cases, monotonicity of X_1 or X_2 might hold within certain strata of a genetic marker C but not in others.

When monotonicity assumptions do hold, we can test for epistatic interactions by testing a condition weaker than that given in (4) above. Suppose for example that X_1 had a monotonic effect on D then if it were the case that D_{10} were 0 then it must also be the case that D_{00} is 0. Thus if X_1 had a monotonic effect on D and there were individuals for whom $D_{11}=1$ and $D_{10}=D_{01}=0$ then we could also conclude for such individuals that $D_{00}=0$ by monotonicity and thus that relation (1) held for such individuals, i.e. that an epistatic interaction as given in Table 1 was present. If some genetic marker C suffices to control for confounding by stratification and admixture then we could test whether there were individuals for whom $D_{11}=1$ and $D_{10}=D_{01}=0$ within stratum C=c by testing

$$p_{11c} - p_{10c} - p_{01c} > 0 \qquad (5)$$

Note that condition (5) is a weaker condition than condition (4); condition (5) does not require subtracting p_{00c}. When we can assume that the effect of X_1 on D is monotonic (i.e. never preventive) then we can test this weaker condition instead. By symmetry, it is also the case that if X_2, rather than X_1, has a monotonic effect on D, then condition (5) can be used to test whether there are individuals with a response pattern like that given in Table 1. We seen then that if *either* X_1 or X_2 has a monotonic effect on D then we can use the weaker condition (5), rather than condition (4), to test for epistasis in the sense of Bateson (1909). Once again, however, condition (5) does *not* correspond to a standard statistical test for interaction.

Finally suppose that *both* X_1 and X_2 have monotonic effects on D. Suppose $X_1=1$ (as compared with $X_1=0$) never prevented the outcome for any individual and $X_2=1$ (as compared with $X_2=0$) never prevented the outcome for any individual. Stated another way, we are supposing that D_{ij} is non-decreasing in i and j. In the appendix we show that if some genetic marker C suffices to control for confounding by stratification and admixture then we could test whether there were individuals for whom $D_{11}=1$ and $D_{10}=D_{01}=D_{00}=0$ within stratum C=c by testing

$$p_{11c} - p_{10c} - p_{01c} + p_{00c} > 0 \qquad (6)$$

Note that condition (6) is weaker than both conditions (4) or (5) because now we are adding back the term p_{00c}. Condition (6) is how interaction is often ordinarily assessed in statistical models; condition (6) essentially examines whether the effects of X_1 and X_2 combined are greater than the sum of the effects of X_1 and X_2 considered separately. However, condition (6) will only imply individuals with the epistatic response pattern in Table 1 if it can be assumed that *both* X_1 and X_2 have monotonic effects on D. In other words, if there are any individuals with C=c for whom the outcome would be present if $X_1=0$ but for whom it would not be present if $X_1=1$ (or similarly for whom the outcome would be present if $X_2=0$ but for whom it would not be present if $X_2=1$) then the monotonicity conditions would be violated and one could not use condition (6) to test for epistatic response patterns. We have seen above that even if these monotonicity assumptions are violated then condition (4) or (5) could be used to test for epistasis in the sense of masking; however, condition (6) which is the usual test for interaction, only gives a test for epistasis, in the sense of masking, under strong monotonicity assumptions for both genetic factors.

More General Settings

Consider now the more setting in which at loci A and B there are three distinct relevant genotypes: *a/a*, *a/A* and *A/A* at locus A and *b/b*, *b/B* and *B/B* at locus B. Now let V_1 and V_2 be variables with three levels indicating the genotype at loci A and B respectively (e.g. $V_1=0$ for *a/a*, $V_1=1$ for *a/A*, $V_1=2$ for *A/A* and $V_2=0$ for *b/b*, $V_2=1$ for *b/B*, $V_2=2$ for *B/B*). Once again, let D be a binary indicator of phenotype, indicating the presence of some dichotomous trait. For each individual in the population let D_{ij} denote what the trait would be if V_1 were i and if V_2 were j. Again let C denote a genetic marker for population substructure and suppose that the marker suffices to control for confounding by stratification and admixture so that $P(D_{ij}=1|C=c) \approx P(D=1|V_1=i,V_2=j,C=c)$.

As before we will let $p_{ijc} = P(D=1|V_1=i,V_2=j,C=c)$. We can then consider a variety of response patterns that would constitute instances of epistasis. For simplicity now assume that the effects of V_1 and V_2 on D are monotonic so that whenever i≥i' we have $D_{ij} \geq D_{i'j}$ and whenever j≥j' we have $D_{ij} \geq D_{ij'}$. Consider the response pattern in Table 2.

Table 2. Example of a table of phenotypes for the effects of genotypes at two loci exhibiting epistasis, with three relevant genetic variants at each locus

Genotype at locus A	Genotype at locus B		
	b/b	*b/B*	*B/B*
a/a	0	0	0
a/A	0	0	0
A/A	0	0	1

By arguments similar to those given above, there must be individuals with genetic marker C=c who have response patterns given by Table 2 if it is the case that

$$p_{22c} - p_{21c} - p_{12c} + p_{11c} > 0 \qquad (7)$$

Note that for the response pattern given in Table 2, the effect of genetic variation at locus A is only apparent when the genotype is *B/B* at locus B (similarly, the effect of genetic variation at locus B is only apparent when the genotype is *A/A* at locus A). Thus the response pattern in Table 2 would be another instance which would be considered epistasis under Bateson's original conception. If control for genetic marker C suffices to control for confounding by stratification and admixture and if condition (7) holds then there must be some individuals with genetic marker C=c who have response patterns given by Table 2. We can once again test for epistasis empirically.

A number of other response patterns constituting instances of epistasis are also possible. Consider, for example, the response pattern in Table 3.

Table 3. Example of a table of phenotypes for the effects of genotypes at two loci exhibiting epistasis, with three relevant genetic variants at each locus

Genotype at locus A	Genotype at locus B		
	b/b	*b/B*	*B/B*
a/a	0	0	0
a/A	0	0	1
A/A	0	0	1

There must be individuals amongst those with genetic marker C=c who have response patterns given by Table 3 if it is the case that

$$p_{12c} - p_{21c} - p_{02c} + p_{01c} > 0. \tag{8}$$

Table 4. Example of a table of phenotypes for the effects of genotypes at two loci exhibiting epistasis, with three relevant genetic variants at each locus			
		Genotype at locus B	
Genotype at locus A	*b/b*	*b/B*	*B/B*
a/a	0	0	0
a/A	0	0	0
A/A	0	1	1

There must be individuals amongst those with genetic marker C=c who have response patterns given by Table 4 if it is the case that

$$p_{21c} - p_{12c} - p_{20c} + p_{10c} > 0. \tag{9}$$

Table 5. Example of a table of phenotypes for the effects of genotypes at two loci exhibiting epistasis, with three relevant genetic variants at each locus			
		Genotype at locus B	
Genotype at locus A	*b/b*	*b/B*	*B/B*
a/a	0	0	0
a/A	0	1	1
A/A	0	1	1

There must be individuals amongst those with genetic marker C=c who have response patterns given by Table 5 if it is the case that

$$p_{11c} - p_{21c} - p_{12c} + p_{00c} > 0. \tag{10}$$

The tests represented by conditions (7)-(10) presupposed that the effects of V_1 and V_2 on D were monotonic (i.e. never preventive). In the appendix, we consider tests for the epistatic interactions given in Tables 2-5 without monotonicity assumptions or when only one of V_1 or V_2 have a monotonic effect on D. We also consider tests when one factor has two levels and the other has

three levels. The basic point here is that the approach described in the previous section to empirically test for epistasis can be employed even when genotypes are considered to have three possible relevant variants rather than two.

Relation to Statistical Models

In this section we will briefly relate the empirical tests for epistasis to standard tests for interactions in statistical models. For simplicity we will again return to the setting in which the relevant genotypes can effectively be considered binary as in Table 1. Similar remarks apply to the more general settings described in the previous section. When two binary genetic variants are considered, a statistical model of the following form is sometimes used to test for a statistical interaction:

$$P(D=1|X_1=x_1, X_2=x_2) = \alpha_0 + \alpha_1 x_1 + \alpha_2 x_2 + \alpha_3 x_1 x_2 \tag{11}$$

To control for confounding by stratification and admixture, one can fit a separate model like (11) within each stratum C=c of some genetic marker. Statistical interaction is then often assessed by testing whether $\alpha_3 > 0$. Testing whether $\alpha_3 > 0$ corresponds to a test of condition (6). We saw above that condition (6) can be used to test for epistasis in the sense of masking only under the strong assumption that both X_1 and X_2 have monotonic effects on the outcome. However, we also saw that even if these monotonicity assumptions do not hold, we could still test for epistatic response patterns; however, we would have to use more stringent conditions like (4) or (5).

We can also express conditions (4) and (5) in terms of the coefficients of the statistical model in (11). We saw above that we could use condition (5) to test for epistasis in the sense of masking if at least one of X_1 or X_2 has a monotonic effect on the outcome. Condition (5) can expressed in terms of the coefficients of statistical model (11) as $\alpha_3 > \alpha_0$. Thus if at least one of X_1 and X_2 had monotonic effects on the outcome then we could test for such epistasis by testing whether $\alpha_3 > \alpha_0$. Even if neither X_1 nor X_2 had monotonic effects (i.e. we make no assumptions about monotonicity) we could still test for epistasis in the sense of masking by testing condition (4). Condition (4) can be expressed in terms of the coefficients of statistical model (11) as $\alpha_3 > 2\alpha_0$. Thus even without making any assumptions about monotonicity we could test for such epistasis by testing whether $\alpha_3 > 2\alpha_0$. These are non-standard tests for interaction but, when satisfied, allow for conclusions to be drawn not just about statistical interaction but about epistatic response patterns. As noted above, these tests are sufficient conditions for epistasis in the sense of masking, but not necessary. If the conditions are satisfied then there are at least some individuals with response patterns manifesting epistasis in Table 1. If the conditions are not satisfied then there may

or may not be individuals with response patterns exhibiting epistasis; we cannot tell from the data.

Multiple Uses of the Word "Epistasis"

Cordell and Clayton (2005) note that although Bateson conceived of epistasis in terms of the masking of the effect of one genetic factor by another, the term "epistasis" soon began to take on a variety of meanings. They note that not long after Bateson, Fisher (1918) used the term "epistacy" to refer to a statistical interaction in the sense of deviation from additive effects such as $\alpha_3 > 0$ in the statistical linear model given above. Cordell and Clayton further argue that Fisher's "epistacy" quickly evolved into "epistasis" so that in the modern genetics literature the two uses of the word coexist creating ambiguity. Cordell (2002, 2009) argues that epistasis in the statistical sense does not in general imply epistasis in original sense of Bateson, the sense of the masking of the effect of one genetic factor by another.

We have seen here, however, that the two uses of the word "epistasis" are not entirely unrelated. In particular, in some very special circumstances epistasis in the statistical sense ($\alpha_3 > 0$) implies epistasis in the sense of the masking the effect of one genetic factor by another ($D_{11}=1$ but $D_{10}=D_{01}=D_{00}=0$). More specifically, epistasis in the statistical sense ($\alpha_3 > 0$) implies epistasis in the sense of the masking the effect of one genetic factor by another only when it can be assumed that both X_1 and X_2 had monotonic effects on the outcome. This is a very strong assumption and one which in many contexts will not hold. When it does not hold, statistical epistasis does not imply epistasis in the sense of masking. However, we have also seen in this section, that there are further relationships between statistical models and data patterns on the one hand and epistasis in the sense of masking on the other; these relations have been previously unrecognized. We have seen that if at least one of X_1 or X_2 have a monotonic effect on the outcome then we can test for epistasis in the masking sense ($D_{11}=1$ but $D_{10}=D_{01}=D_{00}=0$) by testing whether $\alpha_3 > \alpha_0$. Even without any monotonicity assumptions, we can test for epistasis in the masking sense by testing whether $\alpha_3 > 2\alpha_0$. Again, these are stronger conditions than regular tests for statistical interactions.

In a recent review article, Phillips (2008) also discusses the ambiguity in the term "epistasis" and he distinguishes what he considers as three distinct forms of epistasis. Phillips used "statistical epistasis" to refer to a departure from additive effects in a statistical model (or more generally a departure from independent effects on some scale of measurement). Phillips introduced the term "compositional epistasis" to refer to epistasis in Bateson's original sense of the term, i.e. the masking of the effect of an allele at one locus by an allele at another

locus. Finally, Phillips used "functional epistasis" to describe the physical molecular interactions between various proteins (and other genetic elements). Compositional epistasis, as defined by Phillips, need not necessarily imply "functional epistasis" but compositional epistasis is nevertheless arguably a more biological form of interaction than mere "statistical epistasis." The tests described in the previous sections constitute empirical tests for what Phillips referred to as "compositional epistasis."

Testing for Epistatic Interactions in Case-Control Studies

Many analyses of interaction use data from a case-control study. In such case-control studies risks like p_{11c}, p_{10c}, p_{01c}, and p_{00c} cannot in general be estimated but odds ratios for the effects of genetic factors can be estimated. Thus in such studies, logistic regression is often used which for interaction analyses may take the form of:

$$\text{logit} \{ P(D=1 | X_1 = x_1, X_2 = x_2) \} = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_1 x_2 \tag{12}$$

Model (12) to can be used to calculate odds ratios comparing the odds of the outcome when both X_1 and X_2 are present to when both are absent (denoted by OR_{11}), the odds when $X_1=1$ and $X_2=0$ to when both are absent (denoted by OR_{10}) and the odds when $X_1=0$ and $X_2=1$ to when both are absent (denoted by OR_{01}). When the outcome is rare these odds ratios approximate the corresponding relative risks, denoted by RR_{11}, RR_{10}, RR_{01}. Although we cannot test conditions (4) or (5) or (6) above directly using risks we could divide these conditions by p_{00c}. Condition (4) becomes

$$RR_{11c} - RR_{10c} - RR_{01c} - 1 > 0. \tag{13}$$

Condition (5) becomes

$$RR_{11c} - RR_{10c} - RR_{01c} > 0. \tag{14}$$

Condition (6) becomes

$$RR_{11c} - RR_{10c} - RR_{01c} + 1 > 0. \tag{15}$$

Under the assumption that the outcome is rare, these conditions could be tested using the odds ratios from a logistic regression. Thus even in a case-control study one can potentially test for epistasis in settings in which the outcome is rare. The quantity $RR_{11c} - RR_{10c} - RR_{01c} + 1$ is sometimes described as

the "relative excess risk due to interaction" or RERI (Rothman, 1986). The three conditions given above could thus be written respectively as RERI>2, RERI>1, RERI>0. Statistical tests and confidence intervals for this quantity, RERI, are given elsewhere (Richardson and Kaufman, 2009). In a case-control study with a rare outcome, epistasis could be tested by testing RERI>0 if both X_1 and X_2 can be assumed to have monotonic effects, by testing RERI>1 if one of X_1 or X_2 can be assumed to have a monotonic effect, and by testing RERI>2 if no monotonicity assumptions are made. Alternatively, it can also be shown (see the appendix) that if the outcome is rare then (4), (5) or (6) will be satisfied, respectively, if for the coefficients in logistic model (12), we have $\beta_3 > \log(3)$, $\beta_3 > \log(2)$, or $\beta_3 > 0$ provided that the main effects β_1 and β_2 are non-negative. Similar results hold when one or both factors have three levels; see the appendix for additional discussion.

Relation to Sufficient Causation

The tests described above for epistatic response patterns bear a certain relation to tests for synergism in the sense conceived of by Rothman (1976). Rothman conceptualized causation as a series of mechanisms for the outcome each of which involved the conjunction of various factors such that whenever all the factors for a particular mechanism were present the outcome would occur. Such mechanisms or "sufficient causes" might require the absence of presence of two or more particular factors of interest, X_1 and X_2, along with other possibly unknown factors. Rothman conceived of synergism being present whenever there was a sufficient cause that required both X_1 and X_2 to operate. VanderWeele and Robins (2007, 2008) formalized Rothman's sufficient cause framework and introduced the notion of a sufficient cause interaction. A sufficient cause interaction is present if there are individuals who have a response pattern such as that given in Table 6 i.e. if there are individuals for whom $D_{11}=1$ and $D_{10}=D_{01}=0$; D_{00} can be either 1 or 0. VanderWeele and Robins (2008) showed that a sufficient cause interaction implied synergism as conceived of by Rothman.

Table 6. Example of a table of phenotypes for two factors, X_1 and X_2, exhibiting a sufficient cause interaction, implying synergism as conceived by Rothman (1976)

	Value of X_2	
Value of X_1	0	1
0	?	0
1	0	1

The notion of an epistatic response pattern such as that given in Table 1 is stronger than that of a sufficient cause interaction because the response pattern in Table 1 requires that $D_{00}=0$. If at least one of the two genetic factors has a monotonic effect on the outcome D then the concepts of an epistatic interaction ("compositional epistasis") and a sufficient cause interaction between two factors coincide. If neither of the two factors has a monotonic effect on the outcome, then an epistatic interaction is a stronger condition than a sufficient cause interaction. Statistical tests for sufficient cause interactions have been described elsewhere (VanderWeele and Robins, 2007; Vansteelandt et al., 2008; VanderWeele, 2009, VanderWeele et al., in press) and these statistical tests could also be used for epistatic interactions if at least one of the two factors has a monotonic effect on the outcome.

Discussion

Bateson's original conception of epistasis was that the effect of a gene at one locus would be masked for certain values of the genotype at a second locus. In this paper we have derived conditions that can be tested empirically for detecting whether there are individuals whose response patterns manifest epistasis in the sense of masking originally conceived by Bateson. It was shown that only under some very strong assumptions would tests for regular statistical interactions correspond to epistasis in the masking sense of the term. We have, however, further seen that even without such strong assumptions one can still test whether there are individuals for whom the effect of a gene at one locus would be masked for certain values of the genotype at a second locus. The empirical conditions described above for detecting epistasis are quite strong but the conclusions which tests of these conditions allow (conclusions concerning "compositional epistasis") may be of interest in a wide range of studies.

The tests derived required control for a genetic marker, denoted in this paper by C, to control for control for confounding by stratification and admixture so that the associations observed between the genes of interest and the outcome at least approximately correspond to the true effects of these genes. The tests will be valid only to the extent that this approximation holds. Other genetic or environmental factors could be included in C to attempt to better control for confounding. When C contains multiple factors, more sophisticated statistical techniques may be desirable to allow for multivariate control.

In many studies, identified genetic risk factors will be in linkage disequilibrium with the true causal genetic factor; in such cases of linkage disequilibrium the genetic risk factor in the study might be conceptualized as a misclassified version of the true causal factor. Future research will examine the extent to which conclusions about epistasis concerning the true causal genetic

factors can be drawn from identified genetic risk factors in linkage disequilibrium with the true causal factors.

This paper has focused on epistasis for two genetic factors by considering the response pattern tables that exhibit epistasis as conceived of by Bateson. The tests described in this paper may also be of interest, however, in assessing gene-environment interactions. In particular, the tests that have been described could also be used to detect individuals with particular gene-environment interaction response patterns corresponding to Tables 1-5; we might refer to such response patterns as instances of "compositional" gene-environment interaction. As described in this commentary, these will only correspond to the ordinary tests for statistical interactions between genetic and environmental factors in very special cases.

It is hoped that the contributions in this commentary have clarified some of the conceptual relationships between epistasis as conceived of by Bateson and statistical tests using data. It is further hoped that the empirical tests for epistasis derived in this paper will be employed in future analyses of genetic data.

Appendix

Proof that relation (6) suffices for epistasis under monotonicity

Here we prove that under the assumption that both X_1 and X_2 have monotonic effects on D then $p_{11c} - p_{10c} - p_{01c} + p_{00c} > 0$ implies there is an individual for whom $D_{11}=1$ and $D_{10}=D_{01}=D_{00}=0$. Suppose that both X_1 and X_2 have monotonic effects on D so that D_{ij} is non-decreasing in i and j i.e. for no individual does $X_1=1$ (as compared with $X_1=0$) ever prevent the outcome and for no individual does $X_2=1$ (as compared with $X_2=0$) ever prevent the outcome. Under these monotonicity assumptions if there is an individual for whom $D_{11}=1$ and $D_{10}=D_{01}=0$ then it is also the case for that individual that $D_{00}=0$. Suppose there were no individual for whom $D_{11}=1$ and $D_{10}=D_{01}=0$; then whenever $D_{11}=1$ then we must have that either $D_{10}=1$ or $D_{01}=1$. We also have that $D_{10} \geq D_{00}$ and that $D_{01} \geq D_{00}$. Thus if for some individual it is not the case that $D_{11}=1$ and $D_{10}=D_{01}=0$ then we must have that $D_{11} - D_{10} - D_{01} + D_{00} \leq 0$. Thus if there is no individual in some subpopulation with C=c such that $D_{11}=1$ and $D_{10}=D_{01}=0$ then it must be the case that for all individuals with C=c that $D_{11} - D_{10} - D_{01} + D_{00} \leq 0$. From this it would follow that if some genetic marker C suffices to control for confounding by stratification and admixture then if there were no individuals with C=c and with $D_{11}=1$ and $D_{10}=D_{01}=0$ then we would have

$$p_{11c} - p_{10c} - p_{01c} + p_{00c} \leq 0.$$

From this it follows that if we were to find

$$p_{11c} - p_{10c} - p_{01c} + p_{00c} > 0$$

then there must be some individuals with C=c such that $D_{11}=1$ and $D_{10}=D_{01}=0$ and by monotonicity it would also be the case for these individuals that $D_{00}=0$. Thus these individuals would have an epistatic response pattern.

Further tests for epistasis when the genetic factors have three levels

Here we will consider further tests for epistatic response patterns in which, at each locus of interest, there are three distinct relevant genotypes. As in the paper we will let V_1 and V_2 denote variables with three levels indicating the genotype at loci A and B respectively (e.g. $V_1=0$ for *a/a*, $V_1=1$ for *a/A*, $V_1=2$ for *A/A* and $V_2=0$ for *b/b*, $V_2=1$ for *b/B*, $V_2=2$ for *B/B*). Once again, let D be a binary indicator of phenotype, indicating the presence of some dichotomous trait. For each individual in the population, D_{ij} denotes what the trait would be for that individual if V_1 were i and if V_2 were j. We again let C denote a genetic marker for population substructure and suppose that control for the marker suffices to control for confounding by stratification and admixture so that $P(D_{ij}=1|C=c) \approx P(D=1|V_1=i,V_2=j,C=c)$. As in the text, we will let $p_{ijc} = P(D=1|V_1=i,V_2=j,C=c)$.

In the main text we considered tests for response patterns exhibiting epistasis such as those in Tables 2-5; tests were derived that assumed that both V_1 and V_2 had monotonic effects on D i.e. that D_{ij} was non-decreasing in i and j. Here we describe tests for such epistatic response patterns when only one or neither of the genetic factors has a monotonic effect on D.

Suppose first that only V_1 has a monotonic effect on D so that D_{ij} is non-decreasing in i (but possibly not j) for all individuals. Using arguments similar to those in the text, it can be shown that there will be individuals with C=c and with response patterns like those in Table 2 if it is the case that:

$$p_{22c} - p_{21c} - p_{20c} - p_{12c} > 0.$$

Essentially, if $p_{22c} - p_{21c} - p_{20c} - p_{12c} - p_{02c} > 0$ then there must be individuals for whom $D_{22}=1$ but for whom $D_{21}=D_{20}=D_{12}=0$. But for such individuals if $D_{21}=0$ then it must also be the case that $D_{11}=0$ and $D_{01}=0$ by the monotonicity of V_1; similarly by the monotonicity of V_1, if $D_{20}=0$ then it must also be the case that $D_{10}=0$ and $D_{00}=0$; finally, by the monotonicity of V_1, if $D_{12}=0$ then it must also be the case that $D_{02}=0$ We thus have that if $p_{22c} - p_{21c} - p_{20c} - p_{12c} > 0$ then there must be individuals for whom $D_{22}=1$ but for whom $D_{21}=D_{11}=D_{01}=D_{20}=$

$D_{10}=D_{00}=D_{12}=D_{02}=0$ i.e. for whom the phenotype response pattern is given by that in Table 2.

By similar reasoning it can be shown that if V_1 has a monotonic effect on D then there are individuals with response patterns like those in Table 3 if it is the case that:

$$p_{12c} - p_{21c} - p_{20c} - p_{02c} > 0.$$

It is not in general possible to test for epistatic response patterns like those in Tables 4 and 5 if it can only be assumed that V_1 has a monotonic effect on D. This is because it is not possible to test for individuals for whom it is the case that both $D_{22}=1$ and $D_{21}=1$ without making monotonicity assumptions about V_2.

Now consider the case in which it can be assumed that V_2 has a monotonic effect on D but for which it may not be reasonable to suppose that V_1 has a monotonic effect on D. By similar reasoning to that above it can be shown that if V_2 has a monotonic effect on D then there are individual with response patterns like those in Table 2 if it is the case that:

$$p_{22c} - p_{12c} - p_{02c} - p_{21c} > 0.$$

Likewise, if V_1 has a monotonic effect on D then there are individual with response patterns like those in Table 4 if it is the case that:

$$p_{21c} - p_{12c} - p_{02c} - p_{20c} > 0.$$

It is not in general possible to test for epistatic response patterns like those in Tables 3 and 5 if it can only be assumed that V_2 has a monotonic effect on D. This is because it is not possible to test for individuals for whom it is the case that both $D_{22}=1$ and $D_{12}=1$ without making monotonicity assumptions about V_1.

Now consider the case in which no monotonicity assumptions are made. In such settings, it is not in general possible to test for response patterns like those in Tables 3-5. One can still test for response patterns like that given in Table 2. However, to conclude that there are individuals with response patterns like those in Table 2 without making any monotonicity assumptions one would need to test:

$$p_{22c} - p_{21c} - p_{20c} - p_{12c} - p_{11c} - p_{10c} - p_{02c} - p_{01} - p_{00c} > 0.$$

If this condition were satisfied and if it could be assumed that genetic marker C suffices to control for confounding by stratification and admixture then it could be concluded that there were individuals with response patterns like those in Table 2 even without making any monotonicity assumptions.

Tests for epistasis when one factor has two levels and one factor has three

Suppose now that V_1 has two levels and V_2 has three levels indicating the genotype at loci A and B respectively (e.g. , $V_1=0$ for genotype *a/a* or *a/A* and $V_1=1$ for genotype *A/A* and $V_2=0$ for *b/b*, $V_2=1$ for *b/B*, $V_2=2$ for *B/B*). Let D denote an indicator for dichotomous trait, D_{ij} denote what the trait would be for an individual if V_1 were i and if V_2 were j, and C denote a genetic marker for population substructure. The effect of V_1 or V_2 on D is said to be monotonic if D_{ij} is non-decreasing in i or j respectively. We suppose that control for the marker suffices to control for confounding by stratification and admixture so that $P(D_{ij}=1|C=c) \approx P(D=1|V_1=i,V_2=j,C=c)$. We again let $p_{ijc} = P(D=1|V_1=i,V_2=j,C=c)$.

Epistasis, in the sense of masking, would be present if there were individuals for whom

$$D_{12}=1 \text{ but } D_{11}=0, D_{10}=0, D_{02}=0, D_{01}=0, D_{00}=0.$$

Using arguments similar to those above, if the effects of V_1 and V_2 on D are both monotonic then there are individuals with the epistatic response pattern above if

$$p_{12c} - p_{11c} - p_{02c} + p_{01c} > 0.$$

If only the effect V_1 on D can be assumed to be monotonic then there are individuals with the epistatic response pattern above if

$$p_{12c} - p_{11c} - p_{10c} - p_{02c} > 0.$$

If only the effect V_2 on D can be assumed to be monotonic then there are individuals with the epistatic response pattern above if

$$p_{12c} - p_{11c} - p_{02c} > 0.$$

If neither the effect of V_1 or V_2 can be assumed to be monotonic then there are individuals with the epistatic response pattern above if

$$p_{12c} - p_{11c} - p_{10c} - p_{02c} - p_{01c} - p_{00c} > 0.$$

Epistasis, in the sense of masking, would also be present if there were individuals for whom

$$D_{12}= D_{11}=1 \text{ but } D_{10}=0, D_{02}=0, D_{01}=0, D_{00}=0.$$

Using arguments similar to those above, if the effects of V_1 and V_2 on D are both monotonic then there are individuals with the epistatic response pattern above if

$p_{11c} - p_{10c} - p_{02c} > 0.$

If only the effect V_1 on D or if neither the effect of V_1 or V_2 can be assumed to be monotonic then it is not possible to detect this second epistatic response pattern simply using observed outcome probabilities.

Tests for epistasis in case-control studies using logistic models

Suppose that the outcome is rare so that odds ratios approximate risk ratios. Consider model (12) in the text:

logit $\{$ P(D=1|$X_1= x_1,X_2= x_2$) $\} = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_1 x_2$

Suppose it is known a priori that the main effects β_1 and β_2 are non-negative. The fact that $\beta_3 > 0$ implies condition (6) holds at least approximately and that $\beta_3 > \log(2)$ implies condition (5) holds at least approximately have been shown elsewhere (VanderWeele, 2009). To see that $\beta_3 > \log(3)$ implies condition (13), $RR_{11c} - RR_{10c} - RR_{01c} - 1 > 0$, and hence condition (4), $p_{11} - p_{10} - p_{01} - p_{00} > 0$, note that:

$RR_{11c} - RR_{10c} - RR_{01c} - 1$
$\approx OR_{11c} - OR_{10c} - OR_{01c} - 1$
$= \exp(\beta_1 + \beta_2 + \beta_3) - \exp(\beta_1) - \exp(\beta_2) - 1$
$= (1/3)\exp(\beta_1+\beta_2+\beta_3) - \exp(\beta_1) + (1/3)\exp(\beta_1+\beta_2+\beta_3) - \exp(\beta_2) + (1/3)\exp(\beta_1+\beta_2+\beta_3) - 1$
$= \exp(\beta_1)((1/3)\exp(\beta_2+\beta_3)-1) + \exp(\beta_2)((1/3)\exp(\beta_1+\beta_3)-1) + ((1/3)\exp(\beta_1+\beta_2+\beta_3)-1)$

If $\beta_3 > \log(3)$ and also the main effects β_1 and β_2 are non-negative, then each of the terms, $((1/3)\exp(\beta_2 + \beta_3) - 1)$ and $((1/3)\exp(\beta_1 + \beta_3) - 1)$ and $((1/3)\exp(\beta_1 + \beta_2 + \beta_3) - 1)$ will be positive and thus we will have that $RR_{11c} - RR_{10c} - RR_{01c} - 1 > 0$ and consequently also $p_{11} - p_{10} - p_{01} - p_{00} > 0$ i.e. condition (4) will be satisfied.

Using similar arguments and the results of VanderWeele (in press), similar relations can be obtained when one or both of the two exposures have three levels. In such cases, what VanderWeele (in press) defined as "definite interdependence" will correspond to an epistatic response pattern if at least one of the two factors have a monotonic effect on the outcome; otherwise definite interdependence is a weaker condition than an epistatic response pattern.

Suppose V_1 has two levels and V_2 has three levels and we use the regression model:

$$\text{logit } \{P(D=1|V_1=v_1,V_2=v_2)\} = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_1 x_2 + \beta_5 x_1 x_3$$

where $X_1=1$ if $V_1=1$ and 0 otherwise, $X_2=1$ if $V_2 \in \{1,2\}$ and 0 otherwise and $X_3=1$ if $V_2=2$ and 0 otherwise. Suppose further that the outcome is rare and that $P(D=1|V_1=v_1,V_2=v_2)$ is non-decreasing in v_1 and v_2. Then there are individuals with epistatic response pattern $D_{12}=1$ and $D_{11}=0$, $D_{10}=0$, $D_{02}=0$, $D_{01}=0$, $D_{00}=0$ if (i) $\beta_5>0$ and both V_1 and V_2 have monotonic effects on D or if (ii) $\beta_5>\log(2)$ and just V_2 (the factor with three levels) has a monotonic effect on D or if (iii) $\beta_5>\log(3)$ and just V_1 (the variable with two levels) has a monotonic effect on D or if (iv) $\beta_5>\log(5)$ and it is not assumed that either V_1 and V_2 have monotonic effects on D.

Similarly, it is the case that there are individuals with epistatic response pattern $D_{12}= D_{11}=1$ and $D_{10}=0$, $D_{02}=0$, $D_{01}=0$, $D_{00}=0$ if (i) $\beta_4-\beta_3>0$ and both V_1 and V_2 have monotonic effects on D or if (ii) $\beta_4-\beta_3>\log(2)$ and just V_2 has a monotonic effect on D. If only V_1 has a monotonic effect on D or neither V_1 nor V_2 have a monotonic effect on D then it is not in general possible to detect this second type of epistatic response pattern.

Suppose V_1 and V_2 have three levels and we use the regression model:

$$\text{logit } \{P(D=1|V_1=v_1,V_2=v_2)\} = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4$$
$$+ \beta_5 x_1 x_3 + \beta_6 x_1 x_4 + \beta_7 x_2 x_3 + \beta_8 x_2 x_4$$

where $X_1=1$ if $V_1 \in \{1,2\}$ and 0 otherwise and $X_2=1$ if $V_1=2$ and 0 otherwise and similarly $X_3=1$ if $V_2 \in \{1,2\}$ and 0 otherwise and $X_4=1$ if $V_2=2$ and 0 otherwise. Suppose further that the outcome is rare and that $P(D=1|V_1=v_1,V_2=v_2)$ is non-decreasing in v_1 and v_2. Then there are individuals with the epistatic response pattern of that in Table 2 if (i) $\beta_8>0$ and both V_1 and V_2 have monotonic effects on D or if (ii) $\beta_8>\log(3)$ and either V_1 or V_2 have a monotonic effect on D or if (iii) $\beta_8>\log(8)$ and it is not assumed that either V_1 and V_2 have monotonic effects on D.

There are individuals with epistatic response pattern of that in Table 3 if (i) $\beta_6-\beta_7-\beta_2>0$ and both V_1 and V_2 have monotonic effects on D or if (ii) $\beta_6-\beta_7-\beta_2>\log(3)$ and V_1 has a monotonic effect on D. There are individuals with epistatic response pattern of that in Table 4 if (i) $\beta_7-\beta_6-\beta_4>0$ and both V_1 and V_2 have monotonic effects on D or if (ii) $\beta_7-\beta_6-\beta_4>\log(3)$ and V_2 has a monotonic effect on D. There are individuals with epistatic response pattern of that in Table 5 if $\beta_5-\beta_4-\beta_2>0$ and both V_1 and V_2 have monotonic effects on D.

Alternatively, with case-control data with a rare outcome, one could test the conditions in the previous two sections of the appendix somewhat more directly. Each of the conditions in the previous two sections of the appendix could be divided by p_{00} to express the conditions in terms of risk ratios; because the outcome is rare, the risk ratios will be approximated by odds ratios which can be obtained from the logistic regression models. This approach was described in the text for two binary genetic factors but applies also to settings in which one factor has two levels and the other three or to settings in which both factors have three levels.

References

Bateson, W. *Mendel's Principles of Heredity*. Cambridge University Press, Cambridge (1909).

Chung, Y., Lee, S. Y., Elston, R. C. & Park, T. Odds ratio based multifactor-dimensionality reduction method for detecting gene–gene interactions. *Bioinformatics* **23**, 71–76 (2007).

Cordell, H. J. Epistasis: what it means, what it doesn't mean, and statistical methods to detect it in humans. *Hum. Mol. Genet.* **11**, 2463-2468 (2002).

Cordell, H. J. Detecting gene-gene interaction that underlie human diseases. *Nat. Rev. Genet.* **10**, 392-404 (2009).

Cordell, H. J. & Clayton, D. G. A unified stepwise regression procedure for evaluating the relative effects of polymorphisms within a gene using case/control or family data: application to HLA in type 1 diabetes. *Am. J. Hum. Genet.* **70**, 124–141 (2002).

Cordell, H. J., Barratt, B. J. & Clayton, D. G. Case/pseudocontrol analysis in genetic association studies: a unified framework for detection of genotype and haplotype associations, gene–gene and gene–environment interactions and parent-of-origin effects. *Genet. Epidemiol.* **26**, 167–185 (2004).

Cordell, H. J. & Clayton, D. G. Genetic epidemiology 3 - genetic association studies. *Lancet* **366**, 1121-1131 (2005).

Ferreira, T., Donnelly, P. & Marchini, J. Powerful Bayesian gene–gene interaction analysis. *Am. J. Hum. Genet.* **81** (Suppl.), 32 (2007).

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

Gayan, J. et al. A method for detecting epistasis in genome-wide studies using case–control multi-locus association analysis. *BMC Genomics* **9**, 360 (2008).

Hahn, L. W., Ritchie, M. D. & Moore, J. H. Multifactor dimensionality reduction software for detecting gene–gene and gene–environment interactions. *Bioinformatics* **19**, 376–382 (2003).

Hernán, M. A. A definition of causal effect for epidemiological studies. *J. Epidemiol. Comm. Health*, **58**, 265–271 (2004).

Hoffmann, T. J., Lange, C., Vansteelandt, S., & Laird, N. M. Gene-environment interaction tests for dichotomous traits in trios and sibships. *Genet. Epidemiol.*, (2009 Apr 13, Epub ahead of print).

Hoggart, C. J., et al. Control of confounding of genetic associations in stratified populations. *Am. J. Hum. Genet.* **72**, 1492-1504 (2003).

Khoury, M.J. & Flanders, W.D. Nontraditional epidemiologic approaches in the analysis of gene-environment interaction: case-control studies with no controls! *Epidemiol.* **144**, 207-213. (1996).

Kotti, S., Bickeboller, H. & Clerget-Darpoux, F. Strategy for detecting susceptibility genes with weak or no marginal effect. *Hum. Hered.* **63**, 85–92 (2007).

Laird, N. M. & Lange, C. Family-based designs in the age of large-scale gene-association studies. *Nat. Rev. Genet.* **7**, 385-394 (2006).

Lou, X. Y. et al. A combinatorial approach to detecting gene–gene and gene–environment interactions in family studies. *Am. J. Hum. Genet.* **83**, 457–467 (2008).

Martin, E. R., Ritchie, M. D., Hahn, L., Kang, S. & Moore, J. H. A novel method to identify gene–gene effects in nuclear families: the MDR-PDT. *Genet. Epidemiol.* **30**, 111–123 (2006).

Moore, J. H. Computational analysis of gene–gene interactions using multifactor dimensionality reduction. *Expert Rev. Mol. Diagn.* **4**, 795–803 (2004).

Phillips, P. C. Epistasis – the essential role of gene interactions in the structure and evolution of genetic systems. *Nat. Rev. Genet.* **9**, 855-867 (2008).

Piegorsch, W. W., Weinberg, C. R. & Taylor, J. A. Non-hierarchical logistic models and case-only designs for assessing susceptibility in population-based case–control studies. *Stat. Med.* **13**, 153–162 (1994).

Price A. L., et al. Principal components analysis corrects for stratification in genome-wide association studies. *Nat. Genet.* **38**:904-909 (2006).

Pritchard, J. K. & Rosenberg, N. A. Use of unlinked genetic markers to detect population stratification in association studies. *Am. J. Hum. Genet.* **65**, 220-228 (1999).

Pritchard, J. K., et al. Association mapping in structured populations. *Am. J. Hum. Genet.* **67**, 170-181 (2000).

Purcell, S. et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.* **81**, 559–575 (2007).

Richardson D. B. & Kaufman, J. S. Estimation of the relative excess risk due to interaction and associated confidence bounds. *Am. J. Epidemiol.* **169**, 756-760 (2009).

Ritchie, M. D. et al. Multifactor-dimensionality reduction reveals high-order interactions among estrogen-metabolism genes in sporadic breast cancer. *Am. J. Hum. Genet.* **69**, 138–147 (2001).

Rothman, K. J. Causes. *Am. J. Epidemiol.* **104**, 587-592 (1976).

Rothman, K. J. *Modern Epidemiology*. 1st ed. Little, Brown and Company, Boston, MA (1986).

Satten, G. A., Flanders, W.D. & Yang, Q. Accounting for unmeasured population substructure in case-control studies of genetic association using a novel latent-class model. *Am. J. Hum. Genet.* **68**, 466-477 (2001).

VanderWeele, T. J. Sufficient cause interactions and statistical interactions. *Epidemiol.* **20**, 6-13 (2009).

VanderWeele, T. J. Sufficient cause interactions for categorical and ordinal exposures with three levels. *Biometrika* (in press).

VanderWeele, T. J. & Robins, J. M. The identification of synergism in the sufficient-component cause framework. *Epidemiol*. **18**, 329-339 (2007).

VanderWeele, T. J. & Robins, J. M. Empirical and counterfactual conditions for sufficient cause interactions. *Biometrika* **95**, 49-61 (2008).

VanderWeele, T. J., Vansteelandt, S. & Robins, J. M. Marginal structural models for sufficient cause interactions. *Am. J. Epidemiol.* (in press).

Vansteelandt, S., VanderWeele, T. J., Tchetgen, E. J. & Robins, J. M. Multiply robust inference for statistical interactions. *J. Am. Statist. Assoc*., **103**, 1693–1704 (2008).

Weinberg, C. R. & Umbach, D. M. Choosing a retrospective design to assess joint genetic and environmental contributions to risk. *Am. J. Epidemiol*. **152**, 197–203 (2000).

Yang, Q., Khoury, M. J., Sun, F. & Flanders, W. D. Case-only design to measure gene–gene interaction. *Epidemiol*. **10**, 167–170 (1999).

Zhang, Y. & Liu, J. S. Bayesian inference of epistatic interactions in case–control studies. *Nat. Genet*. **39**, 1167–1173 (2007).