# Epistasis: A means or an end for complex traits?

Table of Contents

Epistasis: A means or an end for complex traits? 1

Abstract 1

Introduction 1

Methods for detecting epistasis 2

Overview of empirical evidence for epistasis influencing complex traits 3

Hypothesis-free studies 3

Hypothesis-driven studies 4

To which scientific question(s) is epistasis the answer? 5

The missing heritability? 5

Elucidating putative biological mechanisms? 6

Evolution of complex traits? 8

Improvement in genetic prediction? 8

Discussion 9

Box 1: Why is epistasis theoretically difficult to detect? 10

Box 2: What constitutes a significant epistatic interaction? 11

Glossary 11

References 11

## Abstract

Genome wide association (GWA) studies have become the focus of the statistical analysis of complex traits in humans, successfully shedding light on several aspects of genetic architecture and also biological etiology. Single nucleotide polymorphisms (SNPs) are usually modeled as having linear, cumulative, and independent effects on the phenotype. Though evidently a useful approach, it is often argued that this is not a realistic biological model and that interactions between SNPs should be included. In this review we discuss the relevance of epistasis in the context of GWA studies, recent advances in methodology, evidence of its contribution to complex traits in humans, and potential hazards in the interpretation of statistical interaction terms.

## Introduction

Complex traits or diseases are those that are influenced by multiple environmental and genetic factors. One can make the reasonable generalization that all those diseases that have a significant burden on human health are shown to be complex at the population scale. Indeed, even classically "Mendelian" diseases that tend to be very rare, such as cystic fibrosis, are at some level complex because numerous genetic effects are involved in modifying the breadth and severity of symptoms. Arguably the most important empirical result to emerge from GWA studies over the past decade is an indication of what it means for a trait to be "complex", demonstrating that the mutational target size for any particular complex trait across the genome is very large, and that the additive genetic variation is comprised of very many effects, each of very small effect sizes.

There exist several methods for estimating the proportion of the phenotypic variance of a trait that is attributed to additive genetic effects (narrow-sense heritability (*h*2)). But estimating how much of the phenotypic variation is attributed to all genetic effects (including dominance and epistasis, broad-sense heritability (*H*2)) is not possible for human traits.1 GWA studies are typically performed on traits that have been shown to have a non-zero *h*2, the logic being that there is additive genetic variation in the trait and this can be dissected into additive effects across the genome. In this context, it is already known that the additive genetic component is typically large, and its architecture is complex. Detecting non-additive genetic effects imposes an assumption that beyond the additive component of genetic variation, for which there is empirical evidence, exists more complex components whose existence cannot be verified empirically due to technical limitations. The measurable gap between the variance explained by all known additive effects detected by GWA studies and *h*2 is termed the "missing heritability". The proportion of the genetic variation that might exist beyond additive variance might be termed the "unknown heritability".

In essence, there is no strict hypothesis-driven precedent for searching for epistasis as there is for additive effects, and yet the question of epistasis in complex traits is often at the forefront of debate. In this review we will survey emerging methodology for the detection of epistasis, summarize current examples of epistasis impacting human complex traits from the literature, discuss the contexts in which epistasis may or may not be important, and suggest guidelines on what constitutes credible statistical support from empirical studies and how this should be reported.

## Methods for detecting epistasis

The past five years have seen vibrant and rapid developments of methods for studying epistasis in human complex traits/diseases1-7. The wide range of methods spans from conventional regression-based methods to nature-inspired algorithms (FIG 1). Most methods use SNP-based tests for pair-wise or high order interactions in GWAS data via either an exhaustive search of all SNP combinations or testing of a reduced, preselected set. In addition, methods have been developed to assess interactions between groups of SNPs (e.g. genes) or functional modules (e.g. pathway or network). One of the key recent achievements during this course is that an exhaustive search for interactions becomes a more routine exercise because the computational barrier has been greatly reduced8-12. Here we try to provide an overview of the many method papers published during this period in an intuitive way to help understand the big picture of the developments.

## Overview of empirical evidence for epistasis influencing complex traits

The literature is replete with reports of epistasis influencing human traits, but often the veracity of these claims is difficult to establish. A range of methods and experimental designs have been used, ranging from hypothesis-driven candidate gene testing, to hypothesis-free searches. And indeed epistasis has been reported for a wide spectrum of complex traits and diseases. Here we provide a summary of some of these findings.

### Hypothesis-free studies

The Welcome Trust Case Control Consortium (WTCCC) data has been fruitful in identifying marginal additive effects of modest size, and indeed exhaustive two-locus searches have been applied also. Wan et al2 used BOOST software to perform a search for pairwise interactions in each of the seven traits, and claimed to have identified thousands of significant interactions in total. The vast majority of interactions were between SNPs in the MHC region affecting type 1 diabetes or rheumatoid arthritis, which may be attributable to haplotype effects; and had a multiplicative effect, which may suggest a scale effect. However, some of the multiplicative interactions were between SNPs that have not previously been identified by GWAS, and this lends empirical support to the idea that searching for epistasis may confer increased statistical power to detect marginal effects. Using a different statistical model and more stringent controls for population stratification, Lippert et al3 also performed exhaustive scans for the seven diseases in WTCCC. Their results largely echoed those presented in Wan et al, but again there was no attempt at replication to verify these statistical claims.

Prabhu and Pe’er used their computationally efficient software, SIXPAC, to search for epistasis influencing bipolar disorder, and identified a pair of interacting SNPs that had not previously been shown to have an effect from GWAS. They also attempted to replicate these results, however although the interacting regions showed some evidence for replication, the actual discovery SNPs did not.

The trend that emerges is that there are hints of epistasis being uncovered through exhaustive searches, but as of yet there is no evidence that this approach detects epistatic interactions that are easily interpreted and statistically replicated. Resonating with this conclusion is a study by Lucas et al,4 which restricted the search for epistatic effects influencing risk of myocardial infarction to either SNPs that had weak marginal effects or to SNPs that had known marginal effects for a number of related traits. Though the sample size was reasonably large and the search space was drastically reduced, no statistically significant epistatic signals were uncovered. Similar conclusions have been drawn for other complex traits too, including type 2 diabetes5, BMI6 and serum uric acid levels7.

### Hypothesis-driven studies

An alternative approach to performing exhaustive searches is to overcome the problem of having a very stringent threshold by restricting the search to a few candidate loci. Using this strategy, there are some examples of epistasis with strong statistical support, and many examples with weak statistical support in the literature.

The sheer volume of reports of epistasis is exemplified by an important study by Combarros et el8, where they collated data from over 100 publications that reported epistasis of some form influencing Alzheimer’s disease (and related traits, such as age of onset etc.). These reports comprised multiple experimental designs, sample sizes, and statistical methods, but they demonstrate that, using a standardized statistical test, 27 had interaction terms at the nominal significance level of *p < 0.05* (mostly involving APOE4). Though, ostensibly, this is more than what is expected by chance, Combarros et al pointed out that there are many limitations to how these studies were conducted, notably, most of the studies did not adjust for relevant covariates, nor did they show evidence of replication.

The Epistasis Project was created for this reason, to provide a regularised framework for replication of claims of epistasis in a large, well-controlled study. To date, a few reports of epistasis from candidate gene studies have been discovered9–11 or shown to replicate12 in the Epistasis Project cohort, but perhaps the main conclusion from this work is that one must be cautious in reporting or interpreting epistasis because it appears that the majority are false positives.

Beyond the Epistasis Project other statistically robust examples of epistasis have been shown. For example, Rhinn et al13 designed a study to identify differential gene expression caused by APOE4 and independent of APOE4 in conferring a risk for late onset Alzheimer’s disease. In doing so, they demonstrated two genetic åinteractions where SNPs regulating FYN and RNF219 each decreased the risk of Alzheimer’s in APOE4 non-carriers, but not in APOE4 carriers. This finding was also statistically replicated in independent samples. The strategy of restricting the search to genetic effects that control endophenotypes that are involved in the trait of interest is an attractive idea, because it is expected that genetic effects influencing endophenotypes will be larger than those influencing higher-order traits because it is assumed that they are less polygenic in nature.

An alternative strategy for narrowing the search to overcome large significance thresholds is to only test for epistasis amongst SNPs that have known marginal effects. Though not routine, many GWA studies report that they performed follow-up analysis of epistasis amongst their hits, but nevertheless the number of positive findings remains very low. Strange et al14 looked for epistasis amongst significant marginal effects from a GWA study for psoriasis, and demonstrated that the risk alleles at the HLA-C and ERAP1 loci only conferred effects if they were both present. A similar pattern of epistasis was uncovered using the same strategy by Evans et al15 in a GWA study for ankylosing spondylitis, this time between ERAP1 and a large additive effect at HLA-B27. In both cases, the same patterns replicated in independent samples, and these are perhaps the first statistically robust examples of epistasis influencing human complex traits.

One criticism of these examples of epistasis is that they are evident on the observed scale of the disease trait, but because the interaction is between SNPs with large effects it could be the case that on the liability scale of disease the contribution to risk is purely additive. This is indeed a philosophical quandary when dealing with binary phenotypes, because on the observed scale one is implicitly using an epistatic threshold model, in the sense that an individual will not get the disease unless they have a sufficiently large number of risk alleles to lead to disease.

Given the success of identifying epistasis through endophenotypes, and by restricting the search to SNPs with known marginal effects, it is intuitive that combining these strategies should be fruitful. Becker et al16 searched for epistasis influencing gene expression, restricting the search to be for probes with known cis-acting eQTLs, and looking for interactions only between a cis-acting eQTL and all remaining trans-SNPs. Though it is unclear if any of their results surpassed a significance threshold that accounts for the experiment-wide testing burden, and no replication was attempted, it appears that there was enrichment for probes with nominally significant interactions.

Though we have only focused on a few examples here, the picture that is emerging is that compelling statistical support is absent for the vast majority of reported epistatic interactions. Potentially through some level of ascertainment bias, those epistatic effects that appear to be statistically robust have predominantly large marginal effects, and contribute relatively little in terms of phenotypic variance from interaction terms alone. At this point in time, one can hypothesise that the number of instances of epistasis may indeed be large, but the evidence currently suggests that the impact that epistasis has on complex traits is rather small.

## To which scientific question(s) is epistasis the answer?

### The missing heritability?

By definition, epistasis does not have a direct role to play in the problem of the missing heritability. The missing heritability pertains to the problem that there are insufficient additive effects discovered to account for the additive genetic variance estimated to exist for that trait. SNPs that interact can have marginal additive (and dominance) effects, but the nexus to declaring that they interact is that they have an epistatic effect that does not contribute to additive genetic variation. In principle, the total additive genetic variance accounted for by interacting SNPs can be captured by a standard, suitably powered GWA study searching for additive effects.

The metric of missing heritability emerges from two variables, the tally of genetic effects uncovered and the estimate of the trait's h2. In each case epistasis may have a subtle part to play. Firstly, some theoretical studies postulate that if, for some interacting loci, the epistatic variance is large compared to the additive variance, searching for epistatic effects might be one way to uncover loci with additive variance with more power than searching for the additive effects directly.17 Such instances might arise if the trait has a large mutational target size and is under direct selection.18 This scenario does not posit that epistatic variance contributes to additive variance, rather, searching for epistasis may lead to the identification of variants with small additive effects that would otherwise go undetected in a standard GWAS. Some empirical evidence exists to support this hypothesis,2,19 however it is unlikely to be a widespread phenomenon20 and theory shows that even in complex multi-locus epistatic patterns marginal additive effects are necessary to produce measurable additive variance for a trait.21

Secondly, because the direct estimation of non-additive genetic variation is almost always intractable, its contribution to the resemblance between relatives is unknown.1 Thus, it is possible that under certain experimental designs heritability estimates are inflated through contamination from non-additive variation.22 To what extent is this realistically a problem and how reliably can it be measured? A recent theoretical study gained much attention after demonstrating that redundancy amongst biological pathways could create an illusion of additive genetic variance in twin studies provided.23 Indeed it is a known issue in such study designs that though it is not the most parsimonious model, a combination of non-additive variation and common environmental variation can lead to a significant additive parameter,24 and this has indeed been demonstrated . To overcome this problem one can attempt to use family-based studies to estimate additive effects directly, and one can use contrasts between different types of relatives. For example, full siblings will share 0.5 additive variance and 0.25 dominance variance, while parent-offspring will share 0.5 additive and 0 dominance variance. If the correlation between degree of shared additive variation and phenotypic similarity is high across all types of relatives then this would be strong evidence for heritability estimates being uncontaminated by non-additive variance. Height shows consistent estimates of heritability between twin studies and family studies, but h2 estimates for BMI are much higher from twin studies than from family studies, suggesting that height is probably mostly influenced by additive effects and non-additive effects may play an important role in BMI.

### Elucidating putative biological mechanisms?

In principle, knowledge of genetically interacting regions could be very useful for two reasons. First, like any associations, the actual locations of the variants can shed light on the mechanisms underlying the trait by virtue of being close to relevant genes or genomic features etc, and as discussed in the previous section theory suggests that in some situations, searching for epistasis directly may improve power to detect variants with small marginal effects.18,25,26

Secondly, the pattern by which they interact genetically can also be informative. For example, suppose an interaction is detected for a disease where risk is only conferred at the first polymorphism in the presence of the risk allele at the second polymorphism (for robust empirical examples of this pattern see 13,27,28). This may signify that there is pathway redundancy, and each variant affects independent pathways. An alternative pattern to the one described above, where the risk allele at one locus only has an effect in the absence of the risk allele at another locus, might suggest that both variants are involved in the same pathway because the loss of either variant is sufficient to confer the effect of the loss of the pathway. A potential example of this was shown in an interaction for systemic lupus erythratosus, where not only did the interaction replicate, but the genes involved (BLK and BANK1) were shown to co-localise *in vivo*.29

If epistasis were easily detectable in a hypothesis-free framework, its potential to drive higher biological understanding would be very attractive. But because at this stage it appears that candidate gene approaches are often more likely to lead to the discovery of epistasis then hypothesis-free approaches, so far biological understanding has led to identifying epistasis (e.g. 30,31) more than epistasis has led to improving biological understanding.

An extension of these ideas is to deconstruct complex traits into endophenotypes32. It has been shown theoretically that if the endpoint of a metabolic pathway depends on the rate of expression of different enzymes within that pathway, then even if the rate of expression of each individual enzyme was controlled by purely additive genetic effects, then in many situations the genetic effects on the pathway outcome would appear non-additive.33 Further theoretical studies have shown that epistasis will form a large component of variation in pathway endpoints if there are negative feedback loops controlling the outcome.34 The limiting pathway model23 employs this idea on a broader scale to consider the impact of endophenotypes on the heritability of a higher level trait.

The most common example of epistatic

Nevertheless, the dearth of instances of reproducible epistatic interactions is not strong evidence for its absence, the power to detect epistasis is low and this can be mitigated by ever increasing sample sizes in GWAS. And there is theoretical precedent for epistasis contributing to

* Gjuvsland / keightley

Genetic interactions determining handedness, levy and nagylaki

### Evolution of complex traits?

The evolutionary mechanisms that lead to phenotypic variation in the population are wide ranging. It has bee proposed that one of these mechanisms is epistasis. To what extent does evidence support this notion, and what is the consequence of epistasis influencing the evolution of complex traits?

There is some empirical evidence that epistasis may play a role in phenotypic differences between species through the mechanism of genetic suppression. It has been shown that disease causing non-synonymous mutations in humans exist with no adverse phenotype in other species. For example, in humans a non-synonymous mutation causing the 53rd amino acid of alpha-synuclein from alanine to threonine predisposes to Parkinson’s disease. However, the homologous site in healthy mice and rats carries threonine, implying that there exist substitutions elsewhere that compensate for the effect. Such compensatory molecules can exist on the same molecule or on an entirely different molecule, and one study found 608 such examples of genetic suppression in a survey of 32 human proteins. It is hypothesized that epistasis allows protein evolution to traverse the fitness landscape.

The evolution of new gene function can both precipitate epistasis, and depend on epistasis. There are many examples of sequentially similar genes with crucial functional differences, but in order for the derived gene to have diverged from the ancestral gene, it is assumed that the intermediate steps will have had to have been functionally viable. Noor et al35 showed experimentally that for a pair of closely related bacterial genes that differ at only 9 amino acid positions but have different metabolic functions, a path of single amino acid changes was possible without compromising enzymatic function at any point. It was shown that epistatic effects emerged for both enzymatic functions at each of the 9 steps.

Selection acts directly on marginal additive effects, and if deleterious effects persisted under selection then theoretically this may indicate that they are involved in non-linear interactions.18 To what extent does there exist empirical evidence for this hypothesis? We cannot

* manolis paper
* Intramolecular epistasis and the evolution of a new enzymatic function.
* Hill – mutation selection balance

### Improvement in genetic prediction?

Epistasis may have a role to play in improving genomic prediction in two ways: from a

- Gustavo and Dan Gianola

A genetic interaction influencing systemic lupus erythematosus between polymorphisms in the candidate genes BANK1 and BLK was shown to replicate in independent samples, and it was also shown *in vivo* that there was a protein-protein interaction between the two genes.29 Both BANK1 and BLK have known pleiotropic effects in autoimmune diseases, and Genin et al31 showed that though the candidate SNPs did not have significant additive effects for rheumatoid arthritis, there was a significant interaction between them. This was replicated in several independent European samples, but not in a Japanese sample.

## Discussion

Scale effects, haplotype effects, linked loci, liability vs observed scale

Should we be looking for epistasis?

For:

1. They may explain missing variation

2. They may indicate the evolutionary persistence of additive effects

3. It seems to be widespread in artificial genetic variation

4. Knowing how independent regions interact statistically could shed light on biological mechanisms

Against:

1. The additive model is simple and effective

2. There are no non-additive models in animal breeding but there is extremely good progress in morphological selection

3. Detection of epistatic effects is very hard, and additive models are already underpowered

4. We expect less variance to be contributed by non-additive effects based on Taylor series expansions, narrow sense heritability estimates, etc.

============================================================================

## Box 1: Why is epistasis theoretically difficult to detect?

Higher order interactions

Curse of dimensionality / significance thresholds

Sample size

Replication - Sample LD is larger for r^8 than it is for r^2, ascertainment of high r in discovery leads to low replication in replication - need to use imputed data.

The genetic variance of complex traits is typically comprised of very many polymorphisms, each with a very small effect. Herein lies the most parsimonious explanation for the problem of the missing heritability, and informs that association studies need to increase in size dramatically in order to detect more variants. Supposing that non-additive variance is similarly comprised of numerous small effects, the statistical power to detect them is, in principle, much lower than that of detecting additive effects for a number of reasons.

- Linkage disequilibrium

The variance explained by a SNP detected in a GWAS is unlikely to be equal to the variance explained by the true causal variant that is being tagged by the marker. The additive variance at the observed marker will decrease linearly with decreasing LD r^2 between itself and the causal variant, thus if effect sizes are small then GWAS is dependent upon there being high LD between causal variants and observed SNPs. However, the decrease in dominance variance with LD r^2 is quadratic, thus the dependence on high LD between observed SNPs and unobserved causal variants is much higher when detecting dominance effects. Extending this to two loci, the additive x additive variance is linearly dependent upon sufficient LD at two independent positions (reduces quadratically with LD r^2), and dominance x dominance variance is quadratically dependent at both positions (reduces to the fourth power with LD r^2). The consequence of these constraints is that any given SNP chip has substantially greater coverage of the genome when searching for additive effects than when searching for epistatic effects. To overcome this problem one needs larger sample sizes or more dense genotyping (or sequencing) to identify non-additive effects at the same power as additive effects.

- Curse of dimensionality / significance thresholds

The search for additive effects is performed in a single dimensional search space, that is, our search is constrained to the number of markers in the experiment. In principle, searching for epistasis involves expanding from one dimension to two or more dimensions, thus the parameter space increases exponentially. This problem, where any signal becomes drowned out by the noise, is known as the "curse of dimensionality". There are several strategies that one can use to scan the genome for epistatic effects, but they typically involve expansion of the search space and a higher multiple testing penalty than is required for detecting non-epistatic effects. Therefore, in order to obtain the same power of detection as searching for additive effects, the sample size must increase and/or one depends on the epistatic variance being larger.

-

Model complexity

Replication

Higher order interactions

## Box 2: What constitutes a significant epistatic interaction?

Mustn't be tolerant of false positives just because detection is hard

Replication is necessary

## Glossary

Complex traits

Mutational target size

Additive genetic variance

Marginal effects

Endophenotype

eQTL

## References

1. Visscher, P. M., Hill, W. G. & Wray, N. R. Heritability in the genomics era--concepts and misconceptions. *Nature Reviews Genetics* **9,** 255–66 (2008).

2. Wan, X. *et al.* BOOST: A Fast Approach to Detecting Gene-Gene Interactions in Genome-wide Case-Control Studies. *American journal of human genetics* **87,** 325–340 (2010).

3. Lippert, C. *et al.* An exhaustive epistatic SNP association analysis on expanded Wellcome Trust data. *Scientific reports* **3,** 1099 (2013).

4. Subirana, I. *et al.* Hypothesis-Based Analysis of Gene-Gene Interactions and Risk of Myocardial Infarction. **7,** (2012).

5. Bell, J. T. *et al.* Genome-Wide Association Scan Allowing for Epistasis in Type 2 Diabetes. *Annals of Human Genetics* **75,** 10–19 (2011).

6. Wei, W. H. *et al.* Genome-wide analysis of epistasis in body mass index using multiple human populations. *Eur J Hum Genet* **20,** 857–862 (2012).

7. Wei, W. *et al.* Characterisation of genome-wide association epistasis signals for serum uric acid in human population isolates. *PLoS One* **6,** e23836 (2011).

8. Combarros, O., Cortina-Borja, M., Smith, A. D. & Lehmann, D. J. Epistasis in sporadic Alzheimer’s disease. *Neurobiology of Aging* **In Press, ,**

9. Kolsch, H. *et al.* Interaction of insulin and PPAR-alpha genes in Alzheimer’s disease: the Epistasis Project. *J Neural Transm* **119,** 473–479 (2012).

10. Bullock, J. M. *et al.* Discovery by the Epistasis Project of an epistatic interaction between the GSTM3 gene and the HHEX/IDE/KIF11 locus in the risk of Alzheimer’s disease. *Neurobiol Aging* **34,** 1309 e1–7 (2013).

11. Combarros, O. *et al.* The dopamine beta-hydroxylase -1021C/T polymorphism is associated with the risk of Alzheimer’s disease in the Epistasis Project. *BMC Med Genet* **11,** 162 (2010).

12. Combarros, O. *et al.* Replication by the Epistasis Project of the interaction between the genes for IL-6 and IL-10 in the risk of Alzheimer’s disease. *Journal of neuroinflammation* **6,** 22 (2009).

13. Rhinn, H. *et al.* Integrative genomics identifies APOE ε4 effectors in Alzheimer’s disease. *Nature* **500,** 45–50 (2013).

14. Strange, A. *et al.* A genome-wide association study identifies new psoriasis susceptibility loci and an interaction between HLA-C and ERAP1. *Nature Genetics* **42,** 985–90 (2010).

15. Evans, D. M. *et al.* Interaction between ERAP1 and HLA-B27 in ankylosing spondylitis implicates peptide handling in the mechanism for HLA-B27 in disease susceptibility. *Nature Genetics* **43,** (2011).

16. Becker, J., Wendland, J. R., Haenisch, B., Nothen, M. M. & Schumacher, J. A systematic eQTL study of cis-trans epistasis in 210 HapMap individuals. *Eur J Hum Genet* **20,** 97–101 (2012).

17. Verhoeven, K. J. F., Casella, G. & McIntyre, L. M. Epistasis: obstacle or advantage for mapping complex traits? *PloS one* **5,** e12264 (2010).

18. Hemani, G., Knott, S. & Haley, C. An Evolutionary Perspective on Epistasis and the Missing Heritability. *PLoS Genetics* **9,** e1003295 (2013).

19. Wei, W., Gyenesei, A., Semple, C. A. M. & Haley, C. S. Properties of Local Interactions and Their Potential Value in Complementing Genome-Wide Association Studies. *PLoS One* **8,** e71203 (2013).

20. Hill, W. G., Goddard, M. E. & Visscher, P. M. Data and Theory Point to Mainly Additive Genetic Variance for Complex Traits. *PLoS Genetics* **4,** (2008).

21. Gjuvsland, a B., Vik, J. O., Woolliams, J. a & Omholt, S. W. Order-preserving principles underlying genotype-phenotype maps ensure high additive proportions of genetic variance. *Journal of evolutionary biology* **24,** 2269–79 (2011).

22. Falconer, D. S. & Mackay, T. F. C. *Introduction to quantitative genetics*. (Longman, 1996).

23. Zuk, O., Hechter, E., Sunyaev, S. R. & Lander, E. S. The mystery of missing heritability: Genetic interactions create phantom heritability. *Proceedings of the National Academy of Sciences* (2012). doi:10.1073/pnas.1119675109

24. Evans, D. M., Gillespie, N. a & Martin, N. G. Biometrical genetics. *Biological psychology* **61,** 33–51 (2002).

25. Greene, C. S., Penrod, N. M., Williams, S. M. & Moore, J. H. Failure to Replicate a Genetic Association May Provide Important Clues About Genetic Architecture. *PLoS ONE* **4,** e5639 (2009).

26. Verhoeven, K. J., Casella, G. & McIntyre, L. M. Epistasis: obstacle or advantage for mapping complex traits? *PLoS One* **5,** e12264 (2010).

27. Strange, A. *et al.* A genome-wide association study identifies new psoriasis susceptibility loci and an interaction between HLA-C and ERAP1. *Nat Genet* **42,** 985–990 (2010).

28. Evans, D. M. *et al.* Interaction between ERAP1 and HLA-B27 in ankylosing spondylitis implicates peptide handling in the mechanism for HLA-B27 in disease susceptibility. *Nat Genet* **43,** 761–767 (2011).

29. Castillejo-López, C. *et al.* Genetic and physical interaction of the B-cell systemic lupus erythematosus-associated genes BANK1 and BLK. *Annals of the rheumatic diseases* **71,** 136–42 (2012).

30. Hill, L. D. *et al.* Epistasis between COMT and MTHFR in maternal-fetal dyads increases risk for preeclampsia. *PloS one* **6,** e16681 (2011).

31. Génin, E. *et al.* Epistatic interaction between BANK1 and BLK in rheumatoid arthritis: results from a large trans-ethnic meta-analysis. *PloS one* **8,** e61044 (2013).

32. Gottesman, I. I., Ph, D. & Gould, T. D. Reviews and Overviews The Endophenotype Concept in Psychiatry : Etymology and Strategic Intentions. 636–645 (2003).

33. Keightley, P. D. Models of quantitative variation of flux in metabolic pathways. *Genetics* **121,** 869–76 (1989).

34. Gjuvsland, A. B., Hayes, B. J., Omholt, S. W. & Carlborg, O. Statistical epistasis is a generic feature of gene regulatory networks. *Genetics* **175,** 411–20 (2007).

35. Noor, S. *et al.* Intramolecular epistasis and the evolution of a new enzymatic function. *PloS one* **7,** e39822 (2012).