

OPINION

Open Access

The genetics of infectious disease susceptibility: has the evidence for epistasis been overestimated?

Matthew D Hall* and Dieter Ebert

Abstract

Interactions amongst genes, known as epistasis, are assumed to make a substantial contribution to the genetic variation in infectious disease susceptibility, but this claim is controversial. Here, we focus on the debate surrounding the evolutionary importance of interactions between resistance loci and argue that its role in explaining overall variance in disease outcomes may have been overestimated.

Opinion

Differences amongst individuals in their susceptibility to infection seldom have a simple genetic basis and are often determined by a complex interplay of multiple loci. Characterizing the number, location and effect size of the quantitative trait loci (QTL) underlying this variation informs our understanding of not only the pathways that influence susceptibility, but also the potential coevolutionary dynamics of host and parasites. Of particular interest is the role that epistasis, defined broadly as interactions among loci in determining a phenotype, has in shaping the variation we see in infectious disease susceptibility. For most complex traits, quantitative genetic theory suggests that epistasis is unlikely to contribute substantially to genetic variation [1,2]. However, models of host-parasite co-evolution typically feature some degree of epistasis between resistance loci [3,4], and the results of empirical linkage and association mapping studies suggest that epistatic interactions can explain considerable variation in infectious disease characteristics within natural populations [5,6].

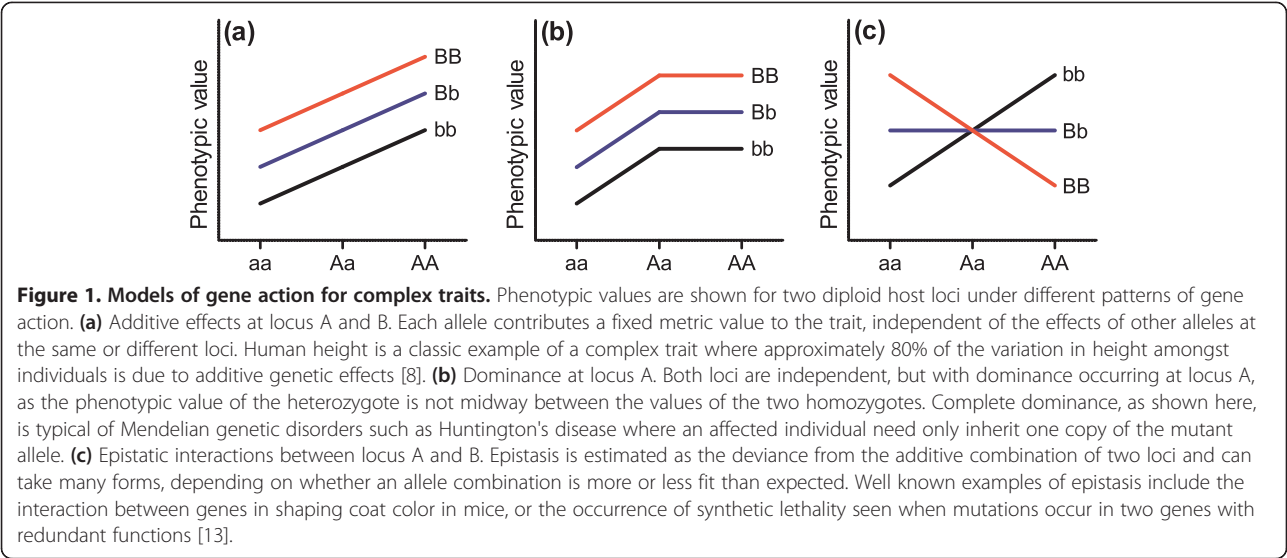
In this article, we will discuss the current state of genetic studies of disease susceptibility, with a particular

focus on the theoretical and empirical support for epistasis. We then ask if the genetic basis of infectious disease susceptibility is different from other traits, or if the evidence for epistatic interactions has been overestimated. To provide the necessary background, we first give a brief overview of the debate surrounding the contribution of epistasis to the genetic architecture of complex traits.

The genetic architecture of complex traits

A quantitative genetic understanding of complex traits is based on partitioning the variation between individuals that is due to their genotypes into additive, dominance and epistatic components [7]. Each component relates to a different form of gene action, with the additive component describing the variance associated with the independent contribution of alleles, dominance describing the variance contributed by interactions between alleles at the same locus, and epistasis referring to the contribution of interactions between alleles at different loci (Figure 1). While the relative contribution of each of these components to genetic variance depends on the underlying allele frequencies within a population, quantitative genetic theory suggests that most of the genetic variation in a population will be due to the additive effect of allelic substitutions [1]. Yet this assertion is not without controversy. Although the additive genetic basis of a trait can be readily estimated using information on the relatedness of individuals (via known breeding designs or pedigrees), only a fraction of this genetic variation has been linked to underlying loci in genome-wide association studies (GWAS; [8], but see [9]). This 'missing heritability', as it has been termed [10], together with the increasing knowledge of modifier genes and gene networks, has led to the suggestion that epistasis may make a substantial contribution to the overall levels of genetic variation [5] and that the contribution of additive variance may have even been overestimated [11].

* Correspondence: matt.hall@unibas.ch
University of Basel, Zoological Institute, Vesalgasse 1, Basel, CH-4051, Switzerland

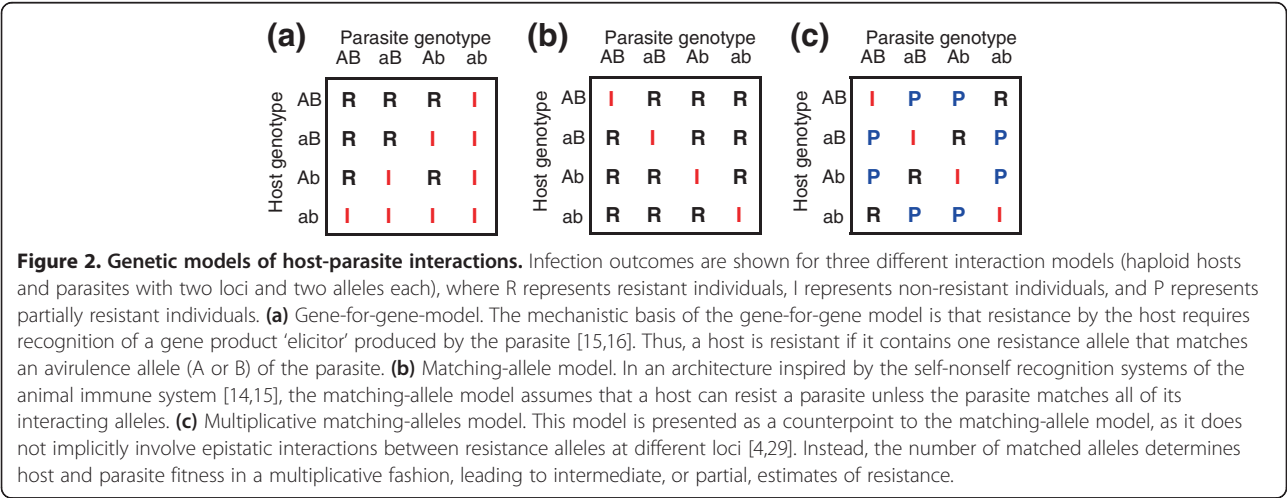


At the heart of this debate is the assumption that the importance of epistasis at the mechanistic level is reflected in the patterns of phenotypic and genetic variation at the level of the population. Within an individual, interactions between genes can result from a wide range of molecular mechanisms and can have positive or negative effects on fitness, depending on whether the resulting phenotype is greater or less than the individual effects of the alleles [12,13]. This functional impact of gene-gene interactions, however, is different from the statistical contribution of epistasis to complex trait variation within a population [13], because the latter depends on the distribution of allele frequencies in that population [7]. If most alleles are at extreme frequencies, then the majority of genetic variation should still be additive, even if there is dominance or epistasis acting at individual loci [1]. In this case, the effect of rare alleles that interact will be negligible as the likelihood that two

rare alleles are present in the same individual is very low. Thus, from a quantitative genetic perspective, it seems unlikely that epistatic interactions will contribute substantially to phenotypic variance unless the alleles are of major effect, and the frequencies of alleles involved in epistatic interactions are intermediate.

Are complex traits resulting from host-parasite interactions different?

How then does the genetic architecture of disease resistance compare to other complex traits? Based on models of host-pathogen coevolution, resistance is commonly predicted to involve multiple host genes and strong interactions between alleles at different host loci (Figure 2). Epistasis in the matching-allele class of models, for example, arises because each multi-locus parasite genotype can only infect a corresponding multi-locus host genotype [14,15]. As natural selection favors parasites that match



the most common host genotype, overrepresented host genotypes become disproportionately unfit, while rarer allele combinations now have a fitness advantage. This type of selection, whereby the fitness of a genotype decreases as its frequency increases, is known as negative frequency-dependent selection. Conversely, in the gene-for-gene model of host-pathogen interactions [16], epistasis can be incorporated via the costs associated with maintaining multiple resistance alleles (for example, [17]). Here, the costs of resistance either accelerate or decelerate with the number of contributing loci [18], aiding the maintenance of polymorphisms in resistance genes. In both cases, epistasis facilitates the rapid co-evolutionary cycles that are a hallmark of host-pathogen theory [19,20], as recombination can now break-up unfavorable allele combinations, allowing oscillations between host and parasite genotypes under negative frequency-dependent selection.

In support of such genetic models, studies of infectious disease susceptibility have commonly documented three key indicators of epistasis. Both linkage and association mapping studies have shown that variation in measures of resistance are often associated with multiple QTL of major effect, that interactions between loci contribute substantially to phenotypic variation, and that evidence for specific candidate loci is often difficult to replicate in other experiments or environments [6,21]. In a meta-analysis of over 500 QTL mapping experiments, for example, Wilfert and Schmid-Hempel [6] found that epistatic interactions were identified in 48 of 62 studies involving genome-wide scans, with most epistatic loci not previously identified using single QTL analyses (123 of 170 loci). Conventional quantitative genetic studies have also characterized the contribution of non-additive genetic components to patterns of susceptibility and resistance. Using reciprocal crosses between four populations, for example, a study of resistance in the red flour beetle (*Tribolium castaneum*) found that epistasis explained significant variation in host survival only upon infection by a parasite, and not under the unexposed and uninfected control conditions [22].

Has the evidence for epistasis been overestimated?

At first glance, the high prevalence of epistasis in mapping studies of disease traits suggests that interactions amongst host resistance genes might indeed contribute substantially to variation in disease susceptibility, as is assumed by models of host-parasite coevolution. What needs to be taken into account, however, is the estimation bias inherent in conventional mapping studies. Minor effect variants, for example, are unlikely to be identified in linkage studies due to a combination of broadly spaced markers and limited sample sizes. Conversely in association mapping, rare alleles will be difficult to detect due to the reliance on linkage disequilibrium between common markers and

common causative variants (see discussions in [23]). Gene frequencies are also altered as part of the design of traditional mapping panels, which typically involve some level of inbreeding, combined with crosses between a few individuals representing phenotypic or even population extremes. In an F2 inter-cross between high and low resistance genotypes, for example, allele frequencies are on average 0.5, even if variants within the mapping panel were at extreme frequencies in the original population. Thus, by concentrating or combining alleles of major effect within and between populations, conventional mapping studies bias allele frequencies towards intermediate values and therefore increase the chance of finding epistasis.

Without information on the effect size and frequency of alleles in natural populations, it is difficult to determine whether epistatic interactions contribute substantially to genetic variation in quantitative susceptibility, or if the contribution of such interactions to individual variation has been overestimated. The lack of success in identifying the same loci across different experiments, for example, could be due to epistasis between resistance loci, or the result of the strong sampling bias and small fraction of genetic variation that is captured using experimental crosses. Nonetheless, identified epistatic interactions can be functionally important. In a number of studies, gene-gene interactions have helped characterize the pathways underlying the mechanisms of resistance (for example, [24-27]). In the mouse, for example, epistatic interactions revealed a new mechanism for resistance to the mouse cytomegalovirus, which involves an interaction between a receptor for natural killer cells and a molecule of the major histocompatibility complex on virus-infected cells [28]. Such studies highlight the functional utility of characterizing epistasis, even if the statistical contribution of each gene-gene interaction to variation in a complex trait remains unclear.

Reconciling quantitative genetic and host-parasite theory

While epistasis is an integral component of many models of disease resistance and antagonistic coevolution (but not all [4,29]; Figure 2), the contribution of epistatic variance to susceptibility remains difficult to evaluate using conventional QTL mapping methods. With the advent of next generation sequencing approaches, however, new insights can be generated into the genetic architecture of susceptibility [23,30]. GWAS, for example, allow for the total genetic variation within a population to be decomposed into the combined effect of all loci acting additively (for example, [9]). The remaining, unexplained genetic variation, therefore, gives an upper limit for how much epistasis could potentially contribute to variation in infectious disease [1]. Observed allele frequencies and effect-size parameters can also be estimated for a range of susceptibility loci (sensu

[31]), and then compared to the expected intermediate allele frequencies predicted by different models of host-parasite coevolution. Yet, higher marker densities and GWAS do not completely resolve the contribution of specific gene-gene interactions to trait variation. Pairwise epistatic interactions are difficult to evaluate using the hundreds or thousands of markers required for conventional QTL mapping studies, let alone using the millions of markers required for GWAS.

Even in model systems where next generation sequencing approaches have been used extensively, we are far from a general understanding of the underlying architecture of resistance and susceptibility. In *Drosophila*, for example, considerable progress has been made in identifying loci underlying resistance to sigma virus transmission and verifying the importance of the resistance alleles in natural populations [32-35]. Yet resistance genes, such as *ref(2)P*, are often strain specific and do not completely account for the genetic variance underlying resistance to multiple virus isolates [36]. Similarly, despite the wide range of infectious diseases that have been studied in humans using high-density genetic maps [37,38], debate is still ongoing as to the distribution of allelic variants, and whether the genetic basis of susceptibility is based on high frequency common variants or the cumulative effects of many rare mutations [30]. As such, we suggest that two key aspects of host-parasite biology will need more consideration as we move forward in the genomics era: first, that our understanding of the genetics of susceptibility will depend on the number of parasite genotypes included in association studies; and second, that the expectation of epistasis may not be appropriate for all measures of resistance.

How we account for the natural genetic diversity of parasites will strongly influence our understanding of the genetic architecture of susceptibility. If the causal parasite is unknown or resistance is assessed using a mix of parasite genotypes, then mechanisms unrelated to resistance could be contributing to variation in infectious disease. Competition between multiple parasite genotypes within the host [39,40] and variation in dose-dependent effects across isolates [41] are all processes that would bias infection estimates. Conversely, if only a single pathogen genotype is used in a mapping study, then the relevance of any candidate loci is difficult to extend beyond the response of the host to that specific genotype. Indeed, where multiple strains of a parasite have been utilized within a mapping or association study, the results suggest that only a subset of identified QTL will confer resistance to all genotypes [6]. A study exploring the association between mosquito immune genes and infection by *Plasmodium falciparum*, for example, revealed that certain candidate loci explained patterns of resistance only for specific parasite isolates [42]. These findings highlight the need to account for the contribution of parasite genetic variation to variation in host

susceptibility, otherwise the genetic architecture of disease susceptibility will be misrepresented.

Careful consideration of the trait used to characterize resistance will also be important for mapping studies. Phenotypes of resistance range from infection rates and parasite loads, through to symptoms of disease such as morbidity and mortality. Underlying each of these measures will be a range of processes involving the ability of a pathogen to penetrate the host, the recognition of parasite proteins by the host, and the subsequent immune response facilitating pathogen replication [43,44]. Thus, the type of trait used to estimate resistance and the timing of a phenotypic assay (early or late in the infection process) could significantly influence the characterization of phenotypic and genetic variance. Estimating resistance based on symptoms of disease, for example, may more closely match classical quantitative genetic theory, whereas the initial ability of a parasite to penetrate a cell or tissue is a better fit for models of host resistance where epistasis features strongly. Indeed, initial infectivity in plants is often highly specific to certain host-parasite combinations, suggesting that susceptibility/resistance may be under control of a few major genes [45]. Although such insights are uncommon in animals, studies are beginning to reveal that initial resistance to certain pathogens may follow a similar pattern [46], with subsequent symptoms of disease being more quantitative [47].

In summary, the contribution of epistasis to phenotypic and genetic variation is a complex issue for studies of host-parasite interactions. Unlike other quantitative traits, where theory points to the largely additive contribution to genetic variation [1], epistasis is a key component of many models of host-parasite interactions. As such, host-parasite research has focused on characterizing epistasis between resistance loci, rather than debating and evaluating the relative contributions of additive and epistatic genetic effects to phenotypic variance. Nonetheless, as more studies characterize the allelic variants underlying quantitative susceptibility in natural populations, the opportunity to reassess the importance of epistasis will help redefine how empirical and theoretical research approaches the genetic architecture of host-parasite interactions.

Acknowledgements

This review was supported by an EU Marie Curie Incoming International Fellowship (PIIF-GA-2009-252417) to MDH and by the Swiss National Science Foundation. We thank three anonymous reviewers, Luc F Bussi re, Matthew C Tinsley and members of the Ebert group for comments on the manuscript.

Published: 15 July 2013

References

1. Hill WG, Goddard ME, Visscher PM: Data and theory point to mainly additive genetic variance for complex traits. *PLoS Genet* 2008, **4**:e1000008.

2. Crow JF: **On epistasis: why it is unimportant in polygenic directional selection.** *Philos Trans R Soc B* 2010, **365**:1241–1244.
3. Peters AD, Lively CM: **Epistasis and the maintenance of sex.** In *Epistasis and the Evolutionary Process*. Edited by Wolf JB, Brodie ED III, Wade MJ. Oxford, UK: Oxford University Press; 2000:99–112.
4. Otto SP, Nuismer SL: **Species interactions and the evolution of sex.** *Science* 2004, **304**:1018–1020.
5. Carlborg O, Haley CS: **Epistasis: too often neglected in complex trait studies?** *Nat Rev Genet* 2004, **5**:618–625.
6. Wilfert L, Schmid-Hempel P: **The genetic architecture of susceptibility to parasites.** *BMCE vol Biol* 2008, **8**:187.
7. Falconer DS, Mackay TFC: *Introduction to Quantitative Genetics*. 4th edition. Harlow, UK: Pearson Education; 1996.
8. Visscher PM: **Sizing up human height variation.** *Nat Genet* 2008, **40**:489–490.
9. Yang J, Benyamin B, McEvoy BP, Gordon S, Henders AK, Nyholt DR, Madden PA, Heath AC, Martin NG, Montgomery GW, Goddard ME, Visscher PM: **Common SNPs explain a large proportion of the heritability for human height.** *Nat Genet* 2010, **42**:565–569.
10. Maher B: **Personal genomes: the case of the missing heritability.** *Nature* 2008, **456**:18–21.
11. Zuk O, Hechter E, Sunyaev SR, Lander ES: **The mystery of missing heritability: genetic interactions create phantom heritability.** *Proc Natl Acad Sci USA* 2012, **109**:1193–1198.
12. Lehner B: **Molecular mechanisms of epistasis within and between genes.** *Trends Genet* 2011, **27**:323–331.
13. Phillips PC: **Epistasis - the essential role of gene interactions in the structure and evolution of genetic systems.** *Nat Rev Genet* 2008, **9**:855–867.
14. Hamilton WD: **Sex versus non-sex versus parasite.** *Oikos* 1980, **35**:282–290.
15. Frank SA: **Recognition and polymorphism in host-parasite genetics.** *Philos Trans R Soc B* 1994, **346**:283–293.
16. Flor HH: **The complementary genic systems in Flax and Flax Rust.** *Adv Genet* 1956, **8**:29–54.
17. Sasaki A: **Host-parasite coevolution in a multilocus gene-for-gene system.** *Proc R Soc B* 2000, **267**:2183–2188.
18. Fenton A, Brockhurst MA: **Epistatic interactions alter dynamics of multilocus gene-for-gene coevolution.** *PLoS One* 2007, **2**:e1156.
19. Peters AD, Lively CM: **The Red Queen and fluctuating epistasis: a population genetic analysis of antagonistic coevolution.** *Am Nat* 1999, **154**:393–405.
20. Otto SP, Michalakis Y: **The evolution of recombination in changing environments.** *Trends Ecol Evol* 1998, **13**:145–151.
21. Kover PX, Caicedo AL: **The genetic architecture of disease resistance in plants and the maintenance of recombination by parasites.** *Mol Ecol* 2001, **10**:1–16.
22. Wegner KM, Berenos C, Schmid-Hempel P: **Nonadditive genetic components in resistance of the red flour beetle *Tribolium castaneum* against parasite infection.** *Evolution* 2008, **62**:2381–2392.
23. Manolio TA, Collins FS, Cox NJ, Goldstein DB, Hindorf LA, Hunter DJ, McCarthy MI, Ramos EM, Cardon LR, Chakravarti A, Cho JH, Guttman AE, Kong A, Kruglyak L, Mardis E, Rotimi CN, Slatkin M, Valle D, Whittemore AS, Boehnke M, Clark AG, Eichler EE, Gibson G, Haines JL, Mackay TFC, McCarrill SA, Visscher PM: **Finding the missing heritability of complex diseases.** *Nature* 2009, **461**:747–753.
24. Lazzaro BP, Scurman BK, Clark AG: **Genetic basis of natural variation in *D. melanogaster* antibacterial immunity.** *Science* 2004, **303**:1873–1876.
25. Williams TN, Mwangi TW, Wambua S, Peto TEA, Weatherall DJ, Gupta S, Recker M, Penman BS, Uyoga S, Macharia A, Mwacharo JK, Snow RW, Marsh K: **Negative epistasis between the malaria-protective effects of alpha+ thalassemia and the sickle cell trait.** *Nat Genet* 2005, **37**:1253–1257.
26. Martin MP, Qi Y, Gao X, Yamada E, Martin JN, Pereyra F, Colombo S, Brown EE, Shupert WL, Phair J, Goedert JJ, Buchbinder S, Kirk GD, Telenti A, Connors M, O'Brien SJ, Walker BD, Parham P, Deeks SG, McVicar DW, Carrington M: **Innate partnership of HLA-B and KIR3DL1 subtypes against HIV-1.** *Nat Genet* 2007, **39**:733–740.
27. Bombliks K, Lempe J, Eppele P, Warthmann N, Lanz C, Dangel JL, Weigel D: **Autoimmune response as a mechanism for a Dobzhansky-Muller-type incompatibility syndrome in plants.** *PLoS Biol* 2007, **5**:e236.
28. Desrosiers M-P, Kielczewska A, Loredi-Osti J-C, Adam SG, Makrigiannis AP, Lemieux S, Pham T, Lodoen MB, Morgan K, Lanier LL, Vidal SM: **Epistasis between mouse Klr and major histocompatibility complex class I loci is associated with a new mechanism of natural killer cell-mediated innate resistance to cytomegalovirus infection.** *Nat Genet* 2005, **37**:593–599.
29. Kouyos RD, Salathé M, Otto SP, Bonhoeffer S: **The role of epistasis on the evolution of recombination in host-parasite coevolution.** *Theor Popul Biol* 2009, **75**:1–13.
30. Hill AVS: **Evolution, revolution and heresy in the genetics of infectious disease susceptibility.** *Philos Trans R Soc B* 2012, **367**:840–849.
31. Park J-H, Gail MH, Weinberg CR, Carroll RJ, Chung CC, Wang Z, Chanock SJ, Fraumeni JF, Chatterjee N: **Distribution of allele frequencies and effect sizes and their interrelationships for common genetic susceptibility variants.** *Proc Natl Acad Sci USA* 2011, **108**:18026–18031.
32. Magwire MM, Bayer F, Webster CL, Cao C, Jiggins FM: **Successive increases in the resistance of *Drosophila* to viral infection through a transposon insertion followed by a duplication.** *PLoS Genet* 2011, **7**:e1002337.
33. Bangham J, Kim K-W, Webster CL, Jiggins FM: **Genetic variation affecting host-parasite interactions: different genes affect different aspects of sigma virus replication and transmission in *Drosophila melanogaster*.** *Genetics* 2008, **178**:2191–2199.
34. Bangham J, Obbard DJ, Kim K-W, Haddrell PR, Jiggins FM: **The age and evolution of an antiviral resistance mutation in *Drosophila melanogaster*.** *Proc R Soc B* 2007, **274**:2027–2034.
35. Wilfert L, Jiggins FM: **Disease association mapping in *Drosophila* can be replicated in the wild.** *Biol Lett* 2010, **6**:666–668.
36. Carpenter JA, Hadfield JD, Bangham J, Jiggins FM: **Specific interactions between host and parasite genotypes do not act as a constraint on the evolution of antiviral resistance in *Drosophila*.** *Evolution* 2012, **66**:1114–1125.
37. Hill AVS: **Aspects of genetic susceptibility to human infectious diseases.** *Annu Rev Genet* 2006, **40**:469–486.
38. Khor CC, Hibberd ML: **Host-pathogen interactions revealed by human genome-wide surveys.** *Trends Genet* 2012, **28**:233–243.
39. Wegner KM, Berenos C, Schmid-Hempel P: **Host genetic architecture in single and multiple infections.** *J Evol Biol* 2009, **22**:396–404.
40. Ben-Ami F, Mouton L, Ebert D: **The effects of multiple infections on the expression and evolution of virulence in a *Daphnia*-endoparasite system.** *Evolution* 2008, **62**:1700–1711.
41. Ben-Ami F, Ebert D, Regoes RR: **Pathogen dose infectivity curves as a method to analyze the distribution of host susceptibility: a quantitative assessment of maternal effects after food stress and pathogen exposure.** *Am Nat* 2010, **175**:106–115.
42. Harris C, Lambrechts L, Rousset F, Abate L, Nsango SE, Fontenille D, Morlais I, Cohuet A: **Polymorphisms in *Anopheles gambiae* immune genes associated with natural resistance to *Plasmodium falciparum*.** *PLoS Pathog* 2010, **6**:e1001112.
43. Schmid-Hempel P: **Parasite immune evasion: a momentous molecular war.** *Trends Ecol Evol* 2008, **23**:318–326.
44. Frank SA, Schmid-Hempel P: **Mechanisms of pathogenesis and the evolution of parasite virulence.** *J Evol Biol* 2008, **21**:396–404.
45. Thompson JN, Burdon JJ: **Gene-for-gene coevolution between plants and parasites.** *Nature* 1992, **360**:121–125.
46. Duneau D, Luijckx P, Ben-Ami F, Laforché C, Ebert D: **Resolving the infection process reveals striking differences in the contribution of environment, genetics and phylogeny to host-parasite interactions.** *BMC Biol* 2011, **9**:11.
47. Hall MD, Ebert D: **Disentangling the influence of parasite genotype, host genotype and maternal environment on different stages of bacterial infection in *Daphnia magna*.** *Proc R Soc B* 2012, **279**:3176–3183.

doi:10.1186/1741-7007-11-79

Cite this article as: Hall MD and Ebert D: The genetics of infectious disease susceptibility: has the evidence for epistasis been overestimated?. *BMC Biology* 2013 **11**:79.