

Review Article

Maternal-Zygotic Epistasis and the Evolution of Genetic Diseases

Nicholas K. Priest¹ and Michael J. Wade²

¹Department of Biology & Biochemistry, University of Bath, Bath BA2 7AY, UK

²Department of Biology, Indiana University, Bloomington, IN 47401, USA

Correspondence should be addressed to Michael J. Wade, mjwade@indiana.edu

Received 10 August 2009; Revised 25 November 2009; Accepted 19 February 2010

Academic Editor: Janet Sinsheimer

Copyright © 2010 N. K. Priest and M. J. Wade. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Many birth defects and genetic diseases are expressed in individuals that do not carry the disease causing alleles. Genetic diseases observed in offspring can be caused by gene expression in mothers and by interactions between gene expression in mothers and offspring. It is not clear whether the underlying pattern of gene expression (maternal versus offspring) affects the incidence of genetic disease. Here we develop a 2-locus population genetic model with epistatic interactions between a maternal gene and a zygotic gene to address this question. We show that maternal effect genes that affect disease susceptibility in offspring persist longer and at higher frequencies in a population than offspring genes with the same effects. We find that specific forms of maternal-zygotic epistasis can maintain disease causing alleles at high frequencies over a range of plausible values. Our findings suggest that the strength and form of epistasis and the underlying pattern of gene expression may greatly influence the prevalence of human genetic diseases.

1. Introduction

Why are some human genetic diseases so common? We expect natural selection to favour the alleles of genes that confer health and hinder alleles of genes that confer disease. Yet, there are a glut of genetically based human diseases, behavioural syndromes, and birth defects that occur at relatively high frequencies [1]. The classic explanation is that diseases are maintained in balance between mutation, which creates the disease, and purifying selection, which removes it [2]. Though this theory is roughly supported by diseases with simple genetic inheritance, it is not designed to apply for diseases with complex patterns of inheritance [3, 4].

Most diseases follow a complex pattern of inheritance [4–6]. One form of complex inheritance occurs when the genetic makeup of mothers influences the disease status of their offspring. This can either happen through maternal genetic effects, whereby the genotype of mothers influences trait expression in offspring, or through maternal-zygotic epistasis, whereby the phenotype of an offspring is dependent on the interaction between genes in the maternal genome

and genes in the offspring genome [7, 8]. It can be viewed as kind of genotype \times environment interaction, in which the maternal genome provides the environment for the offspring genotype. However, unlike standard discussions of genotype \times environment interaction where the environment is fixed, the maternal environment is genetic and can evolve. The most detailed evidence for this type of inheritance comes from work in plant and animal systems [8–15], but there is also direct evidence from humans.

We do not know the full extent to which maternal genetic effects and maternal-zygotic epistasis contribute to human diseases. Our ability to detect complex inheritance is hampered by a range of factors, particularly sampling limitations, population structure, and environmental influences on disease [16–21]. Our ability to detect maternal genetic sources of inheritance is further hampered by the multigenerational complexity of the data sets required to test for it [8, 22–24]. Despite these limitations, there is emerging evidence from association studies that maternal genetic effects and maternal-zygotic epistasis might be common mechanisms of inheritance of human diseases and birth defects [16, 25–31].

The mechanism of inheritance is important to understand because it may affect the prevalence of disease. For simple genetic diseases, prevalence is determined by the mutation-selection balance [2]. However, in maternally expressed diseases half of the alleles are hidden from selection, as the genotype of fathers has no effect on offspring phenotype [32–35]. This means that purifying selection might only be half as effective at removing diseases that result from maternal gene expression. Thus, given the same mutation rate, we might expect maternal genetic diseases to have a substantially higher prevalence than zygotic genetic diseases (diseases caused by gene expression in offspring). The relaxed selective constraint on maternal genetic effects is supported by studies of population variation and diversification of maternal effect genes [34–36].

Epistasis might also contribute to the prevalence of genetic disease. Epistasis is known to confound disease gene mapping [5, 33, 37] because alleles at one locus can mask the effects of expression of alleles at another locus [38–41]. Thus, under specific forms of gene interaction, even genes that confer death in one genetic background might be maintained at higher frequencies than we would expect under traditional, mutation/selection balance, theories of disease evolution.

Previous theories have described general models of maternal genetic effects and maternal-zygotic epistasis in trait expression [7, 8, 34, 36]. However, there are many unresolved questions about their roles in genetic diseases. Do maternal genetic effects increase the incidence of genetic diseases? How do epistatic interactions between maternal and offspring genomes affect disease susceptibility? And, can models of the evolution of complex diseases assist us in identifying the genes which underlie complex diseases?

Here, we address these questions from the perspective of evolutionary population genetics. We present results from population genetic simulations of the evolution of genes affecting viability, whether owing to birth defects or adult diseases. These results are the first to demonstrate that genes that reduce fitness are more likely to be maintained in populations over time and at higher equilibrium frequencies when they are expressed in mothers than when they are expressed in offspring. The results also reveal the specific forms of maternal-fetal epistasis which increase the incidence of alleles with direct and deleterious effects on fetal survival. We show that a gene with a deleterious main effect can become fixed and mean fitness can increase because fixing a genetic background with a strong positive interaction effect may add more to mean fitness than the deleterious main effect removes. We discuss the implications of these findings for how genetic diseases evolve and for detecting genes associated with disease.

2. Materials and Methods

2.1. Statistical Definition of Diseases with Simple Genetic Inheritance. Before we can model the evolution of genetic diseases, we need to develop an appropriate decomposition of genetic effects. From the perspective of population

genetics, a diseased individual has a genetic makeup that encodes lower offspring production (lower fitness) than the average individual in the population. Simple genetic diseases can be described by the additive and dominance effects that their alleles encode. The additive effect of a disease allele is half the fitness difference between the homozygote classes. The dominance effect of an allele is the deviation between the observed fitness of heterozygous individuals and the fitness midpoint between the homozygous classes. For example, the following are the genotypic values of a hypothetical gene “O”, with alleles O and o , which causes lethal birth defects in all homozygous recessive offspring: OO Fitness = 1, Oo fitness = 1, oo fitness = 0. The additive and dominance values for the genotypes in this scenario are as follows: additive value for o allele = -0.5 so that the fitness of the oo genotype equals $1 + 2(-0.5)$; dominance value for the Oo genotype = $+0.5$, so that its fitness equals $1 + (+.50) + 1(-0.5)$. In our simulations, we equate the viability of an offspring genotype with its genotypic value. The following are the genotypic values when the hypothetical gene “O” causes lethal birth defects in 20% of homozygous recessive offspring: OO Fitness = 1, Oo fitness = 1, and oo fitness = 0.8. The additive and dominance values for this scenario are as follows: additive value = -0.1 ; dominance value = $+0.1$.

2.2. Statistical Definition of Diseases with Gene-Gene Interactions. For simple genetic diseases, we can calculate the probability that a given genotype will express the disease from the additive and dominance effects. However, when there are gene-by-gene interactions then the fitness of a given genotype is dependent on the sum of the additive and dominance values plus the effect of the interaction. There are many different classes of epistasis [8, 38–41]. We chose to focus on diploid statistical epistasis, rather than other kinds of gene interactions, because they can be estimated from genetically based epidemiological data and they describe the statistical effect of gene interactions on phenotypes in a similar manner to simple genetic effects [42]. Our goal is to illuminate the evolutionary process in terms of such statistically detectable gene effects.

Complex genetic diseases can be described by additive and dominance effects at one locus, additive and dominance effects at another locus, and all possible two-way interactions among genetic effects. In total, the additive and epistatic effects of a pair of genes (hypothetical genes “A” and “B”) on phenotypes can be decomposed into eight separate orthogonal components (see Table 1). For the sake of clarity, α means additive genetics at the “A” locus, acting on genotypes AA and aa ; D_A means dominance at the “A” locus, acting on the genotype Aa ; β means additive genetics at the “B” locus, acting on genotypes BB and bb ; D_B means dominance at the “B” locus, acting on the genotype Bb ; I_{AB} means interaction between additive alleles at “A” and “B” loci, acting on $AABB$, $aabb$, $aaBB$, and $AAbb$ genotypes; K_{AaB} means interaction between the Aa genotype with BB and bb genotypes; K_{ABb} means interaction between the Bb genotype with AA and aa genotypes; J_{AaBb} means interaction between Aa and Bb genotypes. For example, in Table 1, the fitness of

TABLE 1: The genotypic values of two zygotically expressed interacting genes. Each gene has two alleles with genetic effects decomposed into eight orthogonal components: additive (α and β , dominance (D_A and D_B), additive by additive (I_{AB}), additive by dominance (K_{ABb}), dominance by additive (K_{AaB}), and dominance by dominance (J_{AaBb}).

Genotypes	BB	Bb	bb
AA	$1 + \alpha + \beta + I_{AB}$	$1 + \alpha + D_B + K_{ABb}$	$1 + \alpha - \beta - I_{AB}$
Aa	$1 + \beta + D_A + K_{AaB}$	$1 + D_A + D_B + J_{AaBb}$	$1 - \beta + D_A - K_{AaB}$
aa	$1 - \alpha + \beta - I_{AB}$	$1 - \alpha + D_B - K_{ABb}$	$1 - \alpha - \beta + I_{AB}$

TABLE 2: The genotypic fitness values for a pair of maternally and zygotically expressed interacting genes expressed as deviations to be added to 1. Each gene has two alleles with genetic effects decomposed into eight orthogonal components: zygotic additive (ϕ) and zygotic dominant (D_O), maternal additive (κ) and maternal dominant (D_M), zygotic additive by maternal additive (I_{OM}), zygotic dominant by maternal additive (K_{OoM}), zygotic additive by maternal dominant (K_{OMm}), and zygotic dominant by maternal dominant (J_{OoMm}).

Maternal Genotypes	Offspring Genotypes		
	OO	Oo	oo
MM	$1 + \phi + \kappa + I_{OM}$	$1 + \kappa + D_O + K_{OoM}$	$1 - \phi + \kappa - I_{OM}$
Mm	$1 + \phi + D_M + K_{OMm}$	$1 + D_O + D_M + J_{OoMm}$	$1 - \phi + D_M - K_{OMm}$
mm	$1 + \phi - \kappa - I_{OM}$	$1 - \kappa + D_O - K_{OoM}$	$1 - \phi - \kappa + I_{OM}$

an individual of genotype AABB is the sum of the additive effects of the “A” and “B” loci (α and β) and their additive by additive interaction (I_{AB}).

2.3. Statistical Definition of Diseases with Maternal-Zygotic Interactions. Traditionally, epistasis has been considered for gene combinations within the same individual or genome ($g \times g$ epistasis). However, genes expressed in one individual can also interact with genes expressed in another [8, 32]. The best described form of such between-genome epistasis (often described as $G \times G$ epistasis; here we use the term maternal-zygotic epistasis to describe offspring genotype x mother genotype epistasis) is maternal-zygotic epistasis, which occurs when trait expression in offspring is determined by interactions between genes expressed in a mother and genes expressed in her offspring. Maternal genetic effects and maternal-zygotic epistasis can be described by maternally expressed additive and dominance effects at one locus, additive and dominance effects expressed in offspring, and all possible two-way interactions among genetic effects. In total, the additive and epistatic effects of a pair of hypothetical gene expressed in mothers and offspring (“M” and “O”) that determine diseases in offspring can be decomposed into eight separate components (Table 2). For the sake of clarity, κ means additive genetics at the “M” locus, acting on the offspring of MM and mm mothers; D_M means dominance at the “M” locus, acting on the offspring of Mm mothers; ϕ means additive genetic effect at the “O” locus, acting on the OO and oo offspring; D_O means dominance at the “O” locus, acting on the Oo offspring. I_{MO} means interaction between additive effects of the “M” and “O” loci, acting on OO and oo offspring of MM mothers, and OO and oo offspring of mm mothers; K_{MmB} means interaction between the OO and oo offspring of Mm mothers; K_{MOo} means interaction between the Oo offspring of MM and mm mothers; J_{MmOo} means interaction between Oo offspring of Mm mothers.

Although they look similar, the fitness calculations are more complicated in transgenerational genetic models (Table 2) than physiological epistatic models (Table 1). In transgenerational genetic models the fitness of an offspring depends its mother’s genotype in the previous generation. Thus, two offspring of the same genotype but with different mothers can have different viabilities. For example, in Table 2, the fitness of OoMM offspring from OoMM and OoMm mothers is the sum of the additive effects of the “O” locus (ϕ), additive effects of the “M” locus (κ) and their zygotic additive by maternal additive interaction (I_{OM}). In contrast, the fitness of OoMM offspring from OoMm and OoMm mothers is the sum of the additive effects of the “O” locus (ϕ), dominance effects of the “M” locus (D_M) and the zygotic additive by maternal dominance interaction (K_{OMm}). It is this conditioning on maternal genotype that gives maternal-zygotic gene interaction its unique evolutionary properties [7, 8].

2.4. The Model. Although the interplay between the maternal and zygotic genomes probably involves complex interactions between thousands of genes and gene products, we can begin to understand the fundamental nature of these interactions by simulating maternal-zygotic interactions under greatly simplified conditions. We consider two unlinked loci where one locus, “M,” with two alleles, M and m , is expressed in mothers and influences the fitness of offspring. A second locus, “O,” with two alleles, O and o , is expressed in offspring and influences the fitness of offspring. In our simulations, we also assume for simplicity standard population genetic assumptions, which include that there are diploid males and females in a population of infinite size that mate randomly and produce offspring via sexual reproduction. We assumed that there was no inbreeding depression and that fitnesses were equivalent for males and females. The model also assumes that fitness is strictly determined by the expression

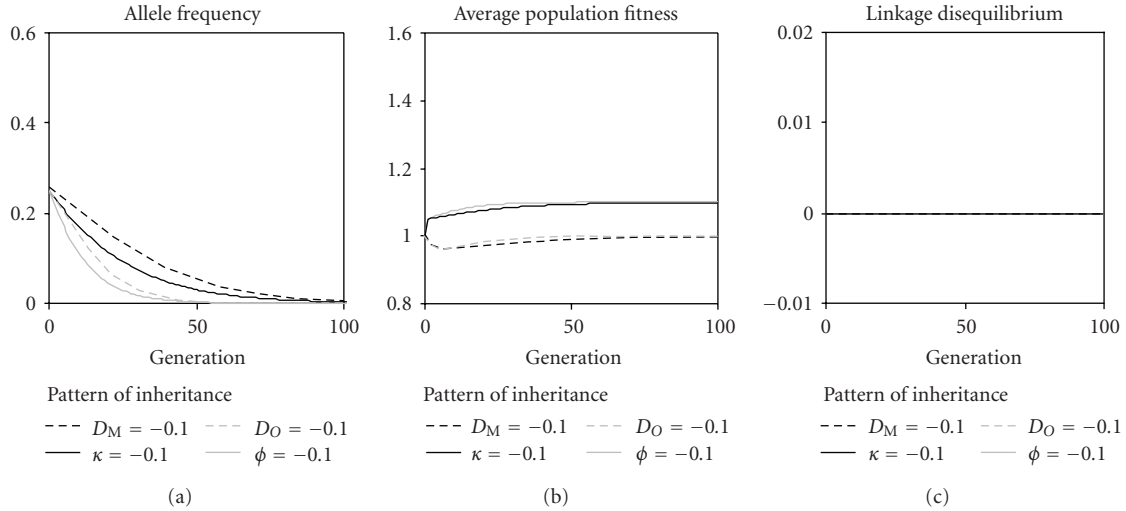


FIGURE 1: The evolution of genetic diseases in the absence of epistasis. The simulations track the evolution of each allele depending on its pattern of inheritance: κ , additive effect at a maternal effect locus; D_M , dominance effect at a maternal effect locus; ϕ , additive effect at a zygotic effect locus; D_O , dominance effect at a zygotic effect locus. Purifying selection is assumed to be -0.1 in all cases.

of alleles and by interactions between alleles at the “M” and “O” loci. The statistical genetic parameterization of epistasis has additional underlying genetic assumptions [42]. Fitness is assigned to offspring after they are born, but before they reproduce.

In every simulation, we set the initial starting frequency of the m and o alleles at 0.25 at two-locus Hardy-Weinberg equilibrium. Because the strength and direction of epistatic effects depends on allele frequency [37, 43] restricting the analysis to an initial allele frequency to 0.25 for both alleles limits general inferences from our simulations; however, previous models have shown that over a large range of initial allele frequencies (between 0 and 0.5) the evolutionary dynamics of maternal-zygotic interactions are determined by the relative strengths of direct selection and epistasis [43]. Our goal was to uncover some of the dynamics of how deleterious genes are maintained in populations under reasonable values of selection and epistasis, not to characterize all dynamics across all allele frequencies, selection coefficients and epistatic interactions. We chose our initial allele frequency because at a frequency of 0.25 it is easy to visualize both increases and decreases in allele frequency.

We undertook simulations to understand the evolution of genes and gene combinations which reduce fitness, our definition of genes that encode birth defects and genetic diseases. The simulations operate by measuring how the frequency of each genotype changes over 100 generations when we assign values of fitness to the sets of the parameters described in Table 2, which are the genetic effects of and epistatic interactions between the “M” and “O” genes. For each generation we also calculated the average fitness of the population and the linkage disequilibrium between the “M” and “O” loci. Linkage disequilibrium (LD), also referred to as gametic phase disequilibrium, is the nonrandom association of alleles at different loci into gametes [44, 45]. In our simulations, a positive LD means that there are a greater

proportion of MO and mo gametes in the population than Mo and mO gametes. Conversely, a negative LD indicates that there are a greater proportion of Mo and mO gametes in the population than MO and mo gametes.

We conducted a total of 56 simulations of the evolution of genetic diseases under simple and complex patterns of inheritance. The following is a list of the assigned parameters used in each simulation. Purifying selection: $\kappa = -0.1$, $\phi = -0.1$, $D_M = -0.1$, and $D_O = -0.1$, each in separate simulations (Figure 1). Maternal additive by zygotic additive epistasis with and without purifying selection: $I_{MO} = -0.4, 0$, and 0.4 with $\phi = 0$ and $\kappa = 0$; $I_{MO} = -0.4, -0.2, 0, 0.2$, and 0.4 with $\phi = -0.1$ and $\kappa = 0$; $I_{MO} = -0.4, -0.2, 0, 0.2$, and 0.4 with $\phi = 0$ and $\kappa = -0.1$ (Figure 2). Maternal dominant by zygotic additive epistasis with and without purifying selection: $K_{MmO} = -0.4, 0$, and 0.4 with $\phi = 0$ and $\kappa = 0$; $K_{MmO} = -0.4, -0.2, 0, 0.2$, and 0.4 with $\phi = -0.1$ and $\kappa = 0$; $K_{MmO} = -0.4, -0.2, 0, 0.2$, and 0.4 with $\phi = 0$ and $\kappa = -0.1$ (Figure 3). Maternal additive by zygotic dominant epistasis with and without purifying selection: $K_{MOo} = -0.4, 0$, and 0.4 with $\phi = 0$ and $\kappa = 0$; $K_{MOo} = -0.4, -0.2, 0, 0.2$, and 0.4 with $\phi = -0.1$ and $\kappa = 0$; $K_{MOo} = -0.4, -0.2, 0, 0.2$, and 0.4 with $\phi = 0$ and $\kappa = -0.1$ (Figure 4). Maternal dominant by zygotic dominant epistasis with and without purifying selection: $J_{MmOo} = -0.4, 0$, and 0.4 with $\phi = 0$ and $\kappa = 0$; $J_{MmOo} = -0.4, -0.2, 0, 0.2$, and 0.4 with $\phi = -0.1$ and $\kappa = 0$; $J_{MmOo} = -0.4, -0.2, 0, 0.2$, and 0.4 with $\phi = 0$ and $\kappa = -0.1$ (Figure 5). In each case parameter values not mentioned were set at zero.

3. Results

3.1. Do Maternal Genetic Effects Increase the Incidence of Genetic Diseases? We found, as we would expect, that genes that encode alleles which reduce fitness are selectively disadvantageous. We found that additive deleterious alleles are

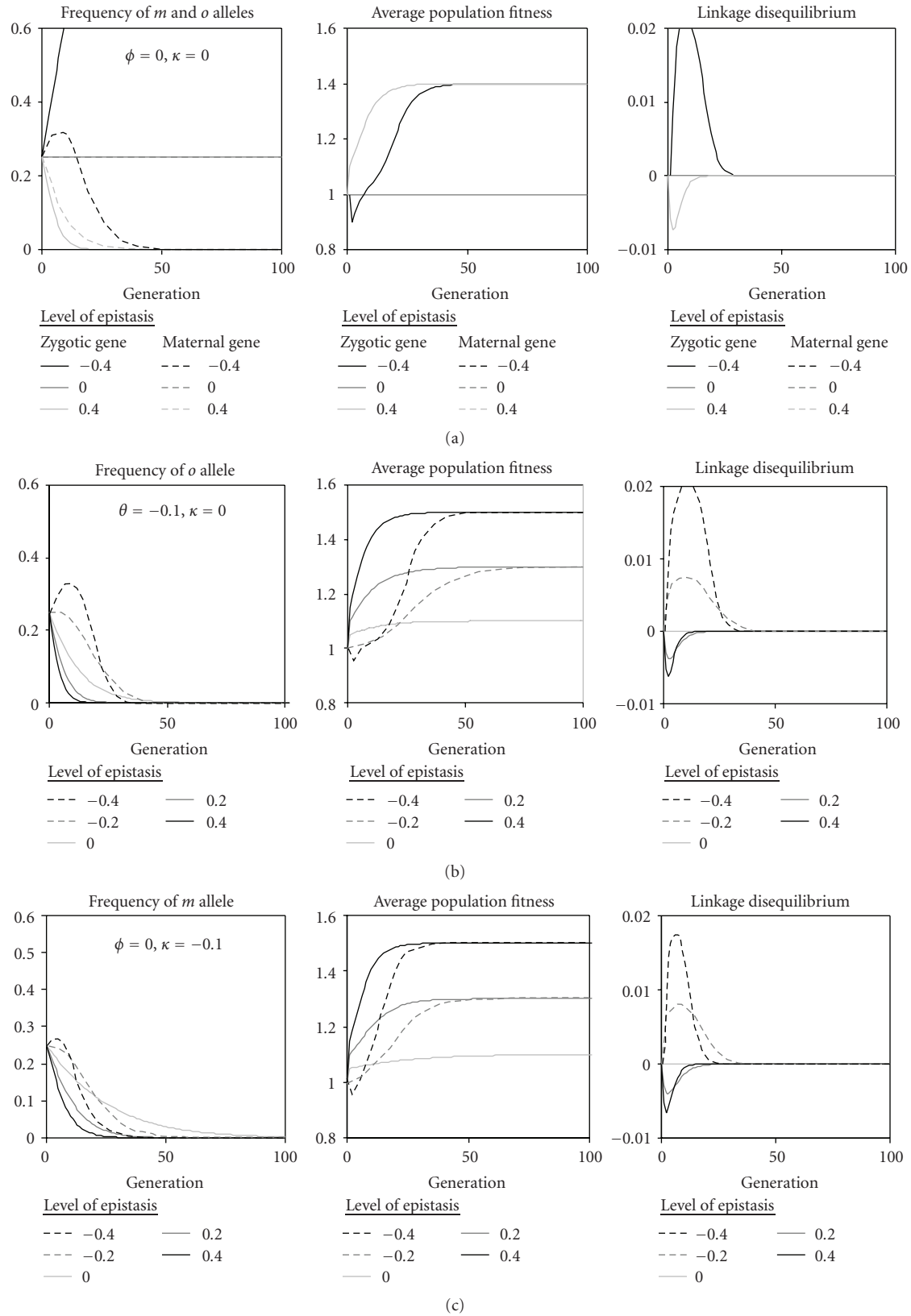


FIGURE 2: The evolution of genetic diseases under I_{MO} , maternal additive by zygotic additive epistasis. (a) Simulations which track the evolution of maternally expressed alleles and zygotically expressed alleles under different levels of epistasis without direct purifying selection. (b) Simulations which track the evolution of zygotically expressed alleles under different levels of epistasis with direct purifying selection on the zygotically expressed allele. (c) Simulations which track the evolution of maternally expressed alleles under different levels of epistasis with direct purifying selection on the maternally expressed allele.

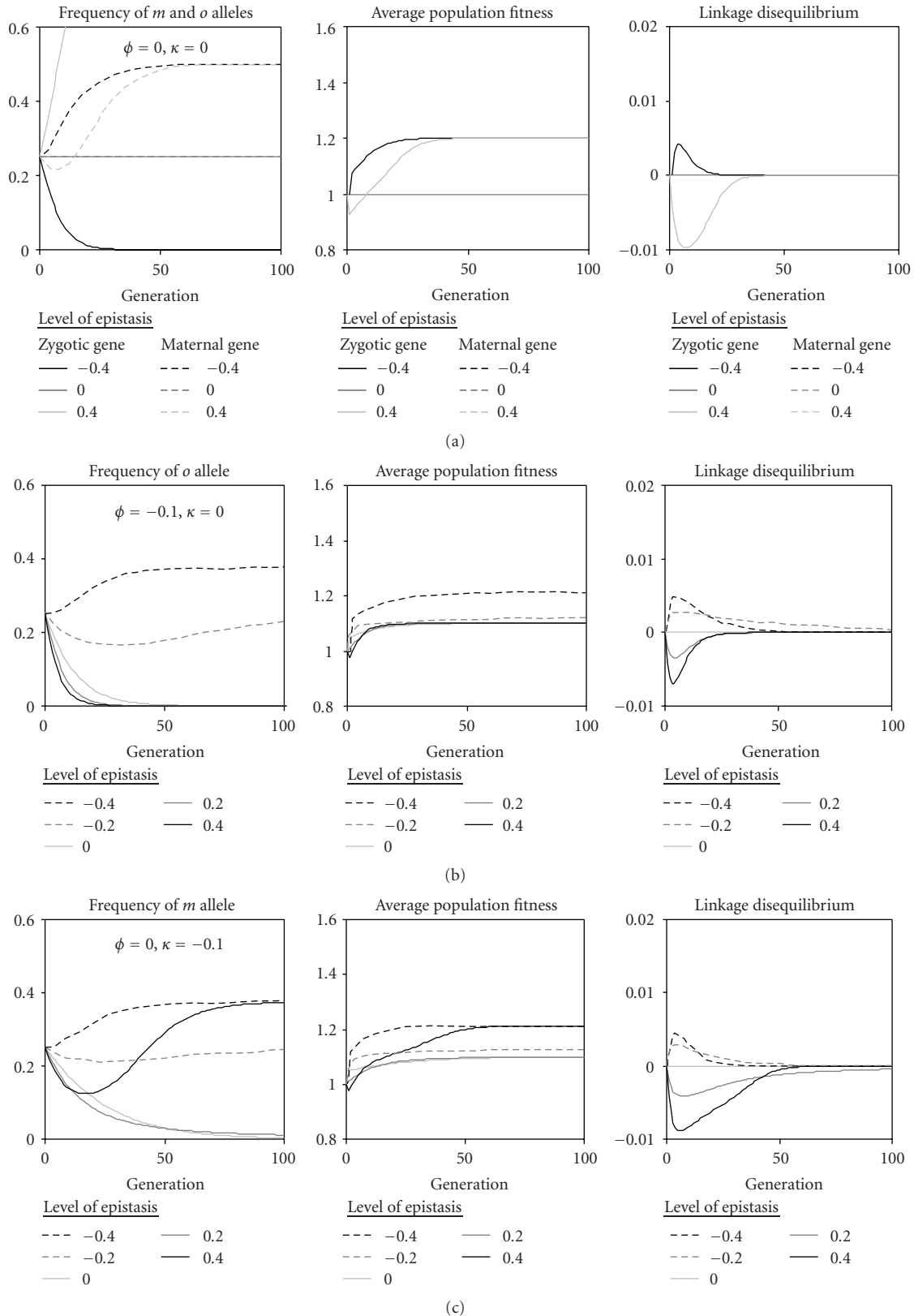


FIGURE 3: The evolution of genetic diseases under K_{OoM} , maternal additive by zygotic dominant epistasis. (a) Simulations which track the evolution of maternally expressed alleles and zygotically expressed alleles under different levels of epistasis without direct purifying selection. (b) Simulations which track the evolution of zygotically expressed alleles under different levels of epistasis with direct purifying selection on the zygotically expressed allele. (c) Simulations which track the evolution of maternally expressed alleles under different levels of epistasis with direct purifying selection on the maternally expressed allele.

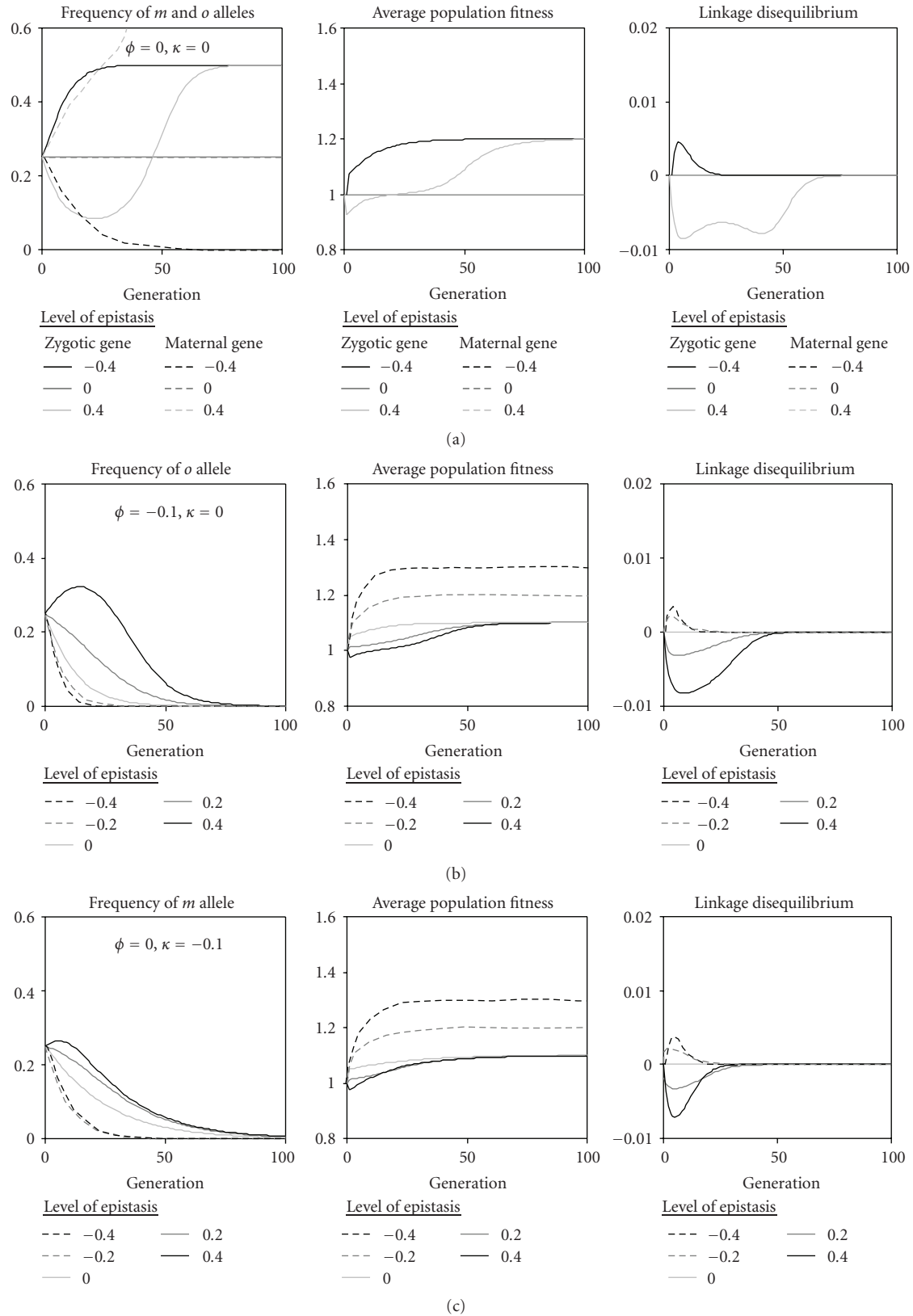


FIGURE 4: The evolution of genetic diseases under K_{OMm} , maternally dominant by zygotic additive epistasis. (a) Simulations which track the evolution of maternally expressed alleles and zygotically expressed alleles under different levels of epistasis without direct purifying selection. (b) Simulations which track the evolution of zygotically expressed alleles under different levels of epistasis with direct purifying selection on the zygotically expressed allele. (c) Simulations which track the evolution of maternally expressed alleles under different levels of epistasis with direct purifying selection on the maternally expressed allele.

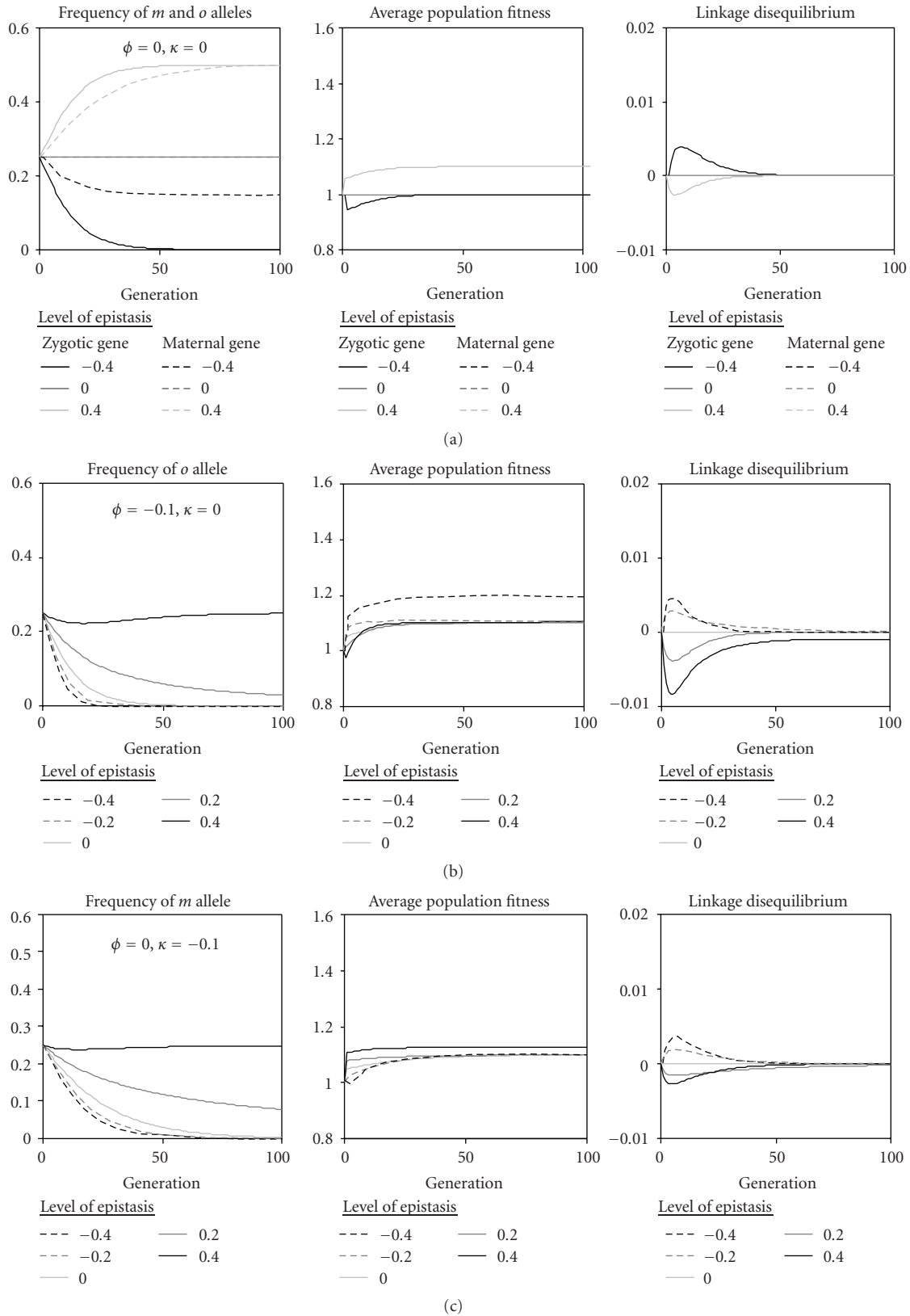


FIGURE 5: The evolution of genetic diseases under J_{OoMm} , maternal dominant by zygotic dominant epistasis. (a) Simulations which track the evolution of maternally expressed alleles and zygotically expressed alleles under different levels of epistasis without direct purifying selection. (b) Simulations which track the evolution of zygotically expressed alleles under different levels of epistasis with direct purifying selection on the zygotically expressed allele. (c) Simulations which track the evolution of maternally expressed alleles under different levels of epistasis with direct purifying selection on the maternally expressed allele.

lost more quickly than alleles with some level of dominance. Dominance, in effect, hides copies of the deleterious allele away from some of the action of purifying natural selection and retards the overall evolutionary process. Complete dominance hides alleles more effectively than the partial dominance we illustrate here. By comparing the maternal and zygotic patterns of gene expression, we found that maternally expressed alleles are always lost more slowly than offspring alleles with the same additive and dominance effects (Figure 1). Maternal gene expression, like dominance, effectively hides copies of the deleterious alleles in fathers, where they are not expressed and hence are hidden from the action of purifying natural selection. Note that average population fitness smoothly increases in all cases as expected when deleterious genotypes are removed by selection and their place is taken by genotypes of higher fitness. There is no LD because we are examining selection against single alleles. These results indicate that genetic diseases caused by gene expression in mothers will have a higher incidence than diseases exclusively caused by gene expression in offspring.

We see the same basic findings when purifying selection acts in conjunction with maternal-zygotic epistasis. Note that, in all cases, purifying selection acting on maternal effect genes is slower than it is on similarly acting zygotic effect genes. This reinforces and extends our general finding for single genes with only additive and dominance effects to more complex genetic architectures: maternal effect disease-causing genes will be more common in populations than similar zygotically expressed genes.

Because selection on genes with only maternal effects is weakened, we note that, in some cases, the evolutionary outcome differs for maternal and zygotic effect genes. In particular, with additive-by-dominance epistasis, strong interaction outweighs the additive, deleterious effect and produces permanent polymorphism of the maternal effect gene but not for the zygotic effect gene (compare the Left Panels of Figures 3(b) and 4(b)). This is the most dramatic case where the evolutionary equilibrium frequency of a maternal effect gene ($p_M \sim 0.36$) exceeds that of a zygotic gene ($p_O \sim 0.0$) of comparable effect.

3.2. How Do Epistatic Interactions between Maternal and Offspring Genomes Affect Disease Susceptibility? Our results show that epistasis between a maternally expressed allele and a zygotically expressed allele can alter the evolution of genetic diseases. In the absence of maternal-zygotic epistasis, purifying selection rapidly removes disease genes with additive or dominant effects (Figure 1). However, in the presence of maternal-zygotic epistasis purifying selection can have a variety of different consequences for the evolution of disease-encoding alleles. It can hasten the loss of the allele (Figures 2(a) and 2(b)). It can slow the loss of the allele (Figures 3(c), 4(b), 4(c), 5(b) and 5(c)). It can increase the frequency of the allele before eventual loss (Figures 2(b), 2(c), 4(b) and 4(c)); and it can result in permanent maintenance of disease-causing alleles (Figures 3(b), 3(c), 5(b) and 5(c)).

For additive-by-additive epistasis, the o allele is deleterious and, on some genetic backgrounds of the “M” gene, its deleterious effect on fitness is intensified, while on others it is diminished. In order to determine the overall effect of the interaction, the fitness of the o allele has to be averaged over all genetic backgrounds at the “M” locus. This averaging produces some counter intuitive results with additive-by-additive epistasis. For example, when $I_{OM} = +0.4$ (I_{OM} in Table 2 and heavy solid line in (Figure 2(a) Left Panel), the fitness of the MMOO genotype is substantially greater than 1.0 when produced by MM mothers (Fitness = $1.3 [1 - 0.1 + 0.4]$), and close to 1.0 when produced by Mm mothers (Fitness = $0.9 [1 - 0.1]$). Because the fitness of the MMOO genotype offspring from MM mothers is high, we might expect selection to favor the spread of M and O alleles; however, the MM genotypes also have lowered fitness in the absence of the o allele, so it is not clear, a priori, how the frequencies of o and m will change each generation. The simulations indicate that under positive I_{OM} epistasis both the m and o alleles are quickly removed from the population because the fitness of mmOO and MMoo intermediate genotypes is low and because oo genotypes have a fitness advantage when they are produced by mm mothers. In fact, the m and o alleles are actually lost more quickly with positive I_{OM} epistasis than with no epistasis (Figure 2(a) Left Panel). As the bad gene combinations are selected against, the mean fitness of the population is increased (Figure 2(a) Top Middle Panel). During this process, Mo and mO gametes are more common than random expectation so that LD is negative (Figure 2(a) Top Right Panel). With fixation of the O allele, LD returns to zero.

When we change the sign of I_{OM} , MMoo, and mmOO genotypes have the highest fitness, and the MMOO and mmoo genotypes have the lowest fitness. Here selection against the O disease susceptibility allele is weaker than with no epistasis because the MMoo and mmOO genotypes have higher fitness which causes the M allele to fix, after which in the MM genetic background, the o allele is no longer favored. Note that changes in mean population fitness under negative I_{OM} do not increase monotonically and that the rate of change in mean population fitness shifts as genotypes are sorted by selection (Figure 2(b) Middle Panel). During this process, OM and om gametes are more common than random expectation so that LD is positive (Figure 2(b) Middle Right Panel). However, with fixation of the o allele, LD returns to zero. These findings show that additive-by-additive epistasis opposite in sign of an allele’s direct effect on fitness will mitigate purifying selection. Intuitively, by synergistically making a bad allele worse, selection will act to remove the allele more quickly, while in the opposite case, selection will be reduced.

Negative maternal dominance by zygotic additive epistasis appears to increase selection against additive zygotically expressed disease genes (Figure 4 Left Panels). Without epistasis, the fitnesses of the “O” locus genotypes, OO, Oo, and oo are 0.9, 1, and 1.1, respectively; however, with K_{MmO} equal to -0.4 the fitnesses are 0.5 and 0.9 for OO, 1 for Oo, and 1.1 and 1.5, for oo, depending on maternal “M” locus genotype. For the “M” locus, on the most common

“O” locus genotypic background, oo, the fitnesses are 1.1 when they are derived from MM or mm mothers and 1.5 when they are derived from Mm mothers. This is the fitness pattern of heterozygote advantage at the “M” locus (i.e., marginal overdominance or balancing selection). As a result, for negative values of K_{OoM} , the zygotic effect disease gene is rapidly lost and the maternal effect gene with no additive effect on its own evolves to a frequency of 0.5 (Figure 4(a) Top Left Panel). Mean population fitness quickly increases for negative values of epistasis (Figure 4 Middle Panels). For positive values of K_{OoM} on the most common “O” locus background there is marginal under-dominance of the fitnesses of the three “M” locus genotypes, MM, Mm, and mm, 1.1, 0.7, 1.1, respectively. Though the O disease causing allele has a substantial fitness advantage in an MM background, because the M allele is lost over time as a result of the marginal under-dominance, the O allele is concurrently lost, though more slowly and with greater levels of negative LD than without epistasis (Figure 4(c) Bottom Right Panel).

With the maternal additive by zygotic additive epistasis equal to -0.4 for mm, the most common “M” locus background, the fitnesses of the three “O” locus genotypes, OO, Oo, and oo are 0.90, 1.4, and 1.1, respectively. This is the fitness pattern of balancing selection at the “O” locus. As a result, for negative values of K_{OoM} , the frequency of the deleterious O allele evolves to a stable intermediate frequency (Figure 5 Left Panels). Mean population fitness is not as high as in previous cases, because at the polymorphic equilibrium, recombination and segregation produce deleterious genotypes in the population every generation (Figure 5 Middle Panels). For positive values of K_{OoM} , there is marginal under-dominance of the fitnesses of the three “O” locus genotypes, OO, Oo, and oo: 0.90, 0.6, and 1.1, respectively when in the mm maternal background. As a result, the O allele is lost more quickly with epistasis than without epistasis. Overall, only moderate levels of LD are generated by positive and negative dominance-by-additive epistasis (Figure 5 Left Panels).

With dominance-by-dominance epistasis and J_{OoMm} equal to $+0.4$ (J_{OoMm} in Table 1), again there is marginal overdominance and heterozygote advantage at the “O” locus and the population achieves a stable intermediate polymorphism. Mean population fitness is lowest at this equilibrium because recombination and segregation continue to produce deleterious offspring genotypes. It is also important to note that negative LD is generated by strong positive dominance-by-dominance epistasis for nearly 200 generations although it is absent at equilibrium (Figure 5(b) Right Panel).

When only maternal-zygotic interaction determines disease risk, the variety of possible evolutionary outcomes is greater for both genes relative to all cases that we discussed above. In particular, stable polymorphism is the most common outcome with any kind of dominance epistasis (Figures 3, 4, and 5 Left Panels). At the polymorphic equilibria, no detectable LD exists in a population (Figures 3, 4, and 5 Right Panels). As a result of the polymorphism, equilibrium disease risk and incidence will be high and, owing to the absence of LD, *nonrandom associations across the population between*

maternal and zygotic alleles, that might point us toward a causal interaction, will be lacking. This might well represent the current state of genetic risk for maternal-zygotic diseases in the US population.

3.3. Can Models of the Evolution of Complex Diseases Assist Us in Identifying the Genes Which Underlie Complex Diseases?

Our simulations provide two important lessons for the characterization of complex diseases. First, maternal genetic diseases are difficult to detect using linkage disequilibrium mapping. Maternal-zygotic epistasis is known to produce half the levels of LD as traditional epistasis [8]. Our results indicate that this finding also applies to maternal genetic disease involved in maternal-zygotic interactions. By comparing the middle and lower panels of Figures 2, 3, 4, and 5, it is clear that zygotic genetic diseases generate lower levels of LD than maternal genetic diseases when they exhibit maternal-zygotic epistasis. The implication of this finding is that studies designed to identify maternal-zygotic epistasis up to a particular effect size will need substantially larger sample sizes than studies designed to identify epistasis with effects of the same magnitude.

Second, epistasis generates patterns that, without knowledge of the exact mechanism of inheritance, appear to be similar to simple genetic effects on disease. Our results illustrate cases where allele frequencies appear to be selected for or against, or can be maintained by balancing selection. Without considering the fate and frequency of the interacting alleles, these patterns could be interpreted as patterns driven by genes with simple genetic effects. But, in each of these cases the pattern is influenced by the fate and frequency of interacting loci (see [33]). This finding indicates that characterizations of simple genetic diseases might need to be qualified with the possibility that gene interactions drive the disease expression. This finding also indicates that larger, more genetically detailed data sets may provide a deeper understanding of the evolution of genetic disease.

4. Discussion

All genetic diseases are affected by the same constellation of evolutionary forces: natural selection, mutation, migration, and random genetic drift; however, the complexity of the pattern of gene expression changes the way these forces work together. The evolutionary equilibrium of simple genetic diseases is thought to be primarily determined by the mutation-selection balance [2]. However, the results of our simulations show that the pattern of gene expression (maternal versus offspring) and the form and level of gene interaction can greatly affect the incidence of genetic diseases.

Disease genes persist longer and at higher frequencies in a population when they are expressed maternally than when they are expressed zygotically. We have shown that genes which reduce fitness evolve differently under each of the four statistical forms of epistasis. We found that genes with additive main effects on increased risk of disease can be selected for and even brought to fixation by epistasis between

maternally and zygotically expressed alleles. This was particularly true when the form of epistasis involves dominance. By tracking gene frequency changes in conjunction with the mean fitness of the population we found that as the frequencies of particular genotypes change in the population, the mean population fitness can shift in nonintuitive ways. We also found that epistasis temporarily generates LD between maternal- and zygotic-effect genes. Thus, even with the simplifying assumptions of infinite population size and random mating, maternal/zygotic epistasis had a large confounding effect on the detection of single genes underlying disease etiology.

4.1. Why Maternal Effects Occur. Maternal genetic effects and maternal-zygotic epistatic interactions on offspring phenotypes arise because maternal genotype and condition guide early embryonic development [7, 12, 13, 46, 47]. For example, it is now well established that birth weight is determined by interactions between the maternal uterine genotype and offspring genotype [8, 10, 11, 26]. Mothers contribute prenatal nutrients, mRNA, and antibodies to offspring as well as postnatal effects, most prominently during lactation [10, 11, 48]. In vertebrate embryos, maternal gene expression establishes the formation of axes and induces developmental genes in the embryo [49]. In mammals, zygotic gene expression becomes the predominant controlling factor in development between the 2 and 8-cell stage, which is called the maternal-zygotic transition (MZT) [50–52].

Perhaps the best known maternal-zygotic interaction in human disease is the Rhesus (Rh) blood factor in pregnancy [27]. Maternally expressed alleles of the RHD gene negatively interact with zygotically expressed alleles of the RHAG gene. If mothers are serologically Rh⁻, then exposure to serologically Rh⁺ molecules in fetal blood can result in the formation of antibodies against the Rh⁺ factor, which can result in rejection of the red blood cells of the baby and subsequent anemia, brain damage and even fetal death [27]. Similarly, maternal, zygotic and maternal-zygotic interaction effects in several genes in the folate-dependent homocysteine pathway are associated with neural tube defects [25]. Maternal genetic effects are also thought to increase the risk of Down's syndrome [28].

4.2. Maternal Condition and Gene by Environment Interaction. Though we emphasize maternal-zygotic epistasis in disease etiology, it should be noted that maternal condition also influences the health of offspring. This type of effect results from genotype-by-environment interactions. In utero exposure to smoking [31] and to alcohol [30] are probably the best known causes of fetal syndromes, but maternal age [53, 54] and in utero exposure to disease [55] may have long-term negative impacts on offspring health [29]. Interactions between the maternal environment and genotype can also affect susceptibility to disease and birth defects. For example, the strength of effect of maternal age on the incidence of obsessive-compulsive behaviours in offspring appears to depend on interactions with the dopamine D1 gene [16].

Of course, the natural world is much more complex than our simulations, although digenic interaction models should capture much of the phenotypic variation owing to epistasis. Historically, human populations have been genetically subdivided and stratified—conditions where epistasis and maternal effect would be more important [7, 21, 33]. In the last few centuries, the development of a global economy has increased migration and as a consequence has brought genotypes together that have independently evolved for thousands of years and created populations that are unlikely to be at evolutionary genetic equilibrium. Such admixture of populations also creates LD and segregation among large blocks of genes. How can we ever hope to uncover the genetic basis of disease given this historical complexity?

4.3. Implications for Identifying Disease Genes. As one can clearly see from our simulations, complex genetic disorders have a confusing pattern of inheritance and a nonintuitive evolutionary trajectory. In the last two decades, there has been increasing interest in uncovering the genetic basis of complex diseases. Though association and linkage disequilibrium studies have identified many putative disease genes, they have often been difficult to confirm in independent population samples. Some investigators have argued that this irreproducibility is largely a consequence of weak statistical power [17, 18, 56]. However, the lack of replication of the association between genes and traits is also the signature of epistasis [33]. In some populations the genetic background will facilitate detection of disease genes, while in others, the genetic background masks detection [19, 33].

These complexities do not mean that we cannot map genetic basis of complex genetic disorders. We need to use and develop methods that identify the contexts that lead to increased risk of disease. Our simulations also show that each of the four forms of epistasis affects levels of linkage disequilibrium and, at equilibrium, LD tends toward zero. This implies that linkage association mapping of disease genes may not be that useful when the disease gene involves epistasis. Single allele-disease associations might occasionally point us in the right direction, but for most complex diseases we need approaches that embrace population subdivision and epistasis. For example, Templeton's nested cladistic analysis of phenotypic associations with haplotypes [57] identifies the specific genetic backgrounds that generate the strongest signal between alleles and disease (or any other trait) in nonexperimental populations. For particular diseases it is now possible to estimate an individual's risk of acquiring disease based on a genetic profile which classifies the individual's ethnicity [58]. There are also experimental methods that are specifically designed to detect gene interactions and maternal-zygotic gene interactions [22–24, 35]. By stratifying the population and searching for genetic sources of disease in each distinct biologically relevant clade—by embracing population subdivision and epistasis—we will be more likely to determine the genetic basis of birth defects and disease [33].

Acknowledgment

The authors acknowledge Grant 2-R01-GM065414-05A1 from the NIH for funding and Y. Brandvain and T. Cruickshank for comments.

References

- [1] A. Hamosh, A. F. Scott, J. S. Amberger, C. A. Bocchini, and V. A. McKusick, "Online Mendelian Inheritance in Man (OMIM), a knowledgebase of human genes and genetic disorders," *Nucleic Acids Research*, vol. 33, pp. D514–D517, 2005.
- [2] D. L. Hartl and R. B. Campbell, "Allele multiplicity in simple Mendelian disorders," *American Journal of Human Genetics*, vol. 34, no. 6, pp. 866–873, 1982.
- [3] D. E. Reich and E. S. Lander, "On the allelic spectrum of human disease," *Trends in Genetics*, vol. 17, no. 9, pp. 502–510, 2001.
- [4] T. A. Thornton-Wells, J. H. Moore, and J. L. Haines, "Genetics, statistics and human disease: analytical retooling for complexity," *Trends in Genetics*, vol. 20, no. 12, pp. 640–647, 2004.
- [5] A. R. Templeton, "Epistasis and complex traits," in *Epistasis and the Evolutionary Process*, J. B. Wolf, E. D. Brodie III, and M. J. Wade, Eds., pp. 41–57, Oxford University Press, Oxford, UK, 2000.
- [6] E. S. Lander and N. J. Schork, "Genetic dissection of complex traits," *Science*, vol. 265, no. 5181, pp. 2037–2048, 1994.
- [7] M. J. Wade, "The evolutionary genetics of maternal effects," in *Maternal Effects as Adaptations*, T. A. Mousseau and C. W. Fox, Eds., pp. 5–12, Oxford University Press, Oxford, UK, 1998.
- [8] J. B. Wolf, "Gene interactions from maternal effects," *Evolution*, vol. 54, no. 6, pp. 1882–1898, 2000.
- [9] D. S. Falconer, "Maternal effects and selection response," in *Genetics Today: Proceedings of the 11th International Congress of Genetics*, vol. 3, pp. 763–774, Pergamon, Oxford, UK, 1965.
- [10] D. E. Cowley, "Prenatal effects on mammalian growth: embryo transfer results," in *The Unity of Evolutionary Biology. Vol 2. Proceedings of the Fourth International Congress of Systematic and Evolutionary Biology*, E. C. Dudley, Ed., pp. 762–779, Dioscorides Press, Portland, Ore, USA, 1991.
- [11] D. E. Cowley, D. Pomp, W. R. Atchley, E. J. Eisen, and D. Hawkins-Brown, "The impact of maternal uterine genotype on postnatal growth and adult body size in mice," *Genetics*, vol. 122, no. 1, pp. 193–203, 1989.
- [12] D. A. Roach and R. D. Wulff, "Maternal effects in plants," *Annual Review of Ecology and Systematics*, vol. 18, pp. 209–235, 1987.
- [13] T. Mousseau and C. Fox, Eds., *Maternal Effects as Adaptations*, Oxford University Press, Oxford, UK, 1998.
- [14] J. H. Nadeau, "Transgenerational genetic effects on phenotypic variation and disease risk," *Human Molecular Genetics*, vol. 18, pp. R202–R210, 2009.
- [15] J. B. Wolf, T. T. Vaughn, L. S. Pletscher, and J. M. Cheverud, "Contribution of maternal effect QTL to genetic architecture of early growth in mice," *Heredity*, vol. 89, no. 4, pp. 300–310, 2002.
- [16] D. E. Comings and J. P. MacMurray, "Maternal age at the birth of the first child as an epistatic factor in polygenic disorders," *American Journal of Medical Genetics*, vol. 141B, no. 1, pp. 1–6, 2006.
- [17] H. Shen, Y. Liu, P. Liu, R. R. Recker, and H.-W. Deng, "Nonreplication in genetic studies of complex diseases—lessons learned from studies of osteoporosis and tentative remedies," *Journal of Bone and Mineral Research*, vol. 20, no. 3, pp. 365–376, 2005.
- [18] K. F. Manly, "Reliability of statistical associations between genes and disease," *Immunogenetics*, vol. 57, no. 8, pp. 549–558, 2005.
- [19] M. J. Sillanpää and K. Auranen, "Replication in genetic studies of complex traits," *Annals of Human Genetics*, vol. 68, no. 6, pp. 646–657, 2004.
- [20] Y. Li, C. M. Coelho, T. Liu, et al., "A statistical model for estimating maternal-zygotic interactions and parent-of-origin effects of QTLs for seed development," *PLoS ONE*, vol. 3, no. 9, article e3131, 2008.
- [21] T. A. Linksvayer and M. J. Wade, "The evolutionary origin and elaboration of sociality in the aculeate hymenoptera: maternal effects, sib-social effects, and heterochrony," *Quarterly Review of Biology*, vol. 80, no. 3, pp. 317–336, 2005.
- [22] A. J. Wilcox, C. R. Weinberg, and R. T. Lie, "Distinguishing the effects of maternal and offspring genes through studies of "case-parent triads,"" *American Journal of Epidemiology*, vol. 148, no. 9, pp. 893–901, 1998.
- [23] E. Wheeler and H. J. Cordell, "Quantitative trait association in parent offspring trios: extension of case/pseudocontrol method and comparison of prospective and retrospective approaches," *Genetic Epidemiology*, vol. 31, no. 8, pp. 813–833, 2007.
- [24] J. S. Sinsheimer, C. G. S. Palmer, and J. A. Woodward, "Detecting genotype combinations that increase risk for disease: the maternal-fetal genotype incompatibility test," *Genetic Epidemiology*, vol. 24, no. 1, pp. 1–13, 2003.
- [25] C. L. Relton, C. S. Wilding, M. S. Pearce, et al., "Gene-gene interaction in folate-related genes and risk of neural tube defects in a UK population," *Journal of Medical Genetics*, vol. 41, no. 4, pp. 256–260, 2004.
- [26] E. B. Robson, "Birth weight in cousins," *Annals of Human Genetics*, vol. 19, no. 4, pp. 262–268, 1955.
- [27] C. Le van Kim, I. Mouro, B. Cherif-Zahar, et al., "Molecular cloning and primary structure of the human blood group RhD polypeptide," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 89, no. 22, pp. 10925–10929, 1992.
- [28] C. A. Hobbs, S. L. Sherman, P. Yi, et al., "Polymorphisms in genes involved in folate metabolism as maternal risk factors for Down syndrome," *American Journal of Human Genetics*, vol. 67, no. 3, pp. 623–630, 2000.
- [29] D. J. P. Barker, Ed., *Fetal and Infant Origins of Adult Disease: Papers*, British Medical Journal, London, UK, 1992.
- [30] K. L. Jones, D. W. Smith, C. N. Ulleland, and A. P. Streissguth, "Pattern of malformation in offspring of chronic alcoholic mothers," *Lancet*, vol. 1, no. 7815, pp. 1267–1271, 1973.
- [31] D. P. Misra, N. Astone, and C. D. Lynch, "Maternal smoking and birth weight: interaction with parity and mother's own in utero exposure to smoking," *Epidemiology*, vol. 16, no. 3, pp. 288–293, 2005.
- [32] M. J. Wade, "The coevolutionary genetics of ecological communities," *Nature Reviews Genetics*, vol. 8, pp. 185–195, 2007.
- [33] M. J. Wade, "Epistasis, complex traits, and mapping genes," *Genetica*, vol. 112–113, pp. 59–69, 2001.

- [34] M. S. Barker, J. P. Demuth, and M. J. Wade, "Maternal expression relaxes constraint on innovation of the anterior determinant, bicoid," *PLoS Genetics*, vol. 1, no. 5, article e57, 2005.
- [35] J. P. Demuth and M. J. Wade, "Maternal expression increases the rate of bicoid evolution by relaxing selective constraint," *Genetica*, vol. 129, no. 1, pp. 37–43, 2007.
- [36] M. J. Wade, N. K. Priest, and T. Cruickshank, "A theoretical overview of maternal genetic effects: evolutionary predictions and empirical tests with mammalian data," in *Maternal Effects in Mammals*, D. Maestriperi and J. Mateo, Eds., pp. 38–63, University of Chicago Press, Chicago, Ill, USA, 2009.
- [37] M. J. Wade, "A gene's eye view of epistasis, selection and speciation," *Journal of Evolutionary Biology*, vol. 15, no. 3, pp. 337–346, 2002.
- [38] S. Wright, "Evolution in Mendelian populations," *Genetics*, vol. 16, pp. 97–159, 1931.
- [39] S. Wright, *Evolution and the Genetics of Populations, Volume 2: Theory of Gene Frequencies*, University Of Chicago Press, Chicago, Ill, USA, 1969.
- [40] P. C. Phillips, "The language of gene interaction," *Genetics*, vol. 149, no. 3, pp. 1167–1171, 1998.
- [41] P. C. Phillips, S. P. Otto, and M. C. Whitlock, "Beyond the average. The evolutionary important of gene interactions and variability of epistatic effects," in *Epistasis and the Voluntary Process*, J. B. Wolf, E. D. Brodie, and M. J. Wade, Eds., pp. 20–38, Oxford University Press, New York, NY, USA, 2000.
- [42] C. C. Cockerham, "An extension of the concept of partitioning hereditary variance for analysis of covariance among relatives when epistasis is present," *Genetics*, vol. 41, pp. 138–141, 1954.
- [43] J. B. Wolf, "Genetic architecture and evolutionary constraint when the environment contains genes," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 100, no. 8, pp. 4655–4660, 2003.
- [44] R. C. Lewontin and K. Kojima, "The evolutionary dynamics of complex polymorphisms," *Evolution*, vol. 14, pp. 450–472, 1960.
- [45] J. F. Crow and M. Kimura, *An Introduction to Population Genetics Theory*, Harper and Row, New York, NY, USA, 1970.
- [46] J. B. Wolf, E. D. Brodie III, J. M. Cheverud, A. J. Moore, and M. J. Wade, "Evolutionary consequences of indirect genetic effects," *Trends in Ecology and Evolution*, vol. 13, no. 2, pp. 64–69, 1998.
- [47] D. A. Roff, "The detection and measurement of maternal effects," in *Maternal Effects as Adaptations*, T. A. Mousseau and C. W. Fox, Eds., pp. 83–96, Oxford University Press, Oxford, UK, 1998.
- [48] H. A. Oksh, T. M. Sutherland, and J. S. Williams, "Prenatal and postnatal maternal influence on growth in mice," *Genetics*, vol. 57, no. 1, pp. 79–94, 1967.
- [49] S. Sidi, C. Goutel, N. Peyrieras, and F. M. Rosa, "Maternal induction of ventral fate by zebrafish radar," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 100, no. 6, pp. 3315–3320, 2003.
- [50] V. L. Jarrell, B. N. Day, and R. S. Prather, "The transition from maternal to zygotic control of development occurs during the 4-cell stage in the domestic pig, *Sus scrofa*: quantitative and qualitative aspects of protein synthesis," *Biology of Reproduction*, vol. 44, no. 1, pp. 62–68, 1991.
- [51] R. M. Schultz, "The molecular foundations of the maternal to zygotic transition in the preimplantation embryo," *Human Reproduction Update*, vol. 8, no. 4, pp. 323–331, 2002.
- [52] C. Vigneault, S. McGraw, L. Massicotte, and M.-A. Sirard, "Transcription factor expression patterns in bovine in vitro-derived embryos prior to maternal-zygotic transition," *Biology of Reproduction*, vol. 70, no. 6, pp. 1701–1709, 2004.
- [53] L. A. Gavrilov and N. Gavrilova, "Human longevity and parental age at conception," in *Sex and Longevity: Sexuality, Reproduction, Parenthood*, J.-M. Robine and T. B. L. Kirkwood, Eds., pp. 7–31, Springer, Berlin, Germany, 2000.
- [54] A. Kemkes-Grottenthaler, "Parental effects on offspring longevity—evidence from 17th to 19th century reproductive histories," *Annals of Human Biology*, vol. 31, no. 2, pp. 139–158, 2004.
- [55] D. Almond, "Is the 1918 influenza pandemic over? Long-term effects of in utero influenza exposure in the post-1940 U.S. population," *Journal of Political Economy*, vol. 114, no. 4, pp. 672–712, 2006.
- [56] E. Lander and L. Kruglyak, "Genetic dissection of complex traits: guidelines for interpreting and reporting linkage results," *Nature Genetics*, vol. 11, no. 3, pp. 241–247, 1995.
- [57] A. R. Templeton, "A cladistic analysis of phenotypic associations with haplotypes inferred from restriction endonuclease mapping or DNA sequencing. V. Analysis of case/control sampling designs: Alzheimer's disease and the apoprotein E locus," *Genetics*, vol. 140, no. 1, pp. 403–409, 1995.
- [58] S. Choudhry, N. E. Coyle, H. Tang, et al., "Population stratification confounds genetic association studies among Latinos," *Human Genetics*, vol. 118, no. 5, pp. 652–664, 2006.