Introduction

Complex phenotypes, including quantitative traits and common diseases, are controlled by many genes and by environmental factors. How do these genes combine to determine the phenotype of an individual? The simplest model is to assume that genes act additively with each other both within and between loci, but of course they may interact to show dominance or epistasis, respectively. A longstanding controversy has existed concerning the importance of these non-additive effects, involving both Fisher [1] and Wright [2]. Estimates of genetic variance components within populations have indicated that most of the variance is additive [3,4]. Increasing knowledge about biological pathways and gene networks implies, however, that gene-gene interactions (epistasis) are important, and some have argued recently that much genetic variance in populations is due to these components than do half sibs. The coefficients of epistatic variance between full sibs and half sibs can be partitioned into additive (\(V_{AA}\)), dominance (\(V_D\)), and multiple locus components (\(V_{I}\)), which itself can be partitioned into two locus (\(V_{AA}, V_{AD}, V_{DD}\)) and multiple locus components (\(V_{AAA}, V_{AAH}, V_{AAA}, \ldots\) ) [3,4,14,15,16,17]. Estimation of additive and non-additive variance components utilises the observed phenotypic similarity of relatives and the expected contribution of additive and non-additive effects to that similarity [3,4]. In addition to dominance and epistatic interaction at the level of variance, we also need to consider the contributions of other genetic effects, such as pleiotropy, linkage disequilibrium, and gene-environment interactions.

Empirical Evidence for Additive and Non-Additive Genetic Variance

Estimation of Genetic Variance. The genetic variance \(V_G\) can be partitioned into additive (\(V_{AA}\)), dominance (\(V_D\)), and a combined epistatic component (\(V_I\)), which itself can be partitioned into two locus (\(V_{AA}, V_{AD}, V_{DD}\)) and multiple locus components (\(V_{AAA}, V_{AAH}, V_{AAA}, \ldots\) ) [3,4,14,15,16,17]. Estimation of additive and non-additive variance components utilises the observed phenotypic similarity of relatives and the expected contribution of additive and non-additive effects to that similarity [3,4]. In addition to interaction and dominance of genes, relatives may resemble each other due to common environmental effects.

In an extremely large data set with very many different kinds of relationships present, it is possible in principle to partition variation into many components using modern statistical methods such as residual maximum likelihood [18] (REML) with the animal model [4,19,20]. In practice it is never possible to estimate many variance components with useful precision, however, not least because there is a high degree of confounding: for example, full sibs have a higher covariance for all single and multi-locus genetic components than do half sibs. Therefore, the coefficients of epistatic components are small (e.g., \(V_{AA}/16\) for half-sibs), so estimates have
Author Summary

Genetic variation in quantitative or complex traits can be partitioned into many components due to additive, dominance, and interaction effects of genes. The most important is the additive genetic variance because it determines most of the correlation of relatives and the opportunities for genetic change by natural or artificial selection. From reviews of the literature and presentation of a summary analysis of human twin data, we show that a high proportion, typically over half, of the total genetic variance is additive. This is surprising as there are many potential interactions of gene effects within and between loci, some revealed in recent QTL analyses. We demonstrate that under the standard model of neutral mutation, which leads to a U-shaped distribution of gene frequencies with most near 0 or 1, a high proportion of additive variance would be expected regardless of the amount of dominance or epistasis at the individual loci. We also show that the model is compatible with observations in populations undergoing selection and results of QTL analyses on F2 populations.

high sampling error and there is little power to distinguish $V_A$ from, say, $V_\Delta$. Selection, assortative mating, and non-genetic covariances also confound estimates. Consequently, there are few accurate estimates of non-additive variance components but there is indirect evidence. For instance, a narrow sense heritability value ($h^2 = V_A / V_P$) of one-half, typical for many traits, implies that dominance, all the vast number of epistatic components, and the environmental component, collectively contribute no more than $V_A$. Similarly if the heritability is only a little less than the repeatability (the phenotypic correlation of repeated measures), all non-additive genetic variances and the permanent environmental variance together comprise this small difference. With these caveats we summarise data of various types.

Laboratory Animals and Livestock. The extensive data on experimental organisms show a range of heritability, higher for morphological than fitness associated traits, averaging as follows [21]: morphology - 0.46, physiology (e.g., oxygen consumption, resistance to heat stress) - 0.33, behaviour - 0.30, and life history - 0.26.

There have been extensive estimates of heritability for traits of livestock. For example, for beef cattle, these averaged: post-weaning weight gain 0.31, market weight for age 0.41, backfat thickness 0.44 [22]. In general for morphological traits, such as carcass fatness, egg weight in poultry or fat and protein content of cow’s milk, a heritability of 0.5 or so is the norm, whereas for growth traits or milk yield 0.25–0.35 is more typical [23]. These estimates of heritability from half-sib correlations could be biased upwards by additive epistatic terms, but they can not account for estimates of heritability over 25%. Furthermore, estimates of realised heritability from response to selection [3] are not biased in that way, because epistatic components do not contribute to long term selection response [24], and estimates of realised heritability range up to 0.5 for fat content of mice, for example [23].

There are a number of cases where it can be shown directly that $V_\Delta$ contributes almost all of $V_G$ and indeed almost all of $V_P$. For bristle number in Drosophila melanogaster, the phenotypic correlation between abdominal segments, which, assuming they are influenced by the same genes, estimates $V_G / V_P$ is only a little higher than the heritability, indicating that $V_\Delta / V_G \sim 0.8$ [26]. For finger ridge count (in humans), estimates of heritability are close to one and consistent from different sorts of relatives [27]. Even for lowly heritable traits such as litter size in pigs, the repeatability is little higher than the heritability, implying that most genetic variance is additive [28]. Whilst there is a clear relationship between heritability and type of trait, it should be noted that low heritability does not imply low genetic variance: the evolvability ($< V_G / mean$) is higher for fitness than morphological traits [29], and even for estimates of fitness itself or traits closely related to it, additive genetic variance is present [30,31].

There are rarely good direct estimates of epistatic or dominance variance because these variance components are usually estimated from full-sibs and therefore confounded with the common environment shared by full sibs. However, if the heritability is high, the space for them is limited.

Experiments on inbreeding depression provide some evidence on the importance of non-additive effects. Inbreeding depression implies directional dominance in gene effects but, for a given rate of inbreeding depression, as the number of loci increases and the gene frequencies move toward 0 or 1, the dominance variance decreases towards zero. Consequently, the importance of inbreeding depression for traits related to fitness is not evidence that the dominance variance is large. The observed linearity of inbreeding depression with inbreeding co-efficient is easiest to explain with directional dominance but not with DD or higher order epistatic effects because these would cause non-linearity unless they happened to exactly cancel each other out.

Twin Studies in Humans. In contrast to studies of sibs and more distant relatives, identical twins can provide estimates of $V_G$. The classical twin design of samples of monozygotic (MZ) and dizygotic (DZ) twin pairs has been used extensively to estimate variance components for a wide range of phenotypes in human populations. The primary statistics from these studies are the correlations between MZ pairs ($r_{MZ}$) and between DZ pairs ($r_{DZ}$).

If twin resemblance due to common environmental factors is the same for MZ and DZ twins then $r_{MZ} = r_{DZ}$ implies that part of the resemblance is due to genetic factors and $r_{MZ} > 2 r_{DZ}$ implies the importance of non-additive genetic effects. Conversely, $r_{MZ} < 2 r_{DZ}$ implies that common environmental factors cause some of the observed twin resemblance. Sophisticated variance component partitioning methods to estimate components of additive, non-additive and common environmental effects are used widely [32], but all rely on the strong assumptions that resemblance due to common environmental effects is the same for MZ and DZ twins. Attempts to test this hypothesis have not found any evidence to reject it [33,34]. Nevertheless, even accepting this assumption about common environmental variance, in the classical twin design there are only two primary statistics and three or more variance components cannot be estimated without making additional assumptions.

We summarised the MZ and DZ correlations for a wide variety of phenotypes from published twin studies from a single productive laboratory in Australia (genepi.qimr.edu.au). The criteria were that each study must have more than 100 MZ and more than 100 DZ pairs and that the study subjects were Australian twins. For non-continuous traits, studies were included only if they reported polychoric or tetrachoric correlations. In total, 86 phenotypes qualified of which 42 were clinical measures of quantitative traits (including, for example, blood pressure, biochemical measures in blood, body-mass-index, height, tooth dimensions; a full list of phenotypes is available upon request). The MZ and DZ correlations are summarised in Table 1. The correlations were not separated according to the sex of the individuals in all studies; but for those that did separate the sexes, the overall MZ and DZ correlations were calculated as an average, weighted by the total number of pairs. The distribution of
\( r_{\text{MZ}} - 2r_{\text{DZ}} \) across all 86 phenotypes is shown in Figure 1. On average the MZ correlation is about twice the DZ correlation across a wide range of phenotypes. If we consider only clinically measured phenotypes and ignore opposite-sex twins then the MZ correlation is clearly less than twice the DZ correlation (Table 1). It is possible but unlikely that the variance due to common environmental factors, assortative mating and non-additive genetic factors exactly cancel each other out by chance. Thus the simplest explanation of the results is that additive variance explains most of the observed similarity of twins and non-additive variance is generally of small magnitude and cannot explain a large proportion of the genetic variance.

Model

Gene Frequency Distributions

In view of the apparent conflict between the observations of high proportions of additive genetic variance (often half or more of the phenotypic variance, and even more of the total genetic variance) and the recent reports of epistasis at quantitative trait loci (QTL) [8], we consider explanations beyond that of simple sampling errors and bias of estimates. We focus particularly on the role that the distribution of gene frequencies may play in the relation between the genetic model and the observed genetic variance components.

Genetic variance components depend on the mean value of each genotype and the allele frequencies at the genes affecting the trait [3,4,17]. Unfortunately the allele frequencies at most genes affecting complex traits are not known, but the distribution of allele frequencies can be predicted under a range of assumptions. This distribution depends on the magnitude of the evolutionary forces that create and maintain variance, including mutation, selection, drift and migration. As the effects on fitness of genes at many of the loci influencing most quantitative traits are likely to be small, we can invoke theory for neutral alleles to serve as a reference point. An important such reference is the frequency distribution under a balance between mutation and random genetic drift due to finite population size in the absence of selection. If mutations are rare, the distribution of the frequency \( p \) of the mutant allele is \( f(p) \propto 1/p \), i.e. approximately L-shaped \([2,35,36]\), with the high frequency at the tail being due to mutations arising recently. The allele which increases the value of the trait may be the mutant or ancestral allele, so its frequency has a U-shaped distribution \( f(p) \propto 1/(p - 1/(1-p)) \approx 1/(p(1-p)) \). As we shall use it often, we define the ‘U’ distribution explicitly by this formula. For loci at which the mutants are generally deleterious, the frequency distribution will tend to be more concentrated near \( p = 0 \) or 1 than for this neutral reference point. As another simple reference point we use the uniform distribution, \( f(p) \propto 1, 1/(2N) \leq p \leq 1/(2N) \), with \( N \) the population size. This approximates the steady state distribution of a neutral mutant gene which has been segregating for a very long time \([2]\), and also has much more density at intermediate gene frequencies than the ‘U’ distribution. Our third reference point is at \( p = 0.5 \), as in populations derived from inbred crosses, and is the extreme case of central tendency of gene frequency.

These analyses assume a gene frequency distribution which is relevant to no selection. For a more limited range of examples we consider the impact of selection on the partition of variance. We consider a limited range of genetic models, some simple classical ones and others based on published models of metabolic pathways or results of QTL mapping experiments.

Uniform: \( f(p) = 1 \), assuming \( N \) is sufficiently large that the discreteness of the distribution and any non-uniformity as \( p \) approaches 0 or 1 can be ignored, i.e. integrated over 0 to 1. This and the ‘U’ gene frequency distributions are, for simplicity, assumed to be continuous.

Neutral mutation model (‘U’): \( f(p) \propto 1/(p(1-p)) \). To standardise the distribution, with population size \( N \) assumed to be large, note that

\[
\int_{1/(2N)}^{1-1/(2N)} \frac{1}{p(1-p)} dp = 2 \left[ \ln \left( \frac{1 - 1/(2N)}{2N} \right) - \ln \left( \frac{1}{2N} \right) \right] \sim 2 \ln(2N)
\]

Thus \( f(p) = \frac{1}{2N(1-p)} \), where \( K = \ln(2N) \).

Genetic variance components are obtained by integration of expressions for the variance as a function of \( p \) for a specific model of the gene frequency distribution. For multiple locus models the distribution of all loci is assumed to be identical and there is no

<table>
<thead>
<tr>
<th>Group</th>
<th>All phenotypes</th>
<th>Clinically measured phenotypes</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>No. traits</td>
<td>( r )</td>
</tr>
<tr>
<td>MZ females</td>
<td>58</td>
<td>0.61</td>
</tr>
<tr>
<td>MZ males</td>
<td>48</td>
<td>0.65</td>
</tr>
<tr>
<td>DZ females</td>
<td>58</td>
<td>0.34</td>
</tr>
<tr>
<td>DZ males</td>
<td>48</td>
<td>0.36</td>
</tr>
<tr>
<td>OS pairs*</td>
<td>46</td>
<td>0.29</td>
</tr>
<tr>
<td>All MZ</td>
<td>86</td>
<td>0.58</td>
</tr>
<tr>
<td>All DZ</td>
<td>86</td>
<td>0.29</td>
</tr>
<tr>
<td>MZ – 2DZ</td>
<td>86</td>
<td>0.00</td>
</tr>
</tbody>
</table>

These show the correlations \( r \) of phenotypes of twins, averaged over ranges of traits estimated in large data sets

*Data from published papers by N.G. Martin and colleagues of the Queensland Institute of Medical Research, Brisbane (www.genepi.edu.au)

*Opposite sex

doi:10.1371/journal.pgen.1000008.t001

Figure 1. Distribution of \( r_{\text{MZ}} - 2r_{\text{DZ}} \) for all traits on human twins. Data are from published papers by N.G. Martin and colleagues of the Queensland Institute of Medical Research, Brisbane (www.genepi.edu.au). Across a wide variety of traits the mean difference between the monozygotic twin correlation and twice the dizygotic twin correlation is close to zero, which is consistent with predominantly additive genetic variance and the absence of a large component of variance due to common environmental effects.

doi:10.1371/journal.pgen.1000008.g001

Table 1. Meta-analysis of MZ and DZ correlations in humans*.
linkage disequilibrium. We focus on the contribution of additive genetic variance ($V_A$) to genotypic variance ($V_G$).

Genotypic Values

Single Locus with Arbitrary Dominance. Consider a single biallelic locus with genotypic values for CC, Cc and cc of $+a$, $d$ and $-a$, respectively (notation of [3]). Then, from [3]

$$V_A = 2p(1-p)[a + d(1-2p)]^2, \quad V_D = 4p^2(1-p)^2d^2 \quad \text{and} \quad V_G = V_A + V_D.$$ 

For the uniform distribution of $p$

$$E(V_A) = \frac{1}{3} \int_0^1 \left[ 2p(1-p)[a + d(1-2p)]^2 \right] dp.$$ 

Hence $E(V_A) = a^2/3 + d^2/15$ and $E(V_D) = 2d^2/15$, giving $E(V_A)/E(V_G) = 1 - 2d^2/(3a^2 + 2d^2)$.

Additive × Additive Model without Dominance or Interactions Including Dominance. A simple additive × additive epistatic model has these genotypic values:

- **CC**: $2a$
- **Cc**: $a$
- **cc**: $0$

Assuming the frequency of B is $p$ and of C is $q$, with linkage equilibrium:

$$\text{Mean} = M = 2ap(1-p)(1-q)$$

The average effect of substitution of allele B is given by [37]

$$(a^2_d + d^2)dp = a^2_d \left( \frac{1}{2q} - d_q \right)$$

and hence

$$V_A = 2a^2_d [p(1-p)(1-q)]^2 + d^2 [q(1-q)(1-2p)]^2 = a^2_d (H_p + H_q - 4H_p H_q),$$
where $H$ is heterozygosity.

The AA epistatic effect is given by $(a^2_d)^2 = a^2_d (H_p + H_q - 4H_p H_q), \quad V_G = a^2_d (H_p + H_q - 4H_p H_q), \quad V_D = a^2_d (H_p + H_q - 4H_p H_q), \quad V_G = a^2_d (H_p + H_q - 4H_p H_q)$.

Uninform: simple integration gives $E(V_A) = 2a^2/9, \quad E(V_{AA}) = a^2_d/9, \quad E(V_{G}) = a^2_d/3$.

Uniform: $E(V_A) = \frac{2a^2_d}{2x} \left[ \frac{1}{2x}p(1-p) \right] dp = a^2_d \left( \frac{1}{2x} - d_q \right)$.

Similarly $E(V_{AA}) = a^2_d/4K$. Hence $E(V_A)/E(V_G) = (2-4/K)/(2-3/K) = 1-1/(2K-3)$, which → 3 for large $K$. The residue, if any, is $V_H$.

Duplicate Factor Model. A simple epistatic model involving all epistatic components for two loci is the following:

- **CC**: $a$
- **Cc**: $a$
- **cc**: $0$

For an arbitrary number ($L$) of loci ($i$), the genotypic value is $a$ except for the multiple ‘recessive’ homozygote, and for one locus it is complete dominance.

$$M = a \left[ 1 - H_p (1-p)^2 \right] \quad \text{and} \quad V_G = M(a-M)$$

For $p = 0.5$ at all loci: $V_A = a^2_d (2p^2q^2 - q^4)2d^2$, $V_G = a^2_d (2p^2q^2 - q^4)2d^2$, and $V_D = 2G/2(2L-2-1)$. For two loci, $V_G = 2G/4(1-L)$. Uniform: $E(V_G) = a^2_d (1/(2k))^L$ and $E(V_A) = a^2_d (1/(2k))^L$.

For two loci, $E(V_A)/E(V_G) = 9/16$ and declines to 0 as $L$ increases.

$\mathcal{U}$: $E(V_G) = \frac{2a^2_d}{2x} \left[ \frac{1}{2x} (1-\frac{1}{x})^L - (1-\frac{1}{x})^L \right] \quad \text{and} \quad E(V_A) = \frac{a^2_d}{2x} \left( \frac{1-\frac{1}{x}}{x} \right)$

For large $N$, with two loci $E(V_A)/E(V_G) \to 4/3$ and for very many loci $E(V_A)/E(V_G) \to 0$.

Complementary Model. Another simple epistatic model involving all components is the following:

- **CC**: $a$
- **Cc**: $a$
- **cc**: $0$
- **BB**: $a$
- **Bb**: $a$
- **bb**: $0$

which can also be defined for multiple loci. For two loci, for example, using similar methods it can be shown that: for $p_i = 0.5, \quad V_A/V_G = 0.56$; with the uniform distribution, $E(V_A)/E(V_G) = 2/3$; and for the ‘U’ distribution $E(V_A)/E(V_G) = 1/(1+1/3x)/(1+x/2k)$.

Analyses of General Models. For two-locus models in which the genotypic values were entered as data, and $V_A$ and $V_G$ calculated as functions of the gene frequencies $p$ and $q$. Bivariate numerical integration was undertaken using Simpson’s rule by computing e.g. $V_G(a)/V_G(0)$ over an $(m+1) \times (n+1)$ grid of equally spaced $p$ and $q$ values, taking $m = 2^{10}$ or higher power of 2 as necessary for adequate convergence. Results were computed for some models of metabolic pathways [38,39] and for some published models obtained from QTL analysis [8].

Results/Discussion

Single Locus Model

Many general points are illustrated by two simple examples, the single locus model with dominance and the two locus model with AA interaction, so we consider these in more detail. For the single locus model with genotypic values for CC, Cc and cc of $+a$, $d$ and $-a$, respectively, $V_A = 2p(1-p)[a + d(1-2p)]^2$ and $V_D = 4p^2(1-p)^2d^2$.

For $d = a$, i.e. complete dominance of $C$, $V_A = 2p(1-p)a^2$ and $V_D = 4p^2(1-p)^2d^2$ and thus: at $p = 0.5, \quad V_A = (2/3) V_G$ if the dominant allele is rare (i.e. $p \to 0$), $V_G \to 8p^2$ and $V_A/V_G \to 1$, and if it is common, $V_G \to 4p^2$ and $V_A/V_G \to 0$. Note, however, that $V_G$ and $V_A$ are much higher when the dominant allele is at low frequency, e.g. 0.1, than are $V_G$ and $V_D$ when the recessive is at low frequency, e.g. $p = 0.9$. Even for an overdominant locus ($a = 0$), all genetic variance becomes additive at extreme gene frequencies. Considering now expectancies ($E$) over the frequency distributions, let $\eta^2 = E(V_A)/E(V_G)$, an equivalent to narrow sense heritability if $V_G = 0$. For the ‘U’ distribution, $\eta^2 = \frac{1}{2} - \frac{1}{x} + \frac{1}{x^2}$. Hence,
for a completely dominant locus, \( \eta^2 = 0.8 \) and \( \eta^2 = 0.75 \) respectively; whereas \( \bar{V}_A/\bar{V}_G = 0.67 \) for \( p = 0.5 \). In summary, the fraction of the genetic variance that is additive genetic decreases as the proportion of genes at extreme frequencies decreases (Table 2).

**Two Locus Additive \times Additive Model**

The genotypic values (see Theory section) for the simple AA model for double homozygotes BBCC and bbcc are +2a and for bbCC and BBcc are 0, and all single or double heterozygotes are intermediate (+o). With linkage equilibrium, \( \bar{V}_A/\bar{V}_G = 1 - H_2H_2/(H_1H_1 + H_2H_2 - 3H_1H_2) \), where the heterozygosities are \( H_1 \) and \( H_2 \) at loci B and C. Thus \( \bar{V}_A/\bar{V}_G \rightarrow 1 \) if either locus is at extreme frequency (i.e. \( p \) or \( q \) \( \rightarrow 0 \) or 1), and equals 0 when \( p = q = 0.5 \). If \( p = q \), for gene frequencies 0.1, 0.2, 0.3 and 0.4, \( \bar{V}_A/\bar{V}_G = 0.88, 0.69, 0.43 \) and 0.14. For the uniform distribution \( \eta^2 = 2/3 \), and for the ‘U’ distribution, the variances are a function of the population size, because more extreme frequencies are possible at larger population sizes. Thus \( \eta^2 = (2 - 4/K)/(2 - 3/K) \), where \( K = \ln(2N) \), so \( \eta^2 \rightarrow 1 \) for large \( K \). Any residue is \( \bar{V}_A \).

These two examples, the single locus and A \( \times \) A model, illustrate what turns out to be the fundamental point in considering the impact of the gene frequency distribution. When an allele (say C) is rare, so most individuals have genotype Cc or cc, the allelic substitution or average effect of C vs. c accounts for essentially all the differences found in genotypic values; or in other words the linear regression of genotypic value on number of C genes accounts for the genotypic differences (see [3], p 117). Hence almost all \( \bar{V}_G \) is accounted for by \( \bar{V}_A \).

**Other Epistatic Models**

With the ‘U’ distribution, most genes have one rare allele and so most variance is additive. Further examples (Table 2) illustrate this point, including the duplicate factor and complementary models where there is substantial dominance and epistasis. These models show mostly \( \bar{V}_A \) for the ‘U’ distribution for a few loci but the proportion of the variance which is additive genetic declines as the number increases. With many loci, however, such extreme models do not explain the covariance of sibs (i.e. any heritability) or the approximate linearity of inbreeding depression with inbreeding coefficient, \( F \), found in experiments [3,4,10,41,42], or the linearity in response to artificial selection [43].

We also analysed a well-studied systems biology model of flux in metabolic pathways [38,39,44] and found again that the expected proportion of \( \bar{V}_G \) that is accounted for by \( \bar{V}_A \) is large (Table 3).

**Examples of Models from Highly Epistatic Published QTL Analyses**

A number of QTL analyses using crosses between populations (some inbred, some selected) have been published in which particular pairs (or more) of loci have been identified to have positive epistatic effects. We also included a previously published model of flux in a metabolic pathway [23,44].

**Table 2.** Summary of expected proportion of \( \bar{V}_G \) that is \( \bar{V}_A \) for different models\(^a\).

<table>
<thead>
<tr>
<th>Genetic model</th>
<th>Distribution of allele frequencies</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>( p = % )</td>
</tr>
<tr>
<td>Dominance without epistasis ( d = %a )</td>
<td>0.89</td>
</tr>
<tr>
<td>Dominance without epistasis ( d = a )</td>
<td>0.67</td>
</tr>
<tr>
<td>Dominance without epistasis ( a = 0 )</td>
<td>0.00</td>
</tr>
<tr>
<td>A ( \times ) A without dominance</td>
<td>0.00</td>
</tr>
<tr>
<td>Duplicate factor 2 loci</td>
<td>0.27</td>
</tr>
<tr>
<td>Duplicate factor 100 loci</td>
<td>0.00</td>
</tr>
<tr>
<td>Complementary 2 loci</td>
<td>0.57</td>
</tr>
</tbody>
</table>

\(^a\)Models defined in Methods section
\(^b\)Population size

doi:10.1371/journal.pgen.1000008.t002

**Table 3.** Examples of expected proportion of \( \bar{V}_G \) that is \( \bar{V}_A \) in models of flux in linear metabolic pathways with a model flux \( J/S(\Sigma(1/E))^{-1} \)\(^c\) for a system with 10 loci in which 8 are invariant wild type and two (B, C) are mutants.

<table>
<thead>
<tr>
<th>Activities</th>
<th>Flux relative to wildtype, ( \bar{A}_{bbcc} = 1 )</th>
<th>( E(\bar{V}_A)/E(\bar{V}_G) )</th>
<th>Distribution of allele frequencies</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \bar{A}_{bb} )</td>
<td>( \bar{A}_{cc} )</td>
<td>( \bar{A}_{bbcc} )</td>
<td>( \bar{A}_{ccbb} )</td>
</tr>
<tr>
<td>1</td>
<td>0.1</td>
<td>0.92</td>
<td>1</td>
</tr>
<tr>
<td>0.5</td>
<td>0.1</td>
<td>0.90</td>
<td>0.91</td>
</tr>
<tr>
<td>0.1</td>
<td>0.1</td>
<td>0.86</td>
<td>0.53</td>
</tr>
<tr>
<td>0.1</td>
<td>0.01</td>
<td>0.85</td>
<td>0.53</td>
</tr>
</tbody>
</table>

Enzyme activities are \( E_i = 1 \) for loci 3 to 8, \( E_{bb} = E_{cc} = 1 \), values of \( E_{bb} \) and \( E_{cc} \) are listed, and heterozygotes are intermediate, e.g. \( E_{cc} = \%(1+E_{bb}) \), assuming gene frequency distributions as in Table 2. Flux modelled as [39].

\(^c\)Uniform

\(^d\)U-shaped with population size of 100

\(^e\)U-shaped with population size of 1000

doi:10.1371/journal.pgen.1000008.e003
substantial epistatic effects [8]. We consider examples of the more extreme cases of epistasis found, obtaining variance components by numerical integration. Results are shown in Table 4, for examples from [8] deliberately chosen as extreme. Even so, the proportion of the genetic variance that is additive is high with the ‘U’ distribution, except in the dominance × dominance example. Further, as these examples were selected by Carlborg and Haley and us as cases of extreme epistasis, it is not unreasonable to assume that the real epistatic effects are smaller than their estimates.

Relaxation of Assumptions

Expectation of a Ratio of Variance Components. The formulae we have given have been for the quantities $E(\mathbf{V}_A)$, $E(\mathbf{V}_G)$ and the ratio $E(\mathbf{V}_A)/E(\mathbf{V}_G)$. The quantity actually observed is $V_i/\mathbf{V}_G = \sum f_i V_i / \sum V_i$, where the expression denotes the sums over loci ($i$) of the additive and total genetic variance contributed by each in the absence of epistasis or linkage disequilibrium, or in the presence of these, sums over relevant sets of loci. As, for any locus, or for their sum, in general $E(V_i/\mathbf{V}_G) \neq E(\mathbf{V}_A)/E(\mathbf{V}_G)$, we need to consider the relevance of the quantities calculated. Whilst it would be possible to obtain approximations using statistical differentiation [4], formulae are complicated and invoke an assumption of small coefficients of variation of the quantities which are transient with random mating, they can be ignored, but LD can persist, and hence the estimated effects at locus C depend on those fitted at B and vice versa. The effect of LD is to reduce the number of haplotypes that segregate in the population so what would be epistatic variance becomes additive or dominance variance. For example, consider the $A \times A$ model and complete LD, i.e. equal frequencies at B and C loci and both loci segregating with only two haplotypes present. Then only Bc and bC haplotypes are present, and genotypic values are 0 for homozygous classes and $a$ for heterozygotes (‘pure’ overdominance), or only BC and bc haplotypes, with genotypic values $2a$ for homozygotes and $a$ for heterozygotes (‘pure’ underdominance). In either case

Table 4. Examples of expected proportion of $V_A$ that is $V_A$ in highly epistatic published QTL analyses assuming gene frequency distributions as in Table 2.

<table>
<thead>
<tr>
<th>Model</th>
<th>Genotypic values</th>
<th>$E(\mathbf{V}_A)/E(\mathbf{V}_G)$</th>
<th>Distribution of allele frequencies</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>BBCC</td>
<td>BbCC</td>
<td>bbCC</td>
</tr>
<tr>
<td>DomEp</td>
<td>4</td>
<td>10</td>
<td>15</td>
</tr>
<tr>
<td>Co-ad</td>
<td>39.0</td>
<td>38.7</td>
<td>35.7</td>
</tr>
<tr>
<td>$D \times D$</td>
<td>4</td>
<td>13</td>
<td>6</td>
</tr>
</tbody>
</table>

*Values obtained from tables or by interpolation from Box 1c–e of Carlborg and Haley [8]: key to their nomenclature: DomEp: Dominant epistasis (complex); Co-ad: Co-adaptive epistasis; $D \times D$: dominance × dominance epistasis.

Influence of Linkage Disequilibrium (LD). In this analysis we have assumed there is Hardy-Weinberg equilibrium (HWE) and linkage equilibrium among the loci. As departures from HWE are transient with random mating, they can be ignored, but LD can persist, and hence the estimated effects at locus C depend on those fitted at B and vice versa. The effect of LD is to reduce the number of haplotypes that segregate in the population so what would be epistatic variance becomes additive or dominance variance. For example, consider the $A \times A$ model and complete LD, i.e. equal frequencies at B and C loci and both loci segregating with only two haplotypes present. Then only Bc and bC haplotypes are present, and genotypic values are 0 for homozygous classes and $a$ for heterozygotes (‘pure’ overdominance), or only BC and bc haplotypes, with genotypic values $2a$ for homozygotes and $a$ for heterozygotes (‘pure’ underdominance). In either case variances are the same as for the dominance case with $a = 0$. Thus LD would lead to attribution of real epistatic variance to additive or dominance variance, and would exacerbate the results obtained from discussions of gene frequency distribution.

Consequences of Multiple Alleles. In these models we have considered solely biallelic loci, appropriate for low mutation rates. Multiallelic loci, in terms of their effects on the trait, can arise from mutations at different structural or control sites. Predictions are complicated by the need to consider $k(k-1)/2$ genotypic values at a $k$ allelic locus, and many further epistatic terms, so we consider two extreme cases. If the alleles all have similar effects, for example due to a knock-out, the effective mutation rate is increased, but it would require very many such sites for the distribution of frequencies of the trait alleles to differ greatly from proportionality to $1/[p(1-p)]$. Such segregation of multiple alleles will be more common in large populations, where in any case the frequency distribution is most extreme, and so the impact is unlikely to be large. A second case is where all alleles have

Table 5. Bias in use of $E(\mathbf{V}_A)/E(\mathbf{V}_G)$ rather than $E(\mathbf{V}_A)/E(\mathbf{V}_G)$ for some models in Table 2 as a function of Numbers of Loci.

<table>
<thead>
<tr>
<th>Uniform distribution</th>
<th>$E(\mathbf{V}_A)/E(\mathbf{V}_G)$</th>
<th>$E(\mathbf{V}_A)/E(\mathbf{V}_G)$ from simulation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Loci*</td>
<td>64</td>
<td>16</td>
</tr>
<tr>
<td>$a=1$, $d=1$</td>
<td>0.750</td>
<td>0.749</td>
</tr>
<tr>
<td>$a=0$, $d=1$</td>
<td>0.333</td>
<td>0.335</td>
</tr>
<tr>
<td>$A \times A$</td>
<td>0.667</td>
<td>0.667</td>
</tr>
<tr>
<td>Dupl. factor</td>
<td>0.562</td>
<td>0.559</td>
</tr>
<tr>
<td>’U’ distribution with $N=1000$</td>
<td>0.549</td>
<td>b</td>
</tr>
<tr>
<td>E(\mathbf{V}_A)/E(\mathbf{V}_G) from simulation</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Loci*</td>
<td>64</td>
<td>16</td>
</tr>
<tr>
<td>$a=1$, $d=1$</td>
<td>0.800</td>
<td>0.798</td>
</tr>
<tr>
<td>$a=0$, $d=1$</td>
<td>0.500</td>
<td>0.502</td>
</tr>
<tr>
<td>$A \times A$</td>
<td>0.918</td>
<td>0.918</td>
</tr>
<tr>
<td>Dupl. factor</td>
<td>0.746</td>
<td>0.743</td>
</tr>
</tbody>
</table>

*Number of loci for non-epistatic cases (complete dominance $a=1$, $d=1$), and overdominance $a=0$, $d=1$), numbers of pairs of loci for two-locus epistatic models ($A \times A$ and duplicate factor).

Not computed as $V_A=0$ in some replicates.
doi:10.1371/journal.pgen.1000008.0004
different effects and dominance interactions. Any allelic substitution then produces a change in the mean and so additive variance is present and for example, contributes more $V_A$ than does the overdominance model at $p = 0.5$.

**Alternative Models.** The analysis we have given for estimating effects of dominance and epistasis is for the classical method using simple averages over genotypes weighted by their frequencies, which are the least squares estimates in the balanced case and the basis for the analysis of variance [14,15,16]. There are alternative parameterisations aimed at exemplifying more clearly the nature of the interactions, including that of ‘physiological epistasis’ [45]. Whilst such alternatives may be of use in the analysis and interpretation of gene or QTL mapping experiments where individual genotypes can be identified or predicted from linked markers, such alternative parameterizations are not feasible in analysis of populations using data solely on the quantitative trait, from which the estimates of genetic variance components and heritability are obtained. Further, as has been pointed out [46], although the estimated effects may differ, the variances explained by different models are generally the same in segregating populations.

**Effects of Selection on Gene Frequency Distributions and Partition of Variance.** The ‘U’ and indeed uniform gene frequency distributions are limiting cases applying in the absence of selection on loci affecting the quantitative trait. The results for a wide range of models can be summarised as follows: gene frequencies that cause $V_A/V_G$ to be small also cause $V_G$ to be small. Consequently, when $V_A$ and $V_G$ are summed over a full range of frequencies, $V_A/V_G$ is large. This conclusion is dependent on the distribution of gene frequencies being symmetrical, so that cases with large $V_G$ and large $V_A/V_G$ are as common as cases with small $V_A$ and small $V_A/V_G$. The impact of selection will depend on how it acts on the trait or traits analysed and also on other aspects of fitness, so we need to consider whether the findings are robust to selection.

Stabilising selection on the trait, such that individuals with phenotype closest to an optimum are most fit, leads to maintenance of the population mean at or close to the optimum, so that mutants are at a disadvantage if they increase or decrease trait values. Consequently, the gene frequency distribution is still U-shaped, but with much more concentration near 0 or 1 [47]. Hence such selection is likely to increase proportions of additive variance. This conclusion would be wrong if there was widespread overdominance at the level of individual genes because this would push gene frequencies to intermediate values. However, the observed inbreeding depression is incompatible with widespread overdominance [48].

Under the neutral mutation or stabilising selection models where gene frequency distributions have extreme U shape, subsequent directional selection will lead to either rapid fixation or increase to intermediate frequency of genes affecting the trait. Even if the distribution of allele frequencies is initially symmetrical, a net increase in variance over generations might thus be expected [49] (Chapter 6). Accelerated responses to artificial selection have been seen, however, in lines founded from natural populations [50]. Calculations show that if genes are analysed independently such an increase in variance with artificial selection can in theory occur following the neutral model only if most gene effects are large (unpublished) or with more extreme frequency distributions following stabilising selection [51]. These ignore the build up of negative gametic disequilibrium through the Balmer effect [52], however, whereas in simulated multi-locus models of Drosophila no increase in variance was found [51]. Linkage effects would be weaker in species with more chromosomes, but selection lines in these have typically not been founded directly from natural populations.

Other types of selection do lead to an asymmetrical distribution of allele frequencies because the unfavourable allele will typically be at a low frequency. We have considered the case of genes whose effect on both the trait measured and on fitness shows complete dominance. Thus recessive and dominant favourable and unfavourable mutants were considered, and their expected contribution to variance computed during their lifetime to fixation or loss, using transition matrix methods. Results are given in Table 6 for population size ($N$) 100 and selective values ($s$) of the homozygote of $0.05$ ($N = 5$), but the qualitative result is not affected by using weaker or stronger selection. Deleterious, recessive mutations show the lowest $V_A/V_G$ but even here it is 0.44 and these cases also show the lowest total variance. Consequently, in a trait affected by a mix of genes with varying types of gene action, $V_A/V_G$ is likely to be well above 0.5.

Thus if the highest and lowest genotypic values correspond to multiple homozygous classes, it is clear that a high proportion of the variance is expected to be additive genetic even with selection. The potential exceptions occur when there is a maximum at intermediate frequencies, such as with an overdominant locus or some of the cases shown in Table 4. Nevertheless, few confirmed cases of clear overdominance/heterozygote superiority have been found (other than sickle cell anaemia) and the patterns in Table 4 are somewhat erratic.

**Effect of Population Size and Bottlenecks.** The theoretical analysis has been undertaken for large populations but much of the experimental data comes from livestock, laboratory animals and humans, all of which have experienced bottlenecks of reduced population size. As has been much explored, bottlenecks of population size are likely to change the proportion of variation that is additive, and for example to increase levels of $V_A$ for recessives at low frequency [53] and to ‘convert’ epistatic into additive variation [54,55,56,57,58], thereby increasing the ratio $V_A/V_G$. For example, for the additive × additive two locus model, the ratio of variances at inbreeding level $F$ in terms of values at $F = 0$ is $V_A/F/V_G(F) = (V_A + 4F V_{AA})/(V_A + 2V_{AA} + 3F V_{AA})$ for any gene frequency (using results of [54], but for loci with dominance or dominance interactions, $V_A/F/V_G(F)$ depends on gene frequency.

<table>
<thead>
<tr>
<th>Model</th>
<th>$s$(het)</th>
<th>$s$(hom)</th>
<th>$a$</th>
<th>$d$</th>
<th>$E(V_A)$</th>
<th>$E(V_A)/E(V_G)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Neutral dominant</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0.388</td>
<td>0.86</td>
</tr>
<tr>
<td>Neutral recessive</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>-1</td>
<td>0.166</td>
<td>0.66</td>
</tr>
<tr>
<td>Neutral random$^a$</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>-1</td>
<td>0.277</td>
<td>0.80</td>
</tr>
<tr>
<td>Deleterious dominant</td>
<td>-0.05</td>
<td>-0.05</td>
<td>1</td>
<td>1</td>
<td>0.145</td>
<td>0.97</td>
</tr>
<tr>
<td>Deleterious recessive</td>
<td>0</td>
<td>-0.05</td>
<td>1</td>
<td>-1</td>
<td>0.052</td>
<td>0.44</td>
</tr>
<tr>
<td>Advantageous dominant</td>
<td>0.05</td>
<td>0.05</td>
<td>1</td>
<td>1</td>
<td>0.375</td>
<td>0.74</td>
</tr>
<tr>
<td>Advantageous recessive</td>
<td>0</td>
<td>0.05</td>
<td>1</td>
<td>-1</td>
<td>0.151</td>
<td>0.71</td>
</tr>
</tbody>
</table>

$^a$e.g., if the mutant gene is completely recessive for the trait and for fitness, $d = -a$ and $s$(hom) = 0.

$^b$Equally likely to be completely dominant or recessive mutants, hence values as in Table 2.

Table 6. Expected variance contributed by mutant genes before fixation for population size 100, specified dominance on the quantitative trait ($a$ vs $d$) and selective (dis)advantage ($s$ in heterozygote and homozygote).
This occurs because the bottleneck leads to the dispersal of gene frequencies and the reduction in mean heterozygosity, so for the AA model, if frequencies are initially intermediate (e.g. 0.5) there is a substantial increase in $V_A/V_G$, whereas if frequencies initially follow the ‘U’ distribution, there is little $V_{AA}$ initially, total variance falls and the level of dispersion and $V_A/V_G$ do not increase appreciably. Indeed, for a population that starts with the gene frequency distribution U-shaped, the loss of heterozygosity is due to fixation. Among the genes that remain segregating the distribution of gene frequencies flattens considerably, and in the absence of new mutation approaches the uniform distribution which has a lower ratio of $V_A/V_G$ than the ‘U’ distribution.

However, despite this, $V_{AA}$ declines faster than $V_A$ because, as loci become fixed, the number of pairs of segregating loci declines faster than the number of segregating loci. Thus it is not obvious what effect the bottlenecks in livestock, laboratory or human populations have had on the ratio $V_{AA}/V_G$. We suspect it has not been large because, if a large reduction in heterozygosity had occurred, these populations would show low genetic variance and there is no indication that this is the case. In any case, the results show that the conclusion that most genetic variance is additive is fairly robust to assumptions about the distribution of gene frequencies, for instance the ‘U’ and uniform distributions both lead to qualitatively the same conclusion.

### Evidence for the Effect of Gene Frequency on Variance Components

A test of the hypothesis that the lack of non-additive variance observed in populations of humans or animals is because gene frequencies near 0.5 are much less common than those more extreme, not because non-additive effects are absent, is to compare variance components among populations with different gene frequency profiles. For crops such as maize and for laboratory animals, estimates can be got both from outbreds and from populations with gene frequencies of one-half derived from crosses of inbred lines. There are a limited number of possible contrasts and linkage confounds comparisons of variation in F$_2$ and later “inter se” generations, however, so it is difficult to partition variation between single locus and epistatic components (e.g. [17] ch. 7).

The most extensive data are on yield traits in maize. The magnitudes of heritability and of dominance relative to additive variance estimated for different kinds of populations in a substantial number of studies (including 24 on F$_2$ and 27 on open-pollinated, i.e. outbreds) have been summarised [59]. Average estimates of $h^2$ were 0.19 for open-pollinated populations, 0.23 for synthetics from recombination of many lines, 0.24 for F$_2$ populations, 0.13 for variety crosses and 0.14 for composites. Estimates of $V_A/V_G$ (from tabulated values of $V_{AA}/V_A$ [59]) were 0.57, 0.55, 0.50, 0.42 and 0.43, respectively, which are inconclusive but indicate relatively more dominance variance at frequencies of 0.5. Analyses of the magnitude of epistasis at the level of effects, rather than variance, do not provide consistent patterns. For example, in two recent analyses of substantial data sets of F$_2$ populations of maize, one found substantial epistasis [60] and the other almost none [61]. In an analysis of a range of traits in recombinant inbred lines, F$_2$ and triple test crosses [62] in Arabidopsis thaliana, there was substantial additive genetic and dominance variance for all traits, with most estimates of $V_{AA}/V_A$ in the range 0.3 to 0.5, essentially no significant additive *x* additive epistatic effects, but several cases of epistasis involving dominance [63].

Although there does appear to be more dominance variance in populations with gene frequencies of one-half than with dispersed frequencies, from these results we cannot reject or accept the hypothesis that there is relatively much more epistatic variance in such populations. One explanation is indeed that there is not a vast amount of epistatic variance in populations at whatever frequency, although another is that maize has unusually small amounts of epistasis. Many additive QTL were identified in an analysis of a line derived from the F$_2$ of highly divergent high and low oil content lines from the long term Illinois maize selection experiment, but with almost no evidence of epistasis or indeed dominance effects [64]. In contrast, an F$_2$ of divergent lines of long-term selected poultry and an F$_2$ from inbred lines of mice showed evidence of highly epistatic QTL effects for body weight [65,66]. We do not claim to understand these different results, but as has been pointed out [67,68], QTL with significant epistatic interaction effects might not represent the majority of QTL with small effects contributing to gene networks.

### Conclusions and Consequences

We have summarised empirical evidence for the existence of non-additive genetic variation across a range of species, including that presented here from twin data in humans, and shown that most genetic variance appears to be additive genetic. There are two primary explanations, first that there is indeed little real dominant or epistatic gene action, or second that it is mainly because allele frequencies are distributed towards extreme values, as for example in the neutral mutation model. Complete or partial dominance of genes is common, at least for those of large effect; and epistatic gene action has been reported in some QTL experiments [8,69]. Detailed analyses in Drosophila melanogaster, using molecular and genetic tools available for it, identify substantial amounts of epistasis, including behavioural traits [70] and abdominal bristle number [71], yet most genetic variation in segregating populations for bristle number appears to be additive (as noted above). But many QTL studies of epistatic gene action suffer from a high degree of multiple testing, increasingly so the more loci and orders of interaction are included, such that they may be exaggerating the amount of epistasis reported. On the assumption that many of the effects are indeed real, we have turned our attention to the second explanation.

The theoretical models we have investigated predict high proportions of additive genetic variance even in the presence of non-additive gene action, basically because most alleles are likely to be at extreme frequencies. If the spectrum of allele frequencies is independent of which are the dominant or epistatic alleles, $V_A/V_G$ is large for almost any pattern of dominance and epistasis because $V_A/V_G$ is low only at allele frequencies where $V_G$ is low, and so contributes little to the total $V_G$. The distribution of allele frequencies is expected to be independent of which are the dominant or epistatic alleles for neutral polymorphisms; but under natural selection the favourable allele is expected to be common and lead to high or low $V_A/V_G$ depending on whether it is dominant (low $V_A$) or recessive (high $V_A$). The equivalent case for epistasis is that all genotype combinations except one is favourable (low $V_{AA}$) vs. only one genotype combination is favourable (high $V_{AA}$).

If genetic variation in traits associated with fitness is due almost entirely to low frequency, deleterious recessive genes which are unresponsive to natural selection, these traits would show low $V_A/V_G$. However, neither the empirical evidence nor the theory supports this expectation. There seems to be substantial additive genetic variance for fitness associated traits [21] and fitness itself [30,31,72]. Although heritabilities for such traits may be low, they show high additive genetic coefficient of variation (evolvability) [29], and the correlation of repeat records is typically little higher than the heritability (e.g., litter size in pigs), indicating that $V_A/V_G$
is one-half or more. In agreement with this, when the life history of deleterious, recessive mutants was modelled, $V_A/V_G$ was found to be 0.44 (Table 6), basically because rare recessives contribute so little variance, albeit most is $V_G$ in non-inbred populations.

We believe we have a plausible gene frequency model to explain the near-absence of non-additive genetic and particularly epistatic variance. What consequences do our findings have? For animal and plant breeding, maintaining emphasis on utilising additive variation by straightforward selection remains the best strategy. For gene mapping, our results imply that $V_A$ is so large we should be able to detect and identify alleles with a significant gene substitution effect within a population. Such variants have been reported from genome-wide association studies in human populations [9,10,11,12,13]. Although there may well be large non-additive gene effects, the power to detect gene-gene interactions in outbred populations is a function of the proportion of variance they explain, so it will be difficult to detect such interactions unless the effects are large and the genes have intermediate frequency.

Thus we expect that the success in replicating reported epistatic effects will be even lower than it is for additive or dominance effects, both because multi-locus interactions will be estimated less accurately than main effects and because they explain a lower proportion of the variance. Finally, if epistatic effects are real, gene substitution effects may vary widely between populations which differ in allele frequency, so that significant effects in one population may not replicate in others.

Acknowledgments

We thank Bernardo Ordas for helpful information and Nick Martin, Naomi Wray, Nick Barton, and referees for comments or discussion on previous versions of the manuscript.

Author Contributions

Conceived and designed the experiments: WH MG PV. Analyzed the data: WH PV. Wrote the paper: WH MG PV.

References