

Dissolving the missing heritability problem

Pierrick Bourrat & Qiaoying Lu

Abstract: Heritability estimates obtained in genome-wide association studies (GWAS) are much lower than those of traditional quantitative methods. This has been called the “missing heritability problem”. By analyzing and comparing these two kinds of methods, we first show that the estimates obtained by traditional methods involve some terms that GWAS do not. Second, the estimates obtained by GWAS do not take into account epigenetic factors transmitted across generations, whilst they are included in the estimates of traditional quantitative methods. Once these two factors are taken into account, we show that the missing heritability problem can be largely dissolved. Finally, we briefly contextualize our analysis within a current discussion on how non-additive factors relate to the heritability estimates in GWAS.

1. Introduction.

One pervasive problem encountered when estimating the heritability of quantitative traits is that the estimates obtained from Genome-Wide Association Studies (GWAS) are much smaller than that calculated by traditional quantitative methods. This problem has been called the missing heritability problem (Turkheimer 2011). Take human height for example. Traditional quantitative methods deliver a heritability estimate of about 0.8, while the first estimates using GWAS were 0.05 (Maher 2008). More recent GWAS methods have revised this number and estimate the heritability of height to be at most 0.45 (Yang et al. 2010; Turkheimer 2011). Yet, half of the heritability is still missing.

In quantitative genetics, heritability is defined as the portion of phenotypic variation in a population that is caused by genetic difference (Downes 2015). Traditionally, this portion is estimated by measuring the phenotypic resemblance of genetically related individuals without identifying at the molecular level (more particularly the DNA level) the genetic causes of phenotypic variation. GWAS have been developed in order to locate the DNA sequences that influence the target trait and estimate their effects, especially for common complex diseases such as obesity, diabetes and heart disease (Visscher et al. 2012; Frazer et al. 2009). As for height, almost 300 000 common DNA variants in human populations that associate with it have been identified by GWAS (Yang et al. 2010). Granted by many that the heritability estimates obtained

by traditional quantitative methods are quite reliable, the method(s) used in GWAS have been questioned (Eichler et al. 2010).

A number of partial solutions to the missing heritability problem have been proposed, with most of them focusing on improving the methodological aspects of GWAS in order to provide a more accurate estimate (e.g., Manolio et al. 2009; Eichler et al. 2010). Some authors have also suggested that heritable epigenetic factors might account for part of the missing heritability. For instance, in Eichler et al. (2000, 488), Kong notes that “[e]pigenetic effects beyond imprinting that are sequence-independent and that might be environmentally induced but can be transmitted for one or more generations could contribute to missing heritability.” Furrow et al. (2011) also claim that “[e]pigenetic variation, inherited both directly and through shared environmental effects, may make a key contribution to the missing heritability.” Others have made the same point (e.g., McCarthy and Hirschhorn 2008; Johannes et al. 2008). Yet, in the face of this idea one might notice what appears to be a contradiction: how can *epigenetic* factors account for the missing heritability, if the heritability is about *genes*?

To answer this question as well as to analyze the missing heritability problem, we compare the assumptions underlying both heritability estimates in traditional quantitative methods and those in GWAS. We argue that a) the heritability estimates of traditional methods include some terms associated with broad-sense heritability (H^2), as opposed to narrow-sense heritability (h^2); b) although GWAS are supposed to get h^2 , h^2 relies on an evolutionary concept of the gene

that can include epigenetic factors while heritability estimates obtained from GWAS do not. With these two points being illustrated, we expect the missing heritability problem to be largely dissolved as well as setting the stage for further discussions.

The remainder of the paper will be divided into three parts. First, we briefly introduce how heritability is estimated in two traditional methods, namely twin studies and parent-offspring regression. We show that the estimates obtained by each methods include *some* non-additive elements and consequently correspond neither to H^2 nor to h^2 , but to a notion in between which we term “broader-sense heritability”. Second, we outline the basic rationale underlying GWAS and illustrate that they estimate heritability by considering solely DNA variants. By arguing that the notion of additive genetic variance does not necessarily refer to DNA sequences but can also refer to epigenetic factors in traditional quantitative methods, we show that the notion of heritability estimated in GWAS is more restrictive than that of traditional quantitative methods, and term this notion “DNA-based narrow-sense heritability”. Finally, in Section 4, based on the conclusions from Section 2 and Section 3, we claim that the gap between the heritability estimates of traditional quantitative methods and those of GWAS can be explained away in two major ways. One consists in recognizing that if non-additive variance was removed from the estimates obtained via traditional methods, they would be lower. The other consists in recognizing that if epigenetic factors were taken into account by GWAS, the heritability estimates obtained would be higher. We conclude Section 4 by showing how our analysis sheds

some light on a discussion about the role played by non-additive factors in the missing heritability problem. Because human height has been “the poster child” of the missing heritability problem (Turkheimer 2011, 232), we will use this example to illustrate each of our points.

2. Heritability in Traditional Quantitative Methods.

According to quantitative genetics, the phenotypic variance (V_P) of a population can be explained by two components, its genotypic variance (V_G) and its environmental variance (V_E).

In the absence of gene-environment interaction and correlation, we thus have:

$$V_P = V_G + V_E \quad (1)$$

From there broad-sense heritability (H^2) is defined as:

$$H^2 = \frac{V_G}{V_P} \quad (2)$$

V_G can further be portioned into the additive genetic variance (V_A), the dominance genetic variance (V_D) and the epistasis genetic variance (V_I). We have:

$$V_P = V_A + V_D + V_I + V_E \quad (3)$$

where V_A is the variance due to hypothetical genes making an equal and additive contribution to the trait studied (e.g., height). V_D is the variance due to interactions between alleles at one locus for diploid organisms, and V_I is the variance due to interactions between alleles from different loci. V_D and V_I together represent the variance due to particular combinations of genes of an organism.

Since genotypes of sexual organisms recombine at each generation via reproduction, dominance and epistasis effects are not transmitted stably across generations, only additive genetic effects are. Therefore, V_A is the variance due to stably transmitted genetic effects. Narrow-sense heritability (h^2) measures to what extent variation in phenotypes is determined by the variation in genes transmitted from parent(s) to offspring (Falconer and Mackay 1996, 123). It is defined as:

$$h^2 = \frac{V_A}{V_P} \quad (4)$$

h^2 is important in breeding studies and is used by evolutionary theorists who are interested in making evolutionary projections of a trait within a population across generations.

To know h^2 , both V_A and V_P must be known. V_P , for most quantitative traits (including height), can be directly estimated by measuring individuals. However, there is no direct way to estimate V_A in traditional quantitative methods. The traditional way to estimate it requires two elements. First, one needs a population-level measure of a phenotypic resemblance of family

relative pairs¹. This measure is obtained by calculating the *covariance* of the phenotypic values for those pairs. The choice of what sort of relatives to use depends on what data is available. The second element is the genetic relation between family pairs. It indicates the percentage of genetic materials the pairs are expected to share. With these two elements, one can estimate how much the genes shared contribute to the phenotypic resemblance. In a large population with different phenotypes, one can then estimate how much the additive genetic difference contributes to phenotypic difference in this population, which estimates h^2 .

For simplicity, traditional quantitative methods usually assume that there is neither gene-environment interaction nor correlation (Falconer and Mackay 1996, 131). Thus the covariance between the phenotypic values (e.g., height) of pairs equals to additive genetic covariance, dominant and epistasis genetic covariance, plus the environmental covariance. A general equation for traditional quantitative methods can be written as follows:

$$Cov(P_1, P_2) = Cov(A_1 + D_1 + I_1 + E_1, A_2 + D_2 + I_2 + E_2) = \\ Cov(A_1, A_2) + Cov(D_1, D_2) + Cov(I_1, I_2) + Cov(E_1, E_2) \quad (5)$$

where indexes “1” and “2” represent the two family members for each pair studied. $Cov(P_1, P_2)$ is the covariance between the phenotypic values of one individual with the other.

¹ Or the mean values of their class (e.g., offspring) depending on the particular method used.

A , D , I and E represent additive effects, dominant effects, epistasis effects and environmental effects respectively.

The most commonly used traditional methods for estimating heritability are twin studies. In these studies one already knows that monozygotic twins share almost 100% of their genetic material while dizygotic twins about 50%. The environment is typically divided into the part of the environment that affects both twins in the same way (the shared environment, C) and the part of the environment that affects one twin but not the other (the unique environment, U) (Silventoinen et al. 2003). Hence, in the absence of interaction and correlation between C and U , we have:

$$E = C + U \quad (6)$$

Assuming epistasis effects to be negligible (a common assumption in twin studies), by inserting Equation (6) into Equation (5), we have:

$$\begin{aligned} Cov(P_{T1}, P_{T2}) = Cov(A_{T1} + D_{T1} + C_{T1} + U_{T1}, A_{T2} + D_{T2} + C_{T2} + U_{T2}) = \\ Cov(A_{T1}, A_{T2}) + Cov(D_{T1}, D_{T2}) + Cov(C_{T1}, C_{T2}) + Cov(U_{T1}, U_{T2}) \quad (7) \end{aligned}$$

where indexes “T1” and “T2” represent the two twins for each twin pair studied.

$Cov(P_{T1}, P_{T2})$ is the covariance between the phenotypic values of one twin with the other.

Because each twin's unique environment by definition is independent of that of the other twin, $Cov(U_{T1}, U_{T2})$ is zero for both monozygotic and dizygotic twins. Given that variance is a special case of covariance when the two variables are identical, and that for monozygotic twins A_{T1} , D_{T1} , and C_{T1} equal to A_{T2} , D_{T2} , and C_{T2} respectively, we can formulate the equation from Equation (7) as follows:

$$Cov_{MT}(P_{T1}, P_{T2}) = V_A + V_D + V_C \quad (8)$$

where $Cov_{MT}(P_{T1}, P_{T2})$ is the covariance between the phenotypic values of monozygotic twin pairs studied.

By contrast, dizygotic twins are expected to share half of their genes, which means that the covariance between the phenotypic values of one twin with the other of dizygotic twin pairs studied ($Cov_{DT}(P_{T1}, P_{T2})$) is expected to be equal to half of the additive genetic variance, a quarter of dominant variance², and all of the shared environmental variance (with $Cov(U_{T1}, U_{T2})$ also to be zero). We have:

$$Cov_{DT}(P_{T1}, P_{T2}) = \frac{1}{2}V_A + \frac{1}{4}V_D + V_C \quad (9)$$

It is classically assumed that V_C in Equation (8) and (9) is the same. That is to say, for both monozygotic and dizygotic twin pairs, it is assumed that the shared environment would act in

² For each given gene with two alleles, the possibility that dizygotic twins have the same genotype is one quarter.

the same way if the pair has been reared together.³ V_C can be cancelled by subtracting Equation (9) from Equation (8). The heritability can then be estimated as follows:

$$h_{b_{TS}}^2 = \frac{2\{Cov_{MT}(P_{T1}, P_{T2}) - Cov_{DT}(P_{T1}, P_{T2})\}}{V_P} = \frac{V_A}{V_P} + \frac{\frac{3}{2}V_D}{V_P} \quad (10)$$

We call $h_{b_{TS}}^2$ broader-sense heritability (the index “b” is for “broader-sense”) from *twin studies*, because the resulting estimate (which is about 0.8 for height) provides an accurate estimate of neither H^2 nor h^2 , although it is closer to H^2 than to h^2 (Falconer and Mackay 1996, 172). That is to say, it corresponds to a definition of heritability that includes *some* elements of broad-sense heritability but not all of it.

Another often used traditional quantitative method to estimate heritability involves a parent-offspring regression. This method also assumes neither gene-environment interaction nor correlation, the covariance between the height of parents (one or the mean of both) and the mean of their offspring (Falconer and Mackay 1996, 164), equals to additive genetic covariance, dominant covariance (the epistasis covariance is assumed to be small and is not included), plus environmental covariance. Hence, Equation (5) can be formulated as follows:

³ This assumption might be problematic because monozygotic twins are often treated more similarly by their parents than are dizygotic twins, and monozygotic twins are more likely to share a placenta than dizygotic twins. The difficulty can be mitigated by using adoption twin studies in which the environments for twins are random on average. But large adoption twins’ data are exceedingly difficult to get (Griffiths 2005).

$$\begin{aligned}
Cov(P_P, P_O) &= Cov(A_P + D_P + I_P + E_P, A_O + D_O + I_O + E_O) = \\
&Cov(A_P, A_O) + Cov(D_P, D_O) + Cov(E_P, E_O)
\end{aligned} \tag{11}$$

where indexes “P” and “O” represent the “parents” and the “offspring”.

Two assumptions are then made. The first one is that there is no dominant effects transmitted from the parents to the offspring assuming the parents are unrelated (Doolittle 2012, 178), which means $Cov(D_P, D_O)$ is nil. Another assumption is that there is no correlation between the parents’ environment and the offspring’s environment so that $Cov(E_P, E_O)$ in Equation (11) is also nil. Given that on average, parents share in expectation 50% of genes with their offspring (parents and offspring share half of their genes), it leaves Equation (11) with a result of half of additive genetic variance ($\frac{1}{2}V_A$). Given V_P , h^2 can be estimated straightforwardly.

But the above two assumptions are problematic. First, the assumption of unrelated parents might be violated because of assortative mating in humans resulting in parents to be more genetically similar than two randomly chosen individuals (Guo et al. 2014). Hence, $Cov(D_P, D_O)$ is likely to be non-nil. Second, because the environments experienced by individuals are likely to be more similar within a family line, $Cov(E_P, E_O)$ might not be nil, either. If we take these two factors into consideration, the covariance of the parents and their

offspring is equal to half of additive genetic variance, *plus* a variance term representing effects due to dominance and similarities between environments. This can be written formally as:

$$Cov(P_P, P_O) = Cov(A_P, A_O) + Cov(D_P, D_O) + Cov(E_P, E_O) = \frac{1}{2}V_A + V_{D\&EC} \quad (12)$$

where $V_{D\&EC}$ represents the variance due to some dominance and environmental correlation effects between the parents and the offspring studied.

The heritability can then be estimated by doubling the parent–offspring covariance in Equation (12) and dividing the total phenotypic variance of the population as follows:

$$h_{bPOR}^2 = \frac{2Cov(P_P, P_O)}{V_P} = \frac{V_A}{V_P} + \frac{2V_{D\&EC}}{V_P} \quad (13)$$

For similar reasons as with the heritability estimates from twin studies, we call h_{bPOR}^2 broader-sense heritability (with the index “b” also being for “broader-sense”) from *parent-offspring regression*. Indeed, although it is often assumed that h_{bPOR}^2 represent h^2 (Falconer and Mackay 1996, 147), the resulting estimate (also about 0.8 for height) is broader than h^2 as it can include a component led by dominance variance and environmental correlation between parent and offspring.

To conclude this section, heritability estimates in both twin studies and parent-offspring regression include an extra term when compared to h^2 , but they do not correspond to H^2 . For this reason we regroup them under the term h_b^2 for “broader-sense heritability”, such that:

$$h_b^2 = h^2 + h_{other}^2 \quad (14)$$

where h_{other}^2 is the part of heritability contributed by the extra component(s) representing non-additive variance.

3. Heritability in GWAS.

Although any two unrelated individuals share about 99.5% of their DNA sequences, their genomes differ at specific nucleotide locations (Aguilar and Istrail 2013). Given two DNA fragments at the same locus of two individuals, if these fragments differ at a single nucleotide, they represent two variants of a Single Nucleotide Polymorphism (SNP). GWAS focus on SNPs across the whole genome that occur in the population with a probability larger than 1% which are called common SNPs. If one variant of a common SNP, compared to another one, is associated with a significant change on the trait studied, then this SNP is a marker for a DNA region (or a gene) that leads to phenotypic variation. For a polygenic trait like height, if we can detect all the SNPs that associate with it, then all the DNA difference makers that determine height difference can be located.

The development of commercial SNP chips makes it possible to rapidly detect common SNPs of DNA samples from all the participants involved in a study. By using a series of statistical tests, it can be investigated at the population level whether each SNP associates with

that target trait. The choice of the statistical tests depends on the data available as well as the trait studied. For quantitative traits like height, the most common approach is to make an analysis of variance table and assess whether the mean height of a group with one variant at one nucleotide is significantly different from the group with another variant of the same SNP⁴ (Bush and Moore 2012). With all the SNPs associated with height being detected, data from the HapMap project, which provides a list of SNPs that are markers for most of the common DNA variants in human populations (Consortium, International HapMap 3 2010), is used to map the associated SNPs with common DNA variants. These mapped DNA variants, to be distinguished from DNA variants that do not affect the target trait, have been called “causal variants” (Visscher et al. 2012).

Based on the readings of SNP chips as well as further independent tests for SNPs, the effects of the associated SNPs (markers for causal DNA variants) on the trait can be calculated. By estimating the phenotypic variance contributed by these SNPs and the total phenotypic variance of the population, the heritability of causal DNA variants can be estimated as the ratio of the phenotypic variance caused by all the associated SNPs compared to the total phenotypic variance of the population (Weedon et al. 2008). Since it is common for biologists to assume

⁴ For categorical (often binary disease/control) traits, the association test used involves measuring an odds ratio, namely the ratio of the odds of disease for individuals having a specific variant of a SNP, and the odds of disease for individuals who have another variant at the same locus. If the odds ratio of a common SNP is significantly different from 1, then that SNP is considered to be associated with the disease (Bush and Moore 2012).

that genes are only made up of pieces of DNA, it is thought that the variance obtained from all the causal DNA variants represent exactly the additive genetic variance, and the heritability estimated by GWAS should match narrow-sense heritability (h^2) (Yang et al. 2010; Visscher et al. 2006). However, the assumption that additive genetic effects are solely based on DNA sequences is problematic when faced with the evidence of epigenetic inheritance.

As was mentioned in Section 2, traditional quantitative methods for estimating heritability are based on measuring phenotypic values and genetic relations without reaching the molecular level. The genes are not defined physically, but functionally as heritable difference makers (Falconer and Mackay 1996, 123). In other words, they are theoretical units defined by their effects on the phenotype. With the discovery of DNA structure in 1953, it was thought that the originally theoretical genes were found in the physical DNA molecules. Since then, biologists commonly refer to genes as DNA molecules and this assumption is also made by researchers of GWAS. As Lu and Bourrat claim, this step was taken too hastily. If there is physical material, other than DNA pieces, that can affect the phenotype and be transmitted stably across generations, then it should also be thought to play the role that contributes to additive genetic effects.

Many studies have provided evidence for epigenetic inheritance⁵, namely the stable transmission of epigenetic modifications across multiple generations and affect organism's traits

⁵ We use the notion of “epigenetic inheritance” in the broad sense that refers to the inheritance

(e.g., Youngson and Whitelaw 2008; Dias and Ressler 2014). A classical example of this is the methylation pattern on the promoter of the agouti gene in mice (Morgan et al. 1999). It shows that mice with the same genotype but different methylation levels display a range of colors of their fur, and the patterns of DNA methylation can be inherited through generations causing heritable phenotypic variations. Epigenetic factors such as self-sustaining loops, chromatin modifications and three-dimensional structures in the cell can also be transmitted over multiple generations (Jablonka et al. 2014). Studies on various species suggest that epigenetic inheritance is likely to be ‘ubiquitous’ (Jablonka and Raz 2009).

The increasing evidence of epigenetic inheritance seriously challenges the restriction of the concept of the gene in the evolutionary sense to be materialized only in DNA. Relying on traditional quantitative methods, it is impossible to distinguish whether additive genetic variance is DNA based or based on other material(s). Some transmissible epigenetic factors, which are not DNA based, might *de facto* be included into the additive genetic variance used to estimate h^2 . This extension of heritable units also echoes to the recent suggestion that genetic (assuming genes to be DNA based) and non-genetic heredity should be unified in an inclusive inheritance theory (Danchin 2013; Day and Bonduriansky 2010).

of phenotypic features via causal pathways other than the inheritance of nuclear DNA (Griffiths and Stotz 2013, 112).

To apply the idea that some epigenetic factors can lead to additive genetic effects, the additive variance of them ($V_{A_{epi}}$) should be added to the additive variance of DNA sequences ($V_{A_{DNA}}$) to obtain V_A . Assuming there is no interaction between $V_{A_{epi}}$ and $V_{A_{DNA}}$, we have:

$$V_A = V_{A_{DNA}} + V_{A_{epi}} \quad (15)$$

Inserting Equation (15) to Equation (4) leads to:

$$h^2 = \frac{V_{A_{DNA}}}{V_P} + \frac{V_{A_{epi}}}{V_P} \quad (16)$$

Here we term the first term on the right side of Equation (16) “DNA-based narrow-sense heritability” (h_{DNA}^2), and the second term “epigenetic-based narrow-sense heritability” (h_{epi}^2), we thus have:

$$h_{DNA}^2 = h^2 - h_{epi}^2 \quad (17)$$

4. Dissolving the Missing Heritability.

As we mentioned it in Introduction, since the first successful GWAS was published in 2005 (Klein et al. 2005), there have been a lot of proposals for methodological improvements in GWAS (Manolio et al. 2009; Eichler et al. 2010). Studies have been conducted according to those proposals that permit to obtain higher heritability estimates. Examples include increasing

the sample sizes which has resulted in more accurate estimates (e.g., Wood et al. 2014), considering all common SNPs simultaneously instead of one by one which has increased the heritability estimates of height from 0.05 to 0.45 (see Yang et al. 2010), and conducting meta-analyses which can lead to more accurate results when compared to single analysis (see Bush and Moore 2012). Biologists have also suggested to search for SNPs with lower frequencies than 1% in order to account for a wider range of possible causal variants (Schork et al. 2009).

Aside from these partial improvements, our analysis reveals two reasons explaining away the missing heritability problem: a) In traditional quantitative methods, the heritability estimates include extra terms which are not presented in GWAS; b) In GWAS, heritability is estimated solely from causal DNA variants, while in traditional quantitative methods the additive effects contributed by epigenetic difference (h_{epi}^2) are *de facto* included in the estimates.

These two reasons can be shown formally. Using our terminology, missing heritability (MH) equals to the estimates obtained by traditional quantitative methods (h_b^2) minus the estimates obtained by GWAS (h_{DNA}^2), which are 0.8 and 0.45 respectively in the case of height.

Thus we have:

$$MH = h_b^2 - h_{DNA}^2 \quad (18)$$

Replacing h_b^2 and h_{DNA}^2 by the right hand side of Equation (14) and (17), we obtain:

$$MH = h_b^2 - h_{DNA}^2 = h^2 + h_{other}^2 - (h^2 - h_{epi}^2) = h_{other}^2 + h_{epi}^2 \quad (19)$$

Which means that the missing heritability results from the part of heritability originating from epigenetic factors stably transmitted across generations, plus the part of heritability originating from non-additives factors.

Our point that part of the missing heritability can be dissolved by considering non-additive effects echoes to the claim that almost all GWAS to date have focused on additive effects might be a reason for the missing heritability (McCarthy and Hirschhorn 2008). Although there is not enough data to confirm that non-additive effects do explain away some part of missing heritability, this claim appears again and again in discussions on the missing heritability problem (see for instance Maher 2008; Frazer et al. 2009; Gibson 2010; Kong 2010; Moore 2010). Yang et al. (2010, 565) disagree with this claim and respond that “[n]on-additive genetic effects do not contribute to the narrow-sense heritability, so explanations based on non-additive effects are not relevant to the problem of missing heritability.”

We agree with Yang et al. (2010) that non-additive effects do not contribute to h^2 . That said, because the heritability estimates obtained from traditional quantitative methods do not strictly correspond to h^2 but include some non-additive elements, non-additive effects cannot be dismissed as irrelevant for the missing heritability problem, though probably they are relevant in a way that both Yang et al. (2010) as well as their opponents did not consider.

5. Conclusion.

We have provided two ways in which the missing heritability problem can be explained away. First, heritability estimates from traditional quantitative methods (h_b^2) are overestimated when compared to h^2 . The resulting estimates would be smaller if the non-additive elements were eliminated. Second, heritability estimates from GWAS (h_{DNA}^2) are underestimated when compared to h^2 because they do not take into account the additive effects of epigenetic factors behaving like evolutionary genes. The resulting estimates would be larger if epigenetic factors were taken into account. We have voluntarily stayed away from the question of whether heritability should be defined strictly relative to DNA sequences or if it should encompass any factors behaving effectively like an evolutionary gene. Our inclination is that there is no principled reason to exclude non-DNA transmissible factors from heritability measures, but our analysis does not bear on this choice.

References:

- Aguiar, Derek, and Sorin Istrail. 2013. "Haplotype Assembly in Polyploid Genomes and Identical by Descent Shared Tracts." *Bioinformatics* 29 (13): i352–i360.
- Bush, William S., and Jason H. Moore. 2012. "Genome-Wide Association Studies." *PLoS Computational Biology* 8 (12): e1002822.
- Consortium, International HapMap 3. 2010. "Integrating Common and Rare Genetic Variation in Diverse Human Populations." *Nature* 467 (7311): 52–58.
- Danchin, Étienne. 2013. "Avatars of Information: Towards an Inclusive Evolutionary Synthesis." *Trends in Ecology & Evolution* 28 (6): 351–358.
- Day, Troy, and Russell Bonduriansky. 2011. "A Unified Approach to the Evolutionary Consequences of Genetic and Nongenetic Inheritance." *The American Naturalist* 178 (2): E18–E36.
- Dias, Brian G., and Kerry J. Ressler. 2014. "Parental Olfactory Experience Influences Behavior and Neural Structure in Subsequent Generations." *Nature Neuroscience* 17 (1): 89–96.
- Doolittle, Donald P. 2012. *Population Genetics: Basic Principles*. Vol. 16. Springer Science & Business Media.
- Downes, Stephen M. 2015. "Heritability." In *Stanford Encyclopedia of Philosophy*, ed. Edward N. Zalta. Stanford, CA: Stanford University.
- Eichler, Evan E., Jonathan Flint, Greg Gibson, Augustine Kong, Suzanne M. Leal, Jason H. Moore, and Joseph H. Nadeau. 2010. "Missing Heritability and Strategies for Finding the Underlying Causes of Complex Disease." *Nature Reviews Genetics* 11 (6): 446–450.
- Falconer, Douglas S., and Trudy F. C. Mackay. 1996. *Introduction to Quantitative Genetics*. 4th edition. Longman: Benjamin Cummings.
- Feil, Robert, and Mario F. Fraga. 2012. "Epigenetics and the Environment: Emerging Patterns and Implications." *Nature Reviews Genetics* 13 (2): 97–109.
- Frazer, Kelly A., Sarah S. Murray, Nicholas J. Schork, and Eric J. Topol. 2009. "Human Genetic Variation and Its Contribution to Complex Traits." *Nature Reviews Genetics* 10 (4): 241–251.
- Furrow, Robert E., Freddy B. Christiansen, and Marcus W. Feldman. 2011. "Environment-Sensitive Epigenetics and the Heritability of Complex Diseases." *Genetics* 189 (4): 1377–1387.

- Griffiths, Anthony JF., Susan R. Wessler, Richard C. Lewontin, William M. Gelbart, David T. Suzuki, and Jeffrey H. Miller. 2005. *An Introduction to Genetic Analysis*. 8th edition. New York: W. H. Freeman.
- Griffiths, Paul, and Karola Stotz. 2013. *Genetics and Philosophy: An Introduction*. Cambridge University Press.
- Guo, Guang, Lin Wang, Hexuan Liu, and Thomas Randall. 2014. "Genomic Assortative Mating in Marriages in the United States." *PloS One* 9 (11): e112322.
- Jablonka, Eva, Marion J Lamb, and Anna Zeligowski. 2014. *Evolution in Four Dimensions: Genetic, Epigenetic, Behavioral, and Symbolic Variation in the History of Life*. Revised edition. MIT Press.
- Jablonka, Eva, and Gal Raz. 2009. "Transgenerational Epigenetic Inheritance: Prevalence, Mechanisms, and Implications for the Study of Heredity and Evolution." *The Quarterly Review of Biology* 84 (2): 131–176.
- Johannes, Frank, Vincent Colot, and Ritsert C. Jansen. 2008. "Epigenome Dynamics: A Quantitative Genetics Perspective." *Nature Reviews Genetics* 9 (11): 883–890.
- Klein, Robert J., Caroline Zeiss, Emily Y. Chew, Jen-Yue Tsai, Richard S. Sackler, Chad Haynes, Alice K. Henning, John Paul SanGiovanni, Shrikant M. Mane, and Susan T. Mayne. 2005. "Complement Factor H Polymorphism in Age-Related Macular Degeneration." *Science* 308 (5720): 385–389.
- Lu, Qiaoying, and Bourrat Pierrick. Forthcoming. "The Evolutionary Gene and the Extended Evolutionary Synthesis." *British Journal for Philosophy of Science*.
- Maher, Brendan. 2008. "Personal genomes: The Case of the Missing Heritability." *Nature News* 456 (7218):18–21.
- Manolio, Teri A., Francis S. Collins, Nancy J. Cox, David B. Goldstein, Lucia A. Hindorff, David J. Hunter, Mark I. McCarthy, Erin M. Ramos, Lon R. Cardon, and Aravinda Chakravarti. 2009. "Finding the Missing Heritability of Complex Diseases." *Nature* 461 (7265): 747–753.
- McCarthy, Mark I., and Joel N. Hirschhorn. 2008. "Genome-Wide Association Studies: Potential next Steps on a Genetic Journey." *Human Molecular Genetics* 17 (R2): R156–165.
- Morgan, Hugh D., Heidi GE Sutherland, David IK Martin, and Emma Whitelaw. 1999. "Epigenetic Inheritance at the Agouti Locus in the Mouse." *Nature Genetics* 23 (3): 314–318.
- Schork, Nicholas J., Sarah S. Murray, Kelly A. Frazer, and Eric J. Topol. 2009. "Common vs. Rare Allele Hypotheses for Complex Diseases." *Current Opinion in Genetics & Development* 19 (3): 212–219.

- Silventoinen, Karri, Sampo Sammalisto, Markus Perola, Dorret I. Boomsma, Belinda K. Cornes, Chayna Davis, Leo Dunkel, Marlies De Lange, Jennifer R. Harris, and Jacob VB Hjelmberg. 2003. "Heritability of Adult Body Height: A Comparative Study of Twin Cohorts in Eight Countries." *Twin Research* 6 (05): 399–408.
- Turkheimer, Eric. 2011. "Still Missing." *Research in Human Development* 8 (3-4): 227–241.
- Visscher, Peter M., Matthew A. Brown, Mark I. McCarthy, and Jian Yang. 2012. "Five Years of GWAS Discovery." *The American Journal of Human Genetics* 90 (1): 7–24.
- Visscher, Peter M., Sarah E. Medland, Manuel AR Ferreira, Katherine I. Morley, Gu Zhu, Belinda K. Cornes, Grant W. Montgomery, and Nicholas G. Martin. 2006. "Assumption-Free Estimation of Heritability from Genome-Wide Identity-by-Descent Sharing between Full Siblings." *PLoS Genet* 2 (3): e41.
- Visscher, Peter M., William G. Hill, and Naomi R. Wray. 2008. "Heritability in the Genomics Era—concepts and Misconceptions." *Nature Reviews Genetics* 9 (4): 255–266.
- Weedon, Michael N., Hana Lango, Cecilia M. Lindgren, Chris Wallace, David M. Evans, Massimo Mangino, Rachel M. Freathy, John RB Perry, Suzanne Stevens, and Alistair S. Hall. 2008. "Genome-Wide Association Analysis Identifies 20 Loci that Influence Adult Height." *Nature Genetics* 40 (5): 575–583.
- Wood, Andrew R, Tonu Esko, Jian Yang, Sailaja Vedantam, Tune H Pers, Stefan Gustafsson, Audrey Y Chu, Karol Estrada, Jian'an Luan, and Zoltán Kutalik. 2014. "Defining the Role of Common Variation in the Genomic and Biological Architecture of Adult Human Height." *Nature genetics* 46 (11): 1173–1186.
- Yang, Jian, Beben Benyamin, Brian P. McEvoy, Scott Gordon, Anjali K. Henders, Dale R. Nyholt, Pamela A. Madden, et al. 2010. "Common SNPs Explain a Large Proportion of the Heritability for Human Height." *Nature Genetics* 42 (7): 565–569.
- Youngson, Neil A., and Emma Whitelaw. 2008. "Transgenerational Epigenetic Effects." *Annual Review of Genomics and Human Genetics* 9: 233–257.