Why the missing heritability might not be in the DNA

Pierrick Bourrat[1)2)], Qiaoying Lu[3)] and Eva Jablonka[4)]*

Finding the missing heritability has become an important challenge for genome wide association studies (GWAS) for nearly a decade [1]. The estimates of heritability obtained from GWAS can only account for a fraction of that evidenced in classical family correlation studies. For instance, while the classical estimates for the heritability of height are of about 0.8, the initial GWAS in 2008 found a heritability of about 0.05 [2]. By considering common single nucleotide polymorphisms (SNPs) simultaneously, Yang and colleagues have generated new heritability estimate of about 0.45 [3]. In recent simulations with rare variants included, they expect that heritability estimates of 60–70% could be obtained in future studies [4], hence reducing the missing heritability to a negligible difference. However, their method has been criticized by Kumar et al. [5], who claim that it gives unstable heritability estimates by over-fitting the data. This has led to a dispute between Yang and colleagues, and Kumar and colleagues, which has not yet been resolved. Apart from this dispute, GWAS studies contain certain assumptions that might be violated, and will encounter particular limitations [6]. We believe, therefore, that it is important to remain open to new approaches to and ideas on finding the missing heritability.

DOI 10.1002/bies.201700067

[1)] Macquarie University – Philosophy, New South Wales, Australia
[2)] The University of Sydney – Unit for the History and Philosophy of Science, Philosophy Department and Charles Perkins Centre, New South Wales, Australia
[3)] Sun Yat-sen University – Philosophy, Guangzhou, China
[4)] Tel-Aviv University – History and Philosophy, Tel-Aviv, Israel

***Corresponding author**: Pierrick Bourrat
E-mail: p.bourrat@gmail.com

Most of the current GWAS for common complex diseases are searching for common variants (with a frequency >0.01 in the population) [7], and heritability is estimated based on those variants. Yang et al. [4] claim that there are three hypotheses that may explain the missing heritability. The first one is that not only common variants, but also rare variants with a frequency <0.01, may contribute to heritability. There is evidence that rare SNPs might have a large effect on phenotype [4]. The second hypothesis is that a substantial portion of heritability may be attributable to common variants with effects that are too small to be detected by current methods. New methods have been developed for detecting rare variants [4], and common variants with small effects could be identified by increasing the sample sizes in studies. By taking both hypotheses into account, the hope is to bridge the gap between the heritability estimates obtained by classical studies and those obtained by GWAS. Finally, a third hypothesis is that the heritability obtained with family studies is overestimated by not eliminating shared environmental effects. Dominance and epistasis effects can also lead to an overestimated heritability, depending on which family members are compared.

We agree that these three mutually compatible hypotheses are important, but we argue that there is another hypothesis concerning a fundamental conceptual confusion that could contribute to the solution of the missing heritability problem. The hypothesis is that the word "genetic" used in classical family studies and the word "genetic" used in GWAS do not refer to the same concept of the gene. In the latter case, "genetic" is exclusively associated with differences in DNA. In classical studies, which were first developed before the knowledge that DNA is a carrier of genetic information, a "genetic" factor refers to a factor that can be transmitted from one generation to the other and that contributes to phenotypic variation – be it DNA-based or based on some other physical material [8]. There is increasing evidence showing that non-DNA (epigenetic) factors can be transmitted reliably from parent to offspring [9], and models have been built to estimate the contributions that non-DNA factors could make. Because GWAS focus solely on DNA, these studies will remain blind to non-DNA factors when there are no correlations between DNA and those factors [10]. The four hypotheses are represented in Fig. 1.

The question as to whether the discrepancy between these two notions of "genetic" is a (partial) reason why heritability is missing can only be answered empirically. Fortunately, Tal and colleagues have devised a method to test whether, and to which extent non-DNA factors contribute to phenotypic resemblance between family members, which is usually measured in terms of correlation [11]. This method is an extension of classical family studies used in quantitative genetics [12]. It starts by considering the number of opportunities for "epigenetic reset" between different family members following meiosis and fertilization of the egg (e.g. one between parent and offspring, two between siblings, three between uncle and nieces, and nephews). With knowledge of the phenotypic covariance between parent and offspring, siblings, and uncle and nephews/nieces, it can be estimated by deduction – given a number of plausible and testable assumptions – to which extent phenotypic variance is explained, respectively, by DNA-factors and non-DNA factors, and, furthermore, the extent to which non-DNA factors are transmitted between generations. This method could also be used to test the third hypothesis of Yang and colleagues, assuming that they consider DNA factors to be the only ones to include in heritability estimates.

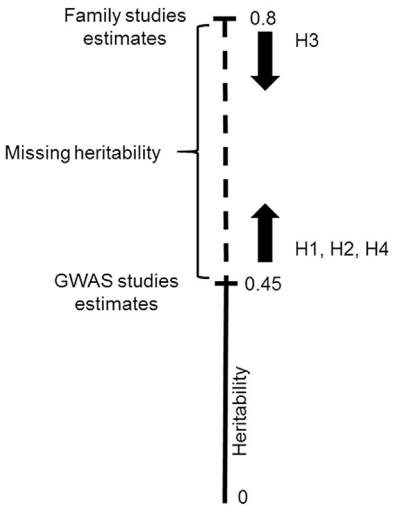

Figure 1. The four major hypotheses (H1, H2, H3, and H4) as to the source of missing heritability. Values for heritability estimates are those for human height. If H1or H2 is verified, heritability estimates obtained from GWAS should increase in future studies because of current methodological limitations. If H3 is verified, heritability estimates obtained from family studies are too high, and do not represent the genuine heritability of the focal trait. Taking confounding factors into consideration might thus reduce the gap between the estimates coming from the two types of methods. Finally if H4 is verified, heritability estimates from GWAS underestimate the genuine heritability of the focal trait by not taking into account non-DNA (epigenetic) sources of heritability, which is accounted for in family studies. Taking those factors into account should thus result in an increased heritability estimate.

To conclude, given that some non-DNA factors can be transmitted stably from parent to offspring, that they can consequently respond to selection, and that they correspond to a legitimate conception of the gene when heritability was initially estimated in quantitative genetics, we believe that these factors should be incorporated in the definition of heritability. Therefore, it is important for proponents of GWAS to recognize this hypothesis and to test it empirically. We believe that having insights into the heritability of non-DNA factors will certainly help to solve the missing heritability problem.

References

1. **Turkheimer E.** 2011. Still missing. *Res Hum Dev* **8**: 227–41.
2. **Maher B.** 2008. Personal genomes: the case of the missing heritability. *Nat News* **456**: 18–21.
3. **Yang J**, **Benyamin B**, **McEvoy BP**, **Gordon S** et al**.** 2010. Common SNPs explain a large proportion of the heritability for human height. *Nat Genet* **42**: 565–69.
4. **Yang J**, **Bakshi A**, **Zhu Z**, **Hemani G** et al**.** 2015. Genetic variance estimation with imputed variants finds negligible missing heritability for human height and body mass index. *Nat Genet* **47**: 1114–20.
5. **Kumar SK**, **Feldman MW**, **Rehkopf DH**, **Tuljapurkar S.** 2016. Limitations of GCTA as a solution to the missing heritability problem. *Proc Natl Acad Sci USA* **113**: E61–70.
6. **Eichler EE**, **Flint J**, **Gibson G**, **Kong A** et al**.** 2010. Missing heritability and strategies for finding the underlying causes of complex disease. *Nat Rev Genet* **11**: 446–50.
7. **Manolio TA**, **Collins FS**, **Cox NJ**, **Goldstein DB** et al**.** 2009. Finding the missing heritability of complex diseases. *Nature* **461**: 747–53.
8. **Lu Q**, **Bourrat P.** 2017. The evolutionary gene and the extended evolutionary synthesis. *Br J Philos Sci* in press https://doi.org/10.1093/bjps/axw035.
9. **Jablonka E**, **Raz G.** 2009. Transgenerational epigenetic inheritance: prevalence, mechanisms, and implications for the study of heredity and evolution. *Q Rev Biol* **84**: 131–76.
10. **Bourrat P, Lu Q.** 2017. Dissolving the missing heritability problem. *Philos Sci* **85.** In press.
11. **Tal O**, **Kisdi E**, **Jablonka E.** 2010. Epigenetic contribution to covariance between relatives. *Genetics* **184**: 1037–50.
12. **Falconer DS**, **Mackay TFC.** 1996. *Introduction to Quantitative Genetics*. New York: Longman.