Twin Research and Human Genetics (2020), 1–2
doi:10.1017/thg.2020.21

CAMBRIDGE UNIVERSITY PRESS

Article

Musings on Visscher et al. (2006)

Peter M. Visscher

Institute for Molecular Bioscience, The University of Queensland, Brisbane, Queensland, Australia

Abstract

The classical twin design relies on a number of strong number of assumptions in order to yield unbiased estimates of heritability. This includes the equal environments assumption — that monozygotic and dizygotic twins experience similar degrees of environmental similarity — an assumption that is likely to be violated in practice for many traits of interest. An alternative method of estimating heritability that does not suffer from many of these limitations is to model trait similarity between sibling pairs as a function of their empirical genome-wide identity by descent sharing, estimated from genetic markers. In this review, I recount the story behind Nick Martin's and my development of this method, our first attempts at applying it in a human population and more recent studies using the original and related methods to estimate trait heritability.

Keywords: Linkage; identity by descent; heritability; height; equal environments assumption

(Received 17 February 2020; accepted 3 April 2020)

In 2004, I attended the (fourth) Australian 'Genemappers' meeting in Perth. I had spent a 3-month sabbatical at the Queensland Institute of Medical Research (QIMR) the previous year, and Naomi (Wray) and I, who were working in Edinburgh at the time, were seriously considering migrating to Australia at some point — hence the interest in attending this relatively small meeting at the other side of the world. I am pretty sure that Nick very generously paid for my travel expenses from Edinburgh to Perth (Australia, not Scotland) so that I could attend the meeting (although there might have been an element of self-interest too!).

At the meeting, Nick gave a short presentation of work that Gu Zhu and he had been doing using results from genetic linkage analyses, using data from microsatellite markers on sibling pairs (mostly dizyotic [DZ] twin pairs). Linkage analyses, that is, the analysis of association between identity-by-descent (IBD) status at genomic loci and complex traits within families, were still popular in those days — they were to be replaced by genome-wide association studies (GWAS) very soon afterward. Interestingly, Gu and Nick were using the IBD estimates for a purpose that differed from the standard locus-by-locus genome scan.

The title of Nick's presentation was 'Biometrical Genetics — with real data!'. They had used the locus-by-locus estimates of IBD to obtain a genome-wide IBD estimate of 'realized relatedness' or 'actual relatedness' in about 900 sibling pairs, and also genome-wide coefficients of dominance. The estimate of genome-wide relationship was obtained by averaging IBD estimates across many (about 3500) points in the genome. The estimate of the mean and standard deviation of additive and dominance relatedness were (0.5, 0.04) and (0.25, 0.04), respectively. The standard deviations are the most interesting parameters in this context and turn out to be spot-on with what is expected under (previously published) theory. Nick also showed results from using trait data on height (and other traits), from fitting and comparing various statistical models using genome-wide and chromosome-wide estimates of realized relationships. (Twin researchers like to perform model testing and model selection, rather than just focus on the estimation of variance components and their standard errors. I have never quite understood this, because the inference from model testing depends on the sample size and can lead to winner's curse. In addition, why would I want to calculate a p value for narrow-sense heritability when we know that all traits that vary in the population will have some genetic variation?). There was a lack of power of the trait-based analyses, but the idea to combine realized relationships with trait data intrigued me very much.

As an aside, Nick has consistently claimed that the idea to estimate realized relationships from marker data and then perform statistical analyses for complex traits came to him while traveling on a bus in Provence. This must be a true story, because Nick hardly ever uses public transport, let alone a bus. Nick has had other famous Road to Damascus moments in his life, not least his 180-degree turn from socialism to conservatism in his early twenties. But let's keep to the scientific eureka moments. The story of 'The Great Provence Insight' was repeated many times after the events, perhaps most infamously when Nick and I (and others) were being interviewed for a major grant proposal in Australia a few years later. The interviewees had absolutely no idea what Nick was talking about scientifically (they were neither geneticists nor quantitative), but may have been envious about his (working) holiday in Provence. In the end, we did not get the grant, but that was most likely because of other issues.

Although Nick was not, to my knowledge, aware of it, theory and empirical applications of the variation in realized relatedness about the expected value (e.g., variation around 0.5 for DZ twins) goes back to the 1970s. In the 1990s, several authors had started to quantify how much of this variation could be captured with genetic markers; for example, in line crosses (I worked on this in mid to

Author for correspondence: Peter Visscher, Email: peter.visscher@uq.edu.au
Cite this article: Visscher PM. Musings on Visscher et al. (2006). *Twin Research and Human Genetics* https://doi.org/10.1017/thg.2020.21

late 1990s) but also in outbred populations. For complex trait data, multiple authors prior to 2004 had suggested to use estimates of relatedness from marker data and subsequently estimate genetic parameters using those estimates. However, those applications were generally in cases where the pedigree is not known — for example, in ecology and evolutionary studies. Therefore, the combination of IBD-based estimation of relatedness and complex trait analyses was novel and opened the door to address a number of interesting scientific questions using a new experimental design.

After joining QIMR in 2005, the first question I was interested in addressing using realized relatedness was the estimation of within-family additive genetic variance using sibling pairs. In a random mating population, 50% of variance is between and 50% is within families. Within-family variance is sometimes called segregation variance or the variance of Mendelian sampling terms (it is because of this variance that children have a path coefficient of $\sqrt{0.5}$ with themselves for the additive term A in an extended twin design). The association between the departures of realized relationship from its expected value (of 0.5) and trait similarity for sibling pairs can be used to estimate within-family additive genetic variance, and therefore heritability. The beauty of this experimental design is that it is free from confounding due to environmental factors and G-E correlations: we are simply comparing how similar sibling pairs are that happen to share, say, 55% of their genome IBD versus those that share, say, 45% of their genome IBD. Estimating variance in this way is the same as performing a linkage analysis with the entire genome (instead of with a single locus). Therefore, it is in theory an extremely nice design to estimate and partition genetic variation. Indeed, Nick and Gu attempted such analyses with height in 2004.

For the 2006 *PLOS Genetics* paper, we used a sample size of 4400 pairs with marker data and 3800 pairs with both marker and data on height — a combination of DZ twins and non-twin siblings. We are now used to huge sample sizes in GWAS, but in 2005, this family-based sample size was probably the largest of its kind in the world. Indeed, it was the availability of data like those that was part of the attraction of moving to Brisbane. I had done the theory of the power of the design and realized that, unfortunately, much larger samples are needed to estimate the variance components accurately — the sampling variance of the estimate of heritability is inversely proportionate to the product of sample size (N pairs) and the variance of relatedness ($\sim 0.038^2$), so the standard error is proportional to $1/(0.038 \times \mathrm{sqrt}(N)) \sim 26/\mathrm{sqrt}(N)$. Our point estimate for height from segregation variance was 0.8, but with a large confidence interval ranging from 0.4 to 0.9.

Despite the large sampling variance, we believed that the paper was a nice proof of concept of a neat experimental design and had great hopes of getting it published in a good journal. We thought that the *American Journal of Human Genetics* (AJHG) was the right journal for it, but the Editor (after consultation with the Editorial Board) did not want to send it out for review. We appealed, twice, but received a rejection every time. The only feedback we received was along the lines of 'we already have a twin/pedigree design to estimate heritability, why do we need another one, in particular if it is not very powerful?'. In other words, they just did not get the novelty. We ended up in *PLOS Genetics* (Visscher et al., 2006), where the referees were quantitative geneticists not working in human genetics, and the paper sailed through. Interestingly, follow-up papers (Visscher et al., 2007 and later Hemani et al., 2013) did get published in AJHG and got a fairly easy time from the referees. A reminder of the stochasticity of the system!

The subsequent papers used the same design to partition genetic variation by chromosome (2007) and included body mass index (BMI) as a trait (2013). The latter paper (Hemani et al., 2013) was on a total of 20,000 sibling pairs and showed clear evidence for 'genomic inflation' from linkage analysis, which is proof (as if we needed it) of the polygenicity of traits like height and BMI.

Recently, the within-family experimental design was extended for complex pedigrees by Young et al. (2018), who applied their method to data from deCODE. They called their method 'Relatedness Disequilibrium Regression', which is a complicated but succinct way of saying that the method estimates the variance of Mendelian segregation effects. There is a renewed interest in estimating variance components using these kinds of designs because it allows the break-up of genotype–environment correlations, which are expected for traits like intelligent quotient (IQ) and educational attainment. Hence, direct additive genetic effects can be estimated from within-family segregation, and these effects can be separated from parental (maternal and paternal) effects.

As with all genetic analyses, there are caveats with the estimation of genetic variance from within-family segregation. Importantly, segregation variance is not affected by nonrandom mating, whereas between-family variance is. Therefore, for traits such as height, IQ and educational attainment, for which there is strong empirical evidence of assortative mating, the estimate of additive genetic variance from within-family estimation is expected to be lower than that inferred from the correlation between relatives, irrespective of parental ('nurturing') effects. Therefore, for traits undergoing assortative mating, the comparison of estimates of additive genetic variance (or heritability) from within and between-family experimental designs can lead to incorrect conclusions.

What's next? Although it seemed inconceivable back in 2006, it is now possible to use the within-family design on sample sizes approaching 100,000 sibling pairs and estimate and partition genetic variance for behavioral and other complex traits with good accuracy and free from confounding factors. Those 100,000 pairs will have genome-wide, single nucleotide polymorphism data from GWAS or whole-genome sequencing (WGS), so in principle joint within and between family analyses could be performed. Quantitative Genetics — with real data!

References

Hemani, G., Yang, J., Vinkhuyzen, A., Powell, J. E., Willemsen, G., Hottenga, J. J., . . . , Visscher, P. M. (2013). Inference of the genetic architecture underlying BMI and height with the use of 20,240 sibling pairs. *American Journal of Human Genetics*, *93*, 775–994.

Visscher, P. M., Medland, S. E., Ferreira, M. A. R., Zhu, G., Cornes, B. K., Montgomery, G. W., & Martin, N. G. (2006). Assumption-free estimation of heritability from genome-wide identity-by-descent sharing between full siblings. *PLOS Genetic*, *2*, 316–325.

Visscher, P. M., Macgregor, S., Benyamin, B., Zhu, G., Gordon, S., Martin N. G. (2007). Genome partitioning of genetic variation for height from 11,214 sibling pairs. *American Journal of Human Genetics*, *81*, 1104–1110.

Young, A. I., Frigge, M. L., Gudbjartsson, D. F., Thorleifsson, G., Bjornsdottir, G., Sulem, P., Kong, A. (2018). Relatedness disequilibrium regression estimates heritability without environmental bias. *Nature Genetics*, *50*, 1304–1310.