Published in final edited form as:
J Hum Genet. 2014 June ; 59(6): 342–345. doi:10.1038/jhg.2014.14.

Testing the Key Assumption of Heritability Estimates Based on Genome-wide Genetic Relatedness

Dalton Conley[*,1], **Mark L. Siegal**[1], **Benjamin Domingue**[2], **Kathleen Mulan Harris**[3], **Matthew McQueen**[2], and **Jason Boardman**[2]

[1] Center for Genomics and Systems Biology, New York University

[2] Institute for Behavioral Sciences, University of Colorado, Boulder

[3] Department of Sociology, University of North Carolina, Chapel Hill

Abstract

Comparing genetic and phenotypic similarity among unrelated individuals seems a promising way to quantify the genetic component of traits while avoiding the problematic assumptions plaguing twin- and other kin-based estimates of heritability. One approach uses a Genetic Relatedness Estimation through Maximum Likelihood (GREML) model for individuals who are related at less than .025 to predict their phenotypic similarity by their genetic similarity. Here we test the key underlying assumption of this approach: that genetic relatedness is orthogonal to environmental similarity. Using data from the Health and Retirement Study (and two other surveys), we show two unrelated individuals may be more likely to have been reared in a similar environment (urban versus non-urban setting) if they are genetically similar. This effect is not eliminated by controls for population structure. However, when we include this environmental confound in GREML models, heritabilities do not change substantially and thus potential bias in estimates of most biological phenotypes is probably minimal.

Ascertaining the proportion of variance in a quantitative trait—such as height or IQ—that is due to genetic variation has long been of interest to a wide range of scientists [1–5]. For human populations, where experimentation is not possible, the workhorse of such analysis has been the twin or extended twin design, where the average relatedness of various kin pairs is correlated with their phenotypic similarity in order to ascertain the effect of shared genotype on a given outcome[6,7]. The reigning critique of this approach is that it is difficult to eliminate the possibility that increased similarity between, say, monozygotic twins as compared to, for example, dizygotic twins is due to more similar environments and not solely their greater genetic similarity [8,9].

Among the recent and novel approaches to overcome this potential environmental confounding are studies that correlate phenotypic similarity with genotypic similarity across the genome among pairs of individuals who are less than 2.5 percent related as computed by identity by state (IBS) and are therefore considered non-kin[10–12]. Simply described, a

[*]Corresponding Author.

Supplementary information is available at the Journal of Human Genetics website

genetic relatedness matrix (GRM) is constructed in which each cell is filled by a measure of 2N gametic correlation between pairs of individuals (the rows and columns) summed across a set of markers that have been pruned for linkage disequilibrium. These values are then used to estimate phenotypic similarity between the pairs. This Genetic Relatedness Estimation through Maximum Likelihood (GREML) approach yields estimates of narrow-sense (additive) heritability (h^2) that are lower than but approaching those obtained from traditional twin-based approaches and has been deployed for diverse phenotypes, including height [13], schizophrenia [10], asthma [14], smoking [15], body mass index [16], educational attainment [17] and political and economic preferences [18].

However, like twin based models, the GREML approach relies on one key assumption about the relationship between genetic similarity and environmental similarity. Although those who share genetic variation may experience more similar environments due to population structure, admixture and, of course, extended family ties, GREML assumes that those who are less related than 2nd cousins share alleles in an essentially random fashion that is itself uncorrelated with environmental similarity. The motivating notion is that at these low levels of relatedness, relative genetic similarity is driven by the randomness of recombination and allele segregation and not by underlying kinship structure. As such, parental relatedness and relevant environmental conditions should be orthogonal to respondent relatedness.

To support this claim that relatedness among these pairs of individuals is random (and thus uncorrelated with potential environmental confounders), Yang et al. (2010) show correlations in relatedness levels between chromosomes in a supplemental table.[11] Their logic is that if the person-wide genetic relatedness measure between individuals (i.e. gametic correlation) was reflecting population structure (and, thus, covaried with environment), pairwise genetic relatedness would be correlated across those individuals' chromosomes. But if the distribution of pairwise relatedness is really just the result of randomization during meiosis, then each chromosome should be independent, demonstrating no correlation. Yang et al. find no single pair of chromosomes for which the p-value of the correlation between the genetic relatedness of those two chromosomes is less than 0.00022, which corresponds to a 0.05 alpha level with a Bonferroni correction for the 231 comparisons they make across the bivariate combinations of the autosomal chromosomes. However, this strikes us as the wrong statistical test: We are not concerned as to whether the relatedness of a specific pair of chromosomes co-varies below a strict Type I error threshold. Rather, we are worried that there is an overall pattern of relatedness in the data and thus should apply a more sensitive test that minimizes Type II error. Along these lines, in Figure 1, we present a histogram of their 231 reported p-values and show that there is indeed an excess of low p-values, particularly below the $p<.10$ threshold as well as a dearth of high p-values ($p>.90$) as compared to a random distribution. Indeed, when we perform a Kolmogorov-Smirnov test on their reported distribution, we find it to deviate from the theoretically expected (uniform) distribution (D^+ = 0.1892, p-value = 7.037e-08). While we do not know the signs of the associated coefficients (since they were not reported by Yang et al.), the overall non-random distribution of correlations suggests that the data fail the test for randomization of alleles across chromosomes.

With this in mind, we do not believe that this core assumption that the environmental similarity between pairs of unrelated persons is uncorrelated with their genetic similarity (below the .025 threshold) has not been adequately interrogated. In the present study, we test the key GREML assumption by asking whether the childhood environments of subjects are more similar if they are more related genetically. If pairs of individuals both experienced an urban (or, by contrast, non-urban) environment growing up this is likely to have the effect of making their formative social and physical environment more similar. Thus, if relatedness predicts environmental similarity in this way, it could confound the premise of GREML-based methods of estimating the genetic component of phenotypes. It makes no difference whether urbanicity is itself causal of the phenotype under consideration; it may be acting merely as a proxy for other, more relevant environmental factors—such as social class, nutritional status and so forth—that are themselves related, through environmental channels, to the offspring phenotype (such as height, BMI or education). That said, a large literature shows that urbanicity is correlated with a range of outcomes studied by geneticists, ranging from mental health [19–21] to immunological response [21,22] to education [23].

Health and Retirement Study (HRS) data allow us to estimate the heritability of urban childhood residence as well as how urban residence during childhood affects GREML estimates of other putatively heritable traits. We used the standard GREML analysis (using GCTA software [12]) to estimate heritability, with population stratification controlled by principal components (PCs) (see Supplementary Materials: Methods). As shown in the first row of Table 1 below, in the HRS sample with two principal components controlled, urban childhood— putatively a childhood environmental variable based on circumstance and parental choices—is indeed highly heritable at 29 percent. Because we suspected that the nonzero heritability might be a result of geographic population structure, we then reran the analysis with 10 and 25 PCs included as controls. These controls attenuated, but did not eliminate, the effect we discovered. Thus, it seems that controls for population structure through deployment of PCs does not adequately address this confounding. We replicated this finding with data from the National Longitudinal Survey of Adolescent Health (Add Health) as well as with another childhood phenotype—maternal education—in Add Health and in the Framingham Heart Study (FHS). Both Add Health and FHS are underpowered to generate statistically precise GREML heritability estimates, but ordinary least square regressions show magnitudes of estimates in line with the HRS results (see Supplementary Materials). Finally, we deployed a more stringent, one percent cut-off for the relatedness matrix, but this, too, was underpowered (also see Supplementary Materials).

Despite the apparent heritability of childhood residence, when we control for this possible confounder in analysis of common human phenotypes of interest—height, BMI and years of schooling—we find that the differences between the “naïve” models and the ones that hold childhood urbancity constant are negligible and not statistically significant. In fact, the only phenotype for which the heritability changes to any noticeable degree is respondent education, which drops by a statistically insignificant two percentage points (p=0.8203) in the model with only two PCs. This makes sense: Of the three phenotypes, we would expect height to be the least influenced by childhood environment, BMI in the middle and education to be the most affected by potential environmental confounds. Because controlling for more PCs did not appear to eliminate the heritability of a putatively environmental

confound—urban childhood—we then tried to see if using a more restrictive relatedness cut-off (.01) would address the "problem." However, when we used this more restrictive cut-off, sample sizes dropped too drastically to yield adequate power. (Results are shown in Supplemental Table S1.)

Our findings have implications not only for GREML analysis of heritability but for genome-wide analysis more broadly. Namely, some scholars have claimed that PCs adequately control for population stratification, especially when data show no evidence of "early take-off" (i.e. across the vast majority of the distribution of p-values, they match what one would expect from chance)[24,25]. Our results suggest that directly modeling error terms as a linear function of relatedness in a sample may be also be necessary to adjust for stratification [26]. Finally, and most importantly, while the key assumption of GREML analysis that the genotype-environment correlation (rGE) is zero is violated, the consequences of that violation appear to be trivial. We cautiously conclude that GREML is a valid estimation technique for heritability but recommend that going forward, researchers test for the violation of this assumption (and robustness to violations) in their own datasets as a standard sensitivity analysis.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

References

1. Breen F, Plomin R, Wardle J. Heritability of food preferences in young children. Physiol. Behav. 2006; 88:443–447. [PubMed: 16750228]
2. Rodgers J, Kohler H, Kyvik K, Christensen K. Behavior genetic modeling of human fertility: Findings from a contemporary Danish twin study. Demography. 2001; 38:29–42. [PubMed: 11227843]
3. Oord, E. Van den A study of genetic and environmental effects on the co-occurrence of problem behaviors in three-year-old-twins. J. Abnorm. Psychol. 2000; 109:360. [PubMed: 11016106]
4. Rodgers J, Rowe D, Buster M. Nature, nurture and first sexual intercourse in the USA: fitting behavioural genetic models to NLSY kinship data. J. Biosoc. Sci. 1999; 31:29–41. [PubMed: 10081235]
5. Allison DB, Kaprio J, Korkeila M, Koskenvuo M, Neale MC, Hayakawa K. The heritability of body mass index among an international sample of monozygotic twins reared apart. Int. J. Obes. 1996; 20:501–506.
6. Plomin R, Owen M, McGuffin P. The genetic basis of complex human behaviors. Science (80-.). 1994; 264:1733–1739.
7. Purcell S. Variance components models for geneenvironment interaction in quantitative trait locus linkage analysis. Twin Res. 2002; 5:572–576. [PubMed: 12573188]
8. Goldberger A. Heritability. Economica. 1979; 46:327–347.
9. Scarr S, Carter-Saltzman L. Twin method: Defense of a critical assumption. Behav. Genet. 1979; 9:527–542. [PubMed: 263638]
10. Purcell S, et al. Common polygenic variation contributes to risk of schizophrenia and bipolar disorder. Nature. 2009; 460:748–752. [PubMed: 19571811]
11. Yang J, Benyamin B, McEvoy B. Common SNPs explain a large proportion of the heritability for human height. Nat. Genet. 2010; 42:565–569. [PubMed: 20562875]
12. Yang J, Lee S, Goddard M, Visscher P. GCTA: a tool for genome-wide complex trait analysis. Am. J. Hum. Genet. 2011; 88:76–82. [PubMed: 21167468]

13. Davies G, et al. Genome-wide association studies establish that human intelligence is highly heritable and polygenic. Mol. Psychiatry. 2011; 16:996–1005. [PubMed: 21826061]
14. Belsky DW, et al. Polygenic risk and the development and course of asthma: an analysis of data from a four-decade longitudinal study. Lancet Respir. Med. 2013; 1:453–461. [PubMed: 24429243]
15. Belsky DW, et al. Polygenic risk and the developmental progression to heavy, persistent smoking and nicotine dependence: evidence from a 4-decade longitudinal study. JAMA Psychiatry. 2013; 70:534–42. [PubMed: 23536134]
16. Belsky DW, et al. Polygenic Risk, Rapid Childhood Growth, and the Development of Obesity. 2012; 166
17. Rietveld C, Medland S, Derringer J, Yang J. GWAS of 126,559 Individuals Identifies Genetic Variants Associated with Educational Attainment. Science (80-.). 2013; 340:1467–1471.
18. Benjamin DJ, et al. The genetic architecture of economic and political preferences. Proc. Natl. Acad. Sci. U. S. A. 2012; 109:8026–8031. [PubMed: 22566634]
19. Krabbendam L, Os J. Van. Schizophrenia and urbanicity: a major environmental influence—conditional on genetic risk. Schizophr. Bull. 2005; 31:795–799. [PubMed: 16150958]
20. Stefanis N, et al. Is the excess risk of psychosis-like experiences in urban areas attributable to altered cognitive development? Soc. Psychiatry. 2004; 39:364–368.
21. Spauwen J, Krabbendam L, Lieb R, Wittchen HU, Van Os J. Evidence that the outcome of developmental expression of psychosis is worse for adolescents growing up in an urban environment. Psychol. Med. 2006; 36:407–415. [PubMed: 16403242]
22. Priftis K, et al. Increased sensitization in urban vs. rural environment–Rural protection or an urban living effect? Pediatr. allergy Immunol. 2007; 18:209–216. [PubMed: 17432999]
23. Jencks, C.; Mayer, S. Inn. poverty United States. 1990. <http://books.google.com/books?hl=en&lr=&id=P7IV4eaGcxwC&oi=fnd&pg=PA111&dq=The+social+consequences+of+growing+up+in+poor+city+neighborhood&ots=Zn2Q7Z3SD4&sig=CxdXMCb0dKjkt1zrCr-6Av_UnwA>
24. Price AL, et al. Principal components analysis corrects for stratification in genome-wide association studies. Nat. Genet. 2006; 38:904–9. [PubMed: 16862161]
25. Price A, Zaitlen N, Reich D, Patterson N. New approaches to population stratification in genome-wide association studies. Nat. Rev. Genet. 2010; 11:459–463. [PubMed: 20548291]
26. Kang HM, et al. Variance component model to account for sample structure in genome-wide association studies. Nat. Genet. 2010; 42:348–54. [PubMed: 20208533]
27. Manichaikul A, et al. Robust relationship inference in genome-wide association studies. Bioinformatics. 2010; 26:2867–2873. [PubMed: 20926424]
28. Purcell S, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. Am. J. Hum. Genet. 2007; 81:559–575. [PubMed: 17701901]

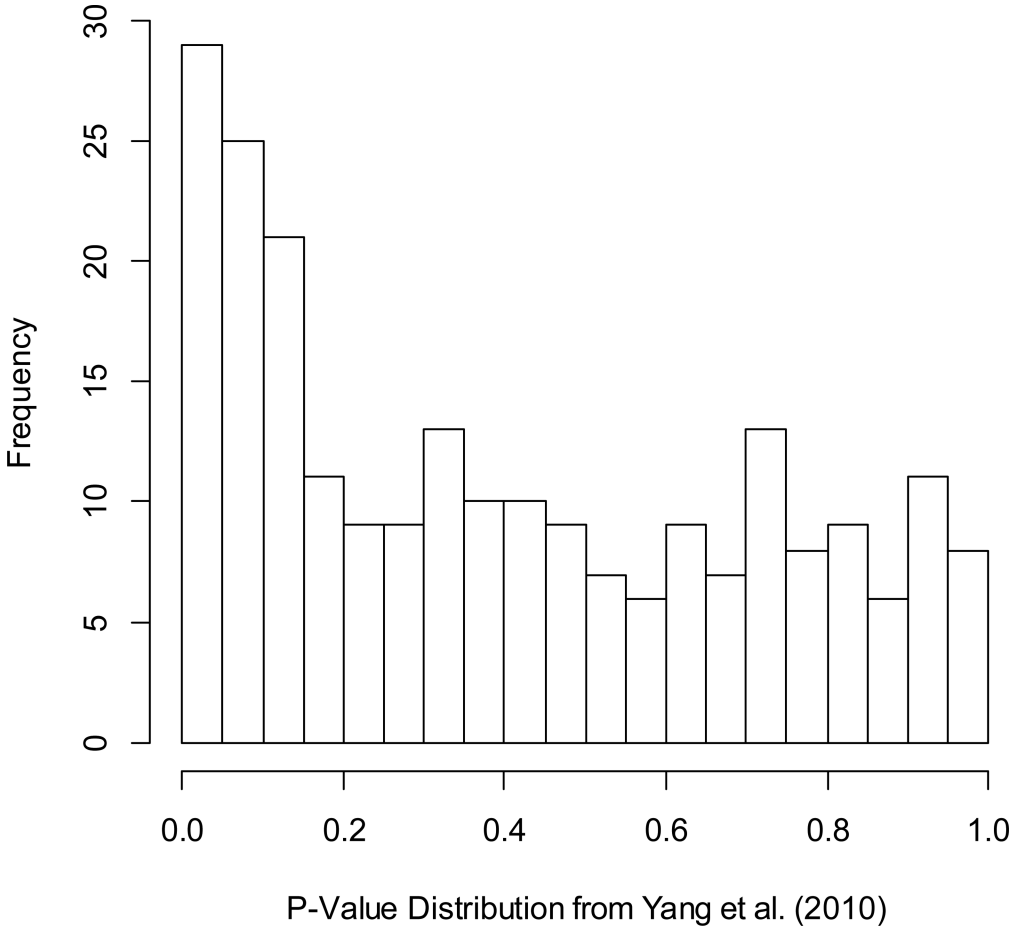

Figure 1.

Histogram of p-values from pairwise-chromosome regressions of relatedness as presented in Supplementary Table 2 of Yang et al. *Nat. Genet.* 42, 565–569 (2010) "Common SNPs explain a large proportion of the heritability for human height." Note excess of low p-values, particularly less than 0.10. This suggests that there is a significant pattern of covariance between independently segregating genomic segments and thus potential non-randomness in overall relatedness (i.e. potential covariance with population structure and thus environmental confounders): Kolmogorov-Smirnov test: (D^+ = 0.1892, p-value = 7.037e-08).

Table 1

GREML heritability estimates for shared childhood urbanicity, height, BMI and education.*

	h^2 No controls (2 PCs) A	h^2 Urban Control (2 PCs) B	|Δ| A - B	h^2 No controls (10 PCs) C	h^2 Urban Control (10 PCs) D	|Δ| C – D	h^2 No Controls (25 PCs) E	h^2 Urban Control (25 PCs) F	|Δ| C – D
Urban Childhd. N=6,439	.29155 [.0574]	n/a	n/a	.14767 [.0622]	n/a	n/a	.13787 [.0626]	n/a	n/a
Height N=6,379	.32489 [.0644]	.32510 [.0644]	.00022 [.0910]	.30397 [.0659]	.30397 [.0659]	.02092 [.0921]	.28410 [.0662]	.28338 [.0662]	.00072 [.0936]
BMI N=6,320	.31300 [.0674]	.31323 [.0675]	.00023 [.0953]	.31300 [.0674]	.3190 [.0678]	.00596 [.0956]	.29836 [.0682]	.29938 [.0682]	.00102 [.0964]
Educ. N=6,414	.17493 [.0650]	.15217 [.0652]	.02276 [.0921]	.1749 [.0650]	.15939 [.0656]	.01554 [.0923]	.14559 [.0660]	.12565 [.0661]	.01994 [.0934]

*Analysis includes white, non-Hispanic respondents in the Health and Retirement Study (HRS) for cryptic relatedness cut-off of 0.025. Two principal components control for population stratification in first set of analyses (A,B), ten PCs in second set of analyses (C,D) and 25 PCs in third set. Standard errors in brackets.