

# Genome-Wide Estimates of Heritability for Social Demographic Outcomes

BENJAMIN W. DOMINGUE,<sup>1</sup>  ROBBEE WEDOW,<sup>2</sup> DALTON CONLEY,<sup>3</sup> MATT MCQUEEN,<sup>4</sup> THOMAS J. HOFFMANN,<sup>5</sup> AND JASON D. BOARDMAN<sup>2</sup>

<sup>1</sup>Graduate School of Education, Stanford University, Stanford, California, USA

<sup>2</sup>Department of Sociology, Institute of Behavioral Science & Institute for Behavioral Genetics, University of Colorado–Boulder, Boulder, Colorado, USA

<sup>3</sup>Department of Integrative Physiology, Institute of Behavioral Science, University of Colorado–Boulder, Boulder, Colorado, USA

<sup>4</sup>Department of Sociology & Center for Genomics and Systems Biology, New York University, New York, New York, USA

<sup>5</sup>Department of Epidemiology & Biostatistics, and Institute for Human Genetics, University of California, San Francisco, San Francisco, California, USA

*An increasing number of studies that are widely used in the demographic research community have collected genome-wide data from their respondents. It is therefore important that demographers have a proper understanding of some of the methodological tools needed to analyze such data. This article details the underlying methodology behind one of the most common techniques for analyzing genome-wide data, genome-wide complex trait analysis (GCTA). GCTA models provide heritability estimates for health, health behaviors, or indicators of attainment using data from unrelated persons. Our goal was to describe this model, highlight the utility of the model for biodemographic research, and demonstrate the performance of this approach under modifications to the underlying assumptions. The first set of modifications involved changing the nature of the genetic data used to compute genetic similarities between individuals (the genetic relationship matrix). We then explored the sensitivity of the model to heteroscedastic errors. In general, GCTA estimates are found to be robust to the modifications proposed here, but we also highlight potential limitations of GCTA estimates.*

## Introduction

Demographic research often describes the factors responsible for variation in population health (Majer et al. 2013; Masters et al. 2014), health behaviors (Pampel and Denney 2011), birth outcomes (Fuller et al. 2014), and mortality (Ross, Masters, and Hummer 2012). Importantly, each of these outcomes has evidenced moderately sized heritability estimates (e.g., Rice et al. 2014; Daw et al. 2013). Not only are most physical health morbidities influenced by genetic factors common to family members (Pilia et al. 2006), but so are health-related lifestyles such as smoking (Boardman et al. 2011), exercise (Bartels et al. 2012; Mustelin et al. 2012), and birth outcomes, including birth weight, gestational age (Clausson, Lichtenstein, and Cnattingius 2000), and even mortality (Wienke et al.

Address correspondence to Benjamin W. Domingue, Stanford University, 520 Galvez Mall, CERAS Rm 510, Stanford CA 94305. E-mail: [ben.domingue@gmail.com](mailto:ben.domingue@gmail.com)

2001). Given that genes influence nearly all of the outcomes of interest to demographers, characterizing the relative contribution of genetic influences to health, health behaviors, birth outcomes, and mortality is critical for demographic researchers.

Heritability is the traditional approach to quantifying genetic influence on a trait. Heritability studies date back to Galton's work in the nineteenth century (e.g., Galton 1869). During the pre-genomic era the workhorse for estimating heritability had been the twin study, which utilizes family pedigrees. Currently there is a proliferation of genome-wide data from unrelated individuals in large, representative, longitudinal data sources such as the Health and Retirement Study, the National Longitudinal Study of Adolescent to Adult Health (McQueen et al. 2015), and many other, more targeted, datasets such as the Framingham Heart Study (Splansky et al. 2007). These studies have begun genotyping respondents and providing information on single nucleotide polymorphisms (SNPs) across the entire human genome. SNPs are common genetic variants and are the most convenient form of genome-wide data available for use by nongeneticists (Guo and Adkins 2008).

Initially, SNP data were the backbone of genome-wide association studies (GWAS), in which specific positions on the human genome are correlated with health phenotypes. This technique generates hundreds of thousands of (and now several million) regression estimates comparing genotype (e.g., zero, one, or two copies of the minor allele of the SNP) to phenotype (e.g., height) for each SNP. Novel genetic associations with many diseases have been found (Welter et al. 2014), but these individual loci only predict a small amount of observed phenotypic variation. For example, the associations identified in a GWAS for educational attainment (Rietveld, Medland, et al. 2013) explain only 0.02 percent of the observed variation.

It is also possible to utilize genetic similarity, based on information from the entire genome, among unrelated persons to decompose overall phenotypic variation into genetic and environmental components. The most common maximum likelihood methods used in these analyses are bundled in GCTA, a suite of software for genome-wide complex trait analysis (Yang et al. 2010; Yang et al. 2011). Although alternative techniques exist for computing such heritabilities (e.g., Ge et al. 2015), GCTA has been widely used and is relatively straightforward.

The key insight embedded in the GCTA approach is that measured SNP-level variation can be used to estimate the genetic similarity between two unrelated individuals, and this estimated genetic similarity can be compared to phenotypic similarity to produce a heritability estimate. A number of scholars are beginning to utilize these techniques. Table 1 contains a range of heritability estimates produced using GCTA that may be of interest to demographers. This is not intended as a comprehensive list of papers using GCTA that have been published, but is rather meant to provide a description of the types of GCTA outcomes that may be of interest to demographers and to illustrate the range of the associated heritability estimates. The estimates are grouped into different categories of phenotypes. One possible expectation might be for anthropometric phenotypes such as height to evince larger heritabilities than behavioral traits such as nicotine use and alcohol consumption. Height, for example, is driven largely by biology (outside of extreme nutritional environments), whereas decisions about nicotine and alcohol use are clearly influenced by peers and broader society. Yet heritability estimates between the two sets of outcomes are frequently quite similar. We also emphasize that heritabilities do not capture fundamental unchanging biological mechanisms but are instead highly contextual. Dating back to at least Feldman and Lewontin's characterization of heritability estimation as "local perturbation analyses" (1975, p. 1163), it has been understood that heritabilities are not fixed, immutable quantities but are contingent upon the social world in which the relevant actors are embedded.

**Table 1**  
Heritability estimates from applications of GCTA

Outcome	$h^2$ (SE)	Sample Size	Reference
<b>Anthropometric Phenotypes</b>			
Height	0.44 (0.09)	2,000	Speed et al. 2012
Height	0.35 (0.12)	3,154	Plomin et al. 2013
Height	0.32 (0.06)	6,379	Conley et al. 2014
Weight	0.42 (0.12)	3,154	Plomin et al. 2013
BMI	0.43 (0.10)	4,233	Boardman et al. 2015
BMI	0.31 (0.07)	6,320	Conley et al. 2014
<b>Medical/Clinical Phenotypes</b>			
Type 1 Diabetes	0.73 (0.06)	2,000	Speed et al. 2012
Type 1 Diabetes	0.28 (0.04)	2,599	Lee et al. 2011
Type 2 Diabetes	0.35 (0.06)	2,000	Speed et al. 2012
Rheumatoid Arthritis	0.57 (0.06)	2,000	Speed et al. 2012
Crohn's Disease	0.61 (0.08)	2,599	Lee et al. 2011
Crohn's Disease	0.54 (0.06)	2,000	Speed et al. 2012
Coronary Artery Disease	0.39 (0.06)	2,000	Speed et al. 2012
Pediatric Obesity	0.37 (0.15)	3,152	Llewellyn et al. 2013
Hypertension	0.42 (0.06)	2,000	Speed et al. 2012
Parkinson's Disease (Early Onset)	0.15 (0.14)	7,096	Keller et al. 2012
Parkinson's Disease (Late Onset)	0.31 (0.07)	7,096	Keller et al. 2012
Parkinson's Disease (All Types)	0.27 (0.05)	7,096	Keller et al. 2012
Parkinson's Disease	0.22 (0.02)	3,426	Do et al. 2011
Multiple Sclerosis	0.3 (0.02)	1,854	Watson et al. 2012
<b>Cognitive Phenotypes</b>			
General Cognitive Ability	0.35 (0.12)	3,154	Plomin et al. 2013
General Cognitive Ability	0.29 (0.05)	6,609	Marioni et al. 2014
Nonverbal Cognitive Ability	0.20 (0.11)	3,154	Plomin et al. 2013
Verbal Cognitive Ability	0.26 (0.11)	3,154	Plomin et al. 2013
Language Ability	0.29 (0.12)	3,154	Plomin et al. 2013
Intelligence (Age 7–12)	0.60 (0.26)	2,875	Trzaskowski et al. 2014a
Intelligence	0.51 (0.02)	3,511	Davies et al. 2011
Intelligence from Childhood to Old Age	0.24 (0.20)	1,940	Deary et al. 2012
IQ (Age 12)	0.32 (0.14)	3,000	Trzaskowski et al. 2014b
IQ (Age 7)	0.28 (0.17)	3,000	Trzaskowski et al. 2014b
<b>Psychological Phenotypes</b>			
Bipolar Disorder	0.59 (0.06)	2,000	Speed et al. 2012
Bipolar Disorder	0.37 (0.04)	2,599	Lee et al. 2011

(Continued)

**Table 1**  
(Continued)

Outcome	$h^2$ (SE)	Sample Size	Reference
ADHD	0.42 (0.13)	1,040	Yang et al. 2013
Adult Antisocial Behavior	0.55 (0.41)	2,172	Tielbeek et al. 2012
Depression	0.19 (0.10)	4,233	Boardman et al. 2015
Major Depressive Disorder	0.32 (0.09)	4,605	Lubke et al. 2012
Behavioral Disinhibition	0.19 (0.16)	3,452	Vrieze et al. 2013
Neuroticism	0.06 (0.03)	12,000	Vinkhuyzen et al. 2012
Borderline Personality Features	0.23 (0.09)	7, 125	Lubke et al. 2014
Callous-Emotional Behavior	0.07 (0.12)	2,930	Viding et al. 2013
Extraversion	0.12 (0.03)	12,000	Vinkhuyzen et al. 2012
Anxiety Related Behaviors	0.01–0.12 (0.12)	2,810	Trzaskowski et al. 2013
<b>Substance Dependency Phenotypes</b>			
Drug Use	0.22 (0.16)	3,452	Vrieze et al. 2013
Drug Dependence	0.36 (0.13)	2,596	Palmer et al. 2015
Dependence Vulnerability	0.33 (0.13)	2,596	Palmer et al. 2015
Problematic Drug Use	0.25 (0.13)	2,596	Palmer et al. 2015
Alcohol Consumption	0.16 (0.16)	3,452	Vrieze et al. 2013
Alcohol Dependence	0.12 (0.16)	3,452	Vrieze et al. 2013
Nicotine Use/Dependence	0.18 (0.16)	3,452	Vrieze et al. 2013
<b>Sociological/Health Behavior/Educational Phenotypes</b>			
Socioeconomic Background	0.18 (0.05)	6,533	Marioni et al. 2014
Socioeconomic Status (Age 2)	0.18 (0.12)	3,000	Trzaskowski et al. 2014b
Socioeconomic Status (Age 7)	0.19 (0.12)	3,000	Trzaskowski et al. 2014b
Subjective Well-Being	0.05–0.10 (0.05–0.10)	11,500	Rietveld, Cesarini, et al. 2013
Reporting Stressful Life Events	0.3 (0.15)	2,578	Power et al. 2013
Self-Rated Health	0.18 (0.10)	4,233	Boardman et al. 2015
Moderate to Vigorous Activity	0.17 (0.09)	4,244	Richmond et al. 2014
Sedentary Time	0.25 (0.09)	4,244	Richmond et al. 2014
Total Physical Activity	0.21 (0.10)	4,244	Richmond et al. 2014
Education	0.21 (0.05)	6,578	Marioni et al. 2014
Education	0.33 (0.10)	4,233	Boardman et al. 2015
Education	0.17 (0.07)	6,414	Conley et al. 2014

Although GCTA holds promise, great care needs to be used in the application of these methods to obtain credible results. This article is meant, in part, to act as a guide for demographers who are potentially new to genetic analyses and are interested in conducting a heritability study. It builds on the work of the GCTA development team (e.g., Visscher, Yang, and Goddard 2010; Yang et al. 2010; Yang et al. 2011) and others (Conley et al. 2014) who are well aware of the need for caution in the application of these methods. We begin by describing the method for an audience with minimal training in genetics. We then present three empirical examples demonstrating the sensitivity of GCTA estimates to certain “twists” in the typical approach to using this model. This work is not meant as a critique of the model but is intended to illuminate how the method works and its potential limitations.

## The Method

The core insight underlying the estimation of heritability in both twin studies and with GCTA is that if genetic variation accounts for some measure of phenotypic variation, then more genetically similar pairs should be more phenotypically similar. Clearly, this depends upon being able to measure genetic similarity. In twin, extended twin, or family studies, the estimation of genetic similarity occurs only between family members and is trivial since the family relationships are known and pairs receive their expected identity by descent (IBD) value (e.g. 0.5 for full siblings and dizygotic twins, 0.25 for half-siblings, etc.). With GCTA, we estimate genetic similarity between all pairs of unrelated individuals (with  $n$  unrelated individuals, there are  $\frac{n^2-n}{2}$  possible pairs) on the set of genetic markers in question.<sup>1</sup> We emphasize that the metric for similarity used in GCTA is just one of many possible metrics (Speed and Balding 2014 describe alternatives). Second, a restricted maximum likelihood (REML) estimate of heritability is computed by comparing phenotypic similarity to genetic similarity. We describe these steps in more detail in the following sections.

### Estimating Genetic Similarity

The genetic similarity  $A_{jk}$  between individual  $j$  and individual  $k$  is estimated as (Equation 3, Yang et al. 2011; Equation 5, Yang et al. 2010)

$$A_{jk} = \frac{1}{N} \sum_i \frac{(x_{ij} - 2p_i)(x_{ik} - 2p_i)}{2p_i(1 - p_i)} \quad (\text{Eq. 1})$$

where  $N$  is the number of available genetic markers,  $i$  indexes these markers,  $x_{ij}$  and  $x_{ik}$  are the number of minor alleles at SNP  $i$  for individuals  $j$  and  $k$ , respectively, and  $p_i$  is the minor allele frequency.<sup>2</sup> Genotypes are effectively standardized so that the sample variance is independent of allele frequency. At this stage, it is important to pause to note the consequences of the fact that the genetic similarity is estimated based on the full set of  $N$  markers and not the subset of causal variants that would, ideally, be of interest. The dilemma is that one does not know the set of true causal variants. The causal variants are unlikely to be

<sup>1</sup>GCTA analyses nearly always focus on SNPs rather than other genetic variants. In this article, genetic markers and variants will be used interchangeably for SNPs.

<sup>2</sup>Diagonal elements of  $A$  (when  $j = k$ ) are inbreeding coefficients. We do not discuss them further here since they are of marginal interest in the estimation of heritability (see Yang et al. 2011 for information on their calculation).

a random sample of markers. In particular, they are likely to be a sample with relatively low minor allele frequencies (see Yang et al. 2011). This has implications, as the quality of the heritability estimate based on  $A_{jk}$  will only be as good as the approximation of  $A_{jk}$  to the genetic similarity on the causal variants. For polygenic traits based on many common variants, heritability estimates based on  $A_{jk}$  should be accurate. However, traits associated with rare variants are not a good target for GCTA analyses (see Zuk et al. 2014 on working with rare variants).

### ***Estimating Heritability***

The model for decomposing phenotypic variation is

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{g} + \boldsymbol{\epsilon} \quad (\text{Eq. 2})$$

where  $\mathbf{X}$  is an optional matrix of covariates,  $\mathbf{g}$  is a vector of random effects, and  $\boldsymbol{\epsilon}$  is a vector of errors each with variance  $\sigma^2_{\boldsymbol{\epsilon}}$ . Standard assumptions regarding  $\boldsymbol{\epsilon}$  apply, namely that it is independent of  $\mathbf{X}$  and  $\mathbf{g}$ . The genetic similarity matrix  $A$  enters here through the assumption that

$$\mathbf{g} \sim \text{Normal}[0, \sigma_g^2 A] \quad (\text{Eq. 3})$$

where  $A$  is the matrix of similarity estimates. Heritability is defined as a ratio of the variance of genetic effects to the total variance:

$$\frac{\sigma_g^2}{\sigma_g^2 + \sigma_{\boldsymbol{\epsilon}}^2} \quad (\text{Eq. 4})$$

Heritability is intuitive in Eq. 4 in the sense that we see it is the fraction of total variance accounted for by genetic random effects.

Eq. (2) is estimated using REML. REML is preferred to normal maximum likelihood (ML) estimation, since it leads to improved estimation of variance components (Harville 1977). In contrast to ML estimation, REML focuses on a likelihood function that is independent of nuisance parameters and should, therefore, provide more reliable variance parameter estimates. Additional details on the estimation technique used here can be found in Gilmour, Thompson, and Cullis (1995).

### ***Key Data Requirements***

GCTA should only be applied to a sample that has already been through a quality control (QC) process, including pruning for missingness, minor allele frequency (MAF) thresholds (e.g., an MAF below 0.05, a common threshold for identifying an SNP as a “common” rather than a “rare” variant), and Hardy-Weinberg equilibrium.<sup>3</sup> Dichotomous traits may require even stricter controls (Lee et al. 2011). Statistical power is an important aspect of

<sup>3</sup>Hardy-Weinberg equilibrium (HWE) occurs when observed genotypes match expected genotypes given a particular minor allele frequency. If the minor allele  $a$  has frequency  $p$ , then the genotype frequencies should be  $p^2$  (for homozygous minor allele  $aa$ ),  $2pq$  for the heterozygotes (e.g.,  $ab$  and  $ba$ ), and  $q^2$  for the homozygous major allele. Deviations from HWE are used to detect genotyping errors, deviations from random mating, and genetic drift.

GCTA (Visscher et al. 2014), and while an online tool<sup>4</sup> is available for this purpose, a rule of thumb is that at least 5,000 respondents are needed to detect heritability less than 0.2 (see Figure 3 in Visscher et al. 2014). It is also important that the data be composed of genetically homogeneous respondents. This is due to the sensitivity of Eq. 1 to population stratification in which allele frequencies may differ across socially defined racial and ethnic groups (i.e.,  $p_i$  in Eq. 1 changes substantially across groups). The significance of this issue is shown quite clearly in Figure S5 of Domingue et al. (2014), which demonstrates that black spouses from the Health and Retirement Study are estimated to have extreme genetic similarities as a result of the fact that the majority of the sample is made up of non-Hispanic whites. That is, small and typically meaningless differences in minor allele frequencies among non-Hispanic black and white populations for certain portions of the human genome translate to excessive levels of similarity among same-race groups that may have important implications for the interpretation of heritability estimates.

Even among a racially homogeneous groups of respondents, there may still be a concern that population stratification is biasing the results. Figure 2 of Nelis et al. (2009) suggests that even among racially homogenous groups, population stratification remains. One standard technique for adjusting for such population stratification is through the inclusion of principal components (Price et al. 2006). Such an approach was taken in the original study of height (Yang et al. 2010), and it is probably prudent to consider such adjustments. However, principal components may also adjust for meaningful differences (in terms of the trait in question) between individuals and thus may lead to underestimated heritabilities. We would thus encourage users to report adjusted and unadjusted estimates of heritability when appropriate (i.e., when the values differ).

### *Caveats*

The GCTA approach bypasses our lack of knowledge regarding the true causal variants by assuming that these causal variants are distributed throughout the genome in such a way that an estimate of genome-wide similarity is a suitable proxy for similarity on the causal SNPs. It is important to note that this logic only applies to certain traits. Alzheimer's is an interesting counterexample. The  $e4$  allele of APOE is well known to be a strong genetic risk for developing Alzheimer's (Genin et al. 2011). For carriers of  $e4$ , the probability of developing the disease is substantially elevated compared to noncarriers, regardless of their overall genetic similarity to fellow carriers. For complex traits that are completely polygenic (e.g., the causal variants are large in number but weak in effect size), it is reasonable to inquire how consistent estimates of genetic similarity are over different sets of markers that might be used to compute heritability. This is the empirical focus of Example 1 (see next section).

Although we still have only limited knowledge about the variants that underlie complex traits, over the last 10 years there has been a large-scale hunt for the genetic variants that underlie specific diseases, traits, and other attributes such as education (Rietveld, Medland, et al. 2013). The key technique in linking phenotype and genotype is the previously discussed GWAS approach. Given that we now have a large number of GWAS results, it is natural to inquire about potential changes in GCTA estimates if estimates were computed based on genetic similarities from SNPs known to be associated with the relevant outcome. We use information from a GWAS on height to inquire about the sensitivity of GCTA estimates to causal variants that are known to underlie a trait in Example 2 (see next section).

<sup>4</sup>See <http://spark.rstudio.com/ctgg/gctaPower/>.

Estimation of Eq. 2 is premised on additional assumptions that one might question. Just as in the case of a simple linear model, one key assumption underlying estimation is that the errors are of constant variance (i.e., homoscedastic). Heteroscedasticity is a common problem in applied settings that typically leads to incorrect estimates of standard errors. Given that GCTA focuses on a ratio containing the estimated error variance, heteroscedasticity could have important implications here. We examine the consequences for heritability estimates if the error term is heteroscedastic in Example 3.

## Examples

The following examples rely upon data from non-Hispanic white adults (born between 1900 and 1970, but with the majority born between 1930 and 1940) in the Health and Retirement Study (HRS).<sup>5</sup> DNA samples were collected via buccal swabs in 2006 and via saliva samples in 2008. Genotype calls were then made based on a clustering of both datasets using the Illumina HumanOmni2.5-4v1 array. Details on this process can be found online at the HRS website. After standard quality control procedures (e.g., removing SNPs that were missing in more than 5 percent of samples, MAF below 1 percent, failure to meet Hardy-Weinberg equilibrium; complete details are available upon request), we retained 1,698,845 SNPs. From this sample of SNPs, the main genetic similarity estimates were computed based on 1,473,658 SNPs (comprising only autosomal SNPs, which were also pruned slightly because of a second MAF filter imposed by GCTA) for 4,950 non-Hispanic whites (those from the full sample of non-Hispanic whites who had no missing data on several key variables). With this sample, we obtained reasonable heritability estimates: cognition 0.23, height 0.40, weight 0.25, and educational attainment 0.33 (all standard errors are 0.09, which is to be expected given Figure 1 of Visscher et al. 2014).<sup>6</sup>

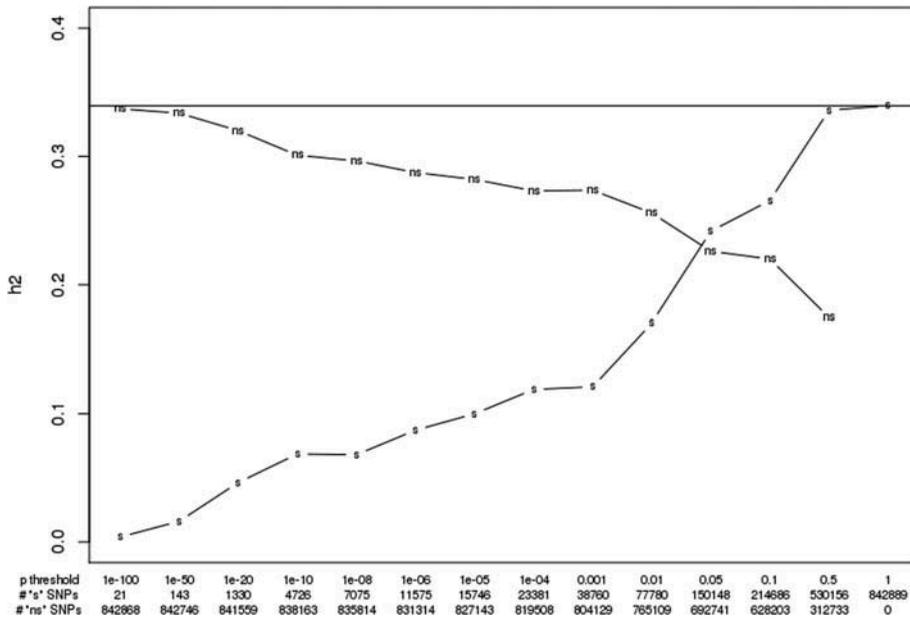
### *Example 1: Sensitivity of Genetic Similarity to the Set of SNPs*

Heritability estimates rely upon genetic similarity, which may be sensitive to the choice of markers. We first chose SNPs that were pruned from the full set to ensure that they were in linkage equilibria (for different thresholds). Linkage disequilibrium arises when genetic markers at nearby locations are correlated as a result of the fact that large segments of DNA are inherited together. Although genetic similarity is frequently computed via sets of markers that have not been pruned for linkage disequilibrium (LD), Speed and Balding (2014, p. 8) note that the use of multiple SNPs in regions of high LD can have consequences for heritability estimates. We also considered randomly chosen sets of markers that were 10 percent, 30 percent, and 50 percent of the full sample of SNPs. Given the underlying philosophy of GCTA, heritability estimates based on reasonably large subsamples of markers should be similar to those based on the full sample of markers. This requires that the different samples of markers produce similarity estimates that are highly correlated.

Table 2 presents the correlations between the genetic similarity estimates (greater than 0.025, as might be used in a heritability analysis) based on the various sets of markers. We focus on the correlations between the similarity estimates from the full set of markers and the similarity estimates from the various subsets (the bolded column). To begin with, consider that the genetic relationship values for all persons  $i$  and  $j$  are correlated

<sup>5</sup>Specifically, the RAND fat files, available at <http://www.rand.org/labor/aging/dataproduct/enhanced-fat.html>.

<sup>6</sup>All variables except educational attainment were taken from Wave 8.



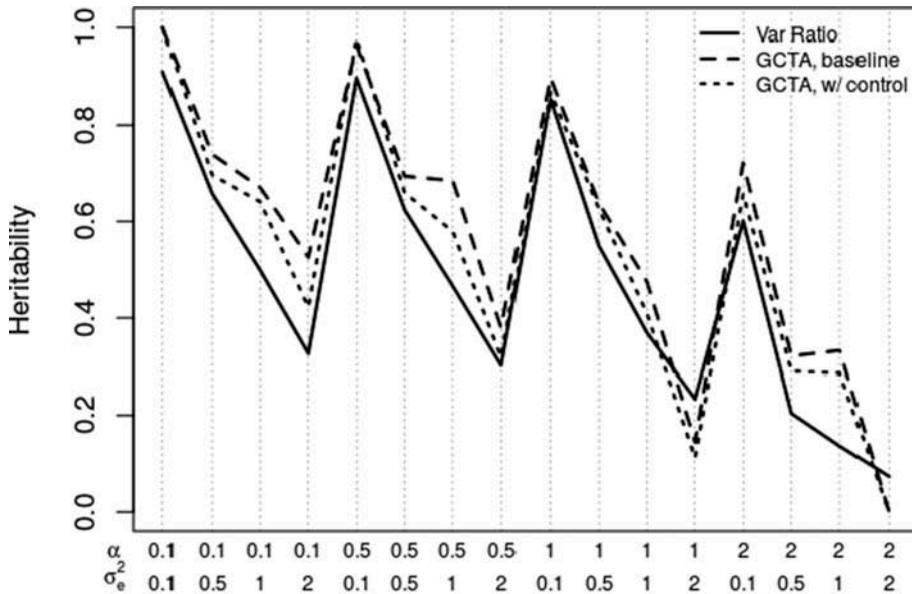
**Figure 1.** GWAS-informed GCTA heritability results for height. The heritabilities are computed for different sets of markers. The “ns” lines are based on only those markers with a  $p$  value above the given  $p$ -value threshold (not significant markers), while the “s” lines are based on only those markers with a  $p$  value below the given  $p$ -value threshold (significant markers) for which the  $p$  value for each marker is taken from a large GWAS on height (Wood et al. 2014). Confidence intervals are omitted since they are consistent. Standard errors for “ns” estimates are less than 0.02 for  $p$ -value thresholds of  $1e-04$  and below, and less than 0.07 for larger thresholds. The standard GRM pruning threshold of 0.025 is not included, since it leads to the exclusion of large numbers of sample participants in cases using relatively small (e.g., fewer than 100,000) markers.

at 0.57 when we examine their genetic similarity based on all SNPs compared to their genetic similarity using only SNPs that are not in LD using the most conservative threshold. However, when we increase the  $r^2$  threshold from 0.01 to 0.2, the correlation jumps from 0.57 to 0.75. Increasing the threshold again to 0.5, the correlation rises to 0.88. These values can be interpreted using a comparison to correlations between the full set of markers and a random subset of markers. When we compute genetic similarities based on random subsets of SNPs, the correlations are generally high ( $> 0.9$ ), except for the 10 percent sample. Nevertheless, even when we only use 10 percent of the SNPs, we present relationship estimates that are correlated with the overall genetic relationship matrix (GRM) at a value of 0.83.

We now turn to the impact of the differences in the estimates of genetic similarity on the estimated heritability. To address this, we computed heritability estimates for height based on the various sets of SNPs (italicized column of Table 2). We dropped any pair with a relationship greater than 0.025, since these are typically excluded in the calculation of heritability. The full set of markers produces an estimate of 0.40, which, it should be noted, is identical to the estimate in the original GCTA paper (Yang et al. 2010). The 10 percent random sample of SNPs produces a substantially lower estimate of 0.31, but the 30 percent and 50 percent samples produce estimates much closer to 40 percent. Interestingly, the sets of markers pruned only modestly for LD produce slightly higher estimates of heritability,

**Table 2**  
Correlations between similarity estimates based on different sets of markers and their resulting heritability estimates for height

	$r = 0.01$	$r = 0.2$	$r = 0.5$	10% Sample	30% Sample	50% Sample	Full	$h^2$	$se$
$r = 0.01$	1.00	0.52	0.53	0.47	0.54	0.56	<b>0.57</b>	0.10	0.06
$r = 0.2$	0.52	1.00	0.88	0.63	0.71	0.73	<b>0.75</b>	0.44	0.10
$r = 0.5$	0.53	0.88	1.00	0.73	0.83	0.86	<b>0.88</b>	0.42	0.11
10% Sample	0.47	0.63	0.73	1.00	0.79	0.81	<b>0.83</b>	0.31	0.07
30% Sample	0.54	0.71	0.83	0.79	1.00	0.93	<b>0.95</b>	0.39	0.08
50% Sample	0.56	0.73	0.86	0.81	0.93	1.00	<b>0.98</b>	0.38	0.08
Full	0.57	0.75	0.88	0.83	0.95	0.98	<b>1.00</b>	0.40	0.09
$N$ SNPs	61,904	331,433	626,182	169,884	509,654	849,422	<b>1,698,845</b>		



**Figure 2.** Comparison of true ratio of genetic to total variance to GCTA estimates (with and without control for the variable associated with the heteroscedastic error) in presence of heteroscedastic errors.

with the exception of the rather extreme case of  $r^2 = 0.01$ . Although the rise isn't large, this effect has also been observed (Vilhjalmsson et al. 2015) in the context of genetic risk scores (indices derived from GWAS studies meant to predict a given phenotype; see Belsky et al. 2013 for a review of this method). When the LD pruning threshold is quite stringent (0.01), the heritability estimate is only 0.1. Thus, by calculating genetic similarity using only SNPs that are independent of one another, we reduce the heritability estimate by roughly three-quarters. However, this estimate is based on a relatively small number of SNPs ( $N = 61,904$ ). The next example continues to examine the sensitivity of heritability results to the choice of SNPs, but with the subsample of SNPs chosen in a different manner.

### **Example 2: Incorporation of GWAS Information**

For some traits, such as height, there is now high-quality information available about which SNPs “matter.” Thus, we can use published GWAS results to decide which SNPs to include in the GRM, and heritability estimates can be limited to markers with significant  $p$  values (Wood et al. 2014). To use the results from this GWAS, we first selected a set of 842,889 SNPs that are in both the GWAS and our genetic database of SNPs. Based on these SNPs, we estimated a GCTA heritability for a height of 0.34, which is reduced from the original estimate of 0.40 using the full set of markers, but still significant ( $SE = 0.077$ ).<sup>7</sup> This is an important observation (consistent with Example 1), because although we eliminated roughly 50 percent of the SNPs, the heritability was only reduced

<sup>7</sup>As noted previously, it is a standard practice in GCTA to remove individuals from pairs with estimated genetic similarities greater than 0.025 (in the metric established by Eq. 1) to ensure that no closely related (e.g., parent–offspring, siblings, etc.) individuals are included. Such individuals may share a common environment that may bias the resulting heritability estimate. However, we do not include such a threshold here because of the fact that the changing number of markers has

by roughly 20 percent. This bolsters support for the GWAS results, but also highlights that much of the information across the genome is not necessary for reliable indicators of heritability.

For a given  $p$ -value threshold, we designate two sets of markers. The first set of markers, those with a  $p$ -value greater than the threshold, are designated “ns” (for “not significant”). These are the markers that are unassociated with height, as judged by the  $p$ -value threshold from the height GWAS (Wood et al. 2014). The second set of markers, those markers with  $p$  values less than the threshold (e.g., those SNPs that are deemed to be associated with height), are designated by “s” (for “significant”). Consider Figure 1. The horizontal line shows the GCTA heritability of 0.34. The other two lines show the GCTA estimates for the “ns” and “s” SNPs using a range of thresholds (the  $p$ -value threshold and the number of markers for each set of SNPs are shown on the  $x$ -axis). At the far left, we start with a threshold of  $1e-100$ . This is an extreme threshold (only 21 SNPs reach such a level of significance), and the heritability computed for the 842,868 “ns” markers is very nearly the original estimate. The estimate of the heritability from the “ns” markers above the threshold stays above 0.25 until the 0.05 threshold. Even after the removal of 312,733 SNPs at the 0.5 threshold, there is still a statistically significant heritability “ns” estimate of 0.17. This is noteworthy, since we have removed any marker remotely associated with height. The fact that GCTA does not explicitly utilize information related to causal SNPs is very clear.

Now let us consider the curve associated with heritability estimates from the “s” markers. The 21 markers that are the most predictive of height produce genetic similarity estimates that lead to a heritability estimate of 0.004. This is not surprising, since collectively these markers predict only a very small amount of variability of height. One can observe a slow rise in the estimated heritability of height as the  $p$ -value threshold is relaxed (so that increasing numbers of SNPs fall in the “s” category). The curves cross around the 0.05 threshold, meaning that similarities in height are better explained by similarities on the 150,148 SNPs below this threshold rather than by the 692,741 SNPs above this threshold.

### Example 3: Heteroscedastic Outcome

In many empirical settings, the assumption of a constant error variance is questionable. To probe the performance of GCTA in such cases, we simulate an outcome in which the variance of the errors is a function of an individual’s height. We generate data using

$$y = 0.5 * \text{height} + \mathbf{g} + \epsilon \quad (\text{Eq. 5})$$

where  $\epsilon_i$  is normally distributed with variance  $e^{(\alpha * \text{height}_i)} * \sigma_\epsilon^2$  (where height is standardized). The degree of heteroscedasticity is controlled via  $\alpha$  (note that when  $\alpha = 0$ , the errors are homoscedastic) such that there is a greater variance in the  $\epsilon_i$  for tall individuals. This has clear implications for the definition of heritability, since Eq. 4 depends on  $\sigma_\epsilon^2$ . We fix  $\sigma_g^2 = 1$  and control the level of heritability via  $\sigma_\epsilon^2$  (increasing this variance decreases heritability, and vice versa). In our simulation, we use the observed ratio of the variability of the genetic component to the total observed variability (these quantities are available only because the data are simulated and thus completely known) as a metric for heritability, but advise the reader that this ratio is not identical to the GCTA definition.

major implications for the number of pairs that fall below this threshold. We did remove 347 individuals from these analyses, such that the original set of genetic similarity estimates are all below the 0.025 threshold.

Indeed, the quantity of interest in GCTA is poorly defined because of the nonconstant error variance. Thus, instead of exact recovery, we focus on the relevant patterns.

Figure 2 compares the variance ratio discussed previously (solid line) to GCTA estimates that do not (dashed line) and do (dotted line) include height as a covariate. The three sets of estimates consistently move together. The GCTA estimate when the control is included tends to be closer to the observed ratio of variances than the estimate without the control, but again we caution that the observed variance here is a somewhat amorphous quantity since the error variance is nonconstant. Importantly, the baseline estimates of heritability from GCTA (estimates that do not include height as a predictor) are robust to heteroscedasticity. There may not always be an identifiable correlate of the error variance, so it is reassuring to know that relatively reliable information regarding heritability can still be recovered in such cases.

## Discussion

The examples considered here help to illustrate two key points about GCTA. First, Examples 1 and 2 illustrate the fact that GCTA is a method for computing heritability based on genome-wide similarity. Example 1 illustrates the relative consistency of results as long as sufficient samples of SNPs are used. Example 2 illustrates the fact that one does not need to include SNPs thought to be causal for GCTA to estimate heritability. Of course, if too many of these SNPs are removed, the estimate may start to suffer (note the decline in the “ns” line toward the right of Figure 2). Second, Example 3 suggests that GCTA estimates are relatively robust to heteroscedasticity. Intuitively, there is reason to be concerned about heteroscedasticity, since GCTA is based on estimates of variance components. While GCTA estimates are likely to overestimate heritability in the presence of heteroscedasticity, the bias does not seem extreme, and relevant information regarding heritability may still be obtained.

This article adds to the evidentiary base regarding GCTA’s performance in the face of violations of the underlying assumptions. An additional concern is that genetic similarity may be associated with environmental similarities. If that is the case, then these environmental similarities could be the true cause of phenotypic similarities between respondents, rather than the genetic similarities studied via GCTA. Other research (Conley et al. 2014) has considered this possibility. The environments studied in that research (e.g., childhood urbanicity and parental education) did not seem to bias GCTA estimates for other, putatively heritable, outcomes such as height. Later research (Conley et al. 2015) has tried to explore this issue further in a more nuanced manner by decomposing the correlation between parent and offspring education levels into genetic and environmental components, but has focused on genetic predisposition toward educational attainment (as determined by an educational polygenic risk score) rather than GCTA heritability.

There are several additional applications of GCTA that this article does not explore. We focus here on two: heritability by environment and bivariate analyses. There is ample reason to think that the relative influence of genotype on phenotype varies across environmental context. GCTA allows one to model the effect of environment on heritability, but the ability to adjust for environmental differences is not a cure-all. The relevant environments may be unknown, unobserved, or poorly measured. Even when there is a promising candidate for the appropriate environment, GCTA analyses suggesting environmental differences must be interpreted with caution. For example, if environmental differences are associated with, say, ethnic differences, then population stratification could be an issue.

In such a case, LD patterns between the causal SNPs and other markers across the two ethnic groups may be different. It could also be the case that genetic or phenotypic variation may be constrained in one environment relative to the other. For that matter, the phenotype could be measured with less fidelity in certain environments. All of these scenarios could potentially lead to HxE findings via GCTA and yet would not necessarily indicate that there is truly a difference in the influence of genotype across environment.

Finally, bivariate GCTA models (Lee et al. 2012) are an interesting method for engaging in genetically informed demographic research. This method yields an estimate of genetic correlation ( $r_G$ ) between two traits that indicates whether an observed correlation between traits, such as height and weight, is due to common genetic factors. For example, Boardman, Domingue, and Daw (2015) used this method to show that a non-negligible proportion of the correlation between education and self-rated health appears to be confounded with genes that influence both traits. Such genetic associations may underlie many variables frequently considered in demographic inquiry, and a failure to account for these associations may lead to forms of omitted variable bias.

## ORCID

Benjamin W. Domingue  <http://orcid.org/0000-0002-3894-9049>

## Funding

This research was supported in part by the following grants from the Eunice Kennedy Shriver National Institute of Child Health and Human Development (NICHD): R01HD060726 and R21HD078031. We also received support from the NICHD-supported University of Colorado Population Center (CUPC R24 HD066613), and Wedow was supported by the National Science Foundation's Graduate Research Fellowship Program (DGE 1144083). The HRS (Health and Retirement Study) is sponsored by the National Institute on Aging (NIA U01AG009740) and is conducted by the University of Michigan.

## References

- Bartels, M., M. H. de Moor, N. Van der Aa, D. I. Boomsma, and E. J. de Geus. 2012. Regular exercise, subjective wellbeing, and internalizing problems in adolescence: Causality or genetic pleiotropy? *Frontiers in Genetics* 3 (4). doi:10.3389/fgene.2012.00004.
- Belsky, D. W., T. E. Moffitt, K. Sugden, B. Williams, R. Houts, J. McCarthy, and A. Caspi. 2013. Development and evaluation of a genetic risk score for obesity. *Biodemography and Social Biology* 59 (1):85–100. doi:10.1080/19485565.2013.774628.
- Boardman, J. D., C. L. Blalock, F. C. Pampel, P. K. Hatemi, A. C. Heath, and L. J. Eaves. 2011. Population composition, public policy, and the genetics of smoking. *Demography* 48 (4):1517–33. doi:10.1007/s13524-011-0057-9.
- Boardman, J. D., B. W. Domingue, and J. Daw. 2015. What can genes tell us about the relationship between education and health? *Social Science & Medicine* 127:171–80. doi:10.1016/j.socscimed.2014.08.001.
- Clausson, B., P. Lichtenstein, and S. Cnattingius. 2000. Genetic influence on birth weight and gestational length determined by studies in offspring of twins. *BJOG: An International Journal of Obstetrics & Gynaecology* 107 (3):375–81. doi:10.1111/bjo.2000.107.issue-3.
- Conley, D., D. Cesarini, C. Dawes, B. Domingue, and J. Boardman. 2015. Is the effect of parental education on offspring biased or moderated by genotype? *Sociological Science* 2:82–105. doi:10.15195/v2.a6.

- Conley, D., M. L. Siegal, B. W. Domingue, K. M. Harris, M. B. McQueen, and J. D. Boardman. 2014. Testing the key assumption of heritability estimates based on genome-wide genetic relatedness. *Journal of Human Genetics* 59 (6):342–45. doi:10.1038/jhg.2014.14.
- Davies, G., A. Tenesa, A. Payton, J. Yang, S. E. Harris, D. Liewald, and I. J. Deary. 2011. Genome-wide association studies establish that human intelligence is highly heritable and polygenic. *Molecular Psychiatry* 16 (10):996–1005. doi:10.1038/mp.2011.85.
- Daw, J., M. Shanahan, K. M. Harris, A. Smolen, B. Haberstick, and J. D. Boardman. 2013. Genetic sensitivity to peer behaviors: 5HTTLPR, smoking, and alcohol consumption. *Journal of Health and Social Behavior* 54 (1):92–108. doi:10.1177/0022146512468591.
- Deary, I. J., J. Yang, G. Davies, S. E. Harris, A. Tenesa, D. Liewald, and P. M. Visscher. 2012. Genetic contributions to stability and change in intelligence from childhood to old age. *Nature* 482 (7384):212–15.
- Do, C. B., J. Y. Tung, E. Dorfman, A. K. Kiefer, E. M. Drabant, U. Francke, and N. Eriksson. 2011. Web-based genome-wide association study identifies two novel loci and a substantial genetic component for Parkinson's disease. *PLoS Genetics* 7 (6):e1002141. doi:10.1371/journal.pgen.1002141.
- Domingue, B. W., J. Fletcher, D. Conley, and J. D. Boardman. 2014. Genetic and educational assortative mating among US adults. *Proceedings of the National Academy of Sciences* 111 (22):7996–8000. doi:10.1073/pnas.1321426111.
- Feldman, M. W., and R. C. Lewontin. 1975. The heritability hang-up. *Science* 190 (4220):1163–68. doi:10.1126/science.1198102.
- Fuller, T. D., C. N. Spracklen, K. K. Ryckman, L. A. Knake, T. D. Busch, A. M. Momany, and J. M. Dagle. 2014. Genetic variation in CYB5R3 is associated with methemoglobin levels in preterm infants receiving nitric oxide therapy. *Pediatric Research* 77 (3):472–476. doi:10.1038/pr.2014.206.
- Galton, F. 1869. *Hereditary genius*. London: Macmillan and Company.
- Ge, T., T. E. Nichols, P. H. Lee, A. J. Holmes, J. L. Roffman, R. L. Buckner, M. R. Sabuncu, and J. W. Smoller. 2015. Massively expedited genome-wide heritability analysis (MEGHA). *Proceedings of the National Academy of Sciences of the United States of America* 112 (8):2479–84. doi:10.1073/pnas.1415603112.
- Genin, E., D. Hannequin, D. Wallon, K. Sleegers, M. Hiltunen, O. Combarros, and C. Van Broeckhoven. 2011. APOE and Alzheimer disease: A major gene with semi-dominant inheritance. *Molecular Psychiatry* 16 (9):903–07. doi:10.1038/mp.2011.52.
- Gilmour, A. R., R. Thompson, and B. R. Cullis. 1995. Average information REML: An efficient algorithm for variance parameter estimation in linear mixed models. *Biometrics* 51:1440–50. doi:10.2307/2533274.
- Guo, G., and D. E. Adkins. 2008. How is a statistical link established between a human outcome and a genetic variant? *Sociological Methods & Research* 37 (2):201–26. doi:10.1177/0049124108324526.
- Harville, D. A. 1977. Maximum likelihood approaches to variance component estimation and to related problems. *Journal of the American Statistical Association* 72 (358):320–38. doi:10.1080/01621459.1977.10480998.
- Keller, M.F., M. Saad, J. Bras, F. Bettella, N. Nicolaou, J. Simón-Sánchez, F. Mittag, F. Büchel, M. Sharma, J.R. Gibbs and C. Schulte, et al. 2012. Using genome-wide complex trait analysis to quantify 'missing heritability' in Parkinson's disease. *Human Molecular Genetics* 21 (22): 4996–5009. doi:10.1093/hmg/ds335.
- Lee, S. H., T. R. DeCandia, S. Ripke, J. Yang, P. F. Sullivan, M. E. Goddard, and M. C. Keller, et al. 2012. Estimating the proportion of variation in susceptibility to schizophrenia captured by common SNPs. *Nature Genetics* 44 (3):247–50. doi:10.1038/ng.1108.
- Lee, S. H., N. R. Wray, M. E. Goddard, and P. M. Visscher. 2011. Estimating missing heritability for disease from genome-wide association studies. *American Journal of Human Genetics* 88 (3):294–305. doi:10.1016/j.ajhg.2011.02.002.

- Llewellyn, C. H., M. Trzaskowski, R. Plomin, and J. Wardle. 2013. Finding the missing heritability in pediatric obesity: The contribution of genome-wide complex trait analysis. *International Journal of Obesity* 37 (11):1506–09. doi:[10.1038/ijo.2013.30](https://doi.org/10.1038/ijo.2013.30).
- Lubke, G. H., J. J. Hottenga, R. Walters, C. Laurin, E. J. De Geus, G. Willemsen, and D. I. Boomsma. 2012. Estimating the genetic variance of major depressive disorder due to all single nucleotide polymorphisms. *Biological Psychiatry* 72 (8):707–09. doi:[10.1016/j.biopsych.2012.03.011](https://doi.org/10.1016/j.biopsych.2012.03.011).
- Lubke, G. H., C. Laurin, A. Amin, J. Hottenga, G. Willemsen, G. van Grootheest, and A. Abdellaoui, et al. 2014. Genome-wide analyses of borderline personality features. *Molecular Psychiatry* 19:923–929.
- Majer, I. M., R. Stevens, W. J. Nusselder, J. P. Mackenbach, and P. H. van Baal. 2013. Modeling and forecasting health expectancy: Theoretical framework and application. *Demography* 50 (2):673–97. doi:[10.1007/s13524-012-0156-2](https://doi.org/10.1007/s13524-012-0156-2).
- Marioni, R. E., G. Davies, C. Hayward, D. Liewald, S. M. Kerr, A. Campbell, and I. J. Deary. 2014. Molecular genetic contributions to socioeconomic status and intelligence. *Intelligence* 44:26–32. doi:[10.1016/j.intell.2014.02.006](https://doi.org/10.1016/j.intell.2014.02.006).
- Masters, R. K., R. A. Hummer, D. A. Powers, A. Beck, S. F. Lin, and B. K. Finch. 2014. Long-term trends in adult mortality for U.S. blacks and whites: An examination of period- and cohort-based changes. *Demography* 51 (6):2047–73. doi:[10.1007/s13524-014-0343-4](https://doi.org/10.1007/s13524-014-0343-4).
- McQueen, M. B., J. D. Boardman, B. W. Domingue, A. Smolen, J. Tabor, L. Killea-Jones, and K. M. Harris. 2015. The national longitudinal study of Adolescent to Adult Health (Add Health) sibling pairs genome-wide data. *Behavior Genetics* 45 (1):12–23. doi:[10.1007/s10519-014-9692-4](https://doi.org/10.1007/s10519-014-9692-4).
- Mustelin, L., J. Joutsä, A. Latvala, K. H. Pietiläinen, A. Rissanen, and J. Kaprio. 2012. Genetic influences on physical activity in young adults: A twin study. *Medicine & Science in Sports & Exercise* 44 (7):1293–301. doi:[10.1249/MSS.0b013e3182479747](https://doi.org/10.1249/MSS.0b013e3182479747).
- Nelis, M., T. Esko, R. Magi, F. Zimprich, A. Zimprich, D. Toncheva, and A. Metspalu. 2009. Genetic structure of Europeans: A view from the North-East. *Plos One* 4 (5):e5472. doi:[10.1371/journal.pone.0005472](https://doi.org/10.1371/journal.pone.0005472).
- Palmer, R. H., L. Brick, N. R. Nugent, L. Bidwell, J. E. McGeary, V. S. Knopik, and M. C. Keller. 2015. Examining the role of common genetic variants on alcohol, tobacco, cannabis and illicit drug dependence: Genetics of vulnerability to drug dependence. *Addiction* 110 (3):530–37. doi:[10.1111/add.12815](https://doi.org/10.1111/add.12815).
- Pampel, F. C., and J. T. Denney. 2011. Cross-national sources of health inequality: Education and tobacco use in the World Health Survey. *Demography* 48 (2):653–74. doi:[10.1007/s13524-011-0027-2](https://doi.org/10.1007/s13524-011-0027-2).
- Pilia, G., W. M. Chen, A. Scuteri, M. Orrù, G. Albai, M. Dei, and D. Schlessinger. 2006. Heritability of cardiovascular and personality traits in 6,148 Sardinians. *Plos Genetics* 2 (8):e132. doi:[10.1371/journal.pgen.0020132](https://doi.org/10.1371/journal.pgen.0020132).
- Plomin, R., C. M. Haworth, E. L. Meaburn, T. S. Price, O. S. Davis, and Wellcome Trust Case Control Consortium 2. 2013. Common DNA markers can account for more than half of the genetic influence on cognitive abilities. *Psychological Science*. 24:562–568. 0956797612457952.
- Power, R. A., T. Wingenbach, S. Cohen-Woods, R. Uher, M. Y. Ng, W. Butler, and M. Ising, et al. 2013. Estimating the heritability of reporting stressful life events captured by common genetic variants. *Psychological Medicine* 43 (9):1965–71. doi:[10.1017/S0033291712002589](https://doi.org/10.1017/S0033291712002589).
- Price, A. L., N. J. Patterson, R. M. Plenge, M. E. Weinblatt, N. A. Shadick, and D. Reich. 2006. Principal components analysis corrects for stratification in genome-wide association studies. *Nature Genetics* 38 (8):904–09. doi:[10.1038/ng1847](https://doi.org/10.1038/ng1847).
- Rice, M. L., S. R. Zubrick, C. L. Taylor, J. Gayán, and D. E. Bontempo. 2014. Late language emergence in 24-month-old twins: Heritable and increased risk for late language emergence in twins. *Journal of Speech, Language, and Hearing Research* 57 (3):917–28.
- Richmond, R. C., G. Davey Smith, A. R. Ness, M. Den Hoed, G. McMahon, and N. J. Timpson. 2014. Assessing causality in the association between child adiposity and physical activity levels: A Mendelian randomization analysis. *PLoS Medicine* 11 (3):e1001618. doi:[10.1371/journal.pmed.1001618](https://doi.org/10.1371/journal.pmed.1001618).

- Rietveld, C. A., D. Cesarini, D. J. Benjamin, P. D. Koellinger, J. E. De Neve, H. Tiemeier, and M. Bartels. 2013. Molecular genetics and subjective well-being. *Proceedings of the National Academy of Sciences* 110 (24):9692–97. doi:10.1073/pnas.1222171110.
- Rietveld, C. A., S. E. Medland, J. Derringer, J. Yang, T. Esko, N. W. Martin, and G. McMahon. 2013. GWAS of 126,559 individuals identifies genetic variants associated with educational attainment. *Science* 340 (6139):1467–71. doi:10.1126/science.1235488.
- Ross, C. E., R. K. Masters, and R. A. Hummer. 2012. Education and the gender gaps in health and mortality. *Demography* 49 (4):1157–83. doi:10.1007/s13524-012-0130-z.
- Speed, D., and D. J. Balding. 2014. Relatedness in the post-genomic era: Is it still useful? *Nature Reviews Genetics* 16:33–44. doi:10.1038/nrg3821.
- Speed, D., G. Hemani, M. R. Johnson, and D. J. Balding. 2012. Improved heritability estimation from genome-wide SNPs. *American Journal of Human Genetics* 91 (6):1011–21. doi:10.1016/j.ajhg.2012.10.010.
- Splansky, G. L., D. Corey, Q. Yang, L. D. Atwood, L. A. Cupples, E. J. Benjamin, and D. Levy. 2007. The third generation cohort of the national heart, lung, and blood institute's framingham heart study: Design, recruitment, and initial examination. *American Journal of Epidemiology* 165 (11):1328–35. doi:10.1093/aje/kwm021.
- Tielbeek, J. J., S. E. Medland, B. Benyamin, E. M. Byrne, A. C. Heath, P. A. Madden, and K. J. Verweij. 2012. Unraveling the genetic etiology of adult antisocial behavior: A genome-wide association study. *PLoS One* 7:1–7.
- Trzaskowski, M., T. C. Eley, O. S. Davis, S. J. Doherty, K. B. Hanscombe, E. L. Meaburn, and R. Plomin. 2013. First genome-wide association study on anxiety-related behaviours in childhood. *Plos One* 8 (4):e58676. doi:10.1371/journal.pone.0058676.
- Trzaskowski, M., J. Yang, P. M. Visscher, and R. Plomin. 2014a. DNA evidence for strong genetic stability and increasing heritability of intelligence from age 7 to 12. *Molecular Psychiatry* 19 (3):380–84. doi:10.1038/mp.2012.191.
- Trzaskowski, M., N. Harlaar, R. Arden, E. Krapohl, K. Rimfeld, A. McMillan, P.S. Dale, and R. Plomin. 2014b. Genetic influence on family socioeconomic status and children's intelligence. *Intelligence* 42:83–88. doi:10.1016/j.intell.2013.11.002.
- Viding, E., T. S. Price, S. R. Jaffee, M. Trzaskowski, O. S. Davis, E. L. Meaburn, and R. Plomin. 2013. Genetics of callous-unemotional behavior in children. *Plos One* 8 (7):e65789. doi:10.1371/journal.pone.0065789.
- Vilhjalmsson, B., J. Yang, H. K. Finucane, A. Gusev, S. Lindstrom, S. Ripke, and G. Genovese, et al. 2015. Modeling linkage disequilibrium increases accuracy of polygenic risk scores. *BioRxiv* 015859.
- Vinkhuyzen, A. A., N. L. Pedersen, J. Yang, S. H. Lee, P. K. Magnusson, W. G. Iacono, and N. R. Wray. 2012. Common SNPs explain some of the variation in the personality dimensions of neuroticism and extraversion. *Translational Psychiatry* 2 (4):e102. doi:10.1038/tp.2012.27.
- Visscher, P. M., G. Hemani, A. A. Vinkhuyzen, G. B. Chen, S. H. Lee, N. R. Wray, and J. Yang. 2014. Statistical power to detect genetic (co) variance of complex traits using SNP data in unrelated samples. *Plos Genetics* 10 (4):e1004269. doi:10.1371/journal.pgen.1004269.
- Visscher, P. M., J. Yang, and M. E. Goddard. 2010. A commentary on 'common SNPs explain a large proportion of the heritability for human height by Yang et al. (2010). *Twin Research and Human Genetics* 13 (06):517–24. doi:10.1375/twin.13.6.517.
- Vrieze, S. I., M. McGue, M. B. Miller, B. M. Hicks, and W. G. Iacono. 2013. Three mutually informative ways to understand the genetic relationships among behavioral disinhibition, alcohol use, drug use, nicotine use/dependence, and their co-occurrence: Twin biometry, GCTA, and genome-wide scoring. *Behavior Genetics* 43 (2):97–107. doi:10.1007/s10519-013-9584-z.
- Watson, C. T., G. Disanto, F. Breden, G. Giovannoni, and S. V. Ramagopalan. 2012. Estimating the proportion of variation in susceptibility to multiple sclerosis captured by common SNPs. *Scientific Reports* 2:1–4. doi:10.1038/srep00770.

- Welter, D., J. MacArthur, J. Morales, T. Burdett, P. Hall, H. Junkins, and H. Parkinson. 2014. The NHGRI GWAS Catalog, a curated resource of SNP-trait associations. *Nucleic Acids Research* 42 (D1):D1001–D1006. doi:10.1093/nar/gkt1229.
- Wienke, A., N. V. Holm, A. Skytthe, and A. I. Yashin. 2001. The heritability of mortality due to heart diseases: A correlated frailty model applied to Danish twins. *Twin Research* 4 (4):266–74. doi:10.1375/1369052012399.
- Wood, A. R., T. Esko, J. Yang, S. Vedantam, T. H. Pers, S. Gustafsson, and U. Lim. 2014. Defining the role of common variation in the genomic and biological architecture of adult human height. *Nature Genetics* 46 (11):1173–86. doi:10.1038/ng.3097.
- Yang, J., B. Benyamin, B. P. McEvoy, S. Gordon, A. K. Henders, D. R. Nyholt, and P. M. Visscher. 2010. Common SNPs explain a large proportion of the heritability for human height. *Nature Genetics* 42 (7):565–69. doi:10.1038/ng.608.
- Yang, J., S. H. Lee, M. E. Goddard, and P. M. Visscher. 2011. GCTA: A tool for genome-wide complex trait analysis. *American Journal of Human Genetics* 88 (1):76–82. doi:10.1016/j.ajhg.2010.11.011.
- Yang, L., B. M. Neale, L. Liu, S. H. Lee, N. R. Wray, N. Ji, and Y. Wang. 2013. Polygenic transmission and complex neuro developmental network for attention deficit hyperactivity disorder: Genome-wide association study of both common and rare variants. *American Journal of Medical Genetics Part B: Neuropsychiatric Genetics* 162 (5):419–30. doi:10.1002/ajmg.b.32169.
- Zuk, O., S. F. Schaffner, K. Samocha, R. Do, E. Hechter, S. Kathiresan, and E. S. Lander. 2014. Searching for missing heritability: Designing rare variant association studies. *Proceedings of the National Academy of Sciences* 111 (4):E455–E464. doi:10.1073/pnas.1322563111.