



# Estimation of Parental Effects Using Polygenic Scores

Jared V. Balbona<sup>1,2</sup> · Yongkang Kim<sup>1</sup> · Matthew C. Keller<sup>1,2</sup>

Received: 7 August 2020 / Accepted: 20 November 2020  
© The Author(s) 2021

## Abstract

Offspring resemble their parents for both genetic and environmental reasons. Understanding the relative magnitude of these alternatives has long been a core interest in behavioral genetics research, but traditional designs, which compare phenotypic covariances to make inferences about unmeasured genetic and environmental factors, have struggled to disentangle them. Recently, Kong et al. (2018) showed that by correlating offspring phenotypic values with the measured polygenic score of parents' nontransmitted alleles, one can estimate the effect of "genetic nurture"—a type of passive gene–environment covariation that arises when heritable parental traits directly influence offspring traits. Here, we instantiate this basic idea in a set of causal models that provide novel insights into the estimation of parental influences on offspring. Most importantly, we show how jointly modeling the parental polygenic scores and the offspring phenotypes can provide an unbiased estimate of the variation attributable to the environmental influence of parents on offspring, even when the polygenic score accounts for a small fraction of trait heritability. This model can be further extended to (a) account for the influence of different types of assortative mating, (b) estimate the total variation due to additive genetic effects and their covariance with the familial environment (i.e., the full genetic nurture effect), and (c) model situations where a parental trait influences a different offspring trait. By utilizing structural equation modeling techniques developed for extended twin family designs, our approach provides a general framework for modeling polygenic scores in family studies and allows for various model extensions that can be used to answer old questions about familial influences in new ways.

**Keywords** Vertical transmission · Genetic nurture · Heritability estimation · Structural equation modeling · Parental effects

## Introduction

Parents share half of their (autosomal) additive genetic effects with their children, causing resemblance between parent and offspring for heritable traits. However, parents also help create and shape their offspring's environment, which may have

an enduring influence for certain traits. If the parental traits that impact their children's environment are themselves heritable, a covariance will develop between the genetic effects underlying those traits and the environmental effects provided by the parents. Educated parents, for example, provide to their offspring not only genes that predispose to higher education, but also a familial environment that is likely conducive to higher education. Thus, offspring who inherit genes that predispose to higher education are also more likely to be influenced by a familial environment that encourages education. This phenomenon is a type of *passive gene–environment (G–E) covariance* that has recently been referred to as *genetic nurture* (Kong et al. 2018). Generally, passive G–E covariance refers to the covariance between the genetic effects on a trait and the parenting environment influenced by that trait, regardless of whether the parenting environment actually influences the offspring (DiLalla and Gottesman 1991). However, passive G–E covariance can also arise from sociocultural influences unrelated to parenting (e.g., the correlation between genes influencing skin pigmentation and

---

Edited by David Evans.

---

**Electronic supplementary material** The online version of this article (<https://doi.org/10.1007/s10519-020-10032-w>) contains supplementary material, which is available to authorized users.

---

✉ Jared V. Balbona  
jared.v.balbona@gmail.com

✉ Matthew C. Keller  
matthew.c.keller@gmail.com

<sup>1</sup> Institute for Behavioral Genetics, University of Colorado at Boulder, Boulder, USA

<sup>2</sup> Department of Psychology & Neuroscience, University of Colorado at Boulder, Boulder, USA

environments conducive to educational opportunity). Here, we adopt the convention that *genetic nurture* refers to a specific type of passive G–E covariance that occurs when the environment provided by parents does directly influence the phenotype of their offspring, leading to covariance between the genes affecting a trait and the rearing environment of the offspring influenced by that parental trait. Genetic nurture is therefore a necessary consequence of the direct effect of a (heritable) parental phenotype on an offspring phenotype, a phenomenon known as *vertical transmission* (VT) in the behavioral genetics literature (Cavalli-Sforza and Feldman b). While passive G–E covariance can also arise from *horizontal transmission*—in which other collateral relatives (typically siblings) influence one another—we focus in this paper on genetic nurture arising from parental influences and discuss horizontal transmission at the end.

The occurrence of genetic nurture has important implications for understanding complex trait genetics. Genetic nurture increases the phenotypic variation in the population over what it would otherwise be, and can bias estimates of genetic or environmental influences. In genome-wide association studies (GWAS's), genetic nurture inflates estimates of SNP associations. This bias can be quantified in part by comparing within-family estimates, which are immune to this inflation, to standard between-family estimates. For example, Lee et al. (2018) found that SNP associations with educational attainment estimated within-families were 40% smaller than those estimated between-families, consistent with the presence of genetic nurture. Similarly, by inflating the associations between traits and their causal variants across the genome, genetic nurture upwardly biases estimates of SNP-heritability, including those from methods that use GWAS summary statistics (e.g., LD-score regression; Bulik-Sullivan et al. 2015) and from methods that use average similarity across genome-wide SNPs (e.g., genomic REML; Yang et al. 2011). Genetic nurture also increases dizygotic and monozygotic twin covariances to the same degree, leading to overestimates of the shared environmental variation in the classical twin design “ACE” models, and (less intuitively) to overestimates of the additive genetic variance in “AE” or “ADE” models that estimate only genetic variation (Coventry and Keller 2005). For these reasons, quantifying genetic nurture is important for interpreting estimates across multiple designs and approaches.

While genetic nurture cannot be estimated in classical twin designs, it can in principle be detected using extended twin family designs (Cloninger et al. 1979). However, doing so has been challenging because no observed statistic directly provides information to estimate it. Rather, genetic nurture must be inferred as a consequence of estimated heritability that co-occurs with estimated VT (Eaves 1976), and in turn VT must be estimated from residual parent-offspring covariance that is higher than expected from

estimated genetic and environmental influences. Nonetheless, factors such as dominance, epistasis, and gene-by-age interactions cannot be simultaneously modeled using these designs. Existing estimates of genetic nurture and VT from the extended twin family literature thus depend strongly on model assumptions and may be biased to the degree they are unmet (Keller et al. 2010).

In the past several years, studies using a Mendelian randomization paradigm (Davey Smith and Ebrahim 2003) have leveraged measured genetic data to examine the extent to which observed parent–offspring similarity is due to a direct causal influence of parents on their children (i.e., due to VT). Typically, these studies have built polygenic scores (PGS's; the predicted genetic scores for a trait based on SNP weights from a GWAS conducted in an independent sample) of a maternal trait (e.g., height) and assessed whether it is related to an offspring trait (e.g., gestational age). To account for shared genetic effects between mother and offspring, these studies have controlled for the offspring PGS (Lawlor et al. 2008; Bonilla et al. 2012; Tyrrell et al. 2016). However, this approach is suboptimal because it can lead to a collider bias (Lawlor et al. 2017). A more elegant approach to controlling for shared genetic effects was thus proposed by Zhang et al. (2015), who built two maternal PGS's of height: one based on the alleles that were transmitted to the offspring and one based on those that were not transmitted. They found that the nontransmitted PGS of maternal height was significantly associated with offspring gestational age. This association is a downwardly biased estimate (due to the PGS imperfectly capturing full trait heritability) of genetic nurture which, as explained above, necessarily implies VT; maternal height is having a causal influence on offspring gestational age, mediated by the maternally provided intrauterine environment (see also Lawlor et al. 2017; Tubbs et al. 2020a). Warrington et al. (2018) and Evans et al. (2019) later built upon this basic idea by incorporating offspring and maternal PGS's into a causal model, allowing recursive relationships between effects that naturally arise in this context (e.g., genetic effects upon which the PGS is based are themselves over-estimated due to genetic nurture) to be properly accounted for.

To date, the most comprehensive and best-known approach for using measured genetic data to examine parental influences is the Kong et al. (2018) investigation into the influences of parental educational attainment. From data on ~22K Icelandic offspring and their parents in the deCODE sample, Kong et al. constructed PGS's for educational attainment using the transmitted and nontransmitted alleles from both mothers and fathers. After controlling for the potential confounding influences of stratification, they found a highly significant relationship between the nontransmitted parental PGS (summed across both parents) and offspring educational attainment. This association is an estimation of genetic nurture that, unlike previous work, incorporates

paternal effects, accounts for assortment, and avoids potential collider bias (Lawlor et al. 2017; Tubbs et al. 2020b).

While a significant advance, their approach was not without its limitations. First, as in the Zhang et al. paper, their estimate of genetic nurture was downwardly biased (again due to the partial predictive ability of the PGS) and they provided no estimate of the magnitude of VT. Secondly, as explained below, primary phenotypic assortative mating (hereafter simply 'AM'; the tendency for parents to chose mates who are similar to themselves) is a competing explanation for the relationship between the nontransmitted PGS and the offspring phenotype. Kong et al. attempted to account for the influence of AM, but because they found evidence that AM for educational attainment in Iceland occurred only in the parental generation (and not before), their approach assumes this specific type of disequilibrium assortment. Third, much of the math presented by Kong et al. was derived from first principles, limiting the generalizability of their approach and not easily allowing it to be extended to other situations.

Thus, building upon this previous work, we present here a *structural equation modeling* (SEM) framework of their basic idea that uses transmitted and nontransmitted PGS's to understand genetic and environmental parental influences. These models can be used to obtain unbiased estimates of total variation due to additive genetic effects ( $V_A$ ), genetic nurture ( $v + w$ ), and the variation accounted for by VT ( $V_F$ ), even when the PGS accounts for a small fraction of trait heritability. Furthermore, using techniques developed for extended twin family designs, we demonstrate how both disequilibrium and equilibrium AM can be tested and accounted for in causal models. Finally, the underlying principles of SEM (described below) allow these models to be easily extended in ways that would otherwise lead to intractable math.

We believe that using PGS's to understand genetic and environmental influences of parents on offspring is an important and exciting advance. The set of models presented here can be used to better utilize PGS's for estimating parental influences. While they can be implemented "out of the box," their greater utility is as examples of a general approach for incorporating PGS's into family causal models in order to test specific hypotheses of interest. To this end, we present the logic and math underlying the models, adopting a tutorial style so that the models presented can be modified and improved upon to best fit the hypotheses being tested.

## Overview of causal modeling

SEM is very useful in the present context for several reasons. First, as mentioned, SEMs are easily extensible and provide a set of rules that can simplify potentially

complex (e.g., recursive) relationships between variables by finding estimates jointly instead of individually. For scenarios such as those involving VT and AM, this can greatly simplify otherwise intractable math. In addition, SEM is advantageous because it encourages a focus on effect sizes rather than  $p$ -values, forces the user to think carefully about the possible causal mechanisms that underlie observed data, and requires that model assumptions be made explicit. Estimates from causal models will be biased to the degree that assumptions are violated, and so it is important to understand how estimates behave when assumptions are unmet in order to properly interpret them; this is usually done by comparing estimates to known parameters in simulated data, as was done for the models below in Kim et al. (2020).

To aid in the understanding of our models, we first review path tracing rules and the overall causal modeling framework we adopt here. Path tracing was developed by Sewell Wright (Wright 1934) as a means for deriving the expected variances and covariances among variables in a particular *causal model*—a set of assumed causal connections between measured and unmeasured variables. SEM can then be used to find estimates that maximize the likelihood of the observed variances and covariances given the assumed causal model. Although sometimes misunderstood, SEM does not test causation but rather assesses the degree to which a set of observed variances and covariances is consistent with a particular causal model (Bollen and Pearl 2013).

A causal model is considered *identified* if there is a unique solution for all the model's parameters. Often, there is insufficient information to estimate all parameters of interest, requiring that the values of some parameters be fixed (typically to 0 or 1) for the model to be identified (though it should also be noted that fixing parameters is itself a model assumption). Even for identified models, certain assumptions about causal relationships between variables can lead to high correlations among estimates, increasing their standard errors and potentially necessitating that one or more parameters be fixed.

Path diagrams pictorially represent causal models and are helpful for deriving the variances and covariances that the models imply. By convention, squares in path diagrams represent observed variables and circles represent unobserved *latent variables*—theorized causes of variation and covariation among observed variables. Single-headed arrows from one variable to another signify causal relationships from the former to the latter, with their associated path coefficients being akin to partial regression coefficients. Double-headed arrows, meanwhile, signify covariances when connecting two variables to each other and variances when connecting variables to themselves. Finally, a straight line with no arrows is called a *copath* and is used to model AM. The copath is a more recent innovation in SEM (Cloninger 1980)

and is not widely used outside the extended twin family literature.

To derive expectations using a path diagram, one must identify all legitimate paths that connect two variables (for expected covariances) or that connect a variable back to itself (for expected variances). These paths can be thought of as chains, with each individual arrow or copath representing a link in a particular chain. The expected value of a chain is equal to the product of all the coefficients associated with each of its links, and the final expected variance or covariance is equal to the sum of all legitimate chains. A chain is considered legitimate if it abides by the following path tracing rules:

1. A chain begins by travelling backwards against the direction of a single or double-headed arrow (from the arrow's head to its tail). However, once a double-headed arrow has been traversed, the direction reverses such that the chain now travels forwards, in the direction of the arrows.
2. A chain must include exactly one double-headed arrow (a variance or a covariance term), which is equivalent to stating that a chain must change directions exactly once. This is necessary because double-headed arrows provide the proper scaling for the coefficients in each chain.
3. All chains must be counted exactly once and each must be unique. However, the order of the links in the chains matters. For example, despite being algebraically equivalent, the chain  $Y_p \rightarrow NT_p \rightarrow T_p \rightarrow Y_p$  is distinct from the chain  $Y_p \rightarrow T_p \rightarrow NT_p \rightarrow Y_p$  in Fig. 1. Both are unique and both must be counted in determining the variance of  $Y_p$ .
4. Copaths may only be traversed once in a given chain, and a chain must be legitimate before traversing the copath. However, once the copath is crossed, the first two rules above reset. A chain must therefore contain exactly one double-headed arrow before traversing the copath, and one double-headed arrow after traversing the copath. Thus, copaths connect two legitimate chains to create a single, longer chain.

To demonstrate the first three rules (the fourth is demonstrated below), we derive the expected  $cov(Y_p, NT_p)$  in Fig. 1, denoted as  $\Omega$  in our models. As noted, deriving the covariance between two terms requires tracing all legitimate chains that begin at one and end at the other. In this case, only two legitimate chains start at  $Y_p$  and end at  $NT_p$  (one could equivalently start at  $NT_p$  and end at  $Y_p$ ). The first travels up the arrow  $Y_p \rightarrow NT_p$  (path coefficient  $\delta$ ), and because all chains require a double-headed arrow, finishes by traversing the double-headed arrow leading back to  $NT_p$  (i.e., the variance of  $NT_p$ , with path coefficient  $\frac{1}{2}$ ). The second travels up the arrow  $Y_p \rightarrow F_p$  (with path coefficient 1) and

then traverses the double-headed arrow  $F_p \rightarrow NT_p$  (i.e., the covariance between  $F_p$  and  $NT_p$ , with path coefficient  $\frac{w}{2}$ ). Thus,  $\Omega = \delta(\frac{1}{2}) + 1(\frac{w}{2}) = \frac{1}{2}(\delta + w)$ .

## Models of parental effects

Although the path tracing procedures described above are simple and algorithmic, the number of unique chains grows rapidly as models become more complicated, making the process error-prone. To simplify chains and reduce the probability of errors, we substitute variables (e.g.,  $\Omega$ ) for chain segments that recur across multiple chains ( $\frac{1}{2}(\delta + w)$ ) when possible. We use parameter subscripts  $p$ ,  $m$ , and  $o$  for paternal, maternal, and offspring, respectively. To reduce redundancy, we use  $[N]T$  to denote either  $NT$  (the nontransmitted) or  $T$  (the transmitted) haplotypic PGS, and we use the  $*$  subscript to denote either  $p$  or  $m$  but not both within a single term. For example,  $cov(Y_*, [N]T_*)$  can be written in the place of  $cov(Y_p, T_p)$ ,  $cov(Y_p, NT_p)$ ,  $cov(Y_m, T_m)$ , or  $cov(Y_m, NT_m)$ , but  $cov(Y_*, T_*)$  does not equal  $cov(Y_p, T_m)$  or  $cov(Y_m, T_p)$ . For ease of comparison, we follow the notation of Kong et al. when possible. For example, the meanings of  $\delta$ ,  $\theta_T$ ,  $\theta_{NT}$ ,  $T_p$ ,  $NT_p$ ,  $T_m$ , and  $NT_m$  are consistent across papers. Finally, the names of other parameters ( $V_F$ ,  $w$ ,  $f$ ,  $V_e$ , and  $\mu$ ) were chosen for consistency with existing extended twin family models. For descriptions of these and other parameters, see Table 1.

To ensure that our parameter derivations are correct, we compared expected equilibrium parameter values to simulated ones from an adapted version of the *GeneEvolve* software (Tahmasbi and Keller 2017), which we modified for efficiency such that causal variants were in linkage equilibrium in the *base population*—the population before any AM or VT has occurred. Because most parameters depend on each other recursively, we found their expected equilibrium values by inputting start values into their expectations derived below and iterated all parameters together in R until their values converged. For all models, the expected equilibrium values of the parameters agreed with their observed equilibrium counterparts from *GeneEvolve*.

Phased whole-genome maternal, paternal, and offspring data are required to detect identical by descent segments, from which both parents' transmitted and nontransmitted alleles are distinguished. We show that these four pieces of information ( $NT_p$ ,  $T_p$ ,  $NT_m$ ,  $T_m$ ) along with  $Y_o$  are sufficient for estimating the full  $V_F$  when there is no AM (Model 0), though data need not be complete for every family as estimates are unbiased by missingness. However, because AM induces covariances between latent and observed genetic scores, estimates of  $V_F$  and genetic nurture will be biased when the PGS accounts for little of the heritability (Model 1). Nonetheless, this bias can be eliminated by modeling the genetic variation not captured by the PGS (Model 2), either

**Table 1** Parameters in models 0, 1, and 2

Parameter	Interpretation
$Y$	Phenotypic score
$\epsilon$	Residual score
$\mu$	Primary phenotypic assortative mating (AM) copath coefficient; $\frac{cov(Y_p, Y_m)}{V_Y^2}$
$F$	Family environmental score arising from $Y_* \rightarrow F_o$ vertical transmission (VT)
$f$	VT path coefficient; the causal effect of $Y_* \rightarrow F_o$
$T_*$	Polygenic score (PGS) of one of the two transmitted haplotypes
$NT_*$	PGS of one of the two nontransmitted haplotypes
$k$	Variance of the haplotypic PGS in the base population (before AM or VT). It is a constant that depends on the scaling of the PGS (see Model 1)
$g$	Increase in the (co)variance of the haplotypic PGS's under AM
$\delta$	effect of haplotypic PGS on $Y$
$w$	Genetic nurture between the PGS and $F$ ; $cov(T_* + NT_*, F_*) = cov([N]T_p + [N]T_m, F_o)$
$\Omega$	Covariance between $Y_*$ and either of that parent's haplotypic PGS's; $cov(Y_*, [N]T_*)$
$\theta_T$	Covariance between $Y_o$ and both of the transmitted haplotypes; $cov(Y_o, T_m + T_p)$
$\theta_{NT}$	Covariance between $Y_o$ and both of the nontransmitted haplotypes; $cov(Y_o, NT_m + NT_p)$
$LT_*$	Latent genetic score (LGS) of one of the two transmitted haplotypes
$LNT_*$	LGS of one of the two nontransmitted haplotypes
$j$	Variance of the haplotypic LGS in the base population. It is a constant defined analogously to $k$ (see Model 2)
$h$	Increase in the (co)variance of the haplotypic LGS's under AM
$i$	Increase in the covariance between the haplotypic PGS's and LGS's under AM
$a$	Effect of haplotypic LGS on $Y$
$v$	Genetic nurture between the LGS and $F$ ; $cov(LT_* + LNT_*, F_*) = cov(L[N]T_p + L[N]T_m, F_o)$
$\Gamma$	Covariance between $Y_*$ and either of that parent's haplotypic LGS's; $cov(Y_*, L[N]T_*)$
$\theta_{LT}$	Covariance between $Y_o$ and the transmitted LGS; $cov(Y_o, LT_m + LT_p)$
$\theta_{LNT}$	Covariance between $Y_o$ and the nontransmitted LGS; $cov(Y_o, LNT_m + LNT_p)$
$V_A$	Full variance due to direct additive genetic effects; $2a^2(j + 2h) + 2\delta^2(k + 2g) + 8ai\delta$
$V_{A_0}$	full variance due to direct additive genetic effects in base population; $2a^2j + 2\delta^2k$
$v + w$	Full genetic nurture (a type of passive G-E covariance)
$V_F$	Full variance due to VT; $2f^2V_Y(1 + \mu V_Y)$
$V_\epsilon$	Residual variance not explained by other factors (i.e., unique environmental variance)
$V_Y$	Phenotypic variance; $V_A + V_F + 2(av + \delta w) + V_\epsilon$

Subscripts:  $m$ : maternal;  $p$ : paternal;  $o$ : offspring;  $*$ :  $m$  or  $p$ ; no subscript:  $m, p$ , or  $o$

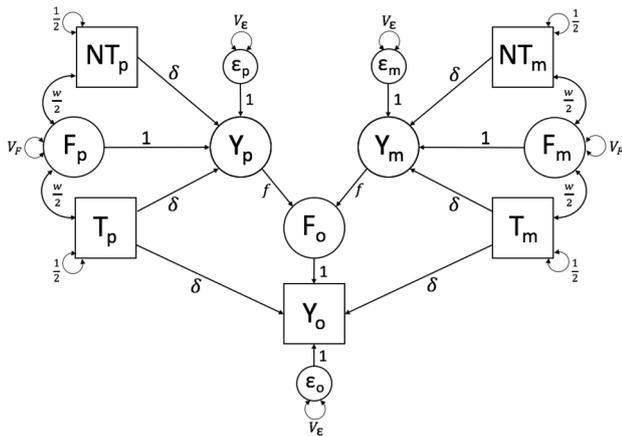
through estimating it by including parental phenotypes in the model or by making an assumption about its value in the base population. Each of these models require assumptions; we describe these in the subsections below, but focus on their influences in Kim et al. (2020).

### Model 0: VT but no AM

Figure 1 shows a path diagram of the simplest model of genetic nurture and so serves as a valuable starting place. It makes two assumptions that distinguish it from later models: (1) there is no AM, and (2) the PGS explains all of the genetic variation in the trait. The first assumption will be unmet for many traits of interest while the latter assumption

is unmet for all traits currently. Nevertheless, when the first assumption is met (no AM), we show below that this simple model can provide unbiased estimates of the full  $V_F$ .

This model estimates five unknown parameters:  $\delta$ , the direct effect of haplotypic PGS on the phenotype after removing the influence of genetic nurture;  $f$ , the direct effect of parental phenotype on the offspring environment (i.e., the VT effect);  $V_F$ , the variance due to VT;  $w$  the genetic nurture effect; and  $V_\epsilon$ , the variance of the residual phenotypic variation. It is worth noting that the values of  $f$  and  $V_F$  are determined given the values of  $\delta$ ,  $w$ , and  $V_\epsilon$ , and so only three of these five estimates are independent. Additionally, the parental phenotypes ( $Y_p$  and  $Y_m$ ), familial environment value arising from VT ( $F$ ), and unique environmental score ( $\epsilon$ ) are latent and are therefore represented by circles. To



**Fig. 1** Path diagram of Model 0, which models the effects of VT, assuming that the PGS explains the full trait heritability and that there is no AM

prevent under-identification, the  $F \rightarrow Y$  and  $\epsilon \rightarrow Y$  paths are fixed to 1. Similarly, the variances of the haplotypic PGS's are constrained to  $\frac{1}{2}$ , which should be true if the full PGS is standardized and there is no AM to induce covariances between haplotypic PGS's.

As previously stated, there are five observed variables in this model—the transmitted and nontransmitted paternal ( $T_p$  and  $NT_p$ ) and maternal ( $T_m$  and  $NT_m$ ) haplotypic PGS's as well as the offspring phenotype ( $Y_o$ )—creating a 5-by-5 observed variance-covariance matrix and leading to 15 unique statistics from which to estimate parameters. Model-fitting software mimics as closely as possible this observed variance-covariance matrix with the one implied by/the maximum likelihood estimates of the model's unknown parameters. While 15 independent statistics are easily sufficient for estimating a model with three unknowns, many of the statistics in this model provide redundant information. The four haplotypic PGS variances and the six covariances between them are assumed to be constants ( $\frac{1}{2}$  and 0, respectively) and provide no information for estimating parameters. The remaining five statistics provide only three independent pieces of information: one from the two covariances between the haplotypic nontransmitted PGS ( $NT_*$ ) and  $Y_o$ , one from the two covariances between the haplotypic transmitted PGS ( $T_*$ ) and  $Y_o$ , and one from the variance of  $Y_o$ . These three independent sources of information are used to estimate three independent parameters ( $\delta$ ,  $w$ , and  $V_e$ ). Thus, this model is just-identified.

Although parental phenotypes are unobserved in this model, it is still useful to define the covariance between haplotypic PGS's and the latent parental phenotypes because this term recurs throughout. We denote this covariance as  $\Omega$  and, as noted above,  $\Omega = \frac{1}{2}(\delta + w)$ . Under this model's assumptions of no sex-specific genetic or VT effects,

$\Omega$  is the same regardless of the PGS's parental origin or whether it is transmitted:  $cov(T_p, Y_p) = cov(NT_p, Y_p) = cov(T_m, Y_m) = cov(NT_m, Y_m)$ . Thus,  $\Omega$  can be used as a substitute for  $\frac{1}{2}(\delta + w)$  in any chain that traverses  $Y_* \rightarrow [N]T_*$  or  $[N]T_* \rightarrow Y_*$  in order to simplify finding other expected values, such as in the two covariances at the core of this model:

$$\begin{aligned} \theta_{NT} &= cov(Y_o, NT_p + NT_m) = 2cov(Y_o, NT_*) = 2f(cov(Y_*, NT_*)) \\ &= 2f(\frac{1}{2}(\delta + w)) = 2f\Omega \end{aligned} \tag{1}$$

$$\theta_T = cov(Y_o, T_p + T_m) = 2cov(Y_o, T_*) = \delta + 2f\Omega \tag{2}$$

Kong et al. emphasized that part of the relationship between  $Y$  and its PGS ( $T_p + T_m$ ) may be due to the confounding influences of genetic nurture. This can be seen in the additional  $2f\Omega$  term in  $\theta_T$  above. Thus, as noted by Kong et al.,  $\theta_T - \theta_{NT} = \delta$  is an estimate of the direct genetic effect of the PGS, controlling for genetic nurture.

This model assumes that parameters have reached equilibrium, which implies that variances and covariances are the same across parental and offspring generations. The equilibrium assumption allows the parameters that change over time ( $V_Y$ ,  $w$ , and  $V_F$ ) to be estimated by constraining their values in the parental generation to their derived values in the offspring generation. For example, at equilibrium, the covariance between  $F_*$  and the haplotypic PGS's in the parental generation ( $\frac{w}{2}$ ) must equal the implied covariance between  $F_o$  and any of the four haplotypic PGS's ( $f\Omega$ , which can be found through path tracing). Thus,

$$\begin{aligned} w &= cov(F_o, T_p + T_m) = cov(F_o, NT_p + NT_m) \\ &= 2cov(F_o, [N]T_*) = 2f\Omega = \theta_{NT} \end{aligned} \tag{3}$$

Note that this estimated value of  $w$  is equal to the estimated value of  $\theta_{NT}$  derived in equation (1), indicating that  $\theta_{NT}$  is a direct estimate of genetic nurture (under the assumption of no AM). Meanwhile, the variance of  $Y_p$  (denoted by  $V_Y$ ) is derived by summing all chains that begin at  $Y_p$  and end back at  $Y_p$ , and is assumed to be equal to the variance of  $Y_m$  and  $Y_o$ :

$$V_Y = var(Y) = 2\delta\Omega + \delta w + V_F + V_e \tag{4}$$

Finally, the expectations for the variances of  $F_p$  and  $F_m$  ( $V_F$ ) can be found by constraining their values to all legitimate chains that connect  $F_o$  back to itself, of which there are two: (1)  $F_o \rightarrow Y_p \rightarrow Y_p \rightarrow F_o$  and (2)  $F_o \rightarrow Y_m \rightarrow Y_m \rightarrow F_o$ . Thus,

$$V_F = var(F_*) = var(F_o) = (fV_Yf) + (fV_Yf) = 2f^2V_Y \tag{5}$$

Note that the variance  $F_o$ —as well as its covariances with the haplotypic PGS's—are not shown in any of the models, as it is already implied through the connections between  $F_o$  and the parental phenotypes; explicitly including  $V_F$  and  $\frac{w}{2}$  in the offspring generation would thus be redundant, resulting

in expectations that are double their true values. Furthermore, note that  $V_F$  is a function of  $V_Y$  but that  $V_Y$  is also a function of  $V_F$ . Similarly,  $\Omega$  is a function of  $w$ , which is a function of  $\Omega$ . These types of recursive relationships are known as *nonlinear constraints*, which describe and constrain such interdependent relationships between parameters in a way that keeps the overall model internally consistent and identified. The implementation of nonlinear constraints is a hallmark of family models, and requires the use of optimizers (such as NPSOL in OpenMx) that can estimate their values iteratively.

Last, we show that when the assumption of no AM is met, this model provides a full estimate of  $V_F$ , regardless of the amount of variance explained by the PGS ( $\delta^2$ ). Given equation (1), as well as the knowledge that  $w = \theta_{NT}$  and  $\delta = \theta_T - \theta_{NT}$ ,

$$\begin{aligned} \theta_{NT} &= f\delta + fw = w \\ &\Rightarrow w(1 - f) = f\delta \\ &\Rightarrow \frac{w}{\delta} = \frac{f}{1 - f} = \frac{\theta_{NT}}{\theta_T - \theta_{NT}} \\ &\Rightarrow f\theta_T - f\theta_{NT} = \theta_{NT} - f\theta_{NT} \end{aligned}$$

Through rearrangement of terms,

$$f = \frac{\theta_{NT}}{\theta_T} \tag{6}$$

Thus, the estimate of  $V_F (= 2f^2V_Y)$  depends on only three observed statistics:  $\theta_{NT}$ ,  $\theta_T$ , and  $V_Y$ . Note that the expectation of  $\theta_{NT}$  contains two parameters,  $w$  and  $\Omega$ , that are functions of one another. Substituting the value of  $w$  recursively thus leads to a geometric series:

$$\begin{aligned} \theta_{NT} &= 2f\Omega = f\delta + fw = f\delta + f(2f\Omega) \\ &= f\delta + f^2\delta + f^2w = \delta \sum_{n=1}^{\infty} f^n \end{aligned} \tag{7}$$

Therefore, because  $\theta_T = \delta + \theta_{NT}$ ,

$$\frac{\theta_{NT}}{\theta_T} = \frac{\delta \left( \sum_{n=1}^{\infty} f^n \right)}{\delta \left( 1 + \sum_{n=1}^{\infty} f^n \right)} \tag{8}$$

As can be seen,  $\delta$  cancels out in the expected value of  $\frac{\theta_{NT}}{\theta_T}$ . Therefore, the point estimates of  $f$  and  $V_F$  are influenced by the magnitude of VT and not by the predictive ability of the PGS (although the standard error of  $V_F$  increases as  $\delta$  decreases; Kim et al. 2020). Finally, because  $-1 < f < 1$ , the geometric series  $\sum_{n=1}^{\infty} f^n$  converges to  $\frac{f}{1-f}$ , and thus

$$\frac{\theta_{NT}}{\theta_T} = \frac{\frac{f}{1-f}}{\left( 1 + \frac{f}{1-f} \right)} = f \tag{9}$$

This demonstrates the same result in Eq. (6) again but from a different approach.

### Model 1: VT and AM

Model 1 assumes that the PGS explains all the trait heritability, as did Model 0, but now incorporates the influences of AM (Fig. 2). As such, Model 1 yields estimates that are unbiased when there is AM, but only to the degree that the PGS captures the heritability of the trait. Given that the PGS's for most traits explain little heritability (e.g., typically < 20%; Torkamani et al. 2018), the utility of this model is mostly didactic.

We model AM using a copath, which follows a special set of path tracing rules, as explained above. The copath is represented as a straight line between  $Y_p$  and  $Y_m$  in Fig. 2, and its path coefficient is denoted  $\mu$ . The expected covariance between mates is all chains  $Y_p \rightarrow Y_m$  or vice-versa. To traverse the copath, a chain must first be legitimate, so it must have already traversed a double-headed arrow. Thus, chains from  $Y_p \rightarrow Y_m$  begin with the sum of all chains that connect  $Y_p$  back to itself (the sum of which =  $V_Y$ , and each of which includes a double-headed arrow) before then crossing the copath ( $\mu$ ). At this point, the other path tracing rules reset, necessitating that each chain traverses another double-headed arrow. Thus, the chains end by traversing all chains from  $Y_m \rightarrow Y_m$  (which also =  $V_Y$ ). Therefore, the covariance between mates is

$$cov(Y_p, Y_m) = V_Y \mu V_Y = \mu V_Y^2 \tag{10}$$

Note that  $\mu = \frac{cov(Y_p, Y_m)}{V_Y^2}$  is neither the covariance (=  $\mu V_Y^2$ ) nor the correlation (=  $\mu V_Y$ ) between mates. AM and VT increase  $V_Y$  over time, and because we assume that the *correlation*

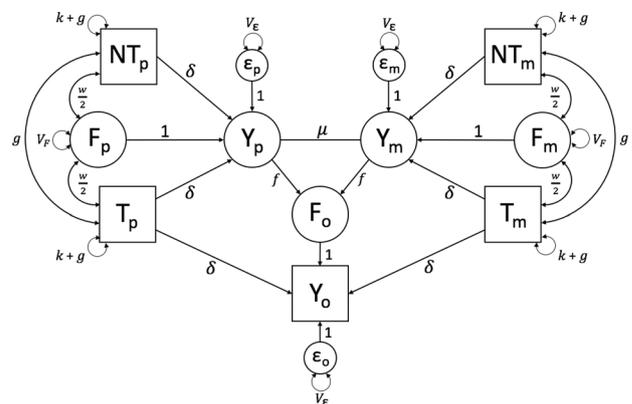


Fig. 2 Path diagram of Model 1, which models the effects of AM and VT, assuming that the PGS explains the full trait heritability

between mates is constant across generations, the value of  $\mu$  correspondingly decreases across generations until equilibrium is reached (which occurs in 5–10 generations). The information to estimate  $\mu$  comes from the six observed covariances between haplotypic PGS's as well as the four observed haplotypic PGS variances.

AM for a trait creates *gametic phase disequilibrium* between causal variants, meaning that trait-increasing alleles tend to coaggregate with other trait-increasing alleles and vice-versa. This occurs because similarity based on mates' phenotypic scores implies similarity of genetic effects across mates as well. Two important consequences of gametic phase disequilibrium are the increase in genetic variation over what it would be in the absence of AM (in the base population), and the increase in genetic covariation between mates and close relatives (Lynch and Walsh 1998).

A single generation of AM leads to covariance between the genetic scores of the maternal ( $[N]T_m$ ) and paternal ( $[N]T_p$ ) haplotypes, which is referred to as a “trans” covariance by Kong et al. and mediated by the copath in Model 1. However, two generations of AM (beginning in the grandparental generation) results in the recombination of alleles on the same haplotype, thus also leading to a “cis” covariance within the parental haplotypes. At equilibrium, after several generations of AM, the cis covariance ( $cov([N]T_*, [N]T_*)$ ) equals the trans covariance ( $cov([N]T_p, [N]T_m)$ ), with both denoted  $g$  in the models. Note, however, that only the cis covariances are explicit in Fig. 2; the trans covariances are implicit, already being accounted for by  $\mu$ . Note too that what is considered a trans covariance in the current generation (e.g., between  $T_p$  and  $T_m$ ) would be considered cis covariances in the next generation, when the offspring has children.

As denoted by the additional  $+g$  terms in the haplotypic PGS variances in Fig. 2, AM increases the variance within haplotypic PGS's to the same degree as the covariance between them. The  $k$  term in the haplotypic PGS variance represents the variance of the haplotypic PGS in the base population, and is not estimated; rather, it is fixed depending on how the user scales the PGS. If the full PGS is standardized in the base population, then  $k = \frac{1}{2}$ . This value of  $k$  is useful because the increase in the variances of haplotypic genetic scores under AM is easily quantified by the degree to which it is greater than  $\frac{1}{2}$ . However, standardizing in the base population is typically impossible in real data, and so is mostly useful only in simulated data or when there is no AM (such as Model 0). In real data, the full PGS ( $T_p + T_m$ ) will typically be standardized in the current generation, in which case  $k = \frac{1}{2} - 2g$ . Finally, if the *haplotypic* PGS is scaled in the current generation to have a variance of  $\frac{1}{2}$ , then  $k = \frac{1}{2} - g$ . So long as the value of  $k$  is consistent with how the PGS is scaled, the estimates of other parameters will not be affected. In all cases, the

variance of the full PGS ( $= var(T_m + T_p) = 2k + 4g$ , which  $= 1 + 4g$  if the full PGS is standardized in the base population)  $= 1$  if the full PGS is standardized in the current population, and  $= 1 + 2g$  if the haplotypic PGS is scaled to have variance  $= \frac{1}{2}$  in the current population.

The increase in genetic (co)variance of the PGS under AM,  $g$ , can be obtained by constraining its value to all chains that connect  $[N]T_p \rightarrow [N]T_m$  or vice-versa. Using  $\Omega$  as a substitute for all chains  $[N]T_* \rightarrow Y_*$ ,

$$g = cov([N]T_*, [N]T_*) = cov([N]T_p, [N]T_m) = \Omega\mu\Omega = \Omega^2\mu \tag{11}$$

Of course, because of the additional variances and covariances between the haplotypic PGS's, the expectation of  $\Omega$  itself is different in this model than it was in Model 0. For Model 1, the expected value is:

$$\Omega = cov([N]T_*, Y_*) = (k + g)\delta + \frac{w}{2}(1) + g\delta = \delta k + 2\delta g + \frac{1}{2}w \tag{12}$$

While accounting for AM makes this model more complicated than Model 0, substituting recurring chain segments drastically simplifies the derivations of parameters. For example,  $\theta_{NT}$ —which is derived by counting all chains  $Y_o \rightarrow NT_p$  and multiplying by 2 (to account for  $Y_o \rightarrow NT_m$ )—includes over 40 chains here as opposed to just 2 in Model 0. However, by using substitutions, this can be reduced to just four chains: (1)  $Y_o \rightarrow Y_p \rightarrow NT_p$  ( $= f\Omega$ , the genetic nurture chain); (2)  $Y_o \rightarrow T_p \rightarrow NT_p$  ( $= \delta g$ , arising from the AM-induced covariance between  $T_p$  and  $NT_p$ ); (3)  $Y_o \rightarrow T_m \rightarrow NT_p$  (also  $= \delta g$ , arising from the AM-induced covariance between  $T_m$  and  $NT_p$ ); and (4)  $Y_o \rightarrow Y_m \rightarrow Y_p \rightarrow NT_p$  ( $= fV_Y\mu\Omega$ , arising from the AM-induced covariance between  $Y_m$  and  $NT_p$ ). Therefore,

$$\theta_{NT} = cov(Y_o, NT_p + NT_m) = 2(f\Omega + \delta g + \delta g + fV_Y\mu\Omega) = 2f\Omega(1 + \mu V_Y) + 4\delta g \tag{13}$$

Similarly,

$$\theta_T = cov(Y_o, T_p + T_m) = 2\delta k + 2f\Omega(1 + \mu V_Y) + 4\delta g \tag{14}$$

Thus,  $\theta_T - \theta_{NT}$  ( $= 2\delta k$ ) is again an estimate of the direct effect of the PGS controlling for genetic nurture and, in this case, for AM.

In the same manner, the estimate of genetic nurture,  $w$ , can be derived by counting two chains  $F_o \rightarrow NT_m$  and multiplying by 2 (to account for  $F_o \rightarrow NT_p$ ): 1)  $F_o \rightarrow Y_m \rightarrow NT_m$  ( $= f\Omega$ , the genetic nurture chain); and 2)  $F_o \rightarrow Y_p \rightarrow Y_m \rightarrow NT_m$  ( $= fV_Y\mu\Omega$ , arising from the AM-induced covariance between  $Y_p$  and  $NT_m$ ). This leads to:

$$\begin{aligned}
 w &= cov(F_o, T_p + T_m) = cov(F_o, NT_p + NT_m) \\
 &= 2cov(F_o, [N]T_*) = 2f\Omega(1 + \mu V_Y)
 \end{aligned}
 \tag{15}$$

In Model 1,  $w$  remains an estimate of genetic nurture with its value being inflated by a factor  $(1 + \mu V_Y)$  under AM. In the Kong et al. notation, direct genetic nurture (denoted  $\eta$ ) refers to the aspect of  $w$  after removing the influence of AM, and thus  $\eta = 2f\Omega$ . Kong et al. also denote  $\phi_\eta$  as the added influence of AM on apparent genetic nurture, and thus  $\phi_\eta = 2f\Omega\mu V_Y$ . We do not further make this distinction between direct and indirect genetic nurture. For completeness, it should be noted that  $\phi_\delta$  (the genetic covariance between  $NT_*$  and  $Y_o$  induced by AM) in Kong et al.'s usage equals  $4\delta g$  here. From this, it follows that  $\theta_{NT}$  is no longer a direct estimate of genetic nurture (and that  $\theta_{NT} \neq w$ ) when there is AM because some of the covariance between  $Y_o$  and  $NT_*$  is now genetic in origin.

Finally, as was the case for  $w$ , the presence of AM causes the expectation of  $V_F$  to be inflated by a factor of  $(1 + \mu V_Y)$ :

$$V_F = var(F_p) = var(F_m) = var(F_o) = 2f^2 V_Y (1 + \mu V_Y) \tag{16}$$

with the value of  $V_Y$  being similarly inflated in Model 1 versus Model 0:

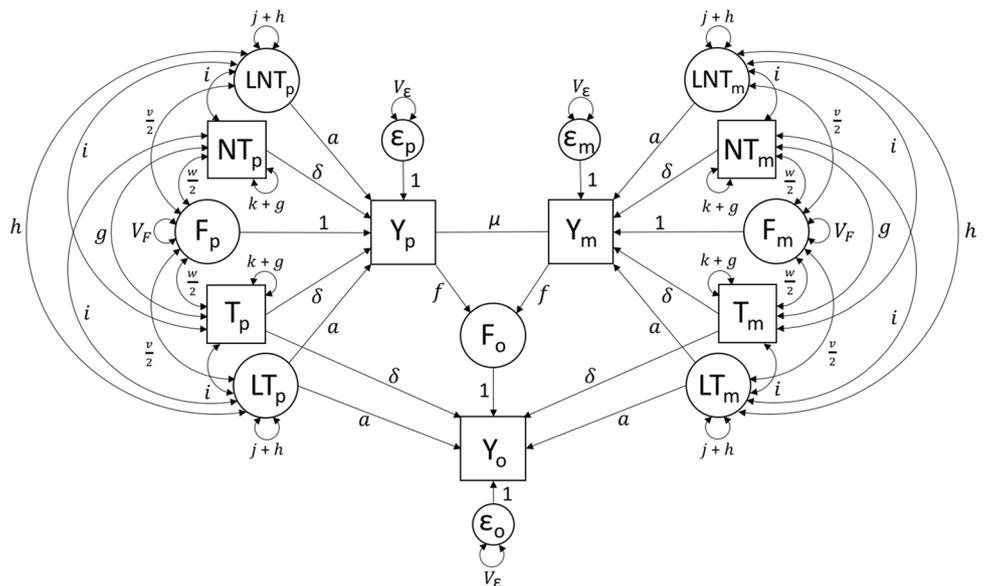
$$\begin{aligned}
 V_Y &= var(Y_p) = var(Y_m) = var(Y_o) \\
 &= 2\delta^2 k + 4\delta^2 g + 2\delta w + V_F + V_e
 \end{aligned}
 \tag{17}$$

### Model 2: VT and AM with latent genetic effects

Model 2 builds on the concepts described above for modeling AM, but unlike Model 1, it provides unbiased estimates when there is AM and the PGS explains little trait heritability. It does this by modeling haplotypic latent genetic scores (LGS's), denoted  $L[N]T_*$  in Fig. 3, that are defined to be statistically orthogonal to the haplotypic PGS's ( $[N]T_*$ ) in the base population. The latent genetic effects can be estimated either by including observed parental phenotypes, or by making an assumption about the base population additive genetic variance (and thus about the value of  $a$ ). Here, we take the first approach by assuming that parental phenotypes are measured (hence the squares used to represent  $Y_p$  and  $Y_m$  in Fig. 3), but discuss the second approach at the end of this section. It should be noted that full information maximum likelihood parameter estimates are unbiased by missingness unless the data is not missing at random (Schafer and Graham 2002). Thus, all three phenotypes need not be measured in every family. Indeed, each family could be made up only of pairs ( $Y_o, Y_p$ ;  $Y_o, Y_m$ ; or  $Y_m, Y_p$ ) and so long as all pairs are observed, parameter estimates would be unbiased, albeit with larger standard errors than with complete data.

Model 2 includes two additional observed variables ( $Y_p$  and  $Y_m$ ), leading to a 7-by-7 observed variance-covariance matrix and 28 unique statistics. The four haplotypic PGS variances and the six covariances between them are used to estimate  $g$  and  $\mu$ . Of the remaining 18 statistics, only six provide information that is not completely redundant to estimating parameters as specified in this model: the three described in Model 0 as well as the covariance between the parental phenotypes, the four covariances between one parent's PGS

**Fig. 3** Path diagram of Model 2, which models the effects of AM, VT, and latent genetics



and the other's phenotype, and the two covariances between each parental phenotype with the offspring phenotype. The parent-offspring covariances are used to estimate the latent genetic path coefficient ( $a$ ), which increases to the degree that  $cov(Y_*, Y_o)$  is higher than expected after accounting for genetic covariance through  $\delta$  and environmental covariance through  $f$ . In addition to  $a$ , there are four additional parameters ( $j, h, v, i$ ) in this model. None of these are estimated. Rather, their values are determined from non-linear constraints, which we turn to in order.

The variance of the haplotypic LGS ( $= j + h$ ) is treated analogously to the variance of the haplotypic PGS ( $= k + g$ ). Like  $k, j$  is defined as the genetic variance of the haplotypic LGS in the base population; however, unlike  $k$  (which is measured and therefore depends upon how the PGS is scaled),  $j$  is the variance of a latent construct and could thereby take any arbitrary value. The simplest choice is to define  $j$  so as to be consistent with  $k$ . Specifically, if the PGS is standardized in the base population (where  $k = \frac{1}{2}$ ), then  $j = \frac{1}{2}$ . If the PGS is standardized at equilibrium to have a variance of 1 (where  $k = \frac{1}{2} - 2g$ ), then  $j = \frac{1}{2} - 2h$ . If the haplotypic PGS is scaled at equilibrium to have a variance of  $\frac{1}{2}$  (where  $k = \frac{1}{2} - g$ ), then  $j = \frac{1}{2} - h$ .

The increase in the variance of the haplotypic LGS due to AM ( $h$ ) can be estimated under the reasonable assumption that the increase in the variance of the LGS from the base to the current population is proportionate to that of the PGS from the base to the current population. This assumption could be violated if the genes that drive the PGS association are more or less correlated with the trait actually being assorted on than the genes underlying the LGS, which seems unlikely. This assumption is equivalent to  $\frac{g}{\delta^2} = \frac{h}{a^2}$ , which leads to

$$h = g \frac{a^2}{\delta^2} \tag{18}$$

Thus,  $h$  and  $g$  are the same only when the PGS and LGS explain the same amount (half) of the total heritability. Furthermore, similar to  $\Omega$ , we define  $\Gamma$  to be the covariance between a parent's phenotype and one of their LGS's ( $\Gamma = cov(L[N]T_*, Y_*)$ ). Using  $\Gamma$  as a shortcut, the expected value of  $h$  can also be found by path tracing

$$h = cov(L[N]T_*, L[N]T_*) = cov(L[N]T_p, L[N]T_m) = \Gamma \mu \Gamma = \Gamma^2 \mu \tag{19}$$

Setting these two values of  $h$  to be equal leads to the non-linear constraint

$$h = g \frac{a^2}{\delta^2} = \Gamma^2 \mu \tag{20}$$

To enable estimation of the covariance between  $F$  and the LGS's (denoted by  $v$ ), we make a similar assumption that the ratio of genetic nurture to direct genetic effects is the same for observed as for latent genetic effects. This assumption could be violated if the genes driving the PGS association are more or less correlated with the trait that VT works through than the genes underlying the LGS, which again seems unlikely. This assumption is equivalent to  $\frac{v}{a} = \frac{w}{\delta}$ , which leads to

$$v = \frac{wa}{\delta} \tag{21}$$

The expected value of  $v$  via path tracing, and the resulting non-linear constraint, are

$$v = cov(F_o, LT_p + LT_m) = cov(F_o, LNT_p + LNT_m) = 2f\Gamma(1 + \mu V_Y) \tag{22}$$

$$\Rightarrow v = \frac{wa}{\delta} = 2f\Gamma(1 + \mu V_Y) \tag{23}$$

For the same reason that AM induces a covariance among PGS's ( $g$ ) and among LGS's ( $h$ ), it also induces a covariance between PGS's and LGS's, which we call  $i$ . From path tracing, the expected value of  $i$  is

$$i = cov(L[N]T_*, [N]T_*) = cov(L[N]T_p, [N]T_m) = cov([N]T_p, L[N]T_m) = \Gamma \mu \Omega = \sqrt{gh} \tag{24}$$

Unlike Model 1, Model 2 yields unbiased estimates of the full  $V_F$ , genetic nurture, and additive genetic variation, even when there is AM and the PGS explains only a fraction of total heritability. Model 2 properly accounts for the additional covariance between the PGS and the offspring phenotype that arises from the AM-induced covariance between PGS and LGS. For example,

$$\theta_{NT} = cov(NT_m + NT_p, Y_o) = 2f\Omega(1 + \mu V_Y) + 4\delta g + 4ai \tag{25}$$

where

$$\Omega = cov([N]T_*, Y_*) = \delta k + 2\delta g + 2ai + \frac{1}{2}w \tag{26}$$

Thus, the PGS-LGS covariance ( $i$ ) inflates both the genetic nurture part of  $\theta_{NT}$  ( $2f\Omega(1 + \mu V_Y)$ ) as well as the genetic part that arises via AM ( $4\delta g + 4ai$ ). When the PGS explains a small portion of the heritability, the covariance between the LGS and the PGS can be much greater than the covariance between the PGS's themselves ( $i \gg g$ ). By not accounting for  $i$ , the observed  $\theta_{NT}$  in Model 1 is inflated over its expected value, upwardly biasing estimates of  $V_F$  and genetic nurture (Kim et al. 2020).

Several other parameters in this model are the latent genetic analogs to parameters related to the observed PGS's, including  $\theta_{LNT} (= cov(Y_o, LNT_p + LNT_m))$ , the analog to  $\theta_{NT}$  and  $\theta_{LT} (= cov(Y_o, LT_p + LT_m))$ , the analog to  $\theta_T$ . Expectations of these and other parameters that have not been derived in this section can be found in the Supplement.

Finally, as noted above, Model 2 can be fit without using parental phenotypes if there exist good estimates of the total heritability in the base population. For a standardized trait, the additive genetic variation in the base population is  $a^2 + \delta^2$ ; thus, by subtraction of the  $\delta^2$  term observed in the data (where  $\delta = \theta_T - \theta_{NT}$ ), one can find the assumed value of  $a$  and substitute it into the model. This leads to unbiased estimates of all parameters whenever the assumed value of  $a$  is correct and downwardly biased estimates of  $w$  and  $V_F$  to the degree that the assumed value of  $a$  is too high (and vice-versa). When using this strategy, it is therefore important to have decent estimates of heritability in the base population that account for the influences of genetic nurture and AM. Estimates from twin studies are biased under these conditions, but estimates from extended twin family models should be much less so (Keller et al. 2010). Kong et al. recognized the confounding influence that the LGS has on parameter estimates when there is AM, and attempted to estimate genetic nurture using assumed values of the base population heritability that came from relatedness disequilibrium regression (RDR; Young et al. 2018). However, estimates of heritability from RDR are actually estimates of the base population genetic variance scaled by the phenotypic variance in the current population, and are therefore biased downwards under AM (Kemper et al. in press). This likely led to overestimates of genetic nurture in Kong et al. and would have led to overestimates of  $V_F$  had it been estimated. Nonetheless, if the equilibrium spousal correlations are known, a simple correction can be applied to RDR estimates of heritability (Kemper et al. in press).

### Accounting for differences in parent vs. offspring phenotypes

In all of these models, we have assumed that the genetic (PGS and LGS) effects are equivalent in parents and offspring. This assumption would be violated if there are gene-by-age or gene-by-cohort effects. Kong et al. provide some evidence of this in their data: the correlation between the PGS and educational attainment was significantly higher among offspring (.22) than among parents (.12). When parental phenotypes are measured in Model 2, or when parental phenotypes are unmeasured but there are independent estimates of the PGS effects in the parental generation, accounting for such effects is possible by estimating two different  $\delta$  values, one for offspring ( $\delta_o$ ) and one for parents

( $\delta_*$ ). While information for estimating  $\delta_o$  still comes directly from  $\theta_T - \theta_{NT}$ , there is no direct estimate of  $\delta_*$ , even though it is informed by  $cov(Y_*, T_* + NT_*)$ . A reasonable assumption, such as equal proportions of direct genetic effects ( $\frac{\delta_*}{cov(Y_*, T_* + NT_*)} = \frac{\delta_o}{cov(Y_o, T_m + T_p)}$ ), should allow estimation of both  $\delta_*$  and  $\delta_o$ , making this model identified (although we have not checked this formally).

It may also be of interest to understand how one parental trait influences a different offspring trait. For example, Kong et al. showed a covariance between the nontransmitted PGS of educational attainment and offspring health, consistent with cross-trait (parental education to offspring health) VT and genetic nurture. Such cross-trait genetic nurture would contribute to apparent genetic correlations that have nothing to do with pleiotropy. To investigate cross-trait VT using the above example under the current framework, one could use the PGS and parental values of educational attainment along with the offspring values of health, and plug them into the above models without modification. This is similar to the approach taken by Kong et al. However, such an approach does not account for AM within (health–health) or across (education–health) traits, nor does it account for within-trait genetic nurture effects. For these reasons, we believe that cross-trait VT and genetic nurture effects are optimally modeled bivariately, using the PGS's and phenotypic values of the two traits in both parents and offspring; including more than two traits would also be possible, but results from such a model would likely be incomprehensibly complex. The parameters from the above models would be 2-by-2 full (in the case of path coefficients) or symmetric (in the case of variance–covariance) matrices. Conveniently, nothing about the derivations in this paper would change except for keeping track of when matrices should be transposed, which obeys an additional path tracing rule (Vogler and Cockerham 1985). This bivariate model would estimate two direct and two cross-trait VT paths and four genetic nurture paths all while accounting for direct genetic effects, pleiotropy, and AM within and between traits. While this may sound like a lot to ask of a model, there is a tremendous amount of unique information contained in the 14-by-14 observed variance-covariance matrix, making this approach powerful if the PGS  $r^2$ 's for both traits are nontrivial. Development of bivariate extensions of the present models is left for future work.

### Testing and modeling different mechanisms of AM

While our models have thus far assumed primary phenotypic AM under equilibrium, other mechanisms of assortment are possible. Indeed, there is considerable power for testing

different mechanisms of AM, which could itself be used as a focus of these models. This power stems from the amount of information in this model relevant to AM. There are a total of 10 observed haplotypic PGS variances or covariances which collectively provide 10 estimates of  $g$  (see Model 1). The  $\chi(1)$  test of whether the average value across all 10  $g$  estimates is significantly greater than 0 tests whether mate similarity (either on the trait in question or a trait genetically correlated to it) has led to genetic covariance, as predicted by primary AM. Additionally, of these 10 estimates of  $g$ , 6 provide information on cis (within-person) genetic covariances and 4 provide information on trans (across-mate) genetic covariances. The  $\chi(1)$  test of whether these two groups of covariances are equal tests whether AM has gone on long enough to lead to equilibrium values of parameters. Significantly higher estimates of trans  $g$  vs. cis  $g$  would suggest that AM is at disequilibrium. Similarly, if trans  $g$  estimates are significantly greater than 0 while cis  $g$  estimates are not, this would support the hypothesis that only a single generation of AM has occurred (which is consistent with what Kong et al. found for educational attainment in Iceland). Furthermore, the ten estimates of  $g$  can be used to derive expected values of  $cov(NT_p + T_p, Y_m)$ ,  $cov(NT_m + T_m, Y_p)$ , and  $cov(Y_p, Y_m)$  to test various models of AM. For example, mate similarity caused by environmental similarity (social homogamy) predicts that observed  $cov(Y_p, Y_m)$  is higher than that implied by  $g$ . On the other hand, primary AM on a trait that is more genetically than phenotypically correlated with the measured trait (a form of genetic homogamy) would predict that observed  $cov(Y_p, Y_m)$  is lower than that implied by  $g$ .

Once the data suggest a particular mechanism underlying mate similarity, it can be modeled using the present framework. For example, in the Supplement, we show how disequilibrium AM (from a single generation of assortment) can be modeled by setting expectations of cis  $g$ ,  $h$ , and  $i$  to zero. Furthermore, social and genetic homogamy can be modeled by assuming that AM occurs on a latent trait,  $\tilde{Y}$ , that is related to  $Y$  through either environmental or genetic pathways, respectively (Keller et al. 2009). This allows the observed  $cov(Y_p, Y_m)$  to differ from the  $cov(Y_p, Y_m)$  implied by  $g$ . Thus, alternative mechanisms of AM can be formally tested using the rich information available from parent and offspring PGS's and phenotypes, and when called for, models can be modified to incorporate alternative mechanisms of AM.

## Discussion

Genetic nurture, the covariance between genes and parentally provided environmental influences, can amplify genetic effects in a way that neither GWAS nor heritability studies have been able to sufficiently account for. While estimating

direct genetic effects after removing their covariance with the environment is important, we argue that the converse—estimating the direct environmental effect after removing its covariance with genetic effects—is at least as important. The models presented above allow for estimates of this direct environmental influences of parents on offspring, and also suggest several important extensions. For example, there exists much more GWAS data with siblings than with parents, and so extending the models above to include siblings, twins, and potentially other collateral relatives remains a next step. Furthermore, as shown by Kong et al. with enough data it is straightforward to estimate differential genetic nurture effects depending on the parent of origin; thus, it is equally straightforward to estimate differential VT effects from fathers vs. mothers. Finally, as noted above, future multivariate extensions to this model would provide insight into how one parental trait influences a different childhood trait.

There are several important caveats and limitations to the present approach. We discuss here the central ones. One very important caveat is that estimated  $V_F$  from these models cannot be considered the full parental effect on this trait. Rather, it is the variance in the trait due to parental influences that are associated with *the specific trait assessed by the PGS*. For example, the vertical transmission variance in a model that uses an educational attainment PGS only estimates how traits genetically related to parental educational attainment influence offspring educational attainment. If other parental traits, such as intelligence, warmth, work ethic, conscientiousness, etc. also influence offspring educational attainment, then the portion of variance due to these and other parental influences that are genetically uncorrelated to educational attainment will be missed.

The above caveat is related to the limitation that, in order to accurately characterize the influence of a parental trait on an offspring trait, sufficiently predictive PGS's (e.g.,  $r^2 > .02$ ) must exist for traits relevant to parenting. Optimally (to use Model 2), these traits should also be measured in parents in the same dataset that the models are applied to. This is, perhaps, the greatest limitation of the current approach: it can only look under the lamppost, at traits analyzed in large GWAS's for which sufficiently predictive PGS's exist. Because it is so easily and frequently collected, educational attainment may be an exception, but a great many traits relevant to parental influences have never been investigated in GWAS. This—the ability to use PGS's to understand how parents influence offspring—is a further motivation to continue to extend GWAS investigations from their traditional focus on health to as many behavioral and psychological traits as possible.

That said, there are many traits that have sufficiently predictive PGS's to answer questions of great interest. For example, does parental liability to major depression, schizophrenia, or externalizing disorder directly influence

the same or different traits in the children (Okbay et al. 2016; Barr et al. 2020)? Does parental socioeconomic status directly impact offspring socioeconomic status, educational attainment, or subjective well-being (Hill et al. 2019)? Does parental smoking influence offspring smoking (Liu et al. 2019)? This latter question is interesting with respect to negative VT, which occurs when higher values of the parental trait lead to lower values of the offspring trait. Negative VT would lead to positive  $V_F$  but to negative genetic nurture, dampening estimated genetic influences from GWAS or heritability studies. While this is probably rare, there is some evidence from extended twin family models that negative VT occurs for smoking (Maes et al. 2006). While this finding has been explained away in the literature as probably being driven by gene-by-age interactions, it is also possible that smoking and other traits associated with teen rebelliousness lose their lustre when parents engage in them. Given that a sufficiently predictive PGS for smoking behaviors exists (Liu et al. 2019) and that Model 2 can be extended to account for gene-by-age effects, a whole-genome dataset that includes parents, offspring, and information on smoking behavior could resolve whether parental smoking directly increases or decreases offspring smoking.

A further caveat to the above approach is that stratification can bias estimates if it is not properly controlled for. For example, if a discovery GWAS for educational attainment does not fully correct for mean differences across ancestry groups, the PGS for educational attainment will predict both educational attainment as well as ancestry. In the models explored above, this stratification effect would increase the covariance between the transmitted and nontransmitted PGS's and the offspring phenotype ( $\theta_{[NT]}$ ), inflating evidence of genetic nurture and  $V_F$ . While stratification is a type of passive G-E covariance that inflates parent-offspring resemblance, the mechanism is due to a factor (ancestry) that is shared between parents and offspring rather than a direct parental-to-offspring influence, and so these should be disambiguated. Therefore, to minimize the effects of stratification, principal components from the genomic relationship matrices should be included as covariates in both the discovery GWAS as well as the causal models discussed above.

Last, we have assumed that the passive G-E covariance we model arises only from VT (genetic nurture) as opposed to horizontal (such as sibling) transmission. For certain traits, such as experimentation with drugs and alcohol, horizontal transmission seems at least as likely as VT. In a model that includes both parents and siblings, there would be sufficient power to differentiate horizontal transmission from VT. In the meantime, it must be kept in mind that estimates of  $V_F$  from these models may also contain environmental influences from siblings or (less likely) from other collateral relatives.

To our knowledge, this is the first treatment of how transmitted and nontransmitted PGS's can be used to estimate the direct effect of parents on their offspring, and the first to account for the influence of different types of AM on these estimates. It builds upon the seminal work by Zhang et al. (2015) and Kong et al. (2018), who recognized that this data could be used to estimate genetic nurture. There has been long-standing interest in how parents influence offspring in fields such as developmental psychology, but the traditional approach of correlating parental behavior with offspring outcomes does not control for the confounding influence of shared genetic effects. Given the skepticism of genetic approaches in fields dedicated to studying parenting, it is perhaps ironic that molecular genetic data provides an excellent way to estimate the direct effect of parents on offspring. Genome-wide data, originally collected to find the specific alleles that underlie health-related traits, has begun to be used for a multitude of purposes never envisioned by early practitioners. As we have argued here, another one of these purposes is to help disentangle how parents influence their children—genetically, environmentally, and in concert.

**Acknowledgements** We thank Richard Border, Hermine Maes, and David Evans for help with ideas and feedback as the project developed. This work was supported by Grant Nos. R01MH100141 (to MCK) and T32MH016880 (to Dr. John Hewitt) and the Institute for Behavioral Genetics. This work utilized resources from the University of Colorado Boulder Research Computing Group, which is supported by the National Science Foundation (Award Nos. ACI-1532235 and ACI-1532236), the University of Colorado Boulder, and Colorado State University.

## Compliance with ethical standards

**Conflict of interest** Jared V. Balbona, Yongkang Kim, and Matthew C. Keller declare that they have no conflicts of interest.

**Human and Animal Rights** This article does not contain any studies with human or animal subjects performed by any of the authors.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

- Barr PB, Salvatore JE, Wetherill L, Anokhin A, Chan G, Edenberg HJ, Kuperman S, Meyers J, Nurnberger J, Porjesz B, Schuckit M, Dick DM (2020) A family-based genome wide association study of externalizing behaviors. *Behav Genet* 50(3):175–183
- Bollen KA, Pearl J (2013) Eight myths about causality and structural equation models. Springer, Dordrecht
- Bonilla C, Lawlor DA, Ben-Shlomo Y, Ness AR, Gunnell D, Ring SM, Smith GD, Lewis SJ (2012) Maternal and offspring fasting glucose and type 2 diabetes-associated genetic variants and cognitive function at age 8: a Mendelian randomization study in the Avon Longitudinal Study of Parents and Children. *BMC Med Genet* 13(1):90
- Bulik-Sullivan BK, Loh P-R, Finucane HK, Ripke S, Yang J, Patterson N, Daly MJ, Price AL, Neale BM (2015) LD Score regression distinguishes confounding from polygenicity in genome-wide association studies. *Nat Genet* 47(3):291–295
- Cavalli-Sforza LL, Feldman MW (1973) Cultural versus biological inheritance: phenotypic transmission from parents to children. (A theory of the effect of parental phenotypes on children's phenotypes). *Am J Hum Genet* 25(6):618–637
- Cloninger CR (1980) Interpretation of intrinsic and extrinsic structural relations by path analysis: theory and applications to assortative mating. *Genet Res* 36(2):133–145
- Cloninger CR, Rice J, Reich T (1979) Multifactorial inheritance with cultural transmission and assortative mating. II. A general model of combined polygenic and cultural inheritance. *Am J Hum Genet* 31(2):176–198
- Coventry WL, Keller MC (2005) Estimating the extent of parameter bias in the classical twin design: a comparison of parameter estimates from extended twin-family and classical twin designs. *Twin Res Hum Genet* 8(3):214–223
- Davey Smith G, Ebrahim S (2003) 'Mendelian randomization': can genetic epidemiology contribute to understanding environmental determinants of disease?\*. *Int J Epidemiol* 32(1):1–22
- DiLalla LF, Gottesman II (1991) Biological and genetic contributors to violence: Widom's untold tale. *Psychol Bull* 109(1):125–129
- Eaves L (1976) The effect of cultural transmission on continuous variation. *Heredity* 37(1):41–57
- Evans DM, Moen G-H, Hwang L-D, Lawlor DA, Warrington NM (2019) Elucidating the role of maternal environmental exposures on offspring health and disease using two-sample Mendelian randomization. *Int J Epidemiol* 48(3):861–875
- Hill WD, Davies NM, Ritchie SJ, Skene NG, Bryois J, Bell S, Di Angelantonio E, Roberts DJ, Xueyi S, Davies G, Liewald DCM, Porteous DJ, Hayward AS, Butterworth AS, McIntosh AM, Gale CR, Deary IJ (2019) Genome-wide analysis identifies molecular systems and 149 genetic loci associated with income. *Nat Commun* 10(1):5741
- Keller MC, Medland SE, Duncan LE (2010) Are extended twin family designs worth the trouble? A comparison of the bias, precision, and accuracy of parameters estimated in four twin family models. *Behav Genet* 40(3):377–393
- Keller MC, Medland SE, Duncan LE, Hatemi PK, Neale MC, Maes HHM, Eaves LJ (2009) Modeling extended twin family data I: description of the cascade model. *Twin Res Hum Genet* 12(1):8–18
- Kemper KE, Yengo L, Zheng Z, Abdellaoui A, Keller MC, Goddard ME, Wray NR, Yang J, Visscher PM (n.d.). Phenotypic covariance across the entire spectrum of relatedness for 86 billion pairs of individuals
- Kim Y, Balbona JV, Keller MC (2020) Bias and precision of parameter estimates from models using polygenic scores to estimate environmental and genetic parental influences
- Kong A, Thorleifsson G, Frigge ML, Vilhjalmsdottir BJ, Young AI, Thorgeirsson TE, Benonisdottir S, Oddsson A, Halldorsson BV, Masson G, Gudbjartsson DF, Helgason A, Bjornsdottir G, Thorsteinsdottir U, Stefansson K (2018) The nature of nurture: effects of parental genotypes. *Science* 359(6374):424–428
- Lawlor D, Richmond R, Warrington N, McMahon G, Davey Smith G, Bowden J, Evans DM (2017) Using Mendelian randomization to determine causal effects of maternal pregnancy (intrauterine) exposures on offspring outcomes: sources of bias and methods for assessing them. *Wellcome Open Research*, 2
- Lawlor DA, Timpson NJ, Harbord RM, Leary S, Ness A, McCarthy MI, Frayling TM, Hattersley AT, Smith GD (2008) Exploring the developmental overnutrition hypothesis using parental-offspring associations and FTO as an instrumental variable. *PLoS Med* 5(3):e33
- Lee JJ, Wedow R, Okbay A, Kong E, Maghziyan O, Zacher M, Nguyen-Viet TA, Bowers P, Sidorenko J, Karlsson Linnér R, Fontana MA, Kundu T, Lee C, Li H, Li R, Royer R, Timshel PN, Walters RK, Willoughby EA, Yengo L, Alver M, Bao Y, Clark DW, Day FR, Furlotte NA, Joshi PK, Kemper KE, Kleinman A, Langenberg C, Mägi R, Trampush JW, Verma SS, Wu Y, Lam A, Zhao JH, Zheng Z, Boardman JD, Campbell H, Freese J, Harris KM, Hayward C, Herd P, Kumari M, Lencz T, Luan J, Malhotra AK, Metspalu A, Milani L, Ong KK, Perry JRB, Porteous DJ, Ritchie MD, Smart MC, Smith BH, Tung JY, Wareham NJ, Wilson JF, Beauchamp JP, Conley DC, Esko T, Lehrer SF, Magnusson PKE, Oskarsson S, Pers TH, Robinson MR, Thom K, Watson C, Chabris CF, Meyer MN, Laibson DI, Yang J, Johannesson M, Koellinger PD, Turley P, Visscher PM, Benjamin DJ, Cesarini D, 23andMe Research Team, Social Science Genetic Association Consortium (2018) Gene discovery and polygenic prediction from a genome-wide association study of educational attainment in 1.1 million individuals. *Nat Genet* 50(8):1112–1121
- Liu M, Jiang Y, Wedow R, Li Y, Brazel DM, Chen F, Datta G, Davila-Velderrain J, McGuire D, Tian C, Zhan X, Choquet H, Docherty AR, Faul JD, Foerster JR, Fritsche LG, Gabrielsen ME, Gordon SD, Haessler J, Hottenga J-J, Huang H, Jang S-K, Jansen PR, Ling Y, Mägi R, Matoba N, McMahon G, Mulas A, Orrù V, Palviainen T, Pandit A, Reginsson GW, Skogholt AH, Smith JA, Taylor AE, Turman C, Willemsen G, Young H, Young KA, Zajac GJM, Zhao W, Zhou W, Bjornsdottir G, Boardman JD, Boehnke M, Boomsma DI, Chen C, Cucca F, Davies GE, Eaton CB, Ehringer MA, Esko T, Fiorillo E, Gillespie NA, Gudbjartsson DF, Haller T, Harris KM, Heath AC, Hewitt JK, Hickie IB, Hokanson JE, Hopfer CJ, Hunter DJ, Iacono WG, Johnson EO, Kamatani Y, Kardina SLR, Keller MC, Kellis M, Kooperberg C, Kraft P, Krauter KS, Laakso M, Lind PA, Loukola A, Lutz SM, Madden PAF, Martin NG, McGue M, McQueen MB, Medland SE, Metspalu A, Mohlke KL, Nielsen JB, Okada Y, Peters U, Polderman TJC, Posthuma D, Reiner AP, Rice JP, Rimm E, Rose RJ, Runarsdottir V, Stallings MC, Stančáková A, Stefansson H, Thai KK, Tindle HA, Tyringsson T, Wall TL, Weir DR, Weisner C, Whitfield JB, Winsvold BS, Yin J, Zuccolo L, Bierut LJ, Hveem K, Lee JJ, Munafò MR, Saccone NL, Willer CJ, Cornelis MC, David SP, Hinds DA, Jorgenson E, Kaprio J, Stitzel JA, Stefansson K, Thorgeirsson TE, Abecasis G, Liu DJ, Vrieze S (2019) Association studies of up to 1.2 million individuals yield new insights into the genetic etiology of tobacco and alcohol use. *Nat Genet* 51(2):237–244
- Lynch M, Walsh B (1998) Genetics and analysis of quantitative traits. *Q Rev Biol* 74(2):225–225
- Maes HH, Neale MC, Kendler KS, Martin NG, Heath AC, Eaves LJ (2006) Genetic and cultural transmission of smoking initiation: an extended twin kinship model. *Behav Genet* 36(6):795–808
- Okbay A, Baselmans BML, Neve J-E, Turley P, Nivard MG, Fontana MA, Meddens SFW, Linnér RK, Rietveld CA, Derringer J, Gratten J, Lee JJ, Liu JZ, Vlaming R, Ahluwalia TS, Buchwald J,

- Cavadino A, Frazier-Wood AC, Furlotte NA, Garfield V, Geisel MH, Gonzalez JR, Haitjema S, Karlsson R, Laan SW, Ladwig K-H, Lahti J, Lee SJ, Lind PA, Liu T, Matteson L, Mihailov E, Miller MB, Minica CC, Nolte IM, Mook-Kanamori D, van der Most PJ, Oldmeadow C, Qian Y, Raitakari O, Rawal R, Realo A, Ruedi R, Schmidt B, Smith AV, Stergiakouli E, Tanaka T, Taylor K, Thorleifsson G, Wedenoja J, Wellmann J, Westra H-J, Willems SM, Zhao W, Amin N, Bakshi A, Bergmann S, Bjornsdottir G, Boyle PA, Cherny S, Cox SR, Davies G, Davis OSP, Ding J, Direk N, Eibich P, Emeny RT, Fatemifar G, Faul JD, Ferrucci L, Forstner AJ, Gieger C, Gupta R, Harris TB, Harris JM, Holliday EG, Hottenga J-J, De Jager PL, Kaakinen MA, Kajantie E, Karhunen V, Kolcic I, Kumari M, Launer LJ, Franke L, Li-Gao R, Liewald DC, Koini M, Loukola A, Marques-Vidal P, Montgomery GW, Mosing MA, Paternoster L, Pattie A, Petrovic KE, Pulkki-Råback L, Quaye L, Rääkkönen K, Rudan I, Scott RJ, Smith JA, Sutin AR, Trzaskowski M, Vinkhuyzen AE, Yu L, Zabaneh D, Attia JR, Bennett DA, Berger K, Bertram L, Boomsma DI, Snieder H, Chang S-C, Cucca F, Deary IJ, van Duijn CM, Eriksson JG, Bültmann U, de Geus EJC, Groenen PJF, Gudnason V, Hansen T, Hartman CA, Haworth CMA, Hayward C, Heath AC, Hinds DA, Hyppönen E, Iacono WG, Järvelin M-R, Jöckel K-H, Kaprio J, Kardia SLR, Keltikangas-Järvinen L, Kraft P, Kubzansky LD, Lehtimäki T, Magnusson PKE, Martin NG, McGue M, Metspalu A, Mills M, de Mutsert R, Oldehinkel AJ, Pasterkamp G, Pedersen NL, Plomin R, Polasek O, Power C, Rich SS, Rosendaal FR, den Ruijter HM, Schlessinger D, Schmidt H, Svento R, Schmidt R, Alizadeh BZ, et al (2016) Genetic variants associated with subjective well-being, depressive symptoms, and neuroticism identified through genome-wide analyses. *Nat Genet* 48(6):624–633
- Schafer JL, Graham JW (2002) Missing data: our view of the state of the art. *Psychol Methods* 7(2):147–177
- Tahmasbi R, Keller MC (2017) GeneEvolve: a fast and memory efficient forward-time simulator of realistic whole-genome sequence and SNP data. *Bioinformatics* 33(2):294–296
- Torkamani A, Wineinger NE, Topol EJ (2018) The personal and clinical utility of polygenic risk scores. *Nat Rev Genet* 19(9):581–590
- Tubbs JD, Porsch RM, Cherny SS, Sham PC (2020a) The genes we inherit and those we don't: maternal genetic nurture and child BMI trajectories. *Behav Genet* 50(5):310–319
- Tubbs JD, Zhang YD, Sham PC (2020b) Intermediate confounding in trio relationships: the importance of complete data in effect size estimation. *Genet Epidemiol* 44(4):395–399
- Tyrrell J, Richmond RC, Palmer TM, Feenstra B, Rangarajan J, Metrustry S, Cavadino A, Paternoster L, Armstrong LL, De Silva NMG, Wood AR, Horikoshi M, Geller F, Myhre R, Bradfield JP, Kreiner-Møller E, Huikari V, Painter JN, Hottenga J-J, Allard C, Berry DJ, Bouchard L, Das S, Evans DM, Hakonarson H, Hayes MG, Heikkinen J, Hofman A, Knight B, Lind PA, McCarthy MI, McMahon G, Medland SE, Melbye M, Morris AP, Nodzenski M, Reichetzeder C, Ring SM, Sebert S, Sengpiel V, for the Early Growth Genetics (EGG) Consortium, et al (2016) Genetic evidence for causal relationships between maternal obesity-related traits and birth weight. *JAMA* 315(11):1129
- Vogler GP, Cockerham CC (1985) Multivariate path analysis of familial resemblance. *Genet Epidemiol* 2(1):35–53
- Warrington NM, Freathy RM, Neale MC, Evans DM (2018) Using structural equation modelling to jointly estimate maternal and fetal effects on birthweight in the UK Biobank. *Int J Epidemiol* 47(4):1229–1241
- Wright S (1934) The method of path coefficients. *Ann Math Stat* 5(3):161–215
- Yang J, Lee SH, Goddard ME, Visscher PM (2011) GCTA: a tool for genome-wide complex trait analysis. *Am J Hum Genet* 88(1):76–82
- Young AI, Frigge ML, Gudbjartsson DF, Thorleifsson G, Bjornsdottir G, Sulem P, Masson G, Thorsteinsdottir U, Stefansson K, Kong A (2018) Relatedness disequilibrium regression estimates heritability without environmental bias. *Nat Genet* 50(9):1304–1310
- Zhang G, Bacelis J, Lengyel C, Teramo K, Hallman M, Helgeland J, et al (2015) Assessing the causal relationship of maternal height on birth size and gestational age at birth: a Mendelian randomization analysis. *PLoS Med* 12(8):e1001865

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.