**TITLE: Partitioning phenotypic variance due to parent-of-origin effects using genomic relatedness matrices**

**RUNNING HEAD: Variance due to parent-of-origin effects using genomic relatedness**

**AUTHOR LIST**

Charles Laurin[1], Gabriel Cuellar Partida[2], Gibran Hemani[1], George Davey Smith[2], Jian Yang[3], and David M Evans[1,2]

[1]MRC Integrative Epidemiology Unit, University of Bristol, United Kingdom

[2]University of Queensland Diamantina Institute, Translational Research Institute, Brisbane, Queensland, Australia

[3]Institute for Molecular Bioscience, Queensland Brain Institute, The University of Queensland, Brisbane, Queensland, Australia

**Corresponding Author:**

Dr David M Evans

Faculty of Medicine

The University of Queensland Diamantina Institute

Brisbane, Queensland

Australia

Tel:    +61 7 344 37051

Fax:    +61 7 3443 6966

Email: d.evans1@uq.edu.au

# Introduction

Parent-of-origin effects (POEs) describe the phenomenon in which the effects of alleles depend upon their parental origin. POEs imply that heterozygote individuals have phenotypes which are distributed differently depending upon which of their alleles were maternally and paternally transmitted (Guilmatre and Sharp 2012; Lawson et al. 2013). The extreme case of POEs is *polar overdominance*, where the two heterozygotes' phenotypes differ in distribution but the two homozygotes share the same distribution (Hoggart et al. 2014). Imprinting, a phenomenon in which one parent's allele is not expressed, is probably the most widely studied example of POE (Peters 2014).

POEs have traditionally been examined in the context of development, where, in mouse models, they have been associated with body size and social behavior (Peters 2014). One evolutionary explanation of POEs concerns genomic conflict between maternal and paternal genes in offspring, with paternal genes encouraging growth and solicitation of maternal care, even at the expense of the mother's health, while maternal alleles are orientated toward success of all offspring, which do not necessarily share paternity (Patten et al. 2014). Other evolutionary explanations for POEs include: different territorial patterns in males and females and coadaptation of maternal and offspring genomes to maximize the efficiency of nurturing behaviors like suckling and grooming (Peters 2014).

Whilst there is considerable support for the importance of POEs in animals (Neugebauer et al. 2010; Lawson et al. 2013), evidence for the existence of POEs in the etiology of complex human traits and diseases is mixed, in part due to the relative paucity of genomic data from families (Kong et al. 2009; Guilmatre and Sharp 2012). Before the genomics era, the children-of-

twins design (Nance and Corey 1976), pedigree analyses (Hall 1990), and parent-offspring regressions (Clemons 2000; Smith et al. 2007) provided some limited evidence for the existence of POEs in human populations. These approaches were not often able to distinguish parental effects (indirect effects of the parental genotype on offspring phenotype) from POEs (interaction between the sex of the transmitting parent and the direct allelic effect in offspring) (Hager et al. 2008). More recently, in the genomics era, genome-wide association studies incorporating parent-of-origin information have been used to identify POEs at individual loci for age at menarche (Perry et al. 2014) Type I diabetes (Wallace et al. 2010), Type II diabetes, basal cell carcinoma, and breast cancer (Kong et al. 2009).

We propose a method, *G-REMLadp*, to estimate the phenotypic variance due to POEs across the genome by applying restricted maximum likelihood (REML) to offspring genome-wide genetic relatedness matrices and phenotype data. Our method involves the construction of a genetic relationship matrix indexing the parental origin of offspring alleles using genome-wide SNP data from parent-child duos or trios. The proposed method is an early adaptation to human genetics of procedures developed for animal breeding (Schaeffer et al. 1989; Spencer 2002; Nishio and Satoh 2015). The genotypic coding for POEs, which we adopt here, is based on the model outlined by Spencer (2002, 2009) and implemented by Nishio and Satoh (2015) in the context of genomic prediction. The genotype coding we used for dominance effects is due to Zhu *et al.*, (2015), who designed it to be orthogonal to the allele count at a locus; it is also orthogonal to the POE coding, which distinguishes our approach from that used by Spencer, Nishio and Satoh, and others. Using these codings, the phased genotype at a locus is coded using three orthogonal components: an additive-coded genotype, a dominance-coded genotype, and a POE-coded genotype.

To estimate the power of *G-REMLadp* to estimate non-null POEs, we provide an approximation using Haseman-Elston regression (Elston et al. 2000; Chen 2014). We also used simulated data to estimate the power and Type I Error rates of *G-REMLadp*, as well as the sensitivity of its variance component estimates to violations of assumptions. We then applied *G-REMLadp* to 38 phenotypes related to body size and obesity, metabolic traits, and IQ in a sample of up to 4689 individuals from a UK-based longitudinal study of childhood health and development.

## Methods

*G-REMLadp* uses REML to fit a linear mixed model incorporating random additive effects, dominance effects, and POEs to human genomic data, with the goal of partitioning phenotypic variance into components reflecting these sources of variation tagged by genome-wide SNP chips. In this model, at each locus, the phased genotype of each individual $\big( X \in aa, \ a_{mo}A_{fa}, \ A_{mo}a_{fa}, \ AA \big)$ is expressed as 3 orthogonal components, which are then standardized. The codings for each of the 3 components are listed in Table 1. The standardized additive-coded genotype is the standardized minor allele (A) count at the locus. The derivation of the standardized dominance-coded genotype is given in Zhu *et al.* (2015); and a flexible equivalent coding is given in Álvarez-Castro (2015). The POE-coded genotype is -1 for the paternal-minor heterozygote $(a_{mo}A_{fa})$, where $a_{mo}$ indicates that the major allele was inherited from the mother and $A_{fa}$ that the minor allele was inherited from the father), as 1 for the maternal-minor heterozygote $(A_{mo}a_{fa})$, and 0 for both homozygotes, and is then standardized.

The mixed model is given in Equation 1, where $\mathbf{y}$ is a vector of phenotypes; $\boldsymbol{\mu}_y$ is a column vector of means; the vector $\mathbf{Tb}$ includes any fixed effects of covariates $\mathbf{T}$; the $\mathbf{Z}$s are $n \times m$ matrices containing the standardised coded genotypes, represented as one individual per row (sample size $n$) and one SNP per column ($m$ markers affect the phenotype; in empirical data, this will be replaced by $m_o$, the number of observed markers, while in power analyses of empirical data, this will be replaced by $m_e$, the number of effective markers); the $\boldsymbol{\beta}$ are $m$-vectors of additive effects, dominance effects and POEs, which are assumed to be independently normally distributed with mean 0 and variances that are inversely proportional to the number of markers affecting the phenotype $\mathrm{Var}(\beta_\alpha) = m^{-1}\sigma_\alpha^2 \mathbf{I}_n$ ($\mathbf{I}$ is the $n \times n$ identity matrix), $\mathrm{Var}(\beta_\Delta) = m^{-1}\sigma_\delta^2 \mathbf{I}_n$, $\mathrm{Var}(\beta_\gamma) = m^{-1}\sigma_\gamma^2 \mathbf{I}_n$, respectively, and $\boldsymbol{\varepsilon}$ is a normal error variable with mean 0 and variance $\sigma_\varepsilon^2 \mathbf{I}_n$ that is independent of the other variables on the right hand side of the equation. Figure 1 gives the phenotypic means at each phased genotype in the upper panel, while its lower panel illustrates how the phased genotype is represented as the three orthogonal standardized coded genotypes.

1:

$$\mathbf{y} = \boldsymbol{\mu}_\mathbf{y} + \mathbf{Tb} + \mathbf{Z}_\alpha\boldsymbol{\beta}_\alpha + \mathbf{Z}_\delta\boldsymbol{\beta}_\Delta + \mathbf{Z}_\gamma\boldsymbol{\beta}_\gamma + \boldsymbol{\varepsilon}$$

The primary concern of this paper is with the partitioning of phenotypic variance components according to Equations 2 and 3, which present the problem in vector and individual-based forms, respectively.

2:

$$\begin{aligned}\mathrm{Var}(\mathbf{y}) \quad &= E\left(\left(\mathbf{y} - \boldsymbol{\mu}_\mathbf{y} - \mathbf{Tb}\right)\left(\mathbf{y} - \boldsymbol{\mu}_\mathbf{y} - \mathbf{Tb}\right)'\right)\\ &= m^{-1}\sigma_\alpha^2 \mathbf{Z}_\alpha{\mathbf{Z}_\alpha}' + m^{-1}\sigma_\delta^2 \mathbf{Z}_\delta{\mathbf{Z}_\delta}' + m^{-1}\sigma_\gamma^2 \mathbf{Z}_\gamma{\mathbf{Z}_\gamma}' + \sigma_\varepsilon^2 \mathbf{I}_n\end{aligned}$$

where the matrices $\mathbf{ZZ}'$ are $n \times n$ matrices giving the cross-products, across all loci, of individuals' coded genotypes. The elements of $\mathrm{Var}(\mathbf{y})$ are given by

3:

$$
\begin{aligned}
\{\mathrm{Var}(\mathbf{y})\}_{ii} \quad &= m^{-1}\left(\sum_{k}^{m}\sigma_{\alpha}^{2}z_{\alpha,\,ik}^{2} + \sigma_{\delta}^{2}z_{\delta,\,ik}^{2} + \sigma_{\gamma}^{2}z_{\gamma,\,ik}^{2}\right) + \sigma_{\varepsilon}^{2}\\
&= \sigma_{\alpha}^{2} + \sigma_{\delta}^{2} + \sigma_{\gamma}^{2} + \sigma_{\varepsilon}^{2}\\
\{\mathrm{Var}(\mathbf{y})\}_{ij} \quad &= m^{-1}\left(\sigma_{\alpha}^{2}\sum_{k}^{m}z_{\alpha,\,ik}z_{\alpha,\,jk} + \sigma_{\delta}^{2}\sum_{k}^{m}z_{\delta,\,ik}z_{\delta,\,jk} + \sigma_{\gamma}^{2}\sum_{k}^{m}z_{\gamma,\,ik}z_{\gamma,\,jk}\right)\\
&= \sigma_{\alpha}^{2}\mathbf{A}_{ij} + \sigma_{\delta}^{2}\boldsymbol{\Delta}_{ij} + \sigma_{\gamma}^{2}\boldsymbol{\Gamma}_{ij}\\
\mathrm{Var}(\mathbf{y}) \quad &= \sigma_{\alpha}^{2}\mathbf{A} + \sigma_{\delta}^{2}\boldsymbol{\Delta} + \sigma_{\gamma}^{2}\boldsymbol{\Gamma} + \sigma_{\varepsilon}^{2}\mathbf{I}_{n}
\end{aligned}
$$

The second equality in Equation 3 is because the mean squares of each individual's standardized genotype codings are expected to be 1 in the absence of inbreeding.

The bold-faced Greek characters on the final line of Equation 3 are used to denote the genetic relationship matrices (GRMs) for the three coded genotypes; the additive GRM is  , the dominance GRM $\boldsymbol{\Delta}$, and the POE-coded GRM is $\boldsymbol{\Gamma}$. These are are averages over all effective markers of the sums of squares and cross-products of individuals' standardized coded genotypes.

## Statistical Methods

### Assumptions

Using REML to estimate the variance components model given in Equation 2 requires that several assumptions (both statistical and genetic) be met in order for inference about the model to be legitimate. In addition to the normality assumptions given above, *G-REMLadp* fitting of the model in Equation 3 assumes Hardy-Weinberg Equilibrium (HWE) and accuracy of phasing. Departures from HWE (such as non-random mating and differential allele frequencies

in male and female parents (Falconer and Mackay 1996)) break the orthogonality of the POE-coded genotype with the additive-coded genotype and the dominance-coded genotype. Non-random mating could decrease the frequency of heterozygotes, while sex-specific allele frequencies cause the frequencies of the $a_{mo}A_{fa}$ and $A_{mo}a_{fa}$ heterozygotes to differ. Accurate phasing is required so that the different heterozygotes are correctly called; if they are not, $Z_\gamma$ is measured with error, diluting estimates of the $Z_\gamma$, $y$ association, hence downward-biasing estimates of $\sigma_\gamma^2$.

## Power analysis via Haseman-Elston regression approximation

The power of G-REML analysis can be approximated in a Haseman-Elston (HE) regression framework where each distinct pair of individuals in the sample is used as the unit of analysis (Elston et al. 2000; Chen 2014; Visscher et al. 2014). In this framework, the outcome variable is the centered cross-product of phenotypes for each pair of unrelated individuals in the analysis (denoted $Y_{ij}$ for individuals $i$ and $j$) and the predictor variables are the GRM entries under additive-coding ($A_{ij}$), dominance-coding ($\Delta_{ij}$), and POE-coding ($\Gamma_{ij}$) of the pair. The coefficients of the predictors in the associated ordinary least squares regression are estimates of the phenotypic variance due to each type of effect. This is similar to unweighted least squares estimation in covariance structure modelling (Browne 1982). Wald tests of significance for the variance components can be made using the assumption that the residuals in the regression are approximately normally distributed (Chen 2014).

The coded genotypes are orthogonal at each locus, meaning that the $\mathbf{A}$, $\mathbf{\Delta}$, and $\mathbf{\Gamma}$ matrices are expected to be uncorrelated with increasing sample size and number of loci (assuming random mating). For example,

4:

$$
\begin{aligned}
\{\mathbf{A\Delta'}\}_{iv} \;&= m^{-2}\mathbf{Z_\alpha}\mathbf{Z_\alpha'}\mathbf{Z_\delta}\mathbf{Z_\delta'} \\
&= m^{-2}\sum_{j=1}^{n}\left\{\left(\sum_{k=1}^{m}z_{\alpha,ik}z_{\alpha,jk}\right)\left(\sum_{w=1}^{m}z_{\delta,jw}z_{\delta,vw}\right)\right\} \\
&= m^{-2}\sum_{j}^{n}\sum_{k}^{m}\sum_{w}^{m}z_{\alpha,ik}\,z_{\alpha,jk}z_{\delta,jw}z_{\delta,vw} \\
&= m^{-2}\sum_{k}^{m}\sum_{w}^{m}z_{\alpha,ik}\,z_{\delta,vw}\left(\sum_{j}^{n}z_{\alpha,jk}z_{\delta,jw}\right) \\
&= m^{-2}\sum_{k}^{n}\sum_{w\neq k}^{n}z_{\alpha,ik}\,z_{\delta,vw}\left(\sum_{j}^{n}z_{\alpha,jk}z_{\delta,jw}\right) \text{ by orthogonality of codings at each locus} \\
&\approx m^{-2}\sum_{k}^{m}\sum_{w\neq k}^{m}z_{\alpha,ik}\,z_{\delta,vw}O_p(n^{-1}) \\
&\approx O_p(n^{-1})
\end{aligned}
$$

The sum over individuals $j$ (in parentheses starting in the fourth line of the equation) will have mean 0 and variance $n^{-1}$ over repeated sampling, which we denote using the $O_p$ notation to imply that its value is bounded in probability by Chebychev's inequality as $n$ increases (Bishop et al. 1975, ch 14); the equality is approximate because the sum over individuals is not strictly uncorrelated with $z_{\alpha,ik}z_{\delta,vw}$. To get this wide bound, we also assume that the product of the additive and dominance codings between individuals $i$ and $v$ averages to be 1 over loci; because the product of the additive and dominance codings for two loci both with minor allele frequency $p$ is $-\sqrt{2}(q/p)^{3/2}$, this assumption might be violated if many rare variants were used in the analysis. In the empirical analysis, the off-diagonal elements of the GRMs were uncorrelated $(r_{\alpha,\gamma}=,r_{\alpha,\delta}=,r_{\gamma,\delta}=)$; Supplementary Figure S2 illustrates this; while Supplementary Table S3 shows that the diagonal elements of the empirical GRMs were close to 1.

Equation 4 means that simple HE regressions of Y on each coded genotype's GRM will yield the same estimates (and associated Wald test statistics) as a multiple regression, and we can analyze the power of simple regressions but fit multiple components in practice. The Wald test statistic of a variance component, POE for example, is given by $\chi^2 \approx F = \hat{\beta}_\gamma^2 / \mathrm{Var}\left(\hat{\beta}_\gamma\right) = \hat{\sigma}_\gamma^4 / \mathrm{Var}\left(\hat{\sigma}_\gamma^2\right)$. By assuming that a given variance component (i.e. the HE slope) is nonzero, the power of this Wald test depends on the noncentrality parameter of the associated statistic. The derivation of the noncentrality parameter is parallel for each type of coded genotype, so we focus on HE regression of POEs, following the example given for additive effects by Visscher *et al.* (2014).

Starting with standardized outcomes and predictor (where *vech* is the operator which transforms a symmetric matrix to a column vector of its lower-diagonal elements (Henderson and Searle 1979)). Simplifying assumptions are required: 1) $Y$ is approximately normal (very unlikely for a cross-product phenotype) so that the Wald test statistic has an $F$ distribution with 1 degree of freedom in the numerator and $1/2\, n(n-1) - 2$ degrees of freedom in the denominator; and 2) the amount of variance in $Y$ explained by $W$ is so small and the sample size so large that the Wald $F$ test statistic is well approximated by a $\chi_1^2$ random variable.

We derive results for a set of $m$ independent loci. The numerator of the Wald test statistic is the square of the estimated regression coefficient $\hat{\beta}_\gamma^2 = \mathrm{Cov}\left(Y_{ij}, \Gamma_{ij}\right)^2 / \mathrm{Var}\left(\Gamma_{ij}\right)^2$. For individuals $i$ and $j$, the expected covariance between $Y$ and $\Gamma$ is

5:

$$
\begin{aligned}
\mathrm{Cov}\left(Y_{ij}, \Gamma_{ij}\right) &= E\left(\sigma_\alpha^2 A_{ij}\Gamma_{ij} + \sigma_\delta^2 \Delta_{ij}\Gamma_{ij} + \sigma_\gamma^2 \Gamma_{ij}^2 + \Gamma_{ij}\left(\beta_\alpha'\beta_\delta m^{-1}\sum_{k=1}^{m} z_{\alpha,ik}z_{\delta,jk} + \beta_\alpha'\beta_\gamma m^{-1}\sum_{k=1}^{m} z_{\alpha,ik}z_{\gamma,jk} \right.\right. \\
&= \sigma_\alpha^2 E\left(A_{ij}\Gamma_{ij}\right) + \sigma_\delta^2 E\left(\Delta_{ij}\Gamma_{ij}\right) + \sigma_\gamma^2 E\left(\Gamma_{ij}^2\right) \\
&= \sigma_\alpha^2 E\left(m^{-1}\sum_{k=1}^{m}\sum_{l=1}^{m} z_{\alpha,ik}z_{\alpha,jk}z_{\gamma,i,l}z_{\gamma,jl}\right) + \sigma_\delta^2 E\left(m^{-1}\sum_{k=1}^{m}\sum_{l=1}^{m} z_{\delta,ik}z_{\delta,jk}z_{\gamma,il}z_{\gamma,jl}\right) + \sigma_\gamma^2 E\left(\Gamma_{ij}^2\right) \\
&= \sigma_\alpha^2 E\left(m^{-1}\sum_{k\neq l} z_{\alpha,ik}z_{\alpha,jk}z_{\gamma,il}z_{\gamma,jl}\right) + \sigma_\delta^2 E\left(m^{-1}\sum_{k\neq l} z_{\delta,ik}z_{\delta,jk}z_{\gamma,il}z_{\gamma,jl}\right) + \sigma_\gamma^2 E\left(\Gamma_{ij}^2\right) \\
&= \sigma_\gamma^2 E\left(\Gamma_{ij}^2\right) = \sigma_\gamma^2 \mathrm{Var}\left(\Gamma_{ij}\right)
\end{aligned}
$$

, recalling that the coded genotypes are orthogonal within a locus. This derivation shows that the HE regression coefficient in the population is $\sigma_\gamma^2$, hence the numerator of the Wald test statistic is $\sigma_\gamma^4$.

The denominator of the Wald $F$ test statistic is the error variance per degree of freedom, divided by the variance of the predictor variable, i.e. $\mathrm{Var}(\epsilon)/\left(\mathrm{Var}\left(\Gamma_{ij} \times df\right)\right)$. Here we use the approximation that $\mathrm{Var}(\epsilon) \approx \mathrm{Var}(Y) = 1$. HE regression uses the distinct pairs of observations as the units of analysis, so the denominator degrees of freedom are $1/2\,n(n-1)$ when estimating the variance component and an intercept term. The variance of the POE GRM calculated at any single locus can be shown by direct calculation to be 1, which is also true for the additive effect and dominance effect GRMs (given assumptions). $\mathrm{Var}\left(\Gamma_{ij}\right)$ is interpretable as the variance of the average of these component GRMs over $m_e$ loci, hence is $m_e^{-1}$ in this simple model. A more accurate approach (for additive effects) is given in Appendix 1 of Goddard (Goddard 2009), in which the variance in relatedness is averaged over the number of effective loci given variation in pedigree as well as linkage disequilibrium. Visscher *et al.* give an empirical estimate of the number of effective loci for GRMs calculated using genome-wide common SNPs in the HapMap3 panel, hence they suggest using $\mathrm{Var}\left(A_{ij}\right) = 2 \times 10^{-5}$ (Visscher et al. 2014). In our model, the denominator of the Wald test statistic is $1/\left(m_e^{-1}(0.5n(n-1) - \right.$

2)) $\approx 2m_e/n^2$. This means the mean Wald test statistic is approximately $n^2\sigma_\gamma^4/2m_e$, which is referred to a $\chi_1^2$ distribution because the denominator degrees of freedom in the $F$-test are very large. This derivation is nearly identical to that used in (Visscher et al. 2014), so it is possible to use their online tool (http://cnsgenomics.com/shiny/gctaPower/) to determine the power to detect a variance component of a particular size.

For $m_e$ loci and sample size $n$, the noncentrality parameters of the $\chi_1^2$ test statistic are:

6:

$$
\begin{aligned}
\lambda_\alpha &= 0.5 m_e^{-1} n(n-1)\sigma_\alpha^4 \\
\lambda_\delta &= 0.5 m_e^{-1} n(n-1)\sigma_\delta^4 \\
\lambda_\alpha &= 0.5 m_e^{-1} n(n-1)\sigma_\gamma^4
\end{aligned}
$$

In our simulations, we used the true $m$, as loci were simulated without LD; in practice, we recommend the approximation $m_e^{-1} \approx 2 \times 10^{-5}$ (Visscher et al. 2014), although its application to dominance effects and POEs is based only on analogy with additive effects and the number of effective loci will differ depending on which SNP panel is used.

## Implementation

*G-REMLadp* requires a set of phased genotypes, each with parent-of-origin assignments. Mitochondrial DNA and X chromosome SNPs are excluded from analysis. In the empirical analysis, we used a *Perl* script to assign parent-of-origin to genotypes which had already been phased, as described below. Given a set of phased, parental-origin-assigned SNPs, we first recoded each genotype to the three-term orthogonal coding given in Table 1, stored this data in the software package *MACH*'s (Li et al. 2010) "dosage" format, and then called *GCTA* in order to generate GRMs for the three variance components. *GCTA*'s --make-grm and --make-grmd (Zhu et al. 2015) make the appropriate GRMs for the additive and dominance effects directly from the

*MACH*-formatted phased genotype. However, for POEs, we recoded each locus (in *R*) by subtracting the paternal minor allele indicator from the maternal indicator and writing this to an appropriately formatted .mldose.gz file, then generated an .mlinfo.gz file from the relevant sources for the SNP data set. We input the .mldose.gz and .mlinfo.gz files to *GCTA*'s MACH dosage function (--dosage-mach-gz) to generate the POE GRM. We then input the additive, dominance, and POE GRMs to *GCTA*, with the phenotype and covariate files, in order to fit the mixed model and estimate variance components. We chose to estimate the variance components without constraining them to be positive so that the null distribution of test statistics would not be a mixture (Visscher 2006). We also used the AI-REML algorithm instead of the Newton-Raphson or EM algorithms in order to fit the model quickly.

**Simulations**

We simulated data according to the *G-REMLadp* model in order to: 1) evaluate the predicted statistical power using the Haseman-Elston regression approximation; 2) test the bias and variance of the method in response to violation of its assumptions; and 3) assess computational requirements. In simulation studies, data were simulated according to Equation 1, and variance components were estimated using both HE regression and *GCTA* software. The goals were to estimate the bias and variance of the variance component estimators, the agreement between HE and *GCTA* estimates, and the empirical power and Type I error rates of the HE test.

The design factors were: ssample size $n = 1000, 2000, 4000, 5000, 7500, m = 500, 1000, 3000, 5000$ SNPs; $\sigma_\alpha^2/\text{Var}(y) = 0, 0.017, 0.033, 0.1$; $\sigma_\delta^2/\text{Var}(y) = 0, 0.017, 0.033, 0.1$; $\sigma_\gamma^2/\text{Var}(y) = 0, 0.017, 0.033, 0.1$; a violation of HWE— simulating maternal and paternal genotypes with different minor allele frequencies (either no

difference, or mothers' allele frequency greater than fathers' by 0.05) so that the POE-coded

genotype is no longer orthogonal to the other two codings (which remain mutually orthogonal);

and a second violation of HWE-simulating parental genotypes to be correlated at $r = 0$ or

$r = 0.25$, on average across all simulated loci, so that the diagonal entries of the GRMs are on

average less than 1 and results might be less stable numerically. The sample size and number of

loci used were constrained to be small due to computational requirements; computing GRMs at

genome scale ($n > 5000, m_e > 10^6$) and fitting models to them required 100% use of at least

16 processors for approximately an hour, which is feasible for single empirical analyses but not

in factorial simulations with thousands of replications. The SNPs were simulated to be

independent with minor allele frequencies uniformly distributed between 0.01 and 0.5. Because

there were relatively few of them, the simulated SNPs had much larger individual effects than

would be expected in empirical data. We chose to simulate violations of HWE to test the

robustness of *G-REMLadp* to violations of its assumptions, and as a way of widening the breadth

of the simulations. Realistically, most samples in which *G-REMLadp* could be applied will have

filtered SNPs for violations of HWE as part of routine QC (Laurie et al. 2010). A total of 7500

replications were simulated, but in the simulation results presented here, these factors were not

completely crossed.

Three measures were used to evaluate the method: 1) empirical relative bias, the average

difference between the simulated variance component and the value estimated in a replication,

given as a percentage of the true value of the simulated variance component (except for null

variance components, where we give the absolute bias); 2) empirical sampling variance, the

variance of estimated variance components over replications; and 3) power/Type I error rate, the

proportion of replications in which the Wald test of a null variance component exceeded the 95% critical value under a simulated non-null/null effect.

We expected power to increase with increasing samples size and variance component size and decreasing number of loci, and expected empirical variance to decrease according to the same pattern. We expected that violation of assumptions would lead to detectable bias.

**Empirical Analysis**

We applied *G-REMLadp* to a sample of up to 4689 individuals gathered as part of the Avon Longitudinal Study of Parents and Children (ALSPAC), a prospective study of health and development beginning at pregnancy. We considered 38 different phenotypes which had been previously associated with POEs or which were related to body size, development, or social functioning.

See the Supplementary Material for further description of the sample, including the quality control procedures that we applied to it. The sample, the longitudinal study and its context, as well as its genotyping, are described in detail in two papers (Boyd et al. 2012; Fraser et al. 2013). The 38 phenotypes, along with estimates of additive effects, dominance effects, and POEs, are listed in Supplementary Table S1. All phenotypes were inverse-normal transformed prior to estimating variance components (Peng et al. 2007); fixed effects of sex and the first four ancestry-informative principal components were modelled. Parental effects and POEs have been previously studied in trios in this sample (Smith et al. 2007), but not genome-wide for the 38 phenotypes.

# Results

## Simulation Results

### Power and sample size

Sample size curves at a Type I error rate of 5%, based on the HE approximation, are given in Figure 2, for POEs responsible for 1%, 3%, 10%, and 15% of the total phenotypic variance explained, tagged by 100000 loci. Figure 2 shows that a sample size of over 10000 genotyped duos or trios, with probands having phased genotypes, is likely necessary to detect the largest conceivable parent-of-origin effect variances and that a sample of 50000 individuals will be needed to detect POEs accounting for $\approx 1 - 3\%$ of phenotypic variance, which was the size observed by Lopes *et al.* (2015). Figure 3 gives the empirical power, which was high because such large effects were simulated and because $m_e$ was low, with a median value of 2000 effective loci.

The test involving HE regression is based on further simplifying and adding assumptions to the *G-REMLadp* procedure, hence discrepancies between the two approaches were expected, with the concern that large discrepancies would render the approximate power calculations unhelpful. Table S2, in the Supplementary Material, shows that the simulated HE-based test statistics lay relatively close to their expected values. Figures S1-S3 illustrate this graphically.

### Type I error rates

There was no evidence that the tests of the POEs had inflated Type I error rates when assumptions were met; under these conditions, the simulated type I error rate was 0.0510. The

HE-based test of the POE effect had slightly elevated Type I error rates (0.0625) under violation of HWE due to dependence of parental genotypes.

## Bias and variance due to violating assumptions

We observed that violating HWE increased the discrepancy between predicted and observed test statistics. However, these results do not indicate whether this was due to increased bias or variance of the estimates, or both. Additionally, because the predicted test statistics were based on many simplifying assumptions, it is possible that the increased discrepancy would not have led to incorrect inference, and it is worth exploring whether violating HWE causes the estimated variance components to be misleading in predictable ways.

Results for simulations with truly null variance components are not shown. The absolute bias was under $3 \times 10^{-3}$ for each type of effect whether or not assumptions were violated. Variances were similar to those for non-null estimates.

Table 2 shows the GCTA results under HWE. Relative bias was under 5% for additive, dominance, and parent-of-origin effects. Results were similar for HE model fitting.

Table 3 shows the GCTA results when the HWE assumptions were violated. Relative bias was under 10% for the additive and dominance effects but, when parental genotypes were correlated was downward biased by about 20% for the parent-of-origin effects. Results were similar for HE model fitting. Biases in estimating dominance effects were smaller but were uniformly positive.

The *GCTA* and HE estimates tended to differ in the presence of correlation between parental genotypes; in this situation, HE regression estimates of POEs had higher variance (and hence higher mean-squared error) than did *GCTA* estimates, which is expected to be the case

even in additive-only models (Chen 2014). As a result, the correspondence between the two methods decreased from a HE-*GCTA* correlation of 0.9911(0.9908,0.9915) with independent parental genotypes to 0.9624(0.9597, 0.9649) with correlated parents. This was not observed for estimates of the two other types of effects.

## Empirical Analysis

Estimates of $\sigma_\alpha^2$, $\sigma_\delta^2$, and $\sigma_\gamma^2$ are given in Figure 4 for the 38 phenotypes. Error bars showing $\pm 1.96$ standard errors are given for each estimate. Based on this, the sample size of $n \approx 5000$ was sufficient to detect established additive effects for anthropometric variables such as height and fat mass. The mean estimate for additive effects was 0.297(0.10). There is evidence for dominance effects on verbal IQ and fat mass in the sample. The mean estimate for dominance effects was 0.031(0.14). The largest (and most reliably estimated) POEs were estimated for age at menarche, FVC, and age at first tooth. The mean estimate for POEs was 0.023(0.10).

Across the 38 phenotypes surveyed, which are listed in Table S1, in the aggregate, additive effects were frequently estimated reliably and away from 0. POEs had standard errors approximately equal to those of additive effects, while estimates of dominance variance showed slightly larger standard errors. This is illustrated using "dodged" (Wilkinson 2006) histograms of variance component estimates in Figure 5, which is broadly similar to that given (for additive and dominance effects) in Figure 2 of Zhu *et al.* 2015 (Zhu et al. 2015).

Two patterns in variance partitions emerged (although the sample size is too small to perform significance tests of differences in heritability estimates across traits). Age at menarche, age at first tooth, CRP, handedness, IL-6, and 2d4d ratio in the right hand showed a pattern of relatively equal additive, dominance, and parent-of-origin effect variance estimates which were

relatively far from 0 (though not usually greater than 2 standard errors from 0). A second pattern was seen in the IQ measures and fat mass, where additive and dominance effect estimates were approximately equal and nonzero. In the remaining 30 phenotypes, if an effect was detectable, it was an additive effect.

# Discussion

The most important findings from our simulations are that: 1) *G-REMLadp* does not seem to be inherently biased in estimating variance due to additive effects, dominance effects, and POEs, and 2) that substantial correlation between parental genotypes is necessary to bias *G-REMLadp* estimates. We did not investigate the effect that linkage disequilibrium might have on our results; if local levels of linkage disequilibrium are associated with POE size, an LDAK implementation of *G-REMLadp* may help address this (Speed et al. 2012). The empirical and simulation results agree that a sample size under 10000 is insufficient to generate precise *G-REMLadp* estimates. Hence, POEs of the size observed by Lopes *et al.* (Lopes et al. 2015) ($\approx$ 2% of variance explained) are likely to require sample sizes close to 50000 in order to resolve properly (their samples were about 4500 purebred pigs, hence $m_e$ far below that for humans on the HapMap 3 panel, accordingly their pattern of results is closer to our simulated findings than to our empirical ones). Nevertheless, it may be interesting following up the ALSPAC results for age at menarche (a phenotype known to be influenced by imprinted loci (Perry et al. 2014)), age at first tooth, and 2d4d ratio results in a larger study with trios. Replicating the dominance results for fat mass would also be of interest, as Zhu *et al.* (2015) found no significant dominance for skinfold thickness; similarly, replicating a dominance heritability component of verbal IQ would

substantiate venerable claims for non-additive effects on cognitive performance (Devlin et al. 1997; Plomin and Deary 2015).

However, the highest-profile claim (Devlin et al. 1997) for non-additive effects on IQ involves maternal effects. The evidence for this claim comes from findings of increased similarity in IQ in the children of female twins relative to the children of male twins (Nance and Corey 1976). Maternal (or paternal) effects are defined as indirect effects of the parent's genotype on offspring phenotype, usually via parental influence on offspring's environment (Wolf and Wade 2009). These are distinct from the POEs discussed here, because POEs are direct effects of the offspring's genotype on its phenotype. In the presence of maternal (paternal) effects, the mother's (father's) genotype explains some of the residual phenotypic variation remaining after regressing the offspring's phenotype on its genotype. However, methods designed to detect POEs by modelling different distributions of the two heterozygotes' phenotypes (including *G-REMLadp*) cannot distinguish POEs from certain patterns of maternal/paternal effects (Hager et al. 2008). It is also possible that there are maternal effects that cause imprinting (epigenetic changes due to intrauterine environment or family environment created by parental behavior). In order to distinguish POEs from maternal effects, one can fit a model with both maternal and POEs or restrict the analysis to offspring of heterozygous mothers. Given the sample size considerations in our results, the former approach is likely to be more successful; *G-REMLadp* could be implemented in OpenMx and used as an addition to an M-GCTA model (Eaves et al. 2014; Kirkpatrick and Neale 2015; Neale et al. 2015). However, POEs and maternal/paternal effects represent departure from Mendelian inheritance; a reliable nonzero $\sigma_\gamma^2$ value detected by *G-REMLadp* is worthy of followup even if it represents a mixed bag of POEs and indirect effects.

Other than the insufficient sample size in the ALSPAC analyses, an additional limitation of this project was that the simulations and empirical data were not closely matched; less idealized, more informative simulations could have been performed. For example, simulated phenotypes could've been based on the empirical distributions of the ALSPAC phenotypes, and these, coupled with real GRMs (with imperfect HWE because of sampling error, if not actual violations), could've resulted in more thorough evaluation of *G-REMLadp*. Pairwise relatedness values could also have been simulated from the empirical distributions of ($A_{ij}, \Delta_{ij}$, and $\Gamma_{ij}$) or smooth approximations to them (Lee and Chow 2014). This approach is more feasible computationally under the HE regression framework than under maximum likelihood. *G-REMLadp* is not limited to genome-wide analyses; future work could also make use of relatedness at known imprinted loci, though this would require sufficiently large samples to account for the lower degree of tagging of phenotypic variation. A 2015 paper by Baran *et al.,* in *Genome Research*, supplies an atlas of such locations in the genome which affect expression (Baran et al. 2015). A related future direction (as suggested by Lopes *et al.*) would be to emphasize the connection between imprinting and attempts to detect POEs, for example by applying the method in the EWAS, rather than GWAS, context, i.e. identifying genome-wide eQTL effects on age at menarche and first tooth, and 2d4d ratio (Zhu et al. 2016).

In summary, *G-REMLadp* offers the ability to easily partition phenotypic variance according to three types of inherited effects. With large studies of parent-child trios/duos, it is possible to fit *G-REMLadp* models and detect effects without bias. Finally, failure to model POEs has been identified as one possible source of missing heritability (Kong et al. 2009; Eichler et al. 2010). For any phenotypes with missing heritability that is uncovered by modelling POEs,

the epigenetic and evolutionary implications of POEs lead to hypotheses of distinctive etiologies

and genomic architectures.

# Acknowledgements

# References

Álvarez-Castro JM (2015) Dissecting genetic effects with imprinting. Models and Estimation of Genetic Effects 35.

Baran Y, Subramaniam M, Biton A et al (2015) The landscape of genomic imprinting across diverse adult human tissues. Genome research 25:927–936.

Bishop Y, Feinberg S, Holland P (1975) Discrete muhwariate analysis: Theory and practice. MIT Press, Cambridge. MA, USA

Boyd A, Golding J, Macleod J et al (2012) Cohort profile: The 'children of the 90s'-the index offspring of the avon longitudinal study of parents and children. International journal of epidemiology dys064.

Browne MW (1982) Covariance structures. Topics in applied multivariate analysis 72–141.

Chen G-B (2014) Estimating heritability of complex traits from genome-wide association studies using ibs-based haseman–Elston regression. Frontiers in genetics 5:107.

Clemons T (2000) A look at the inheritance of height using regression toward the mean. Human biology 447–454.

Devlin B, Daniels M, Roeder K (1997) The heritability of iq. Nature 388:468–471.

Eaves LJ, Pourcain BS, Smith GD et al (2014) Resolving the effects of maternal and offspring genotype on dyadic outcomes in genome wide complex trait analysis (m-gcta). Behavior genetics 44:445–455.

Eichler EE, Flint J, Gibson G et al (2010) Missing heritability and strategies for finding the underlying causes of complex disease. Nature Reviews Genetics 11:446–450.

Elston RC, Buxbaum S, Jacobs KB, Olson JM (2000) Haseman and elston revisited. Genetic epidemiology 19:1–17.

Falconer DS, Mackay TF (1996) Introduction to quantitative genetics, 4th edn. Pearson Education, Edinburgh Gate, Harlow, Essex CM20 2JE, England

Fraser A, Macdonald-Wallis C, Tilling K et al (2013) Cohort profile: The avon longitudinal study of parents and children: ALSPAC mothers cohort. International journal of epidemiology 42:97–110.

Goddard M (2009) Genomic selection: Prediction of accuracy and maximisation of long term response. Genetica 136:245–257.

Guilmatre A, Sharp A (2012) Parent of origin effects. Clinical genetics 81:201–209.

Hager R, Cheverud JM, Wolf JB (2008) Maternal effects as the cause of parent-of-origin effects that mimic genomic imprinting. Genetics 178:1755–1762.

Hall JG (1990) Genomic imprinting: Review and relevance to human diseases. American journal of human genetics 46:857.

Henderson HV, Searle SR (1979) Vec and vech operators for matrices, with some uses in jacobians and multivariate statistics. The Canadian Journal of Statistics / La Revue Canadienne de Statistique 7:65–81.

Hoggart CJ, Venturini G, Mangino M et al (2014) Novel approach identifies snps in slc2a10 and kcnk9 with evidence for parent-of-origin effect on body mass index. PLoS Genet 10:e1004508.

Kirkpatrick R, Neale M (2015) Combining genomic-relatedness-matrix restricted maximum likelihood (greml) with structural equation modeling in openmx. In: BEHAVIOR genetics. SPRINGER 233 SPRING ST, NEW YORK, NY 10013 USA, pp 665–665

Kong A, Steinthorsdottir V, Masson G et al (2009) Parental origin of sequence variants associated with complex diseases. Nature 462:868–874.

Laurie CC, Doheny KF, Mirel DB et al (2010) Quality control and quality assurance in genotypic data for genome-wide association studies. Genetic epidemiology 34:591–602.

Lawson HA, Cheverud JM, Wolf JB (2013) Genomic imprinting and parent-of-origin effects on complex traits. Nature reviews Genetics 14:609–617.

Lee JJ, Chow CC (2014) Conditions for the validity of snp-based heritability estimation. Human genetics 133:1011–1022.

Li Y, Willer CJ, Ding J et al (2010) MaCH: Using sequence and genotype data to estimate haplotypes and unobserved genotypes. Genetic epidemiology 34:816–834.

Lopes MS, Bastiaansen JW, Janss L et al (2015) Estimation of additive, dominance, and imprinting genetic variance using genomic data. G3: Genes| Genomes| Genetics 5:2629–2637.

Nance WE, Corey LA (1976) Genetic models for the analysis of data from the families of identical twins. Genetics 83:811–826.

Neale MC, Hunter MD, Pritikin JN et al (2015) OpenMx 2.0: Extended structural equation and statistical modeling. Psychometrika 1–15.

Neugebauer N, Räder I, Schild H et al (2010) Evidence for parent-of-origin effects on genetic variability of beef traits. Journal of animal science 88:523–532.

Nishio M, Satoh M (2015) Genomic best linear unbiased prediction method including imprinting effects for genomic evaluation. Genetics Selection Evolution 47:32.

Patten M, Ross L, Curley J et al (2014) The evolution of genomic imprinting: Theories, predictions and empirical tests. Heredity 113:119–128.

Peng B, Robert KY, DeHoff KL, Amos CI (2007) Normalizing a large number of quantitative traits using empirical normal quantile transformation. In: BMC proceedings. BioMed Central, p S156

Perry JR, Day F, Elks CE et al (2014) Parent-of-origin-specific allelic associations among 106 genomic loci for age at menarche. Nature 514:92–97.

Peters J (2014) The role of genomic imprinting in biology and disease: An expanding view. Nature Reviews Genetics 15:517–530.

Plomin R, Deary IJ (2015) Genetics and intelligence differences: Five special findings. Molecular psychiatry 20:98–108.

Schaeffer L, Kennedy B, Gibson J (1989) The inverse of the gametic relationship matrix. Journal of Dairy Science 72:1266–1272.

Smith GD, Steer C, Leary S, Ness A (2007) Is there an intrauterine influence on obesity? Evidence from parent–child associations in the avon longitudinal study of parents and children (alspac). Archives of disease in childhood 92:876–880.

Speed D, Hemani G, Johnson MR, Balding DJ (2012) Improved heritability estimation from genome-wide snps. The American Journal of Human Genetics 91:1011–1021.

Spencer HG (2002) The correlation between relatives on the supposition of genomic imprinting. Genetics 161:411–417.

Spencer HG (2009) Effects of genomic imprinting on quantitative traits. Genetica 136:285–293.

Visscher PM (2006) A note on the asymptotic distribution of likelihood ratio tests to test variance components. Twin research and human genetics 9:490–495.

Visscher PM, Hemani G, Vinkhuyzen AA et al (2014) Statistical power to detect genetic (co) variance of complex traits using snp data in unrelated samples. PLoS Genet 10:e1004269.

Wallace C, Smyth DJ, Maisuria-Armer M et al (2010) The imprinted dlk1-meg3 gene region on chromosome 14q32. 2 alters susceptibility to type 1 diabetes. Nature genetics 42:68–71.

Wilkinson L (2006) The grammar of graphics. Springer Science & Business Media

Wolf JB, Wade MJ (2009) What are maternal effects (and what are they not)? Philosophical Transactions of the Royal Society of London B: Biological Sciences 364:1107–1115.

Zhu Z, Bakshi A, Vinkhuyzen AA et al (2015) Dominance genetic variation contributes little to the missing heritability for human complex traits. The American Journal of Human Genetics 96:377–385.

Zhu Z, Zhang F, Hu H et al (2016) Integration of summary data from gwas and eQTL studies predicts complex trait gene targets.

# Figure 1

**Phenotypic mean by phased genotype (top) Phased genotype represented by 3 standardised codings (bottom)**: (TOP)The dark line connects phenotypic means at different phased genotypes under a model including dominance and parent-of-origin effects in addition to additive effects. The dashed line connects phenotypic means under a model with additive and dominance effects only; the dominance deviation makes it so that the slope between the major-homozygote and the heterozygotes is steeper than that between the heterozygotes and the minor-homozygote (See Supplementary Figure S1). The difference due to parent-of-origin effects is illustrated by the dotted line marked $\gamma$. $a$: the major allele; $A$: minor (reference) allele; $a_{mo}$: allele from mother; $a_{fa}$: allele from father; $\alpha^*$: additive allele-substitution effect; $\delta$: dominance deviation; $\alpha = \alpha^* - (1 - 2p)\delta$, half the difference between homozygotes' phenotypic means; $\gamma$: the difference between the phenotypic mean of all heterozygotes and the phenotypic mean of paternal-minor heterozygotes. (BOTTOM) For a single locus with minor allele frequency $p$, the standardised additive coding (0,1,2 standardised to $-\sqrt{2p/q}$, $(q - p)/\sqrt{2pq}$, $\sqrt{2p/q}$) of the phased genotype is given by the medium-weight line. The standardised dominance coding is the dashed line (0,$2p$,$4p - 2$ standardised to $-p/q$, $1$, $-q/p$). The parent-of-origin effect coding is the thick line (0,-1,1,0 standardised to $0$, $-1/\sqrt{2pq}$, $1/\sqrt{2pq}$, $0$).

Phenotype $Y$

$2\alpha$

$\alpha + \delta + \gamma$

$\alpha + \delta$

$\alpha + \delta - \gamma$

0

$\gamma$

Model: $Y = X_a \alpha^* + X_d \delta + X_p \gamma$

Coding $X_*$

$\sqrt{2q/p}$

$1/\sqrt{2pq}$

1

$(q - p)/\sqrt{2pq}$

0

$-p/q$

$-\sqrt{2p/q}$

$-1/\sqrt{2pq}$

$-q/p$

Additive

POE

Dominance

$a_{mo}a_{fa}$    $a_{mo}A_{fa}$    $A_{mo}a_{fa}$    $A_{mo}A_{fa}$    Genotype
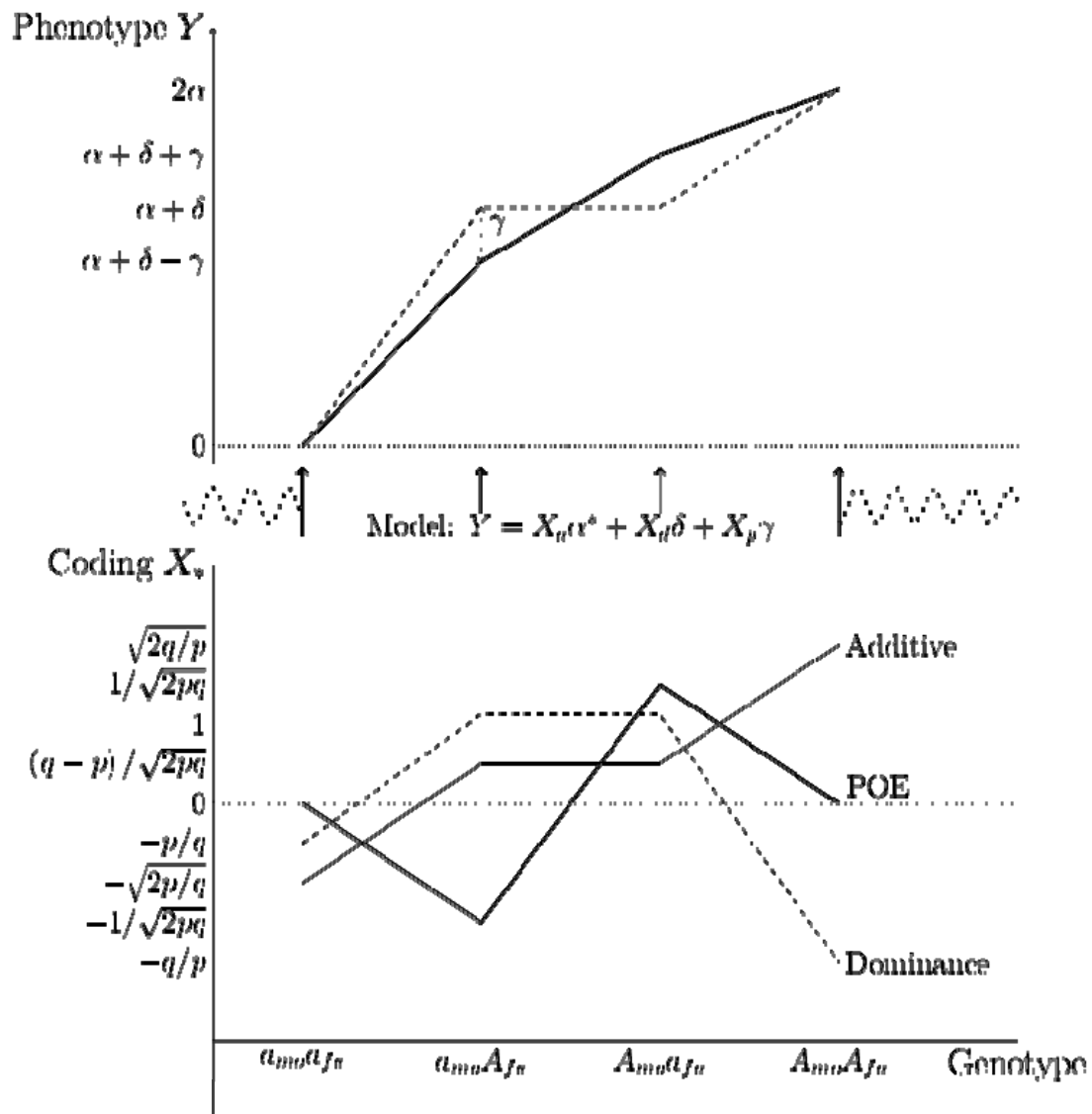
# Figure 2

**Expected power to detect parent-of-origin effects tagged by $m = 10000$ loci, by sample size**: Curves generated using Haseman-Elston regression approximation to the Wald test of nonzero effect. %PoE: proportion of phentoypic variance attributable to parent-of-origin effects; Sample size: number of unrelated individuals with phased genotypes in the sample that would be used to estimate POEs; Power: proportion of test statistics generated under a non-null distribution with noncentrality parameter as given in Equation 4 that would exceed a 95% critical value in the Wald test of a null variance component.
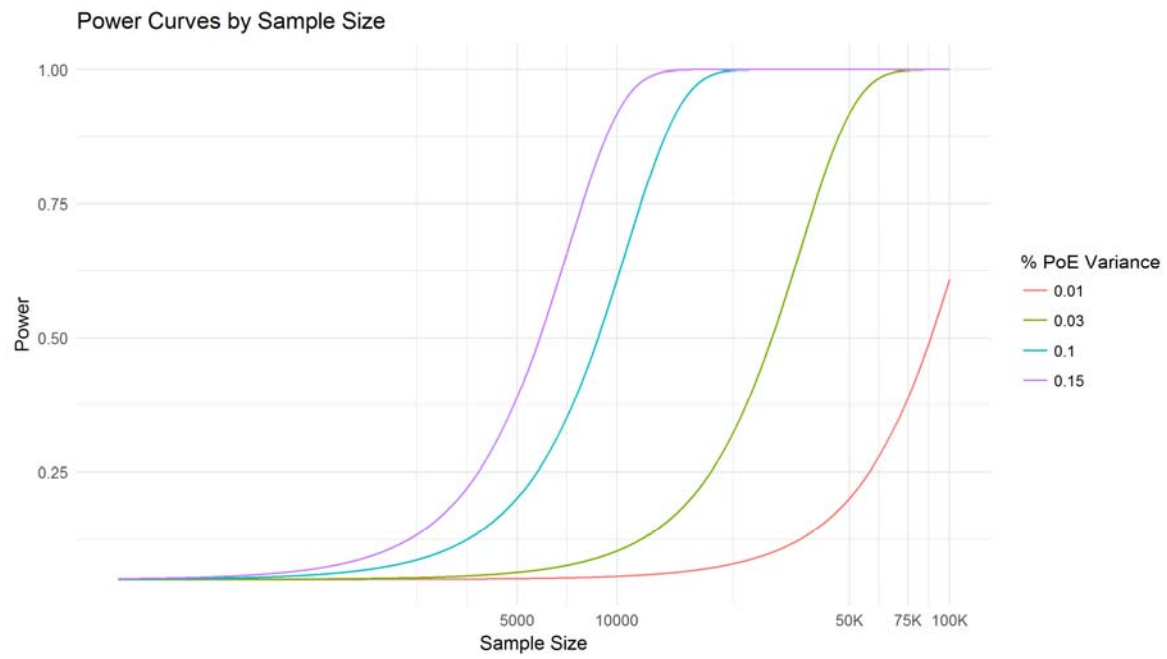
# Figure 3

**Empirical power curves in simulated data**: Curves generated using proportions of significant Wald tests in simulated data with all assumptions specified in the text satisfied. %PoE: proportion of phentoypic variance attributable to parent-of-origin effects in the model used to generate simulated data; Sample size: number of simulated individuals with parent-of-origin determined at all loci; Power: proportion of replicates with Wald tests having $p$-values $< 0.05$. Note that the number of effective loci in this figure is an order of magnitude lower than in Figure 1.

Empirical power in Simulated Data
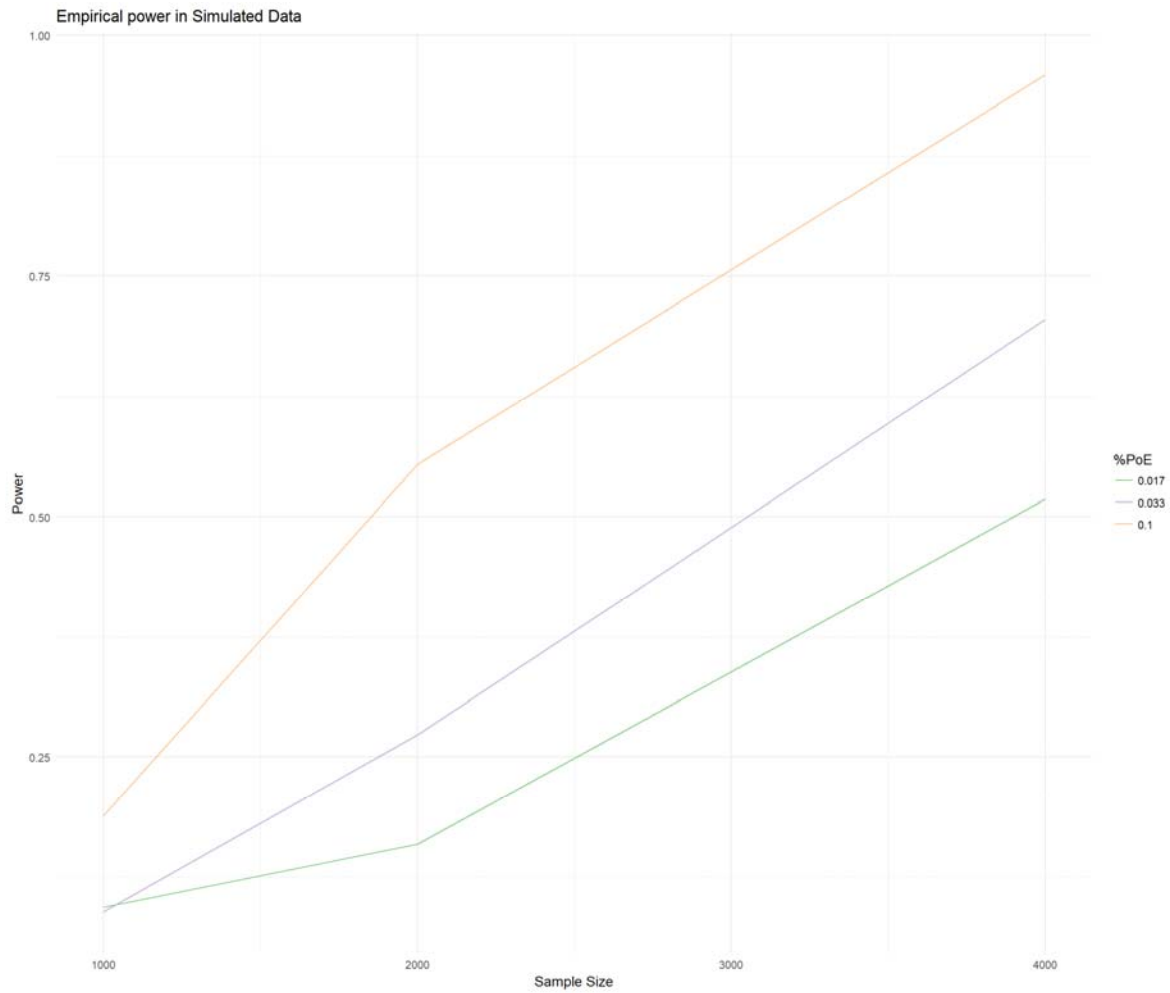
# Figure 4

**Forest plot of variance component estimates of 19 phenotypes in ALSPAC**: The $x$-axis, labelled "Estimate," is the proportion of phenotypic variance attributable to each effect type (PoE: parent-of-origin, squares; Dom: dominance, triangles; Add: additive, circles). To avoid bias, variance estimates were not constrained, hence negative variance component estimates are possible, for example if too many components are fit. Estimates are marked by the shapes, lying in 95% confidence intervals. A vertical line is plotted at 0.

# Figure 5

**Histograms of the estimates of additive, dominance, and parent-of-origin variance components for 38 ALSPAC traits.** Add: Estimated proportions of phenotypic variance attributable to additive effects; Dom: Estimated proportions of phenotypic variance attributable to dominance effects; PoE: Estimated proportions of phenotypic variance attributable to parent-of-origin effects. To avoid biased variance estimates, they were not constrained to be positive; hence, for POEs, which the study was underpowered to detect, there are many negative estimates.
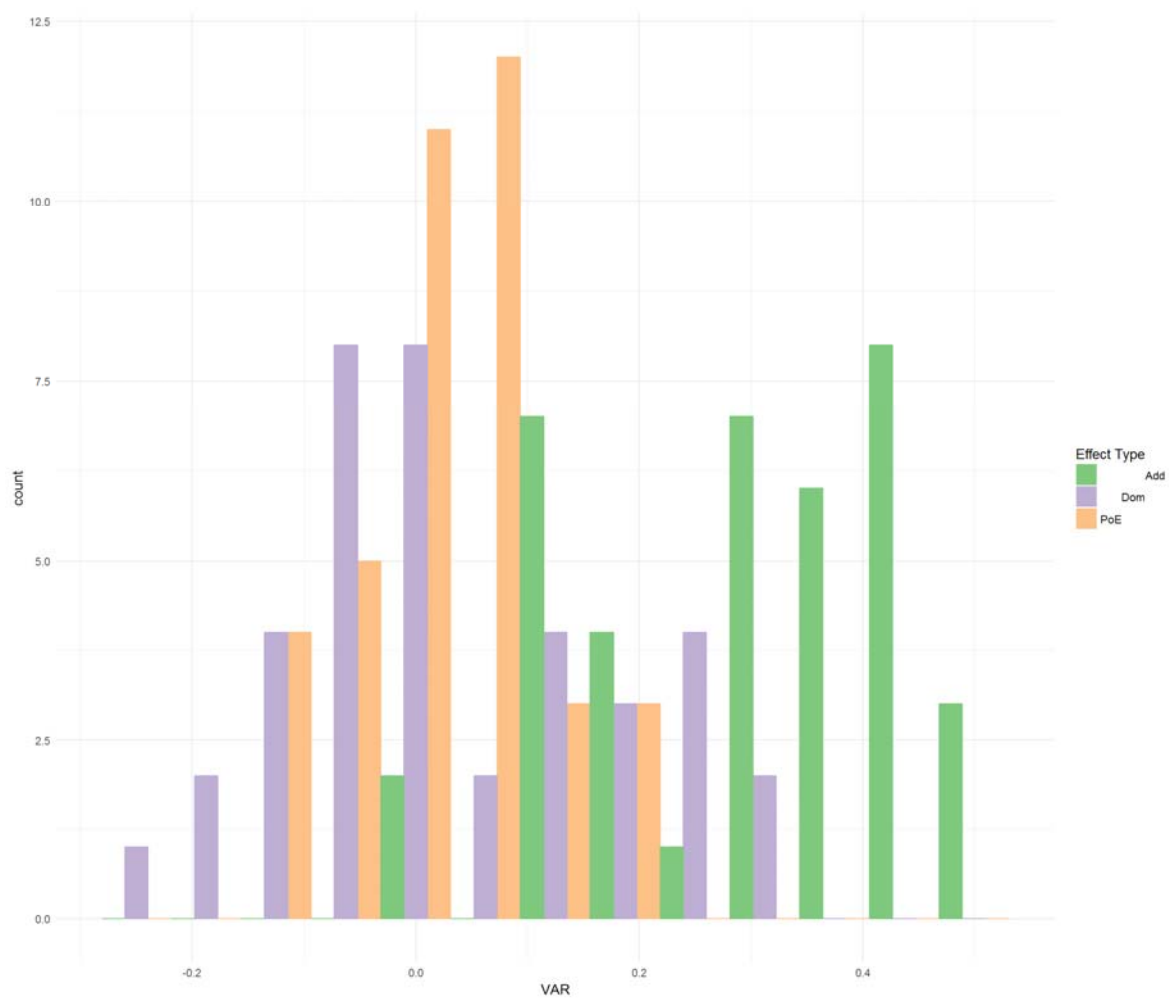
# Figure 5

## Table I: Recoding a phased genotype using three orthogonal terms.

| Phased Genotype | Freq | Add Code | Std Add | Dom Code | Std Dom | POE Code | Std POE |
|---|---|---|---|---|---|---|---|
| $a_{mo}a_{fa}$ | $q^2$ | 0 | $-\sqrt{2p/q}$ | 0 | $-p/q$ | 0 | 0 |
| $a_{mo}A_{fa}$ | $pq$ | 1 | $(q-p)/\sqrt{2pq}$ | $2p$ | 1 | -1 | $-1/\sqrt{2pq}$ |
| $A_{mo}a_{fa}$ | $pq$ | 1 | $(q-p)/\sqrt{2pq}$ | $2p$ | 1 | 1 | $1/\sqrt{2pq}$ |
| $A_{mo}A_{fa}$ | $p^2$ | 2 | $\sqrt{2q/p}$ | $4p-2$ | $-q/p$ | 0 | 0 |

*The minor allele is A, with frequency $p$, major allele $a$ with frequency $q = 1-p$. $A_{mo}$: maternally inherited A allele; $A_{fa}$: paternally inherited A allele. "Freq": expected frequency of the genotype under Hardy-Weinberg equilibrium. "Add Code": additive coding of the phased genotype. "Std Add": standardized additive coding based on mean $2p$ and variance $2pq$. "Dom Code": dominance coding of the phased genotype. "Std Dom": standardized dominance coding based on mean $2p^2$ and variance $4p^2q^2$. "POE Code": parent-of-origin effect coding of the phased genotype. "Std POE": standardized parent-of-origin effect coding based on mean $0$ and variance $2pq$.*

## Table II: Relative bias and variance for G-REMLadp variance components when HWE was met, for non-null simulated effects

| N | $\sigma^2_\alpha$ | $\sigma^2_\delta$ | $\sigma^2_\gamma$ | #Reps | $\hat{\sigma}^2_\alpha$ Bias | $\hat{\sigma}^2_\alpha$ Var | $\hat{\sigma}^2_\delta$ Bias | $\hat{\sigma}^2_\delta$ Var | $\hat{\sigma}^2_\gamma$ Bias | $\hat{\sigma}^2_\gamma$ Var |
|---|---|---|---|---|---|---|---|---|---|---|
| 1000 | 0.017 | 0.017 | 0.017 | 280 | 6.000 | 1.1e-03 | 9.200 | 8.9e-04 | 7.5 | 1.1e-03 |
| 1000 | 0.033 | 0.033 | 0.033 | 290 | -0.330 | 1.9e-03 | -5.200 | 1.8e-03 | -5.9 | 1.8e-03 |
| 1000 | 0.100 | 0.100 | 0.100 | 290 | 3.000 | 7.1e-03 | 0.900 | 5.7e-03 | 2.6 | 6.3e-03 |
| 2000 | 0.017 | 0.017 | 0.017 | 300 | -13.000 | 2.9e-04 | -0.390 | 2.6e-04 | 2.5 | 2.6e-04 |
| 2000 | 0.033 | 0.033 | 0.033 | 300 | -0.810 | 5.7e-04 | -5.700 | 4.9e-04 | 3.7 | 6.3e-04 |
| 2000 | 0.100 | 0.100 | 0.100 | 300 | 3.700 | 1.6e-03 | 1.600 | 1.7e-03 | -1.1 | 1.6e-03 |
| 4000 | 0.017 | 0.017 | 0.017 | 300 | 3.000 | 6.5e-05 | 0.074 | 6.8e-05 | -4.8 | 7.5e-05 |
| 4000 | 0.033 | 0.033 | 0.033 | 300 | 0.088 | 1.7e-04 | -1.900 | 1.5e-04 | -2.4 | 1.4e-04 |
| 4000 | 0.100 | 0.100 | 0.100 | 300 | -2.600 | 4.8e-04 | -1.800 | 4.4e-04 | 1.5 | 4.5e-04 |

*Proportion of phenotypic variance explained: by additive effects given in the $\sigma^2_\alpha$ column, by dominance effects given in the $\sigma^2_\delta$ column, and by parent-of-origin effects given in the $\sigma^2_\gamma$ column. N: number of simulated phased genotypes, #Reps: number of simulated replications, "Bias" is percent relative bias of estimates across simulated replications, "Var" is variance of estimates across simulated replications.*

# Table III: Relative bias and variance for G-REMLadp variance components when HWE was violated, for non-null simulated effects.

| Parent Corr | MAF Diff | N | $\sigma_\alpha^2$ | $\sigma_\delta^2$ | $\sigma_\gamma^2$ | #Reps | $\hat{\sigma}_\alpha^2$ Bias | $\hat{\sigma}_\alpha^2$ Var | $\hat{\sigma}_\delta^2$ Bias | $\hat{\sigma}_\delta^2$ Var | $\hat{\sigma}_\gamma^2$ Bias | $\hat{\sigma}_\gamma^2$ Var |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 0.00 | 0.05 | 1000 | 0.017 | 0.017 | 0.017 | 300 | 10.00 | 1.0e-03 | -2.90 | 9.7e-04 | 17.00 | 1.0e-03 |
| 0.00 | 0.05 | 1000 | 0.033 | 0.033 | 0.033 | 300 | 3.20 | 2.4e-03 | 3.20 | 2.4e-03 | 2.00 | 2.0e-03 |
| 0.00 | 0.05 | 1000 | 0.100 | 0.100 | 0.100 | 300 | -1.00 | 6.6e-03 | 0.29 | 6.4e-03 | 0.31 | 5.7e-03 |
| 0.00 | 0.05 | 2000 | 0.017 | 0.017 | 0.017 | 300 | -8.40 | 2.5e-04 | 2.10 | 3.1e-04 | -0.11 | 2.8e-04 |
| 0.00 | 0.05 | 2000 | 0.033 | 0.033 | 0.033 | 300 | 2.80 | 5.4e-04 | 4.70 | 5.2e-04 | 0.90 | 4.8e-04 |
| 0.00 | 0.05 | 2000 | 0.100 | 0.100 | 0.100 | 300 | 2.20 | 1.6e-03 | -0.52 | 1.6e-03 | 0.69 | 1.6e-03 |
| 0.00 | 0.05 | 4000 | 0.017 | 0.017 | 0.017 | 300 | 1.10 | 7.2e-05 | -4.00 | 6.9e-05 | -5.60 | 8.1e-05 |
| 0.00 | 0.05 | 4000 | 0.033 | 0.033 | 0.033 | 300 | -1.90 | 1.5e-04 | 2.50 | 1.8e-04 | -2.50 | 1.5e-04 |
| 0.00 | 0.05 | 4000 | 0.100 | 0.100 | 0.100 | 290 | -2.10 | 4.9e-04 | 1.70 | 5.6e-04 | -0.20 | 4.3e-04 |
| 0.25 | 0.00 | 1000 | 0.017 | 0.017 | 0.017 | 300 | 3.60 | 8.2e-04 | 9.10 | 1.6e-04 | -8.40 | 1.1e-03 |
| 0.25 | 0.00 | 1000 | 0.033 | 0.033 | 0.033 | 290 | 5.70 | 1.6e-03 | 5.80 | 2.5e-04 | -20.00 | 2.0e-03 |
| 0.25 | 0.00 | 1000 | 0.100 | 0.100 | 0.100 | 290 | 1.70 | 5.6e-03 | 2.10 | 9.8e-04 | -25.00 | 8.3e-03 |
| 0.25 | 0.00 | 2000 | 0.017 | 0.017 | 0.017 | 300 | -0.80 | 2.0e-04 | -2.10 | 5.6e-05 | -33.00 | 2.8e-04 |
| 0.25 | 0.00 | 2000 | 0.033 | 0.033 | 0.033 | 300 | 6.00 | 3.9e-04 | -0.27 | 1.0e-04 | -21.00 | 6.5e-04 |
| 0.25 | 0.00 | 2000 | 0.100 | 0.100 | 0.100 | 300 | 3.60 | 1.5e-03 | 4.50 | 3.8e-04 | -25.00 | 2.2e-03 |
| 0.25 | 0.00 | 4000 | 0.033 | 0.033 | 0.033 | 300 | 0.68 | 1.2e-04 | 1.20 | 4.6e-05 | -21.00 | 1.3e-04 |
| 0.25 | 0.00 | 4000 | 0.100 | 0.100 | 0.100 | 300 | 1.30 | 3.6e-04 | 1.70 | 1.4e-04 | -23.00 | 4.7e-04 |
| 0.25 | 0.05 | 1000 | 0.017 | 0.017 | 0.017 | 130 | -24.00 | 8.4e-04 | 21.00 | 3.6e-04 | -21.00 | 1.3e-03 |
| 0.25 | 0.05 | 1000 | 0.033 | 0.033 | 0.033 | 160 | -3.50 | 1.7e-03 | 0.64 | 5.6e-04 | -45.00 | 2.5e-03 |
| 0.25 | 0.05 | 1000 | 0.100 | 0.100 | 0.100 | 220 | -0.91 | 5.2e-03 | 7.10 | 2.6e-03 | -31.00 | 8.2e-03 |
| 0.25 | 0.05 | 2000 | 0.017 | 0.017 | 0.017 | 280 | -5.70 | 1.9e-04 | 0.26 | 1.1e-04 | -28.00 | 3.2e-04 |
| 0.25 | 0.05 | 2000 | 0.033 | 0.033 | 0.033 | 300 | 0.25 | 3.8e-04 | 6.00 | 2.8e-04 | -26.00 | 5.4e-04 |
| 0.25 | 0.05 | 2000 | 0.100 | 0.100 | 0.100 | 300 | 5.20 | 1.4e-03 | 1.20 | 7.4e-04 | -21.00 | 1.8e-03 |
| 0.25 | 0.05 | 4000 | 0.033 | 0.033 | 0.033 | 300 | -0.87 | 1.3e-04 | 4.40 | 7.8e-05 | -23.00 | 1.6e-04 |
| 0.25 | 0.05 | 4000 | 0.100 | 0.100 | 0.100 | 300 | 2.20 | 3.1e-04 | 3.50 | 2.5e-04 | -23.00 | 4.8e-04 |

*Parent Corr: average correlation of parental genotypes, MAF Diff: average difference in parental MAFs. Proportion of phenotypic variance explained by additive effects is given in the $\sigma_\alpha^2$ column, by dominance effects in the $\sigma_\delta^2$ column, and by parent-of-origin effects in the $\sigma_\gamma^2$ column. N: number of simulated phased genotypes, #Reps: number of simulated replications, "Bias" is percent relative bias of estimates across simulated replications, "Var" is variance of estimates across simulated replications.*