UCLA
UCLA Electronic Theses and Dissertations

Title

Statistical Models for Detecting Transgenerational Genetic Effects

Permalink

https://escholarship.org/uc/item/0jm511f9

Author

Clark, Michelle Marie

Publication Date

2016

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA

Los Angeles

Statistical Models for Detecting Transgenerational

Genetic Effects

A dissertation submitted in partial satisfaction

of the requirements for the degree

Doctor of Philosophy in Biostatistics

by

Michelle Marie Clark

2016

ABSTRACT OF DISSERTATION

Statistical Models for Detecting Transgenerational Genetic Effects

by

Michelle Marie Clark

Doctor of Philosophy in Biostatistics

University of California, Los Angeles, 2016

Professor Janet S. Sinsheimer, Chair

Genome-wide association studies (GWAS) have successfully discovered a number of genes that control disease susceptibility and variation in quantitative traits. Despite the large number of genes found to be associated with human diseases and complex traits, a limited amount of the total heritability is explained by these discoveries. One hypothesis is that some of this missing heritability is due to transgenerational effects, effects of genetic factors in one generation that affect the phenotypes in a subsequent generation without Mendelian transmission of alleles.

The ability to detect transgenerational effects in humans is mainly limited to maternal effects when using epidemiological data. Furthermore, currently available methodologies lack approaches to identify associations between maternal-fetal genotype (MFG) interactions and quantitative traits for arbitrary family structures. To address this issue, I present the Quantitative-MFG (QMFG) test, a linear mixed effect model in which maternal and offspring genotypes are considered fixed effects and residual familial correlations are random effects. This approach handles pedigrees of virtually any size, common or unusual scenarios of MFG incompatibility, and additional covariates. Another attractive feature of the QMFG test is the ability to easily

extend the approach to multiple loci. With simulation studies, I demonstrate the statistical validity of the QMFG analysis method and show that if a standard model, which considers only offspring genotypes, is fit to data with an MFG effect, associations can be missed or misattributed. To allow other researchers to determine if there is evidence of MFG effects in their own data, I have developed and implemented subroutines as part of the software program Mendel, which is freely available. The QMFG test may provide another approach to uncovering sources of missing heritability in association studies.

This dissertation of Michelle Marie Clark is approved.

Thomas R. Belin

Ronald S. Brookmeyer

Eric M. Sobel

Janet S. Sinsheimer, Committee Chair

University of California, Los Angeles

2016

To my mother, father, and sisters

for their unending love and support

every step of the way;

to my husband and best friend, Griffin,

for always opening my eyes to

the humor and joy of life;

and

to Daisaku Ikeda,

whose encouragement reminded me of

my limitless potential

when I needed it the most.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

x

ACKNOWLEDGEMENTS

VITA

2010   B.S. in Mathematics received from Chapman University, CA.

2010 - 2011   Teaching Assistant, Department of Biostatistics, University of California, Los
              Angeles, CA.

2011 - 2012   Graduate Student Researcher, Department of Epidemiology, University of
              California, Los Angeles, CA.

2012 - 2016   Graduate Student Researcher, Department of Human Genetics, University of
              California, Los Angeles, CA.

2014   Advancement to Candidacy, Department of Biostatistics, University of California,
              Los Angeles, CA.

PUBLICATIONS

K.C. Paul, R. Rausch, M.M. Creek, J.S. Sinsheimer, J.M. Bronstein, Y. Bordelon, B. Ritz (2016) APOE, MAPT, and COMT and Parkinson's Disease Susceptibility and Cognitive Symptom Progression. *J Parkinsons Dis*, doi: 10.3233/JPD-150762.

I. Janowitz Koch, M.M. Clark, M.J. Thompson, K.A. Deere-Machemer, J. Wang, L. Duarte, G.E. Gnanadesikan, E.L. McCoy, L. Rubbi, D.R. Stahler, M. Pellegrini, E.A. Ostrander, R.K. Wayne, J.S. Sinsheimer, B.M. vonHoldt (2016). The concerted impact of domestication and transposon insertions on methylation patterns between dogs and gray wolves. *Mol Ecol*, doi: 10.1111/mec.13480.

M.M. Clark, J. Blangero, T.D. Dyer, E.M. Sobel, J.S. Sinsheimer (2016). The Quantitative-MFG Test: A linear mixed effect model to detect maternal-offspring gene interactions. *Ann Hum Gen*, 80: 63–80.

T.C. Ngun, N.M. Ghahramani, M.M. Creek, S.M. Williams-Burris, H. Barseghyan, Y. Itoh, F.J. Sánchez, R. McClusky, J.S. Sinsheimer, A.P. Arnold, E. Vilain (2014) Feminized behavior and brain gene expression in a novel mouse model of Klinefelter Syndrome. *Arch Sex Behav* 43: 1043-57.

J.S. Sinsheimer, M.M. Creek (2013) "Statistical Approaches for Detecting Transgenerational Genetic Effects in Humans", Epigenetics and Complex Traits (eds. A.K. Naumova & C.M.T. Greenwood). New York: Springer Science+Business Media, 247-264.

E.A. Dinsdale, R.A. Edwards, B.A. Bailey, I. Tuba, S. Akhter, K. McNair, R. Schmieder, N. Apkarian, M. Creek, E. Guan, M. Hernandez, K. Isaacs, C. Peterson, T. Regh, V. Ponomarenko (2013) Multivariate analysis of functional metagenomes. *Front Genet* 4: 41.

M. Creek, C. Rakovski (2010) Haplotype Variety Analysis of Human Populations: an Application to HapMap Data. *e-Research: A Journal of Undergraduate Work* 1: 73-81.

Chapter 1

Introduction to Transgenerational Effects and Quantitative Trait

Association Methods

Complex human diseases and traits are believed to result from a combination of genetic and environmental factors. Genome-wide association studies (GWAS) have revealed numerous loci associated with human diseases and complex traits, providing insights into their etiology (Lander, 2011, Hirschhorn, 2009). Despite these successes, previously reported findings explain a relatively small proportion of heritability, that is, a small fraction of the phenotype variability attributed to genetic variation (Maher, 2008, Manolio *et al.*, 2009). This "missing heritability" has implications when it comes to prevention, diagnosis, and treatments of diseases and their associated quantitative traits. Since previous GWAS focus on common variants, researchers postulate that the missing heritability may be due the effects of rare variants that are poorly detected, the combination of multiple variants of smaller effect (polygenic effects), and/or epigenetics. Another possibility is that transgenerational genetic effects are involved and are not accounted for by standard GWAS analyses. Parent of origin effects are one example of transgenerational effects. Because transgenerational effects involve the effects of variants present in previous generations, a standard GWAS which takes into account only an individual's own genotype, could fail to detect significant associations.

Another transgenerational genetic effect involves interactions between mother and offspring genotypes that affect the offspring's disease risk or phenotype. These interactions, also called maternal-fetal genotype (MFG) interactions, have been shown to be associated with complex diseases, even those that do not manifest until adulthood. The term MFG

1

incompatibility sometimes used in reference to these effects comes from the idea that depending on the combination of maternal and offspring genotypes, a maternal response against the fetus or a fetal response against the mother is possible. Such incompatibilities can alter the conditions in which the offspring develops during pregnancy and have the potential to impact both perinatal (Liang *et al.*, 2010, Li *et al.*, 2014, Lupo *et al.*, 2014, Procopciuc *et al.*, 2014, Li *et al.*, 2016) and adult-onset offspring traits (Stubbs *et al.*, 1985, Hollister *et al.*, 1996, Dahlquist *et al.*, 1999, Juul-Dam *et al.*, 2001, Cannon *et al.*, 2002, Palmer *et al.*, 2002, Newton *et al.*, 2004, Insel *et al.*, 2005, Palmer *et al.*, 2008, Freedman *et al.*, 2011). In principle, there could be paternal-offspring gene interactions (PFG incompatibility), although there is less biological support for these interactions than for maternal-offspring gene interactions.

For binary traits, Sinsheimer *et al.* (2003) modified Weinberg's log-linear method (Weinberg *et al.*, 1998) for estimating maternal and offspring effects in case-parent trios to detect maternal-fetal genotype incompatibility at a candidate locus. Childs *et al.* (2010, 2011) further developed the extended maternal-fetal genotype incompatibility (EMFG) test, which examines both maternal and offspring genotypes as risk factors for disease and allows for arbitrary pedigree structures including those with multiple generations and multiple affected individuals. The EMFG test jointly models maternal genotype effects, offspring genotype effects, and maternal-offspring genotype interactions using a retrospective likelihood that incorporates the pedigree structure, in which the offspring's genotype is the outcome and the phenotype is considered a predictor (for a detailed discussion see, for example, Kraft & Thomas, 2000). The EMFG test handles multi-allelic genes, including non-codominant genes and several tightly linked genes, and can also incorporate potential offspring or maternal-related confounders. Many traits, however, are naturally continuous or can't easily be categorized into a binary affection

status. In such a scenario, a test using a quantitative trait as the outcome could be useful in identifying risk genes.

Kistner *et al*. (2004, 2005, 2006) and Wheeler and Cordell (2007) developed methods of detecting quantitative trait association using a retrospective likelihood for case-parent trios. If y_c is the offspring's quantitative trait and g_c, g_r, g_s are the offspring, mother, and father genotypes, respectively, then the retrospective likelihood proposed by Wheeler and Cordell for a single pedigree is

$$P(g_c|g_r, g_s, y_c) = \frac{P(y_c|g_c)P(g_c|g_r, g_s)}{\sum_{g_c^* \in S_M'} P(y_c|g_c^*)P(g_c^*|g_r, g_s)}$$

where the summation of all possible offspring genotypes based on the observed parental genotypes is represented by $\sum_{g_c^* \in S_M'}$. Wheeler and Cordell show that the contribution of a trio to the likelihood may be assumed to be of the form

$$\frac{\exp\left(\beta_{g_c} y_c + \alpha_{M g_c}\right)}{\sum_{g_c^* \in S_M'} \exp\left(\beta_{g_c^*} y_c + \alpha_{M g_c^*}\right)}$$

where β_{g_i}' represents genotype effects, $\alpha_{M g_i}'$ are nuisance parameters for non-Mendelian effects and population stratification. In this case, the likelihood can be calculated using conditional logistic regression. This approach for testing quantitative trait association can further be extended for multi-locus haplotypes, parent-of-origin effects, maternal genotype and mother-child interaction effects, and gene-environment interactions (Wheeler & Cordell, 2007). The likelihood presented by Kistner *et al*. is identical with the exception of distinguishing between possible heterozygote offspring genotypes (1/2 and 2/1) in the summation of the denominator when assuming parent-of-origin effects. Thus, the number and interpretation of nuisance parameters modeling non-Mendelian effects and population stratification vary between these two methods. Although these methods that can handle quantitative outcomes using a retrospective

likelihood when the data consist of case-parent trios, they cannot accommodate arbitrary family structures.

As an alternative to a retrospective likelihood, using a prospective approach models the phenotype given the offspring genotype and assumes either that the quantitative trait is normally distributed or the central limit theorem is applicable. One way to conduct association testing with quantitative traits is with a measured genotype analysis (Lange, 2002), which is a form of quantitative trait association analysis that uses a linear mixed effects model. In this application, the linear mixed effects model takes into account familial correlations through partitioning the variance and hence is also called variance component modeling in the genetic literature. This model assumes that what is not the effect of a gene or genes (possibly many) acting in Mendelian fashion is typically environmentally induced. Classically, genetic effects are modeled as many genes acting approximately equally and independently (polygenes). The additive genetic variance captures the effect of alleles at these genes as if they were acting independently, deviation from allelic independence leads to dominance genetic variation. The genetic correlation between two relatives' trait values depends on the expected distribution of genes shared identically by descent (IBD) among them. Two genes are identical by descent if they have the same ancestral origin. Environmental variation is postulated to come in two forms: shared and independent among the pedigree members. Shared (common) environmental variation captures the additional correlations among family members that remain unexplained by IBD sharing. Adopted relatives and other unrelated individuals living in the same household help to distinguish these correlations from genetic correlations as do relatives living in different environments (see Zhou *et al.*, 2011). Independent environmental variation is any residual

variation in the trait values after accounting for genetic and shared environmental variations. It is always included because measurement errors make the model fit imperfect.

A measured genotype analysis models the likelihood of a trait using the multivariate Gaussian distribution. For a pedigree with observed trait vector y and mean vector v the loglikelihood is

$$L = -\frac{1}{2}\ln|\Sigma| - \frac{1}{2}(y - v)^t \Sigma^{-1}(y - v)$$

where the covariance matrix Σ is given by

$$\Sigma = 2\sigma_a^2\Phi + \sigma_d^2\Delta_7 + \sigma_h^2 H + \sigma_e^2 I.$$

When the effect of the offspring's genotype is of interest, $v = \mu + \beta g_c$, where g_c denotes a vector of risk allele counts for each offspring's genotype. The additive, dominance, household and environmental genetic variances summed over all genes are denoted by $\sigma_a^2, \sigma_d^2, \sigma_e^2$ and σ_e^2, respectively. The global kinship coefficient matrix is given by Φ and is the probability that, at an arbitrary autosomal locus, two genes chosen at random one from i and j match IBD (Lange 2002). The matrix of the probabilities that two individuals i and j share both genes IBD at an arbitrary autosomal locus is denoted by the condensed identity coefficient Δ_7 and captures deviation for strict additivity of the alleles. See Table 1 for examples of these coefficients for various familial relationships. Derivations of these quantities can be found in Lange (2002). The household indicator matrix $H= (h_{ij})$ has entries 0 or 1 depending on whether subjects i and j are in the same household. In theory multiple household matrices could be included, for example to define different household membership over time, although in most cases only one is included (Zhou *et al*., 2011). Environmental contributions and measurement errors are included with the identity matrix I.

Table 1. Examples of kinship and condensed identity coefficient Δ_7

Relationship	Φ	Δ_7
Parent-Offspring	1/4	0
Half Siblings	1/8	0
Full Siblings	1/4	1/4
First Cousins	1/16	0
Double First Cousins	1/8	1/16

With measured genotype analyses, likelihood ratio tests (LRTs) can be used to determine the significance of estimated parameters. LRTs are able to handle complex null and alternative hypotheses but the necessity to iteratively maximize the likelihood under both the null and alternative models for each SNP can be computationally intensive unless when only a few candidate genes are tested and pedigrees are small. LRTs therefore, are not practical for genome-wide gene discovery. As an alternative, score tests can be used for inference.

Score tests lack multiple iterations and can handle various pedigree structures and sizes. In the case of GWAS mapping, Zhou *et al.* (2015) proposed the use of score testing to quickly screen SNPs and identify the most significant SNPs to test using the LRT to obtain parameter estimates. In general, the score statistic is given by

$$S(\theta) = dL(\theta)J(\theta)^{-1}\nabla L(\theta)$$

where $\nabla L(\theta)$ is the gradient of the loglikelihood with the parameter vector θ, $dL(\theta)$ is the first differential of the loglikelihood, and $J(\theta)$ is the expected information matrix. The expected information matrix is block diagonal, $d_{\beta}^2 L(\theta)$ denotes the second differential with respect to the vector of regression coefficients and $d_{\sigma}^2 L(\ )$ denotes the second differential with respect to the vector of variance parameters.

$$J(\theta) = \begin{pmatrix} E[-d_{\beta}^2 L(\theta)] & 0 \\ 0 & E[-d_{\sigma}^2 L(\theta)] \end{pmatrix}$$

Under the null model, $\nabla L_\sigma(\hat\theta) = 0$ so only the elements associated with the regression coefficients are needed and thus, the quantities necessary for calculating the score test are $\sum_{i=1}^{n} \nabla_\beta L_i(\theta)$ and $\sum_{i=1}^{n} E[-d_\beta^2 L(\theta)]$ for n pedigrees. Let $\nu = A\beta$ denote the mean vector. The design matrix under the alternative hypothesis can be written as a combination of the shifted allele count vector for pedigree i (a_i) where each entry is -1, 0, or 1 and the design matrix under the null (N_i), i.e. $A_i = (a_i, N_i)$. It follows that

$$\sum_{i=1}^{n} \nabla_\beta L_i(\theta) = \begin{pmatrix} \sum_{i=1}^{n} a_i^t \Omega_i^{-1}(y_i - A_i\hat\beta) \\ \sum_{i=1}^{n} N_i^t \Omega_i^{-1}(y_i - A_i\hat\beta) \end{pmatrix}$$

$$\sum_{i=1}^{n} E[-d_\beta^2 L(\theta)] = \begin{pmatrix} \sum_{i=1}^{n} a_i^t \Omega_i^{-1} a_i & \sum_{i=1}^{n} a_i^t \Omega_i^{-1} N_i \\ \sum_{i=1}^{n} N_i^t \Omega_i^{-1} a_i & \sum_{i=1}^{n} N_i^t \Omega_i^{-1} N_i \end{pmatrix}$$

Since $\sum_{i=1}^{n} N_i^t \Omega_i^{-1}(y_i - A_i\hat\beta) = 0$ under the null model, the score test statistic for each SNP becomes

$$S = R^t \left[Q - W^t \left(\sum_{i=1}^{n} N_i^t \Omega_i^{-1} N_i \right)^{-1} W \right]^{-1} R$$

where

$$Q = \sum_{i=1}^{n} a_i^t \Omega_i^{-1} a_i \,, R = \sum_{i=1}^{n} a_i^t \Omega_i^{-1}(y_i - A_i\hat\beta), W = \sum_{i=1}^{n} N_i^t \Omega_i^{-1} a_i$$

Zhou *et al.* (2015) suggest precomputing $\Omega_i^{-1}, \Omega_i^{-1} N_i$, and $\Omega_i^{-1}(y - A_i\hat\beta)$ for each pedigree and $\sum_{i=1}^{n} N_i^t \Omega_i^{-1} N_i$ at the maximum likelihood estimates under the null hypothesis to speed up computation.

This dissertation addresses the absence of methods and available software for testing for transgenerational effects, specifically MFG incompatibility, on quantitative traits using arbitrary family structures, including a number of large families. In Chapter 2, we provide thorough descriptions and common examples of transgenerational effects and provide a review of existing statistical methods for such effects (first published as Sinsheimer and Creek (2013)). In Chapter 3, we present the Quantitative-MFG (QMFG) test, a linear mixed effect modeling approach to test for offspring genotype, maternal genotype, and interaction effects jointly, marginally, and conditionally on quantitative traits. This prospective method can handle pedigrees of virtually any size, simultaneously, and easily allows for the inclusion of additional covariates (Clark *et al*., 2016). We apply both LRTs and score tests for inference and demonstrate that the QMFG test may provide the key to uncovering sources of missing heritability in association studies. The flexibility of the QMFG test to handle more complex MFG interactions involving multiple loci is presented in Chapter 4. These chapters are written as stand-alone articles for publication and hence, there is some overlap when reviewing MFG interactions and past methodology. Chapter 5 proposes possible extensions to our approach and serves as a guideline for future research.

References

Cannon, M., Jones, P.B. & Murray, R.M. (2002) Obstetric complications and schizophrenia: historical and meta-analytic review. *Am. J. Psychiatry,* 159**,** 1080-92.

Childs, E.J., Palmer, C.G., Lange, K. & Sinsheimer, J.S. (2010) Modeling maternal-offspring gene-gene interactions: the extended-MFG test. *Genet Epidemiol,* 34**,** 512-21.

Childs, E.J., Sobel, E.M., Palmer, C.G. & Sinsheimer, J.S. (2011) Detection of intergenerational genetic effects with application to HLA-B matching as a risk factor for schizophrenia. *Hum Hered,* 72**,** 161-72.

Clark, M.M., Blangero, J., Dyer, T.D., Sobel, E.M. & Sinsheimer, J.S. (2016) The Quantitative-MFG Test: A Linear Mixed Effect Model to Detect Maternal-Offspring Gene Interactions. *Ann Hum Genet,* 80**,** 63-80.

Dahlquist, G.G., Patterson, C. & Soltesz, G. (1999) Perinatal risk factors for childhood type 1 diabetes in Europe. The EURODIAB Substudy 2 Study Group. *Diabetes Care,* 22**,** 1698-702.

Freedman, D., Deicken, R., Kegeles, L.S., Vinogradov, S., Bao, Y. & Brown, A.S. (2011) Maternal-fetal blood incompatibility and neuromorphologic anomalies in schizophrenia: Preliminary findings. *Prog. Neuropsychopharmacol. Biol. Psychiatry,* 35**,** 1525-9.

Hirschhorn, J.N. (2009) Genomewide association studies--illuminating biologic pathways. *N Engl J Med,* 360**,** 1699-701.

Hollister, J.M., Laing, P. & Mednick, S.A. (1996) Rhesus incompatibility as a risk factor for schizophrenia in male adults. *Arch. Gen. Psychiatry,* 53**,** 19-24.

Insel, B.J., Brown, A.S., Bresnahan, M.A., Schaefer, C.A. & Susser, E.S. (2005) Maternal-fetal blood incompatibility and the risk of schizophrenia in offspring. *Schizophr. Res.,* 80**,** 331-42.

Juul-Dam, N., Townsend, J. & Courchesne, E. (2001) Prenatal, perinatal, and neonatal factors in autism, pervasive developmental disorder-not otherwise specified, and the general population. *Pediatrics,* 107**,** E63.

Kistner, E.O., Infante-Rivard, C. & Weinberg, C.R. (2006) A method for using incomplete triads to test maternally mediated genetic effects and parent-of-origin effects in relation to a quantitative trait. *Am J Epidemiol,* 163**,** 255-61.

Kistner, E.O. & Weinberg, C.R. (2004) Method for using complete and incomplete trios to identify genes related to a quantitative trait. *Genet. Epidemiol.,* 27**,** 33-42.

Kistner, E.O. & Weinberg, C.R. (2005) A method for identifying genes related to a quantitative trait, incorporating multiple siblings and missing parents. *Genet. Epidemiol.,* 29**,** 155-65.

Kraft, P. & Thomas, D.C. (2000) Bias and efficiency in family-based gene-characterization studies: conditional, prospective, retrospective, and joint likelihoods. *Am J Hum Genet,* 66**,** 1119-31.

Lander, E.S. (2011) Initial impact of the sequencing of the human genome. *Nature,* 470**,** 187-97.

Lange, K. (2002) *Mathematical and statistical methods for genetic analysis.* New York: Springer.

Li, M., Erickson, S.W., Hobbs, C.A., Li, J., Tang, X., Nick, T.G., Macleod, S.L., Cleves, M.A. & National Birth Defect Prevention Study (2014) Detecting maternal-fetal genotype interactions associated with conotruncal heart defects: a haplotype-based analysis with penalized logistic regression. *Genet. Epidemiol.,* 38**,** 198-208.

Li, M., Li, J., Wei, C., Lu, Q., Tang, X., Erickson, S.W., Macleod, S.L. & Hobbs, C.A. (2016) A
    Three-Way Interaction among Maternal and Fetal Variants Contributing to Congenital
    Heart Defects. *Ann Hum Genet*, 80**,** 20-31.

Liang, M., Wang, X., Li, J., Yang, F., Fang, Z., Wang, L., Hu, Y. & Chen, D. (2010) Association
    of combined maternal-fetal TNF-alpha gene G308A genotypes with preterm delivery: a
    gene-gene interaction study. *J. Biomed. Biotechnol.*, 2010**,** 396184.

Lupo, P.J., Mitchell, L.E., Canfield, M.A., Shaw, G.M., Olshan, A.F., Finnell, R.H., Zhu, H. &
    National Birth Defects Prevention Study (2014) Maternal-fetal metabolic gene-gene
    interactions and risk of neural tube defects. *Mol. Genet. Metab.*, 111**,** 46-51.

Maher, B. (2008) Personal genomes: The case of the missing heritability. *Nature*, 456**,** 18-21.

Manolio, T.A., Collins, F.S., Cox, N.J., Goldstein, D.B., Hindorff, L.A., Hunter, D.J., Mccarthy,
    M.I., Ramos, E.M., Cardon, L.R., Chakravarti, A., Cho, J.H., Guttmacher, A.E., Kong,
    A., Kruglyak, L., Mardis, E., Rotimi, C.N., Slatkin, M., Valle, D., Whittemore, A.S.,
    Boehnke, M., Clark, A.G., Eichler, E.E., Gibson, G., Haines, J.L., Mackay, T.F.,
    Mccarroll, S.A. & Visscher, P.M. (2009) Finding the missing heritability of complex
    diseases. *Nature,* 461**,** 747-53.

Newton, J.L., Harney, S.M., Wordsworth, B.P. & Brown, M.A. (2004) A review of the MHC
    genetics of rheumatoid arthritis. *Genes Immun.,* 5**,** 151-7.

Palmer, C.G., Mallery, E., Turunen, J.A., Hsieh, H.J., Peltonen, L., Lonnqvist, J., Woodward,
    J.A. & Sinsheimer, J.S. (2008) Effect of Rhesus D incompatibility on schizophrenia
    depends on offspring sex. *Schizophr. Res.,* 104**,** 135-45.

Palmer, C.G., Turunen, J.A., Sinsheimer, J.S., Minassian, S., Paunio, T., Lonnqvist, J., Peltonen, L. & Woodward, J.A. (2002) RHD maternal-fetal genotype incompatibility increases schizophrenia susceptibility. *Am. J. Hum. Genet.,* 71**,** 1312-9.

Procopciuc, L.M., Caracostea, G., Zaharie, G. & Stamatian, F. (2014) Maternal/newborn VEGF-C936T interaction and its influence on the risk, severity and prognosis of preeclampsia, as well as on the maternal angiogenic profile. *J. Matern. Fetal Neonatal Med.,* 27**,** 1754-60.

Sinsheimer, J.S. & Creek, M.M. (2013) Statistical Approaches for Detecting Transgenerational Genetic Effects in Humans. In: *Epigenetics and Complex Traits* A.K. Naumova & C.M.T. Greenwood (eds.) *Epigenetics and Complex Traits.* Springer New York.

Sinsheimer, J.S., Palmer, C.G. & Woodward, J.A. (2003) Detecting genotype combinations that increase risk for disease: maternal-fetal genotype incompatibility test. *Genet Epidemiol,* 24**,** 1-13.

Stubbs, E.G., Ritvo, E.R. & Mason-Brothers, A. (1985) Autism and shared parental HLA antigens. *J. Am. Acad. Child Psychiatry,* 24**,** 182-5.

Weinberg, C.R., Wilcox, A.J. & Lie, R.T. (1998) A log-linear approach to case-parent-triad data: assessing effects of disease genes that act either directly or through maternal effects and that may be subject to parental imprinting. *Am J Hum Genet,* 62**,** 969-78.

Wheeler, E. & Cordell, H.J. (2007) Quantitative trait association in parent offspring trios: Extension of case/pseudocontrol method and comparison of prospective and retrospective approaches. *Genet. Epidemiol.,* 31**,** 813-33.

Zhou, H., Blangero, J., Dyer, T.D., Chan, K.K., Sobel, E. & Lange, K. (2015) Fast Genome-Wide QTL Association Mapping on Pedigree and Population Data. arXiv:1407.8253v2.

Zhou, J.J., Pelka, S., Lange, K., Palmer, C.G. & Sinsheimer, J.S. (2011) Dissecting prenatal,

postnatal, and inherited effects: ART and design. *Genet Epidemiol,* 35**,** 437-46.

Chapter 2

Statistical Approaches for Detecting Transgenerational Genetic Effects in Humans

Janet S. Sinsheimer and Michelle M. Creek

Abstract Transgenerational genetic effects occur when the genes of one generation influence the phenotype of subsequent generations without Mendelian transmission of alleles, possibly through inherited epigenetic effects. The evidence for transgenerational genetic effects in humans comes predominantly from genetic epidemiology studies, which thus presents a number of statistical challenges to their analysis and interpretation. In this chapter, we outline some of the genetic epidemiologic study designs and statistical analysis approaches that have been used to detect these effects and discuss their strengths and weaknesses.

11.1 Introduction

Genetic epidemiology concentrates on disease risks due to a subject's own genes and environment. Although we gain much etiological insight from these studies, many genetic determinants of disease remain undiscovered. One possibility is that transgenerational genetic effects play a role in their etiology. Transgenerational genetic effects occur when the genes of one generation influence the phenotype of subsequent generations without Mendelian transmission of alleles (Fig. 11.1), possibly through inherited epigenetic effects (Gluckman et al. 2007; Nadeau 2009). Most commonly, these transgenerational genetic effects are parental genes having an effect on their offspring's phenotype, but more distant ancestors can have effects.

J.S. Sinsheimer (✉)
Department of Human Genetics, David Geffen School of Medicine at UCLA,
Los Angeles, CA 90095, USA
e-mail: janet@mednet.ucla.edu

M.M. Creek
Department of Biostatistics, UCLA Fielding School of Public Health,
Los Angeles, CA 90095, USA
e-mail: mmcreek@gmail.com

14
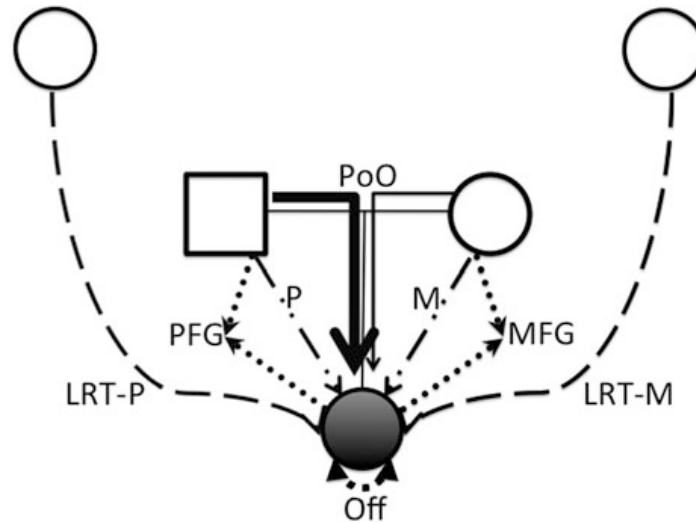
Fig. 11.1 Schematic depiction of potentially detectable effects. Disease risk in the offspring (denoted by the *dark circle*) can be due to maternal (M), paternal (P), offspring (Off), parent of origin (PoO), long-range transgenerational (LRT-M maternal side, LRT-P paternal side), maternal–fetal genotype incompatibility (MFG), or paternal–fetal genotype incompatibility (PFG) effects. These effects are not mutually exclusive. The effects could be genetic or environmental in origin. When genetic in origin, they have direct effects or are epigenetic or environmentally mediated

Less commonly considered, an offspring's genotype could also elicit a phenotype in his/her mother. In this chapter, we explore statistical approaches for detecting transgenerational genetic effects in humans with genetic epidemiological data. With a few exceptions, these studies provide only indirect evidence consistent with epigenetic phenomena and, in some of these cases, the underlying explanations for transgenerational genetic effects will not involve modifications to DNA or chromatin. However, we argue these studies provide an excellent starting point for hypothesis generation and for further investigations leading to more direct tests for epigenetic effects.

To clarify what we mean by transgenerational genetic effects, we first provide descriptions of some common ones before going on to describe appropriate study designs and analysis approaches to detect them. For notational convenience we drop the "genetic" and from now on refer to them collectively as transgenerational effects. Figure 11.1 illustrates some transgenerational effects discussed in this chapter including maternal effects (M), paternal effects (P), parent of origin effects (PoO), maternal–fetal genotype (MFG) incompatibility, paternal–fetal genotype (PFG) incompatibility, and long-range transgenerational effects from the maternal or paternal side (LRT-M or LRT-P).

Maternal genetic effects and paternal genetic effects. There are observations that particular maternal genotypes are strongly associated with offspring phenotypes, regardless of what alleles the offspring inherits (Kistner and Weinberg 2004; Wheeler and Cordell 2007; Weinberg et al. 1998). Although discussed less often, paternal genotypes are also associated with offspring phenotype. A possible mechanism for

these effects includes parental genotype influences on the offspring's environment. This environmental effect could have no effect on the offspring's DNA or chromatin or it could induce epigenetic modifications. Alternatively, parental genotype could directly affect the offspring's epigenome or the effects could be due to maternally derived mitochondrial DNA.

Parent of origin effects. PoO effects occur when the degree of association of an allele with an offspring's phenotype depends on the sex of the transmitting parent (in Fig. 11.1, the effects of the paternally transmitted allele prevail). Although silencing through methylation or histone modification, commonly referred to as imprinting, is one form of PoO effects, other mechanisms can also lead to these effects (Guilmatre and Sharp 2012). These mechanisms include mutational transmission bias and oocyte RNAs or proteins.

MFG incompatibility. The effects of maternal genes on the offspring's disease risk may vary depending on the offspring's genotype (Fig. 11.1). MFG incompatibilities are gene interactions that produce adverse effects on the developing fetus. These gene–gene interactions differ from typical ones because maternal genes interact with offspring genes. MFG incompatibilities are involved in complex diseases, even adult onset diseases where the effects may not evident until long after the MFG incompatibility initiated event has occurred and subsided (Palmer et al. 2002, 2006; Sinsheimer et al. 2003). In principle, there could be paternal–offspring gene interactions (PFG incompatibility), although there is less biological support for these interactions than for MFG incompatibility.

Like PoO and maternal effects, the mechanisms by which MFG or PFG incompatibilities occur could be methylation or chromatin modification but other mechanisms are possible. The prototypical MFG incompatibility is *RHD* incompatibility, which can lead to erythroblastosis, liver damage, hypoxia, or death from hemolytic disease of the newborn (HDN) (Guyton 1981). The biological mechanism underlying *RHD*-induced HDN is well known (Stratchen and Reed 2003) and we provide a simplified description. Alleles at the *RHD* locus are classified into two types, *D* and *d*. The *D* allele codes for an antigen on the erythrocyte surface and the *d* allele is a null allele. *RHD*-induced HDN occurs when a mother with a null allele homozygous genotype (*d/d*) mounts an IgG alloimmune response to her *d/D* offspring's erythrocytes, damaging their ability to carry oxygen and releasing bilirubin. Maternal–fetal *ABO* incompatibility leads to HDN by a similar mechanism (Guyton 1981). *RHD* and *ABO* incompatibilities are implicated as risk factors for complex diseases (Cannon et al. 2002; Dahlquist et al. 1999; Hollister et al. 1996; Insel et al. 2005; Juul-Dam et al. 2001; Kraft et al. 2004; Palmer et al. 2002; Stubbs et al. 1985). Although *RHD* incompatibility involves the same locus in mother and offspring, MFG incompatibilities can also occur between one locus in the mother and another locus in the offspring (Chen et al. 2009).

Long-range transgenerational (LRT) effects. It is difficult to distinguish epigenetic effects from shared environment unless the transgenerational effect persists over multiple generations but the environmental exposure does not. Environmental exposures, even if they are short lived, can effect three generations without

involving specific inherited epigenetic factors. If a pregnant woman is exposed to an environmental stimulus, she, her fetus, her gametes, and her fetus' gametes can be directly affected without involving epigenetic modifications. Likewise environment can affect two generations when a man is exposed because his gametes can be affected. LRT genetic effects caused by inherited epigenetic effects are well documented in model organisms (Nadeau 2009), and there is evidence of their role in common diseases in humans (e.g., Benyshek et al. 2001; Klip et al. 2002).

11.2   Study Designs

In this section, we discuss epidemiological study designs used to detect transgenerational effects. Most statistical approaches to study these effects have been designed for bivariate, qualitative traits. Therefore, when we discuss specific study designs and analyses, we concentrate on these bivariate traits and refer to cases and controls. When methods for continuous traits are commonplace we discuss them in the analysis section along with the appropriate modifications to the study design.

Case-mother, control-mother (*CMCM*). The CMCM design allows detection of offspring genetic effects, maternal genetic effects, and their interactions by comparing the genotype distributions of affected individuals and their mothers to the genotype distributions of unaffected individuals and their mothers (Ainsworth et al. 2011). This design is an extension of the popular case–control design of genome-wide association and is subject to the same limitations, such as confounding from population substructure.

Case-parent trios (*CPTs*). CPTs were first popularized in genetics to avoid the problems of population substructure that originally plagued case–control genetic analysis (e.g., Laird and Lange 2010). Although this advantage is largely eliminated by methods that control for ancestry in case–control studies (e.g., Edwards and Gao 2012), CPTs are popular for detecting transgenerational effects associated with disease (Cordell 2004; Cordell et al. 2004; Laird and Lange 2010). Using CPTs expands the genetic models that can be considered over using CMCMs. For example one can test for PoO and paternal effects. For a bi-allelic locus there are 15 possible maternal–paternal–offspring genotype combinations. Case-mother and case-father duos can be included along with the CPTs by treating the duos as trios with randomly missing data (Sinsheimer et al. 2003; Weinberg et al. 1998). These models can be modified to include control-mother duos or parents of unaffected offspring (Vermeulen et al. 2009) but then population stratification comes back into play.

Nuclear families. The CPT design can be extended to include unaffected and affected siblings of the case (Kraft et al. 2004). These extensions provide additional power and increase the genetic models that can be considered but may require additional modeling assumptions or else be biased. As we discuss in the statistical analysis section, the study design dictates the questions that can be posed as well as the assumptions imposed.

General pedigree data. Intuitively, the ideal study design for detecting transgenerational effects allows simultaneous analysis of unrelated individuals, small pedigrees, and large pedigrees. The inclusion of large, multigenerational pedigrees provides a way to study a variety of complex patterns and detect LRT effects. Being able to analyze all family members is highly efficient. Breaking up large pedigrees into subsets can introduce bias (Childs et al. 2010, 2011). Pedigrees provide a way to model phenotype data in the absence of genotype data (see the statistical analysis section). One disadvantage is that, depending on the research question, using multigenerational families requires more restrictive modeling assumptions to be computationally feasible.

11.3 Statistical Analysis Approaches for Detecting Transgenerational Effects

We briefly outline some statistical analysis approaches. Because the approaches depend on the available data, we group them by data type: (1) phenotype data only, (2) phenotype and genotype data, and (3) phenotype, genotype, and epigenetic data.

11.3.1 Approaches Using Only Phenotype Data

Prior to the wide spread availability of genotype data, evidence supporting the existence of transgenerational effects in humans came from the inference of phenotypic inheritance patterns inconsistent with Mendelian inheritance. These approaches generally require large pedigrees to be effective but, with marked environmental exposures, following matrilineal or patrilineal lines provides evidence of transgenerational effects (e.g., Gluckman et al. 2007).

Indirect evidence for transgenerational effects can, in principle, be obtained from analyzing pedigrees with complex segregation analyses (e.g., Khoury et al. 1993). These analyses use correlations among family members' phenotypes to infer the existence of major genes acting in a Mendelian manner, polygenes, shared environment, and independent environment (residual effects). Generational differences and birth order effects can be inferred. The number of effects inferred is dependent on the variety of relationships and so, in general, large pedigrees are needed to adequately explore transgenerational effects. The biggest difficulties with this approach are the equivalence or near equivalence of sets of models and the inability to prove any model to be true.

Variance component analysis (e.g., Lange 2002) and its related approach, path analysis (e.g., Thomas 2004) have been used to separate genetic sources of phenotypic variation from other sources. In the absence of genetic marker data, what is not the effect of a gene or genes (possibly many) acting in a Mendelian fashion is typically assumed to be environmentally induced. These approaches postulate trait
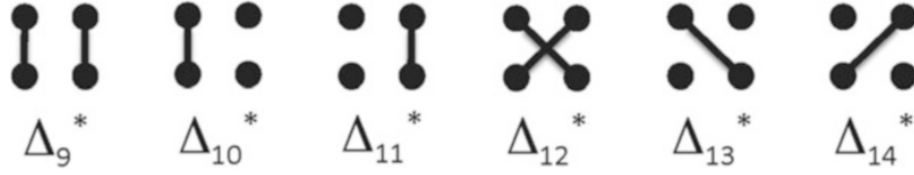
Fig. 11.2 Detailed identity states used in the variance component analyses. Six detailed identity states describe the IBD sharing between two non-inbred individuals i and j. Four *circles* in a block represent the genes from i and j at a single locus. Individual i's genes are on the *top* and j's genes are on the *bottom*. Maternally derived genes are on the *left*, paternally derived genes on the *right*. *Lines* between genes represent genes that are identical by descent. The probability of observing state k is denoted by Δ_k^*

variation is partitioned. Classically, genetic effects are modeled as many genes acting approximately equally and independently (polygenes). The additive genetic variance captures the effect of alleles at these genes as if they were acting independently, deviation from allelic independence leads to dominance genetic variation. The genetic correlation between two relatives' trait values depends on the expected distribution of genes shared identically by descent (IBD) among them (Fig. 11.2). Environmental variation is postulated to come in two forms: shared and independent among the pedigree members. Shared environmental variation captures the additional correlations among family members that remain unexplained by IBD sharing. Adopted relatives and other unrelated individuals living in the same household help to distinguish these correlations from genetic correlations. Independent environmental variation is any residual variation in the trait values after accounting for genetic and shared environmental variations. It is always included because measurement errors make the model fit imperfect.

Although it is possible to use variance component models to test for transgenerational effects with only phenotypes, this approach has not been pursued to any appreciable extent. One reason is that shared environment and transgenerational effects are often confounded, making inference of transgenerational effects difficult. Parent of origin effects provide an exception (Gorlova et al. 2007; Zhou et al. 2011). PoO effects lead to a difference in parent–offspring correlations depending on the parent's sex and thus are accommodated by partitioning the additive genetic variance into two separate effects. When there are no shared environmental effects but there are parent of origin effects, the variance covariance matrix for family phenotypes Y can be written as:

$$\begin{aligned}
\text{Var}(Y) = {} & (\Delta_9^* + \Delta_{10}^*)\sigma_{\text{ma}}^2 + (\Delta_9^* + \Delta_{11}^*)\sigma_{\text{pa}}^2 \\
& + (2\Delta_{12}^* + \Delta_{13}^* + \Delta_{14}^*)cov_{\text{mpa}} + \Delta_9^*\sigma_{\text{d}}^2 + \Delta_{12}^*cov_{\text{d}} + I\sigma_{\text{e}}^2,
\end{aligned} \tag{11.1}$$

where σ_{ma}^2 is the maternal additive genetic variance, σ_{pa}^2 is the paternal additive genetic variance, cov_{mpa} is the additive covariance of the maternal and paternal alleles, σ_{d}^2 is the dominance genetic variance, cov_{d} is the dominance covariance, I is the identity matrix, σ_{e}^2 is the independent environmental variance, and Δ_i is the probability of IBD state i (Fig. 11.2).

Testing for maternal genetic effects with variance component models presents a problem, especially when attempting to dissect prenatal and postnatal effects from maternally inherited effects. This dissection is important because if prenatal effects are not properly modeled, heritability estimates are biased and may lead to false inference of PoO effects (Zhou et al. 2011). In the absence of measured predictors (e.g., genotypes), it is impossible to estimate all three effects in traditional nuclear families because the mother providing the genetic material (genetic mother) is the same person carrying the child (gestational mother) and the same person raising the child (postnatal mother). Although animal experimentation provides opportunities to dissect these effects by embryo transplantation and cross fostering (e.g., Nadeau 2009), in humans the options are limited. Adoption studies have been used to separate out postnatal effects but prenatal and maternal inherited effects are still confounded. Comparing the offspring of sisters to the offspring of brothers separates maternal inheritance from prenatal and postnatal effects (Robson 1955). However, recent advances in assisted reproductive technologies (ART) provide ways to separate all these effects (Thapar et al. 2007; Zhou et al. 2011).

Because ART uses sperm donation, egg donation, or gestational surrogacy depending on approach, children's genetic parents can be different from their prenatal or postnatal parents. By comparing phenotypes within and between these families, it is possible to separate maternal genetic, prenatal, and postnatal effects (Zhou et al. 2011). The key is to modify equation (11.1) by adding in a prenatal household matrix H_{pre}, where offspring born to the same gestational mother are indicated and a postnatal household matrix H_{post} where members with common environmental exposures are indicated.

$$
\begin{aligned}
\text{Var}(Y) = {} & (\Delta_9^* + \Delta_{10}^*)\sigma_{\text{ma}}^2 + (\Delta_9^* + \Delta_{11}^*)\sigma_{\text{pa}}^2 + (2\Delta_{12}^* + \Delta_{13}^* + \Delta_{14}^*)cov_{\text{mpa}} \\
& + \Delta_9^*\sigma_{\text{d}}^2 + \Delta_{12}^*cov_{\text{d}} + I\sigma_{\text{e}}^2 + H_{\text{pre}}\sigma_{\text{pre}}^2 + H_{\text{post}}\sigma_{\text{post}}^2.
\end{aligned}
\tag{11.2}
$$

Phenotypes from all the parents and offspring in ART families can be used to estimate these effects. Known risk factors are included as fixed covariates. Zhou (Zhou et al. 2011) demonstrates this approach with nuclear families, and note that variance component models can, in principle, use ART pedigrees of any complexity.

11.3.2   Approaches Using Phenotype and Genotype Data

When genotypes are available, the possibilities for detecting transgenerational effects improve. Genotypes provide a causal anchor and allow dissection of genetic effects from environmental effects. Tests of the association of parental genotypes with offspring phenotype and interactions between parent and offspring genotypes are possible. Direct evidence for the epigenetic mechanisms underlying these transgenerational effects is not obtained from these studies; however, support for persistent, shared environment can be reduced or eliminated.

Table 11.1 Joint maternal–offspring genotype relative risks

Genotype		Model			
Mother[a]	Offspring	Joint maternal—offspring[b]	Main effects and interactions	*RHD* incompatibility[c]	NIMA[d]
2/2	2/2	δ_{22}[e]	$S_2 R_2$[f]	1	R_2
2/2	1/2	δ_{21}	$S_2 R_1 \gamma_{21}$	1	R_1
1/2	2/2	δ_{12}	$S_1 R_2$	1	R_2
1/2	1/2	δ_{11}	$S_1 R_1$	1	R_1
1/2	1/1	δ_{10}	S_1	1	γ_{10}
1/1	1/2	δ_{01}	$R_1 \gamma_{01}$	γ_{01}	R_1
1/1	1/1	δ_{00}	1	1	1

[a]The first two columns denote genotypes of the mother and her offspring
[b]Columns 3 and 4 denote two different parameterizations of the most general model of maternal–offspring genotype effects that can be used with a bi-allelic locus
[c]Column 5 models *RHD* incompatibility
[d]Column 6 models NIMA and offspring main effects
[e]δ_{ij} is the joint effect of i maternal 2 alleles and j offspring 2 alleles
[f]S_i is the main effect of i maternal 2 alleles R_j is the main effect of j offspring 2 alleles and γ_{ij} is the additional interaction effect

Before embarking on a discussion of the specific analysis approaches, we note that most statistical approaches assume that the genotype is a SNP. Thus we will also focus on a single bi-allelic locus. Readers should be aware however there are a few methods that allow multi-allelic and multi-locus genotype data (e.g., Chen et al. 2009; Childs et al. 2011; Hsieh et al. 2006a; Sinsheimer et al. 2003).

CMCM data can be summarized in a two factor contingency table and analyzed as chi-square or using a Fisher exact test. Analysis of CMCM data can incorporate covariates affecting disease susceptibility by using logistic regression. The combinations of mother–offspring genotypes represent levels of one factor and case–control status represents levels of a second factor. Thus these data can be used to test models regarding maternal genotype main effects, offspring genotype main effects, and their interactions. Under the null hypothesis of no effect of this locus on disease susceptibility, the two factors are independent and the genotype frequencies for cases should be the same as the genotype frequencies for controls.

The number of levels for the first factor depends on whether the maternal and offspring SNPs are at the same locus or two distinct loci. When the maternal and offspring loci are distinct, there are nine possible maternal–offspring combinations. Chen et al. (2009) proposed a likelihood ratio test that allows the inclusion of mother's and offspring's genotypes at both these loci and showed that including both increases overall information and thus increases power. When considering the same bi-allelic locus for mother and offspring, the number of maternal–offspring genotype combinations is seven (Table 11.1). Assumptions regarding the mechanisms by which these genotypes lead to disease or how the maternal and offspring genotypes interact result in restrictions on the levels, further reducing the number of independent parameters. If paternal–offspring interactions are suspected,

the CMCM design can be changed to a case-father control-father (CFCF) design and analyzed in the same manner.

Inference from CMCM studies is sensitive to population stratification. Although this can be corrected by accounting for maternal population history, CMCM designs are limited in the hypotheses that can be tested. One alternative is to use case-parent trios. Several analysis approaches have been used to analyze trios depending on the research questions under consideration.

Parent of origin effects can be detected using the transmission disequilibrium test (TDT) (Spielman et al. 1993; Terwilliger and Ott 1992). The TDT is a form of conditional logistic regression that uses a retrospective design where the genotype of the offspring is the dependent variable (Sham and Curtis 1995; Sinsheimer et al. 2000; Thomas 2004). The TDT is a test of linkage and association between a genetic marker and a disease locus. When used with CPTs, the null hypothesis is no linkage or no association and a heterozygous parent is equally likely to pass on either of their alleles to their offspring. If there are linkage and association, one allele will appear to be transmitted more often to an affected offspring than the other. By comparing a model allowing for separate maternal and paternal transmissions to the standard TDT where the maternal and paternal transmissions are the same, the existence of PoO effects can be tested.

The TDT can be further modified to examine parent–offspring genotype interactions. For example non-inherited maternal antigen (NIMA) effects, a form of MFG incompatibility postulated to occur in rheumatoid arthritis (Harney et al. 2003; Hsieh et al. 2006b), can be tested by comparing the proportion of cases whose genotypes are incompatible with their mother's genotype to the proportion of cases whose genotypes are incompatible with their father's genotype (Harney et al. 2003). The assumption underlying this analysis is that NIMA is a plausible risk factor for a complex disease, but non-inherited paternal antigens (NIPA) are not. One major deficit of this design is that it is not possible to simultaneously check for offspring genotype effects, and maternal genotype effects are confounded with MFG incompatibility. The design also requires a substantial number of fathers be genotyped to have reasonable power.

The TDT gains no information from parents with homozygous genotypes, limiting power. Weinberg (Weinberg et al. 1998) proposed a log-linear model as an alternative and tested for offspring genetic main effects, maternal genetic main effects, and parent of origin effects. Sinsheimer (Sinsheimer et al. 2003) recognized the log-linear model could be extended to allow for maternal–offspring gene interaction at a single locus. Like the TDT, these log-linear models use cases and their parents in a retrospective design in which the genotypes are the dependent variables and no controls are necessary. Sinsheimer's MFG test maximizes the equivalent multinomial likelihood to the log-linear model in order to estimate parameters, and thus easily accommodates maternal–offspring and paternal–offspring dyads as incomplete trios.

The MFG (and equivalently the log-linear) test is very flexible, allowing many inherited disease risk scenarios to be modeled (Ainsworth et al. 2011; Hsieh et al. 2006a, b, 2007; Minassian et al. 2006). When using CPTs and a single

bi-allelic locus, there are 15 possible offspring–maternal–paternal genotype combinations. Under the null model of no genetic effects on the phenotype, Mendelian transmission holds and the number of independent parameters reduces to eight, one less than the number of maternal–paternal genotype combinations (mating types). If one assumes the sex of the parent is irrelevant in determining the probability of the mating types (the symmetric mating assumption), then the nine combinations reduce to six. If random mating with regards to the locus holds, then the mating types can be parameterized in terms of the three genotype frequencies and leads to two independent parameters to estimate. The number of independent parameters under the null further reduces to one if Hardy Weinberg Equilibrium is assumed.

Maternal and offspring genotype effects are estimated as genotype relative risks along with mating-type frequencies. Table 11.1 presents the same mother–offspring combinations and genotype relative risks for a bi-allelic locus as can be modeled with CMCM data. Columns 3 and 4 present two mathematically equivalent parameterizations for the most general model of maternal-offspring effects with a bi-allelic locus. Although column 3, the joint risk model, has seven parameters, the maximum number of maternal–offspring parameters that can be estimated is six because one of these joint risks is the referent with value one. Column 4 is parameterized in terms of maternal main effects, offspring main effects, and two MFG incompatibilities. We also present two examples of restrictions. The model in column 5 represents *RHD* incompatibility without offspring or maternal main effects. Column 6 represents NIMA effects along with offspring genotype effects (Hsieh et al. 2007). All of these models are available for testing in the MFG option of the Mendel Statistical Genetics Software Package (Lange et al. 2013).

The log-linear and equivalent multinomial approaches can also test for the existence of PoO effects in the possible presence of maternal and offspring effects (Ainsworth et al. 2011; Weinberg et al. 1998). These authors caution against over-parameterization and discuss the problem of multiple interpretations.

Although CPTs have much to offer, many families have multiple-affected offspring and including only one of these offspring is inefficient. In order to use any number of affected siblings per family, Kraft et al. (2004) used a conditional retrospective likelihood approach. This approach finds the likelihood of the genotypes conditional on the affection status of the siblings and can estimate the offspring, maternal (or paternal) genotype effects, and their interactions by including these effects in the penetrance function. Families where one or both parents have missing genotypes are included in the likelihood by summing over all possible genotypes for the missing parents. Unaffected siblings are treated as phenotype unknown. The genotypes of these unaffected or phenotype unknown offspring can be included in the likelihood to help infer the possible genotypes for missing parents without introducing any bias provided the disease is not too common (Hsieh et al. 2006a). If the locus under study is causal, is unlinked to other causal loci and there are no gene–environment interactions, then the penetrance functions of the offspring are independent conditional on their own genotype and that of their mothers. The maximum likelihood estimates of the relative risks and the

$$L(\mathbf{G} \mid \mathbf{D}) = \frac{\sum_{g_1} \cdots \sum_{g_n} \prod_j \text{Prior}(g_j) \prod_i \text{Pr}(G_i \mid g_i) \prod_{[c,\,r,\,s]} \text{Pr}(D_c \mid g_c, g_r) \text{Trans}(g_c \mid g_r, g_s)}{\sum_{g_1} \cdots \sum_{g_n} \prod_j \text{Prior}(g_j) \prod_{[c,\,r,\,s]} \text{Pr}(D_c \mid g_c, g_r) \text{Trans}(g_c \mid g_r, g_s)}$$

Fig. 11.3 The retrospective likelihood for a single, arbitrary pedigree. With multigenerational pedigrees, the number of mating types becomes impractically large so founders' genotypes are used under an assumption of random mating. Founder j's genotype frequency = Prior(g_j). These frequencies are estimated along with the other parameters in the likelihood. With CPTs or two generation nuclear families, Prior(g_j)'s are replaced with mating-type frequencies MT(g_r, g_s). Pr ($G_i \mid g_i$) = 1 if the proposed genotype for any pedigree member i, g_i, is consistent with the observed genotype G_i, and 0 otherwise. When G_i is missing, Pr($G_i \mid g_i$) = 1. Pr($D_c \mid g_c, g_r$) is the offspring c's disease probability dependent on both their and their mother's genotype. For computational ease, Pr($D_c \mid g_c, g_r$), is calculated with Trans($g_c \mid g_r, g_s$), the transmission probability for offspring, mother, and father triples (c,r,s). The denominator sums over all possible ordered (phased) genotypes for the n family members. The likelihoods of independent pedigrees multiply. When there are only CPTs, the denominator is constant and is not relevant to the inference. The likelihood of the study samples is then proportional to a 15-mer multinomial

mating-type frequencies are obtained by solving score equations of the sample log likelihood and the standard errors of the estimates are derived through the observed information matrix (Lange 2002). Null hypotheses are tested using likelihood ratio test statistics.

Besides allowing more data to be used, an advantage of using nuclear families is that prior exposure effects can be tested. In this case, the genotypes of unaffected or phenotype unknown siblings fulfill an additional role of defining prior exposure. Kraft et al. (2004) used nuclear families to test whether risk of schizophrenia increased for offspring who were *RHD* incompatible when their older sibling was also *RHD* incompatible and found support for this hypothesis.

This conditional retrospective likelihood approach can be extended for use with large pedigrees. Like the nuclear family test, the extended MFG incompatibility (EMFG) test examines both maternal and offspring genotypes as risk factors for disease. The EMFG test jointly models maternal genotype effects, offspring genotype effects, and maternal–offspring genotype interactions using a retrospective likelihood (see Fig. 11.3 for mathematical details). Childs et al. (2010, 2011) developed this approach to allow any pedigree to be used including those with multiple generations and multiple-affected individuals. To reduce the number of nuisance parameters, the EMFG test replaces mating types with founder genotypes and assumes random mating with respect genotypes among the founders. The EMFG test handles multi-allelic loci, including non-codominant loci and several tightly linked loci, and can also incorporate potential offspring-related confounders. The EMFG test likelihood uses the classic formulation of the pedigree likelihood (Ott 1974) and modifies it by (1) conditioning on the phenotypes and (2) using penetrance functions that depend on both the offspring and maternal genotypes (Fig. 11.3). Each pedigree has its own conditional likelihood and these conditional likelihoods multiply. Unaffected family members are treated as phenotype unknown. Although EMFG is an affected-only analysis, the genotypes of the

unaffected or phenotype unknown offspring are used when there are missing parental genotypes.

The conditional likelihood used by Kraft et al. (2004) and Childs et al. (2010, 2011) has additional attractive features. Under relatively mild assumptions regarding the conditional independence of environmental exposure and offspring genotypes given parental genotypes, the effects of environmental covariates can be incorporated into the models. Serotypes and other non-codominant markers can be used by treating the genotypes underling these phenotypes as missing data (Minassian et al. 2006). The conditional likelihood approach also has some disadvantages. When the locus under study is not the causal locus but is linked to the causal locus, the variance in the parameter estimates are underestimated, which leads to false-positive results unless a robust variance estimator is used (Kraft et al. 2005).

Although these single locus bi-allelic analyses provide insights, biological inference is limited. For example, the models discussed in the previous paragraphs assume that there are joint maternal–offspring genotype effects but no paternal genotype effects. With the same data, we could have equally plausibly tested for joint paternal–offspring genotype effects or main effects of maternal, paternal, and offspring genotypes. In fact there can be multiple mathematically equivalent parameterizations that have different biological interpretations. Although null hypotheses may be rejected, the statistical analyses cannot provide insights into which biological interpretation is the correct one. Thus it is important, when using these models for gene discovery, not to take the results of any parameterization too literally and recognize a number of alternative, equally plausible explanations may exist (see Sinsheimer et al. 2003 and Ainsworth et al. 2011 for details).

11.3.3 *Approaches Using Phenotype, Genetic, and Epigenetic Data*

Currently epigenetic data are scarce in epidemiological studies, particularly at the genome-wide level. The most commonly available genome-wide epigenetic data are DNA methylation profiles (Cortessis et al. 2012). Studies collecting these methylation profiles are still small in scale, chiefly because of the expense. The majority of studies use samples from unrelated individuals. Studies of relatives have mainly consisted of twin studies (Bell and Spector 2012; Bocklandt et al. 2011). The predominant use of twin studies is due to (1) the strong tradition of using twins in heritability studies, which provides a wealth of readily available analysis tools; (2) the expense of using full pedigree data; and (3) changes in DNA methylation profile over the course of an individual's lifetime making comparisons of relatives across generations more complicated than using twins.

The data are often expressed as the fraction of a specific CpG site that is methylated (see Laird 2010 for a review of technologies). In statistical analyses,

this fraction, called the beta value, is sometimes treated as an outcome (the ultimate phenotype of interest), and sometimes treated as an intermediate phenotype associated with an outcome. When treating the beta value as an outcome, all the existing quantitative trait analysis approaches, both for data from unrelated individuals or related individuals, can be used including penalized regression (Bocklandt et al. 2011). The heritability of beta values can be calculated by using pedigrees as well as by using twins (Bjornsson et al. 2008). One potential complication with pedigree data is the strong age dependence of the beta values at many DNA methylation sites (Bjornsson et al. 2008; Bocklandt et al. 2011); however in analogy to age dependence for clinical outcomes age can be included as a covariate (e.g., Watanabe et al. 1999; Kangas-Kontio et al. 2010).

A beta value can also be an intermediate phenotype (like a biomarker) of an outcome. Again there are statistical genetic methods that can use unrelated or related individuals and treat beta values as intermediate phenotypes in association studies (Cortessis et al. 2012). One question following from these association studies is: are these epigenetic changes causal or are they responses to the clinical phenotype? Statistical approaches for inferring causality such as Mendelian randomization (Thomas and Conti 2004), genetical genomics (Li et al. 2005), and structural equation modeling (Morris et al. 2010) provide frameworks for answering this question. Using genetic loci associated with the beta values can anchor the causal direction. These three statistical approaches are somewhat related and for space considerations, we focus on Mendelian randomization as it has been used most frequently for the epigenetic explorations.

In the epigenetic context, Mendelian randomization resolves the question of directionality between beta value and an outcome by examining the effect of introducing a genetic covariate, a proxy, into the analysis (Thomas and Conti 2004). The assumption is that this proxy is directly related to the beta value, but it is only indirectly related to the outcome. Thus the magnitude of the true causal effect of methylation at the CpG site on the outcome is the ratio of the magnitude of effect of the genotype on the outcome divided by the magnitude of the effect of genotype on the beta value.

Because other measured covariates such as age, sex, body mass index, or specific biomarkers like lipid levels can be associated with both the beta value and outcome, it may be hard to discern causality. For example, suppose there is an association of age with the clinical phenotype, and there is an association of age with the beta value at a specific CpG site. Is the age effect for the phenotype manifested through DNA methylation? One promising approach to answering this question is two-step Mendelian randomization (Relton and Davey Smith 2012). In this context, the biomarkers, age, etc. constitute exposures. In the first step, a genetic proxy associated with the exposure is used to determine the causality of the exposure for the beta value. In the second step, a different genetic proxy, independent of the first proxy and associated with the beta value, is used to determine the causality of DNA methylation at the CpG site for the outcome.

Of particular relevance to understanding transgenerational effects is that Mendelian randomization can be applied to family data (e.g., Morris et al. 2009). Two-step Mendelian randomization can also span generations. Relton and Davey Smith (2012) discuss the example of maternal alcohol use during pregnancy as the exposure, offspring methylation fraction at a particular CpG site as the intermediate phenotype and offspring cognition as the outcome. In this case, an appropriate genetic proxy for alcohol consumption is the mother's genotype at an associated locus and an appropriate genetic proxy for the beta value is the offspring's genotype at another locus, unlinked and independent of the first locus.

11.4 Discussion

Epidemiological study designs and statistical genetic approaches make it possible to detect transgenerational effects in humans. Table 11.2 summarizes the study samples presented and the nature of transgenerational effects that can be determined using them. Determining the correct form of the transgenerational effects using the epidemiological studies is difficult but the more genetic and epigenetic information available, the better the chances of differentiating between the possibilities. Researchers need to be mindful that even detailed epigenetic data are of limited value if the study design is inadequate. The model complexity cannot exceed what is possible given the study sample. For example, none of these transgenerational effects can be tested if the study sample is limited to unrelated cases and controls.

Table 11.2 Examples of study samples and research questions

| Study sample | Genetic effect | | | | | | |
	Offspring	Maternal	Paternal	Parent of origin	MFG	Long-range transgenerational	Prenatal, postnatal, maternal inherited
Case–control	Yes	No	No	No	No	No	No
CMCM	Yes	Yes	No	No	No	No	No
CFCF	Yes	No	Yes	No	No	No	No
CPT	Yes	Yes	Yes	Yes	Yes	No	No
Nuclear families	Yes	Yes	Yes	Yes	Yes	No	No
Extended Pedigrees	Yes	Yes	Yes	Yes	Yes	Yes	Yes[a]
Families using ART	Yes	Yes	Yes	Yes	Yes	Yes	Yes

CMCM case-mother, control-mother study sample, *CFCF* case-father, control-father study sample, *CPT* case-parent trio study sample
[a]Yes, if adopted offspring and offspring of sisters are included

Researchers should remember when analyzing data under particular hypotheses that more than one parameterization with different biological interpretations are mathematically equivalent or nearly equivalent. They should also remember that violation of the underlying (and sometime unstated) modeling assumptions may lead to rejection of the null hypothesis without the alternative hypothesis actually being true. For example, violation of the symmetric mating assumption will lead to false inference of maternal effects when analyzing genotype data with CPTs (Sinsheimer et al. 2003). When possible, these modeling assumptions should be checked. Independent mechanistic data from functional studies, in vitro or using model organisms, will be needed to move beyond these associations and resolve these alternative explanations. Despite these caveats, epidemiological data still provide us with strong evidence in support of the existence of transgenerational genetic effects in humans and their roles in complex disease. Moreover they generate hypotheses for further research into the mechanisms of these transgenerational effects.

References

Ainsworth HF, Unwin J, Jamison DL, Cordell HJ (2011) Investigation of maternal effects, maternal-fetal interactions and parent-of-origin effects (imprinting), using mothers and their offspring. Genet Epidemiol 35:19–45

Bell JT, Spector TD (2012) DNA methylation studies using twins: what are they telling us? Genome Biol 13:172

Benyshek DC, Martin JF, Johnston CS (2001) A reconsideration of the origins of the type 2 diabetes epidemic among Native Americans and the implications for intervention policy. Med Anthropol 20:25–64

Bjornsson HT, Sigurdsson MI, Fallin MD, Irizarry RA, Aspelund T, Cui H, Yu W, Rongione MA, Ekstrum TJ, Harris TB, Launer LJ, Eiriksdottir G, Leppert MF, Sapienza C, Gudnason V, Feinberg AP (2008) Intra-individual change over time in DNA methylation with familial clustering. JAMA 299:2877–2883

Bocklandt S, Lin W, Sehl ME, Sanchez FJ, Sinsheimer JS, Horvath S, Vilain E (2011) Epigenetic predictor of age. PLoS One 6:e14821

Cannon M, Jones P, Murray R (2002) Obstetric complications and schizophrenia: historical and meta-analytic review. Am J Psychiatry 159:1080–1092

Chen J, Zheng H, Wilson ML (2009) Likelihood ratio tests for maternal and fetal genetic effects on obstetric complications. Genet Epidemiol 33:526–538

Childs EJ, Palmer CG, Lange K, Sinsheimer JS (2010) Modeling maternal-offspring gene-gene interactions: the extended MFG test. Genet Epidemiol 34:512–521

Childs EJ, Sobel EM, Palmer CG, Sinsheimer JS (2011) Detection of intergenerational genetic effects with applications to HLA-B matching as a risk factor for schizophrenia. Hum Hered 72:161–172

Cordell HJ, Barratt BJ, Clayton DG (2004) Case/pseudocontrol analysis in genetic association studies: a unified framework for detection of genotype and haplotype associations, gene-gene and gene-environment interactions, and parent-of-origin effects. Genet Epidemiol 26:167–185

Cordell HJ (2004) Properties of case/pseudocontrol analysis for genetic association studies: effects of recombination, ascertainment, and multiple affected offspring. Genet Epidemiol 26:186–205

Cortessis VK, Thomas DC, Levine AJ, Breton CV, Mack TM, Siegmund KD, Haile RW, Laird PW (2012) Environmental epigenetics: prospects for studying epigenetic mediation of exposure-response relationships. Hum Genet 131:1565–1589

Dahlquist GG, Patterson G, Soltesz G (1999) Perinatal risk factors for childhood type 1 diabetes in Europe. The EURODIAB substudy 2 study group. Diabetes Care 22:1698–1702

Edwards TL, Gao X (2012) Methods for detecting and correcting for population stratification. Curr Protoc Hum Genet Chap1:1.22.1–14

Gluckman PD, Hanson MA, Beedle AS (2007) Non-Genomic transgenerational inheritance of disease risk. BioEssays 29:145–154

Gorlova OY, Lei L, Zhu D, Weng SF, Shete S, Zhang Y, Li WD, Price RA, Amos CI (2007) Imprinting detection by extending a regression-based QTL analysis method. Hum Genet 122:159–174

Guilmatre A, Sharp AJ (2012) Parent of origin effects. Clin Genet 81:201–209

Guyton AC (1981) Textbook of medical physiology. W. B. Saunders, Philadelphia, PA

Harney S, Newton J, Milicic A, Brown MA, Wordsworth BP (2003) Non-inherited maternal HLA alleles are associated with rheumatoid arthritis. Rheumatology 42:171–174

Hollister J, Laing P, Mednick S (1996) Rhesus incompatibility as a risk factor for schizophrenia in male adults. Arch Gen Psychiatry 53:19–24

Hsieh H, Palmer CGS, Sinsheimer JS (2006a) Allowing for missing data at highly polymorphic genes when testing for maternal, offspring and maternal–fetal genotype incompatibility effects. Hum Hered 62:165–174

Hsieh HJ, Palmer CG, Harney S, Chen HW, Bauman L, Brown MA, Sinsheimer JS (2007) Using the maternal–fetal genotype incompatibility test to assess non-inherited maternal HLA-DRB1 antigen coding alleles as rheumatoid arthritis risk factors. BMC Proc 1:S124

Hsieh HJ, Palmer CG, Harney S, Newton JL, Wordsworth P, Brown MA, Sinsheimer JS (2006b) The v-MFG test: investigating maternal, offspring and maternal–fetal genetic incompatibility effects on disease and viability. Genet Epidemiol 30:333–347

Insel BJ, Brown AS, Bresnahan MA, Schaefer CA, Susser ES (2005) Maternal–fetal blood incompatibility and the risk of schizophrenia in offspring. Schizophr Res 80:331–342

Juul-Dam N, Townsend J, Courchesne E (2001) Prenatal, perinatal, and neonatal factors in autism, pervasive developmental disorder-not otherwise specified, and the general population. Pediatrics 107:e63

Kangas-Kontio T, Kakko S, Tamminen M, von Rohr P, Hoeschele I, Juvonen T, Kere J, Savolainen MJ (2010) Genome scan for loci regulating HDL cholesterol levels in Finnish extended pedigrees with early coronary heart disease. Eur J Hum Genet 18:604–613

Khoury MJ, Beaty TH, Cohen BH (1993) Fundamentals of genetic epidemiology. Oxford University Press, Oxford

Kistner EO, Weinberg CR (2004) Method for using complete and incomplete trios to identify genes related to a quantitative trait. Genet Epidemiol 27:33–42

Klip H, Verloop J, van Gool JD, Koster ME, Burger CW, van Leeuwen FE, OMEGA Project Group (2002) Hypospadias in sons of women exposed to diethylstilbestrol in utero: a cohort study. Lancet 35:1102–1107

Kraft P, Hsieh H-J, Cordell HJ, Sinsheimer J (2005) A conditional-on-exchangeable parental-genotypes likelihood that remains unbiased at the causal locus under multiple-affected-sibling ascertainment. Genet Epidemiol 29:87–90

Kraft P, Palmer CGS, Woodward AJ, Turunen JA, Minassian S, Paunio T, Lonnqvist J, Peltonen L, Sinsheimer JS (2004) *RHD* maternal–fetal genotype incompatibility and schizophrenia: extending the MFG test to include multiple siblings and birth order. Eur J Hum Gen 12:192–198

Laird NM, Lange C (2010) The fundamentals of modern statistical genetics. Springer Verlag, New York, NY

Laird PW (2010) Principles and challenges of genomewide DNA methylation analysis. Nat Rev Genet 11:191–203

Lange K (2002) Mathematical and statistical methods for genetic analysis, 2nd edn. Springer Verlag, New York, NY

Lange K, Papp JC, Sinsheimer JS, Sripracha R, Zhou H, Sobel EM (2013) Mendel: the Swiss army knife of genetic analysis programs. Bioinformatics 29:1568–1570

Li H, Lu L, Manly KF, Chesler EJ, Bao L, Wang J, Zhou M, Williams RW, Cui Y (2005) Inferring gene transcriptional modulator relations: a genetical genomics approach. Hum Mol Genet 14:1119–1125

Minassian SL, Palmer CGS, Turunen JA, Paunio T, Lonnqvist J, Peltonen L, Woodward AJ, Sinsheimer JS (2006) Incorporating serotypes into family based association studies using the MFG test. Ann Human Gen 70:541–553

Morris NJ, Gray-McGuire C, Stein CM (2009) Mendelian randomization in family data. BMC Proc 15:S7, S45

Morris NJ, Elston RC, Stein CM (2010) A framework for structural equation models (SEM) in general pedigrees. Hum Hered 70:278–286

Nadeau JH (2009) Transgenerational genetic effects on phenotypic variation and disease risk. Hum Mol Genet 18:R202–R210

Ott J (1974) Estimation of the recombination fraction in human pedigrees: efficient computation of the likelihood for human linkage studies. Am J Hum Genet 26:588–597

Palmer CGS, Hsieh H, Reed EF, Lonnqvist J, Peltonen L, Woodward JA, Sinsheimer JS (2006) HLA-B maternal–fetal genotype matching increases risk of schizophrenia. Am J Hum Genet 79:710–715

Palmer CGS, Turunen JA, Sinsheimer JS, Minassian S, Paunio T, Lonnqvist J, Peltonen L, Woodward JA (2002) *RHD* maternal–fetal genotype incompatibility increases schizophrenia susceptibility. Am J Hum Genet 71:1312–1319

Relton CL, Davey Smith G (2012) Two-step epigenetic Mendel randomization: a strategy for establishing the causal role of epigenetic processes in pathways to disease. Int J Epidemiol 41:161–176

Robson EB (1955) Birth weight in cousins. Annals Hum Genet 262–8

Sham PC, Curtis D (1995) An extended transmission/disequilibrium test (TDT) for multi-allele marker loci. Ann Hum Genet 59:323–336

Sinsheimer JS, Blangero J, Lange K (2000) Gamete-competition models. Am J Hum Genet 66:1168–1172

Sinsheimer JS, Palmer CGS, Woodward AJ (2003) The maternal–fetal genotype incompatibility test: detecting genotype combinations that increase risk for disease. Genet Epidemiol 24:1–13

Spielman RS, McGinnis RE, Ewens WJ (1993) Transmission test for linkage disequilibrium: the insulin gene region and insulin dependent diabetes mellitus (IDDM). Am J Hum Genet 52: 506–516

Stratchen T, Reed A (2003) Human molecular genetics, 3rd edn. Garland Science/Taylor & Francis Group, Oxford

Stubbs EG, Ritvo ER, Mason-Brothers A (1985) Autism and shared parental HLA antigens. J Am Acad Child Psych 24:182–185

Terwilliger JD, Ott J (1992) A haplotype based 'haplotype relative risk' approach to detecting allelic associations. Hum Hered 42:337–346

Thapar A, Harold G, Rice F, Ge X, Boivin J, Hay D, van den Bree M, Lewis A (2007) Do intrauterine or genetic influences explain the foetal origins of chronic disease? A novel experimental method for disentangling effects. BMC Med Res Methodol 7:25

Thomas DC (2004) Statistical methods in genetic epidemiology. Oxford University Press, Oxford

Thomas DC, Conti DV (2004) Commentary: the concept of 'Mendelian Randomization'. Int J Epidemiol 33:21–25

Vermeulen SH, Shi M, Weinberg CR, Umbach DM (2009) A hybrid design: case-parent triads supplemented by control-mother dyads. Genet Epidemiol 33:136–144

Watanabe RM, Valle T, Hauser ER, Ghosh S, Eriksson J, Kohtamèki K, Ehnholm C, Tuomilehto J, Collins FS, Bergman RN, Boehnke M (1999) Familiality of quantitative metabolic traits in Finnish families with non-insulin-dependent diabetes mellitus. Finland-United States Investigation of NIDDM Genetics (FUSION) Study investigators. Hum Hered 49:159–168

Wheeler E, Cordell HJ (2007) Quantitative trait association in parent offspring trios: extension of case/pseudocontrol method and comparison of prospective and retrospective approaches. Genet Epidemiol 31:813–833

Weinberg CR, Wilcox AJ, Lie RT (1998) A log-linear approach to case-parent-triad data: assessing effects of disease genes that act either directly or through maternal effects and that may be subject to parental imprinting. Am J Hum Genet 62:969–978

Zhou JJ, Pelka S, Lange K, Palmer CG, Sinsheimer JS (2011) Dissecting prenatal, postnatal and inherited effects: ART and design. Genet Epidemiol 35:437–446

Chapter 3

The Quantitative-MFG Test: A Linear Mixed Effect Model to Detect Maternal-Offspring Gene Interactions

Michelle M. Clark[1], John Blangero[2], Thomas D. Dyer[2], Eric M. Sobel[3] and Janet S. Sinsheimer[1,3,4]*

[1]*Department of Biostatistics, Fielding School of Public Health, University of California, Los Angeles, CA, USA*
[2]*South Texas Diabetes and Obesity Institute, School of Medicine, University of Texas, Rio Grande Valley, Brownsville, TX, USA*
[3]*Department of Human Genetics, David Geffen School of Medicine, University of California, Los Angeles, CA, USA*
[4]*Department of Biomathematics, David Geffen School of Medicine, University of California, Los Angeles, CA, USA*

SUMMARY

Maternal-offspring gene interactions, aka maternal–fetal genotype (MFG) incompatibilities, are neglected in complex diseases and quantitative trait studies. They are implicated in birth to adult onset diseases but there are limited ways to investigate their influence on quantitative traits. We present the quantitative-MFG (QMFG) test, a linear mixed model where maternal and offspring genotypes are fixed effects and residual correlations between family members are random effects. The QMFG handles families of any size, common or general scenarios of MFG incompatibility, and additional covariates. We develop likelihood ratio tests (LRTs) and rapid score tests and show they provide correct inference. In addition, the LRT's alternative model provides unbiased parameter estimates. We show that testing the association of SNPs by fitting a standard model, which only considers the offspring genotypes, has very low power or can lead to incorrect conclusions. We also show that offspring genetic effects are missed if the MFG modeling assumptions are too restrictive. With genome-wide association study data from the San Antonio Family Heart Study, we demonstrate that the QMFG score test is an effective and rapid screening tool. The QMFG test therefore has important potential to identify pathways of complex diseases for which the genetic etiology remains to be discovered.

Introduction

Maternal and offspring gene interaction, also termed maternal–fetal genotype (MFG) incompatibility, occurs when the effects of maternal genes on the offspring's phenotype vary depending on the offspring's genotype. The possibility of joint maternal and offspring effects needs to be studied, especially when investigating genetic factors of developmental disorders and their associated quantitative traits. Previous studies have found that MFG interactions are associated with preterm birth, conotruncal heart defects, neural tube defects, and preeclampsia (see as examples, Liang et al., 2010, Li et al.,

2014, Lupo et al., 2014, Procopciuc et al., 2014). Additionally, MFG incompatibilities have been implicated as risk factors in complex adult onset diseases, such as schizophrenia, where the effects are not evident until long after the MFG incompatibility initiated event has occurred and subsided (see as examples, Stubbs et al., 1985, Hollister et al., 1996, Dahlquist et al., 1999, Juul–Dam et al., 2001, Cannon et al., 2002, Palmer et al., 2002, Newton et al., 2004, Insel et al., 2005, Palmer et al., 2008, Freedman et al., 2011). To date, studies have not looked at the role of MFG incompatibility on the quantitative traits related to these adult onset diseases. Although there are a number of methods for investigating MFG incompatibility as a risk factor for disease (see Sinsheimer & Creek, 2013 for a review of these methods), the proposed methods to investigate the effects of maternal and offspring genes on quantitative traits typically rely on retrospective likelihoods and are limited to case–parent trios (Kistner & Weinberg, 2004, Kistner & Weinberg, 2005, Wheeler & Cordell, 2007). Moreover,

*Corresponding author: Janet S. Sinsheimer, Department of Human Genetics, David Geffen School of Medicine at University of California Los Angeles, Los Angeles, 5357C Gonda, Los Angeles, California 90095, USA. Tel: (310) 825-8002; Fax: (310) 825-8685; E-mail: janetS@mednet.ucla.edu

the retrospective likelihood design is not easily generalized to arbitrary family structures, multiple markers, or multivariate traits (Kraft et al., 2004) and parameter interpretation can be challenging.

One way to conduct association testing with quantitative traits using pedigree data is in a measured genotype analysis (Boerwinkle et al., 1986, Lange, 2002). This method of testing uses a linear mixed model (LMM) in which the genotypes are fixed effects and familial correlations are taken into account through partitioning the variance. Hence, the LMM is also called variance component modeling in the genetic literature. We have developed the quantitative-MFG (QMFG) test, an extension to the LMM where the joint maternal and offspring effects including MFG incompatibilities are fixed effects, familial correlations are variance components, and the outcome is a trait with residuals that are reasonably modeled as normally distributed (Lange, 2002). This approach handles pedigrees of virtually any size, both general and specific scenarios of MFG incompatibility, multivariate traits, and covariates in a straightforward manner. Another advantage of this approach is the ability to quickly test genome-wide association study (GWAS) pedigree data for joint maternal and offspring effects including MFG incompatibility via the use of the score test.

Materials and Methods

The QMFG Test

Recall that, for a single pedigree, the general multivariate normal loglikelihood for a LMM is

$$L = -\frac{1}{2} \ln |\Omega| - \frac{1}{2} (\gamma - \nu)^t \Omega^{-1} (\gamma - \nu)$$

with observed trait vector γ, mean vector ν, and covariance matrix Ω (see for example, Lange, 2002 for details regarding the variance component model in classic genetic applications). We propose an extension to this model where the maternal-offspring genotypes are fixed effects. In the QMFG test, $\nu = A\beta$, where A is the design matrix consisting of indicator variables for the MFG combinations of interest and β is the column vector of corresponding regression coefficients. In our applications, ν always includes a grand mean μ so there is one entry of β that equals μ and one column of A is all ones. Consider the effects of a single SNP with a reference allele and a variant allele. When modeling the joint effects of maternal and offspring genotype effects, let $\beta_{a_m a_c}$ denote the difference in the offspring's quantitative trait value from the grand mean for a mother with a_m variant alleles and an offspring with a_c variant alleles at a given SNP. Because there are seven possible mother-offspring genotype combinations for a biallelic locus (see Table 1), in the general MFG incompati-

Table 1 QMFG model parameterizations. G_{Mat} and G_{Off} denote the maternal and offspring genotypes, respectively.

G_{Mat}	G_{Off}	General QMFG model	RHD effects	NIMA and offspring effects
1/1	1/1	β_{00}	β_{00}	β_{00}
1/1	1/2	β_{01}	β_{00}	$\beta_{.1}$
1/2	1/1	β_{10}	β_{00}	β_{10}
1/2	1/2	β_{11}	β_{00}	$\beta_{.1}$
1/2	2/2	β_{12}	β_{00}	$\beta_{.2}$
2/2	1/2	β_{21}	β_{21}	$\beta_{.1}$
2/2	2/2	β_{22}	β_{00}	$\beta_{.2}$

bility case for one SNP, the vector of regression coefficients is $\beta^t = (\mu, \ \beta_{00}, \ \beta_{01}, \ \beta_{10}, \ \beta_{11}, \ \beta_{12}, \ \beta_{21}, \ \beta_{22})$ and the additional columns of the design matrix A are indicator variables corresponding to each of the seven possible MFG combinations. Note that just as additional covariates such as age and sex can be incorporated in the standard measured genotype analysis (Boerwinkle et al., 1986), they can be included in the fixed effect portion of the QMFG model as additional entries in the β vector and additional columns of the A matrix. To avoid nonidentifiability, one of the parameters for the MFG effects should be made the reference state or, equivalently, the sum of MFG parameters should be set to some constant. In our analyses, parameter β_{00}, denoting zero copies of the variant allele in both mother's and offspring's genotypes, is always set to zero and hence, at most six MFG parameters are estimated along with the grand mean.

We continue to treat familial correlations as random effects by partitioning the residual variance. Here, we define the partition of the covariance matrix as $\Omega = \sum_{i=1}^{k} \sigma_i^2 \Gamma_i$, where k is the number of variance components included in the model. Often in genetic studies a very simple version of this matrix with only two components is used, one representing the additive genetic effects and one representing environmental random effects. In this model, the familial correlations are assumed to be due to small and approximately equal effects of alleles at a number of genes each acting independently. The additive genetic and environmental variances are denoted by σ_a^2 and σ_e^2, respectively. The design matrix Γ_a corresponding to σ_a^2 is twice the global kinship coefficient matrix Φ. Each element Φ_{ij} is the probability that, at a randomly chosen autosomal locus, an allele chosen at random from subject i and an allele chosen at random from subject j match identically by descent. When i equals j the alleles are chosen with replacement. The environmental contribution is multiplied by the identity matrix I since the environment is assumed to affect each subject independently. The environmental variance σ_e^2 is always included even when there are thought to be no

environmental factors to insure that the matrix is positive definite. Under this simple model, $\Omega = 2\sigma_a^2 \Phi + \sigma_e^2 I$.

As in other LMM scenarios, likelihood ratio tests (LRTs) can be used here to determine the significance of MFG parameters. Asymptotically, the LRT statistic follows a χ^2 distribution with degrees of freedom equal to the difference in the number of parameters under the null and alternative models. In addition to the LRT, we can use score tests to rapidly screen markers (Chen & Abecasis, 2007). The score statistic is given by

$$S(\theta) = dL(\theta) J(\theta)^{-1} \nabla L(\theta),$$

where $\nabla L(\theta)$ is the gradient of the loglikelihood with the parameter vector θ, $dL(\theta)$ is the first differential of the loglikelihood, and $J(\theta)$ is the expected information matrix. Zhou et al. (2015) pre-compute and store key quantities for a fast score test for individual SNPs. In particular, for family i under the alternative hypothesis, the design matrix A_i can be written as (a_i, N_i) where N_i is the design matrix under the null. Additional covariates are included in the matrix N_i. The array a_i conveys the genotypes at the SNP of interest. Let the residual for family i be $r_i = y_i - N_i\hat{\beta}$, where $\hat{\beta}$ are the maximum likelihood estimates (MLEs) of the fixed effects under the null (in which no SNPs are included in the model). The score statistic for n families then reduces to

$$S = R^t \left[Q - W^t \left(\sum_{i=1}^{n} N_i^t \Omega_i^{-1} N_i \right)^{-1} W \right]^{-1} R,$$

where

$$Q = \sum_{i=1}^{n} a_i^t \Omega_i^{-1} a_i, \quad R = \sum_{i=1}^{n} a_i^t \Omega_i^{-1} r_i, \quad W = \sum_{i=1}^{n} N_i^t \Omega_i^{-1} a_i.$$

Thus, the quantities Ω_i^{-1}, $\Omega_i^{-1} r_i$, $N_i^t \Omega_i^{-1}$, and $\sum_{i=1}^{n} N_i^t \Omega_i^{-1} N_i$ can be computed once under the null model and then reused for the analysis of each SNP. This makes the calculation of the score test statistic for each SNP simple and rapid.

We extend this fast calculation of the score test to MFG incompatibility by replacing a_i, which previously was a vector conveying the variant allele counts at the SNP of interest for family i, with the matrix X_i. In its most general form, each column of matrix X_i represents one of the possible MFG combinations and is composed of zeros and ones, indicating which MFG combination defines the joint mother-offspring genotype for each offspring within the pedigree at a particular SNP. For example, for the pedigree in Figure 1

$$X_i = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \end{pmatrix}$$

and for the pedigree in Figure 2
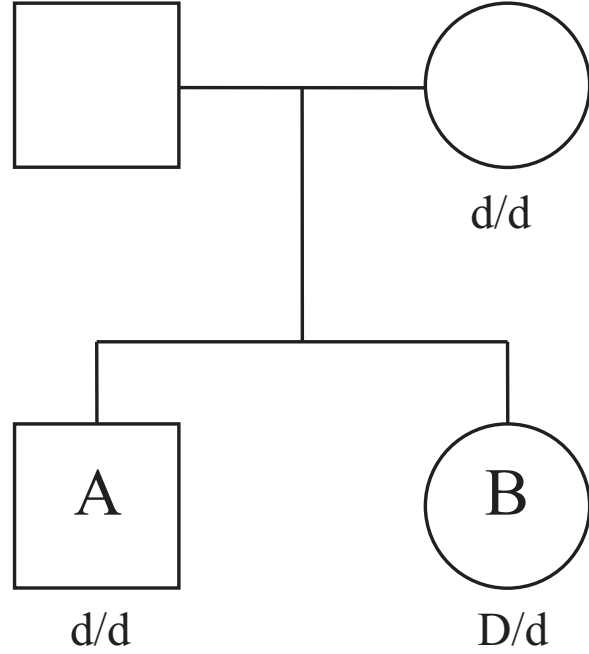
Figure 1 Pedigree depiction of the Rhesus factor D (RHD) scenario. The mother has two variant alleles (d/d). Offspring A has two variant alleles (d/d). Offspring B has one variant allele (D/d) and is therefore RHD incompatible with the mother.

$$X_i = \begin{pmatrix} 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

In these examples, the top row corresponds to offspring A and the bottom row corresponds to offspring B and each column corresponds to the MFG combinations in Table 1. The columns of X_i can be combined depending on the restrictions imposed on the parameters and therefore the number of columns equals the number of parameters to be estimated. As with the LRT, the score test statistic asymptotically follows a χ^2 distribution with degrees of freedom determined by the difference between the number of parameters in the null and alternative models.

There are advantages and disadvantages to both the LRT and the fast score test. For example, dealing with complex null and alternative hypotheses is better handled with the LRT. However, to calculate the LRT iterative maximization of the likelihood under both the null and alternative models is required, which can be computationally intensive when using large numbers of extended pedigrees and markers. We evaluate the ability of both tests to make correct QMFG inferences. Another advantage of the LRT is that in maximizing the likelihood we obtain parameter estimates. We also evaluate the accuracy and precision of these estimates.

SE+/SE−

A

B

SE+/SE+   SE−/SE−

Figure 2 Pedigree depiction of the noninherited maternal antigen (NIMA) scenario. The mother has one variant allele encoding a shared epitope (SE+/SE−). Offspring A has two variant alleles encoding a shared epitope (SE+/SE+). Offspring B has no variant alleles encoding a shared epitope (SE−/SE−) and therefore illustrates the NIMA scenario.

Examples of MFG Incompatibility

To better illustrate how the QMFG model works, we present two well-known examples of MFG incompatibility. Although these two examples are typically framed in terms of disease, they could easily be imagine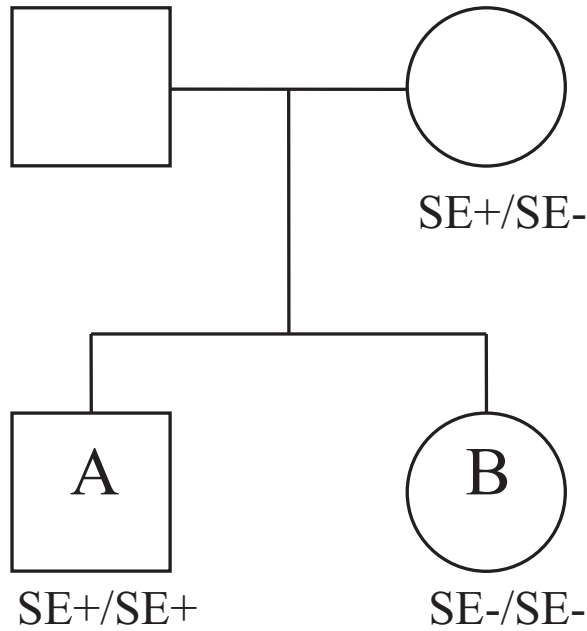d to be operating on associated quantitative traits. First, we consider the case where the mother reacts to antigens created by the offspring. The prototypical example is RHD incompatibility (Fig. 1), which occurs if the mother is homozygous for the variant allele "d" (RHD-negative) and the offspring is heterozygous (RHD-positive). This can lead to hemolytic disease of the newborn (Levine et al., 1941), which is associated with high levels of bilirubin resulting from the breakdown of the fetus's red blood cells (Lee et al., 2009). In our simulation study we estimate parameter β_{21}, which denotes the expected change in the quantitative trait value of the offspring when mother and offspring are RHD incompatible, and the other six MFG parameters form the reference group, i.e., $\beta_{00} = \beta_{01} = \beta_{10} = \beta_{11} = \beta_{12} = \beta_{22} = 0$ (Table 1, column 4).

The second example we investigate is the case where the offspring's immune system reacts to an antigen that has its origins in the mother's genotype. It is inspired by rheumatoid arthritis (RA) and HLA-DRB1, where noninherited maternal antigens (NIMA) have been implicated in offspring disease susceptibility (van der Horst-Bruinsma et al., 1998, Harney et al., 2003, Newton et al., 2004). As an example of an associated quantitative trait, anti–CCP antibodies are important markers for diagnosis and prognosis in RA since they are highly specific and sensitive (Visser et al., 2002, Silveira et al., 2007) and therefore would be an interesting quantitative trait to investigate using the QMFG test. A pedigree depiction of NIMA is shown in Figure 2 and is characterized by a mother that has one variant allele encoding a shared epitope (SE-positive) and an offspring that has none (SE-negative). There is strong evidence that there is an effect when the offspring has one or more variant alleles regardless of the mother's genotype (Gregersen et al., 1987, Jawaheer & Gregersen, 2002) and thus offspring effects must be included in the model. This model allows us to show how more complex restrictions can be handled. As shown in Table 1, column 5, the effects of interest are the NIMA effect (β_{10}) and the offspring genotype effects ($\beta_{.1} = \beta_{01} = \beta_{11} = \beta_{21}$ and $\beta_{.2} = \beta_{12} = \beta_{22}$).

Mendel Software

We implement the QMFG test by modifying the statistical genetics software package Mendel (Lange et al., 2013). When using SNPs, one allele is considered the reference allele and the other is the derived, variant allele. In order to implement the measured genotype analysis option for MFG incompatibility, the Mendel code was updated to extract the variant allele counts for mother and offspring from the genotypes included in the pedigree files. The reference allele is by default the more frequent allele but this can be changed if the user specifies in the Mendel control file. Once variant allele counts are determined, the LRT option is run by internally including a new covariate for each offspring that indicates which of the seven possible maternal-fetal gene–gene combinations describes the offspring's and his mother's genotypes. This enables MFG incompatibility parameters to be estimated and the likelihood to be calculated within the variance component analysis option in Mendel. The user can place restrictions on parameter estimates thus allowing for specific forms of MFG incompatibility such as offspring antigen–maternal antibody (referred in this document by the prototypical example RHD incompatibility) or maternal antigen–offspring antibody (referred in this document by the prototypical example NIMA). To program the QMFG score test, we used much of the machinery in the existing ped-GWAS option in Mendel that implements an LMM-based fast score test for GWAS on pedigree data with quantitative traits (Zhou et al., 2015). We forced the existing algorithm to include

in its model the seven possible MFG combinations and any user-specified restrictions on these combinations.

Simulation of Pedigrees

To evaluate the type I error or power of the QMFG tests and, for the LRT implementation, parameter estimation, we simulate data under ten specific scenarios (A-J) using the parameters shown in Table 2. Simulation A data are under the null of no genetic effects. Simulations B, E, and H involve conditions consistent with the effects of RHD incompatibility. Simulations C, D, F, and I are consistent with the effects of NIMA with and without the additional effects of offspring alleles and maternal effects. Simulation G is another possible scenario where each variant allele in the mother or offspring has the same effect on the phenotype and the effects are additive and independent; it is a special case of a scenario where there are both maternal and offspring main effects but no interaction. We use Simulation G when we want to evaluate the properties of fitting the general model. Simulation J involves offspring effects only and is used to investigate model misspecification. For power analyses, we vary selected parameters of particular interest in these simulation scenarios. The simulation design consists of 2000 repetitions of 1000 three-generational pedigrees (except when studying the effect of family structure or sample size), a biallelic locus, and a quantitative trait. Every three-generational pedigree is comprised of a nuclear family with two offspring, each of which have a partner and child of their own, and therefore the extended family contains a total of four founders and four offspring. Unless otherwise specified, the variant allele frequency is 0.40. Genotypes are simulated using Mendel's gene dropping option. Additional effects include a grand mean (intercept) $\mu = 40$ and variance components $\sigma_a^2 = 1$ and $\sigma_e^2 = 5$ (residual heritability $h^2 = 0.167$) unless otherwise specified. A univariate quantitative trait is simulated for all offspring by modifying the trait simulation option of Mendel (Lange et al., 2013).

Assessing the Statistical Properties of the Tests

We use a variety of full and reduced models (Table 3, Models 1–9) to analyze the simulated data, the exact choice depending on our alternative and null hypotheses. All models fit to the data estimate a grand mean (μ) and both variance components (σ_a^2 and σ_e^2). To quantify the degree of bias in type I error, genomic control values (λ) are reported (Devlin & Roeder, 1999) and 95% confidence bounds are included on Q–Q plots. Confidence bounds are based on the standard errors of the order statistics of the comparison distribution (Fox, 2008).

Power is defined as the rejection rate, which is the proportion of simulations in which the statistical test rejects the null

model in favor of the alternative. If not otherwise specified, we used a per test significance level of 0.001 when determining power. Standard errors of power estimates are calculated using $SE = \sqrt{\frac{p(1-p)}{N}}$ where p is the proportion of rejected tests and N is the number of repetitions. The proportion of variation explained is calculated as the ratio of phenotypic variation due to the effects of interest and the total phenotypic variation, and is based on the true parameter values and allele frequencies with which the data are simulated.

The San Antonio Family Heart Study

To show the feasibility of using this approach on a real pedigree-based GWAS dataset, we use data from the San Antonio Family Heart Study (SAFHS). These data have described elsewhere (Mitchell et al., 1996) but we briefly describe the subset of the data we use. The complete study data consist of 3637 subjects in Mexican American families of various sizes. High-density lipoprotein (HDL) levels were measured at up to three time points. For the first time point, 1397 individuals were phenotyped (Table S1). Of the subjects that were phenotyped, 1043 also have genotype data. To reduce the computational time used to impute missing genotypes for irrelevant members of the pedigree, we first trim the data to include only subjects with a quantitative trait measurement and their connecting relatives (Lange & Sinsheimer, 2004). Because the QMFG test is an offspring-only analysis, we are interested in the subset of 855 offspring in the data set that have both phenotype and genotype data. If an offspring's mother is completely ungenotyped before imputation, they were not used. Our analysis therefore involves the 419 offspring from 43 families with phenotype, genotype, and maternal genotype information. In this subset of data, the largest family has 176 members and five generations while the smallest family has eight members and three generations. Next, we use Mendel's imputation option to fill in all missing genotypes for subjects who have some existing SNP data (Ayers & Lange, 2008). Standard imputation programs, which do not take pedigree data into account, inevitably produce impossible maternal-offspring genotype combinations. We remove SNPs that have a maternal and offspring genotype combination of either 2/2 and 1/1, respectively, or 1/1 and 2/2, respectively. Additionally, we filter out any SNPs that have a minor allele count less than 10.

Results

RHD Incompatibility

We first compare type I error rates of the LRT and score test for a simple example of maternal-fetal gene

Table 2 Examples of QMFG data simulation scenarios.

	Simulation scenario	Constraints	Number of MFG parameters simulated	Simulation parameters						
				β_{00}	β_{01}	β_{10}	β_{11}	β_{12}	β_{21}	β_{22}
A	No genetic effects	$\beta_{ij} = 0$ for all $i = 0, 1, 2$ and $j = 0, 1, 2$	0	0	0	0	0	0	0	0
B	RHD effect	$\beta_{ij} = 0$ for all i, j unless $i = 2$ & $j = 1$	1	0	0	0	0	0	0.55^*	0
C	NIMA & offspring effects	$\beta_{00} = 0$ $\beta_{01} = \beta_{11} = \beta_{21}$ $\beta_{12} = \beta_{22}$	3	0	0.18	0.60^*	0.18	0.36	0.18	0.36
D	NIMA, offspring, & maternal effects	$\beta_{00} = 0$	6	0	0.18	0.60	0.36	0.54	0.54	0.72
E	RHD effect	$\beta_{ij} = 0$ for all i, j unless $i = 2$ & $j = 1$	1	0	0	0	0	0	1.75	0
F	NIMA & offspring effects	$\beta_{00} = 0$ $\beta_{01} = \beta_{11} = \beta_{21}$ $\beta_{12} = \beta_{22}$	3	0	0.60	1.90	0.60	1.20	0.60	1.20
G	Count model	$\beta_{00} = 0$	6	0	0.17	0.17	0.34	0.51	0.51	0.68
H	RHD effect	$\beta_{ij} = 0$ for all i, j unless $i = 2$ & $j = 1$	1	0	0	0	0	0	0.70	0
I	NIMA effect	$\beta_{ij} = 0$ for all i, j unless $i = 1$ & $j = 0$	1	0	0	0.70^*	0	0	0	0
J	Offspring effects	$\beta_{00} = \beta_{10} = 0$ $\beta_{01} = \beta_{11} = \beta_{21}$ $\beta_{12} = \beta_{22}$	2	0	0.27	0	0.27	0.54	0.27	0.54

*Values varied for power analysis

Table 3 QMFG full and reduced models.

	Models	Number of MFG parameters estimated	MFG parameters estimated
1	No genetic effects model	0	None
2	RHD effect model	1	β_{21}
3	NIMA and offspring effects model	3	$\beta_{10}, \beta_{01} = \beta_{11} = \beta_{21}, \beta_{21} = \beta_{22}$
4	Offspring effects model	2	$\beta_{01} = \beta_{11} = \beta_{21}, \beta_{12} = \beta_{22}$
5	NIMA effect model	1	β_{10}
6	NIMA and dominant offspring effects model	2	$\beta_{10}, \beta_{01} = \beta_{11} = \beta_{21} = \beta_{12} = \beta_{22}$
7	NIMA, offspring, and maternal effects model	5	$\beta_{01}, \beta_{10}, \beta_{11}, \beta_{12}, \beta_{21}, \beta_{22} = \beta_{12} + \beta_{21} - \beta_{11}$
8	General model	6	$\beta_{01}, \beta_{10}, \beta_{11}, \beta_{12}, \beta_{21}, \beta_{22}$
9	Additive offspring effect model	1	$2\beta_{01} = 2\beta_{11} = 2\beta_{21} = \beta_{12} = \beta_{22}$

interactions that mimics RHD incompatibility. In this example, heterozygous offspring whose mothers are homozygous for the variant allele differ in their trait value from other offspring. Figure 3A shows the results when data are simulated under the null hypothesis of no genetic effects (Simulation A) and the parameter β_{21} is tested for significance. For the LRT, Model 1 reflects the null hypothesis and Model 2 reflects the alternative hypothesis, resulting in a one degree of freedom (df) test. The QMFG score test provides almost identical results (Figure S1A). Based on the confidence bounds on the Q-Q plot, we conclude there is little bias in the type I error for either test (genomic control value $\lambda = 1.065$). Next,
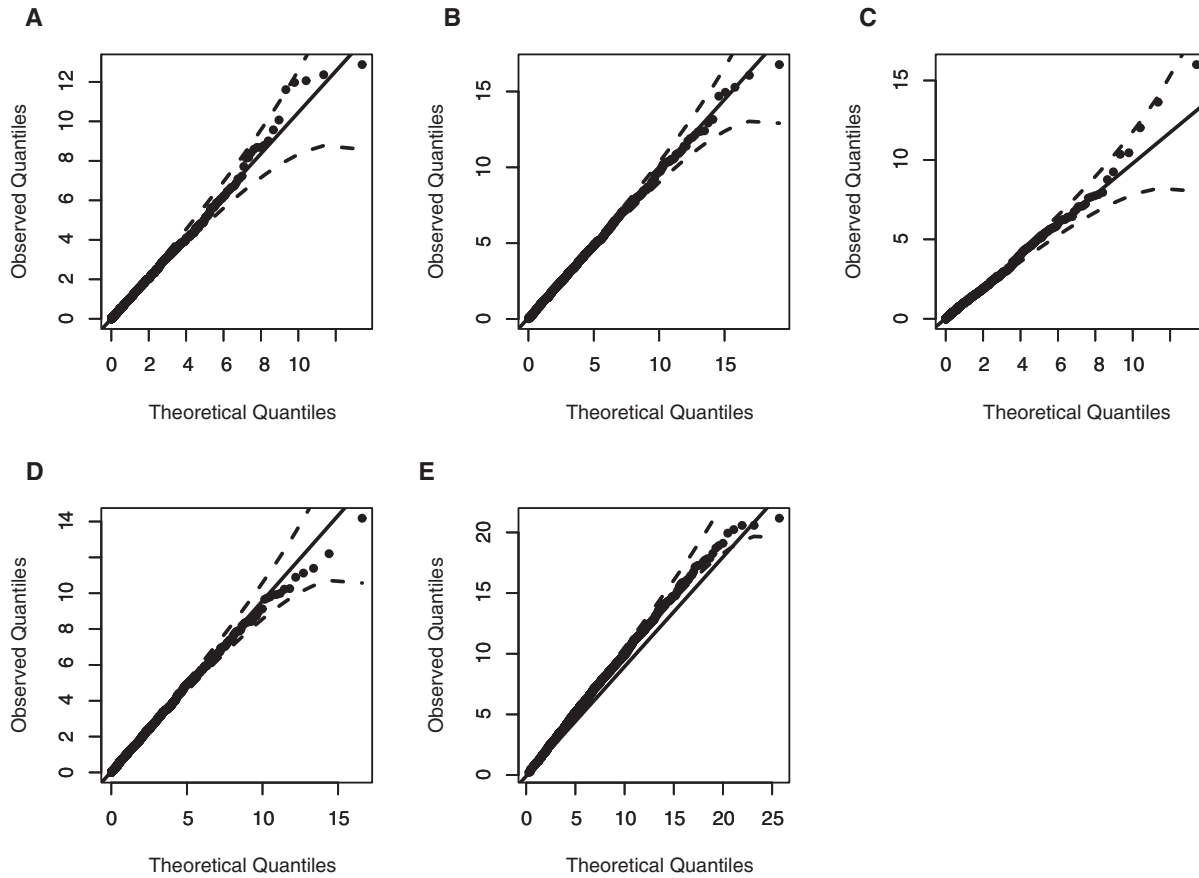
Figure 3 Q-Q plots for LRT using Simulation A data. Genotypes and quantitative traits for each replication were simulated for 1000 pedigrees under the null of no MFG effects and were tested with the LRT for (A) RHD effects (df = 1, λ = 1.065), (B) NIMA or offspring effects (df = 3, λ = 0.989), (C) NIMA effects in the presence of offspring genotype effects (df = 1, λ = 1.048), (D) offspring effects in the presence of NIMA effects (df = 2, λ = 0.959), and (E) any MFG effects (df = 6, λ = 1.086).

data are simulated to provide a moderately small RHD effect (Simulation B). In Simulation B, the only effect associated with the offspring's phenotype is an expected increase of 0.55 units when mother and offspring are incompatible, corresponding to 0.0044 of the variance explained by the RHD effect. Figure 4A provides boxplots of the parameter estimate bias over 2000 replicates when the underlying MFG mechanism, RHD incompatibility, is suspected and consequently the correct model, Model 2, is fit to the data. As desired, the QMFG method produces unbiased parameter estimates. Keeping all other simulation conditions the same, we then varied the true RHD effect between 0 and 0.7, resulting in the true proportion of variation explained by the MFG effect ranging from 0 to 0.007. Figure 5A shows the LRT and score test power curves corresponding to these variations for the one degree of freedom test (solid lines) testing the signifi-

cance of the RHD parameter (β_{21}). When the significance level is 0.001, the curve illustrates that 80% power is achieved when the proportion of variation explained by the RHD effect is approximately equal to 0.004 for both the LRT and score test.

Noninherited Maternal Antigen (NIMA) Effects

The case of NIMA provides a more complex model under which to investigate the properties of the QMFG test due to the added offspring allelic effects. For this more complicated case of MFG incompatibility, we start by comparing the LRT and score test type I error rates. Model 1 corresponds to the null hypothesis of no genetic effects and Model 3 corresponds to the alternative hypothesis of offspring or NIMA effects. By fitting Models 1 and 3 to data simulated with no genetic
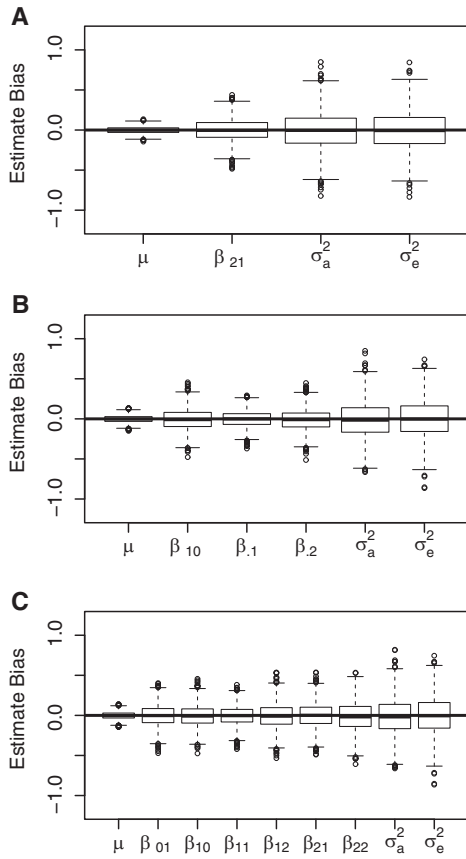
A

B

C

Figure 4 Parameter estimate bias. Genotypes and quantitative traits for each replication were simulated for 1000 pedigrees using (A) Simulation B data with an RHD effect of 0.55 ($\mu = 40$, $\beta_{21} = 0.55$, $\sigma_a^2 = 1$, $\sigma_e^2 = 5$), (B) Simulation C data with a NIMA effect of 0.60 and an additive offspring allelic effect of 0.18 ($\mu = 40$, $\beta_{10} = 0.60$, $\beta_{.1} = 0.18$, $\beta_{.2} = 0.36$, $\sigma_a^2 = 1$, $\sigma_e^2 = 5$), and (c) Simulation G data with a variant allele count effect of 0.17 ($\mu = 40$, $\beta_{01} = 0.17$, $\beta_{10} = 0.17$, $\beta_{11} = 0.34$, $\beta_{12} = 0.51$, $\beta_{21} = 0.51$, $\beta_{22} = 0.68$, $\sigma_a^2 = 1$, $\sigma_e^2 = 5$). Boxplots show bias of parameter estimates, additive genetic variance, and environmental variance over 2000 replications. A horizontal line is drawn at zero bias.

effects (Simulation A), the three degrees of freedom LRT results in the Q-Q plot in Figure 3B ($\lambda = 0.989$). Figure S1B shows that the score test for the significance of the same three parameters (β_{10}, $\beta_{.1}$, $\beta_{.2}$) is a suitable substitute for the LRT as it yields nearly identical P-values. As mentioned previously, using the LRT we can test the significance of NIMA in the presence of offspring effects or the offspring effect in the

presence of NIMA. Figure 3C is the Q-Q plot ($\lambda = 1.048$) of the results of testing the significance of the NIMA parameter while allowing for offspring genotype effects (Model 3 vs. Model 4) and Figure 3D is the Q-Q plot ($\lambda = 0.959$) of the results for offspring parameters in the presence of NIMA effects (Model 3 vs. Model 5). Together these three Q-Q plots (Fig. 3B–D) and confidence bounds demonstrate that our type I error rates are correct.

In Simulation C, a quantitative trait is simulated for offspring with a NIMA effect of 0.6 and an additive offspring allelic effect of 0.18 per allele, corresponding to 0.0058 of the variance explained by the combination of MFG and offspring effects. Mimicking a situation where there is *a priori* evidence for a particular model (in this case, NIMA and offspring effects), we fit Model 3 to the simulated data thus estimating three parameters (β_{10}, $\beta_{.1}$, $\beta_{.2}$) in addition to the variance components when using the LRT. The bias of each parameter estimate is shown in Figure 4B. As a valid method should, the QMFG method generates bias centered at zero. We further evaluate power over 2000 simulation replicates for varying levels of proportion of variation explained for this model (Model 3). The power curves in Figure 5B are the results of jointly testing for NIMA or offspring effects (solid lines) when the simulated NIMA effect sizes range from 0 to 0.7 and the offspring effects are $\beta_{.1} = 0.18$ and $\beta_{.2} = 0.36$. These simulation parameters are consistent with proportion of variation explained ranging from 0.002 to 0.008. The power to detect NIMA or offspring effects (three degrees of freedom test) is approximately 80% when proportion of variation explained reaches 0.0055 for both the LRT and the score test. Figure S2 displays the power of the LRT for testing the significance of the NIMA effect in the presence of offspring genotype effects (one degree of freedom test) using the same collection of data. Power is approximately 80% when the proportion of variation explained by the NIMA effect reaches 0.002.

The NIMA analyses mentioned above are free of constraints on the relationship between offspring allelic effect parameters $\beta_{.1}$ and $\beta_{.2}$, that is, a genotypic model. Offspring effects may act in an additive, recessive, or dominant manner. If there is *a priori* evidence to suggest any of these models, it is possible to impose restrictions on these parameters for the models that are fit, therefore reducing the degrees of freedom and increasing power. To demonstrate the ability of the QMFG LMM to handle such a situation, we simulate data with a NIMA effect varying from 0 to 0.6 and dominant offspring effects equal to 0.1 and analyze the data estimating again the NIMA and offspring effects, this time imposing the additional constraint $\beta_{.1} = \beta_{.2}$. Figure S3 shows the power curves resulting from testing for a NIMA or dominant offspring effect. For the LRT, this involves using Model 1 corresponding to the null hypothesis and Model 6 corresponding to the alternative
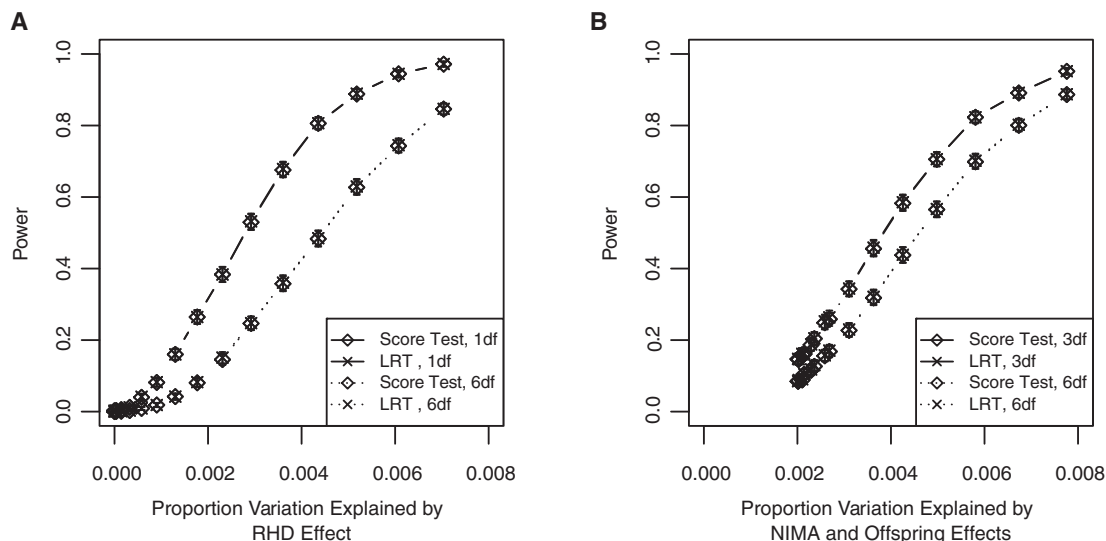
A

B

Figure 5 RHD incompatibility and NIMA power curves. (A) Solid lines show power for the one degree of freedom test for an RHD effect (β_{21}) using the LRT and score test. Dotted lines show power when there is no specific MFG hypothesis (so all six MFG parameters are tested) using the LRT and score test. RHD effect sizes range from 0 to 0.7. Error bars represent approximate 95% confidence intervals. (B) Solid lines show power for the three degrees of freedom test for NIMA or offspring effects ($\beta_{10}, \beta_{.1}, \beta_{.2}$) using the LRT and score test. Dotted lines show power when there is no specific MFG hypothesis (so all six MFG parameters are tested) using the LRT and score test. NIMA effect sizes β_{10} range from 0 to 0.7 and offspring effects are fixed ($\beta_{.1} = 0.18$, $\beta_{.2} = 0.36$). The power is not 0.001 and the proportion of variation explained is not zero when $\beta_{10} = 0$ because $\beta_{.1}, \beta_{.2}$ are not zero. Error bars represent approximate 95% confidence intervals.

hypothesis, resulting in a two degrees of freedom test. The power to detect these effects remains high. There is greater than 80% power to detect NIMA or dominant offspring effects when the proportion of the variance explained by these effects is 0.006.

Thus far, we have considered scenarios where just MFG and offspring effects are present. It is possible that in addition to these effects, there are maternal effects. To further demonstrate the flexibility of the QMFG method, we simulate data with both maternal and offspring main effects such that each variant allele further increases the offspring's trait by 0.18 (Simulation D). In this scenario, the proportion of variation explained by the NIMA, offspring, and maternal effects is 0.009. For this specific scenario, we fit Model 1 corresponding to the null of no genetic effects and Model 7 corresponding to our alternative hypothesis. Model 7 is a five parameter model that is mathematically equivalent to genotypic offspring main effects, genotypic maternal main effects and a NIMA effect. Parameterized through using the coefficients for seven possible maternal–offspring genotype combinations, Model 7 requires an additional constraint $\beta_{22} + \beta_{11} - \beta_{12} - \beta_{21} = 0$ in addition to $\beta_{00} = 0$. In other words, one of the parameters, $\beta_{11}, \beta_{12}, \beta_{21}, \beta_{22}$, is completely determined by the other three and so when using a LRT, five MFG parameters are

tested (e.g. $\beta_{01}, \beta_{10}, \beta_{11}, \beta_{12}, \beta_{21}$). Figure S4 shows that there is no parameter estimate bias when we fit Model 7 to the data. The power of the LRT for this five degrees of freedom test is 0.96 (SE = 0.004).

Effects of Allele Frequency and Variance Parameters

In the previous sections, the variant allele had a frequency of 0.4. For RHD incompatibility we consider "d" to be the variant allele and for NIMA we consider the alleles encoding a shared epitope (SE+) to be variant. To evaluate the effect of allele frequency on our QMFG method, we compare our power results for Simulations B and C while varying the variant allele frequency from 0.1 to 0.9. Figures S5A and S5B show the impact of changing the variant allele frequency keeping the other simulation parameters the same for Simulations B and C, respectively. The power is maximized when the frequency of the "d" allele is 0.67 for data simulated under RHD incompatibility. Under NIMA and offspring effects, power is maximized when the SE+ allele frequency is 0.33.

We also investigate the performance of the QMFG test when the values for additive genetic and environmental

40

variance are changed, increasing the residual heritability. We recreate the power curves for RHD incompatibility as well as for NIMA and offspring effects, this time changing both the additive genetic and environmental variance simulation values to 3 ($\sigma_a^2 = 3$, $\sigma_e^2 = 3$, $h^2 = 0.50$). Repeating the same one degree of freedom test for an RHD effect with these new variance values, we see an increase in power (Fig. S6A) over our previous results (Fig. 5A). Repeating the same three degrees of freedom test for NIMA or offspring effects using data simulated with these new variance values, we also see an increase in power (Fig. S6B) over our previous results (Fig. 5B).

Effects of Family Structure and Sample Size

To address the impact of family structure on our MFG tests, we simulate 4000 trios, keeping the total number of offspring at 4000, with quantitative traits with the same simulation parameters in Simulations B and C. With linear mixed models the additive and environmental variances are confounded when using a single individual per pedigree so instead of estimating σ_a^2 and σ_e^2 separately, we estimate their sum. Figure S7A shows there is no parameter estimate bias when fitting Model 2 to data simulated under RHD incompatibility (Simulation B). Power for either the LRT or the score test to detect an RHD effect (Model 2 vs. Model 1) is not significantly altered. It is 0.81 (SE = 0.009) for three-generational families and 0.79 (SE = 0.009) for trios. Parameter estimates also remain unbiased (Fig. S7B) when using trios with a quantitative trait simulated with NIMA and offspring effects (Simulation C) when fitting Model 3. Power with the three degrees of freedom test for NIMA or offspring effects (Model 3 vs. Model 1) is not significantly changed. It is 0.82 (SE = 0.009) for three-generational families and 0.85 (SE = 0.008) for parent-offspring trios.

It is important to also consider the effect of sample size on the statistical properties of the tests. Here, we run simulations with a smaller sample of 400 offspring from 100 three-generational families. As shown in Figure S8, type I error rates remain unaffected by a reduction in sample size. To achieve equivalent power to 1000 families, we need to increase the RHD effect to 1.75. Fitting Model 2 to 100 three-generational families simulated with quantitative traits given Simulation E parameters, results in no bias for grand mean, RHD, and environmental variance parameters (Fig. 6). Additive variance is slightly underestimated. The power to detect an RHD effect, which accounts for 0.042 of the trait variance in this scenario, is 0.78 (SE = 0.009). To achieve equivalent power to 1000 families, we need to increase effect sizes for NIMA and offspring effects (Model 3) to $\beta_{01} = 1.90$, $\beta_{.1} = 0.60$, $\beta_{.2} = 1.20$ when generating data for 100 three-generational families (Simulation F). We
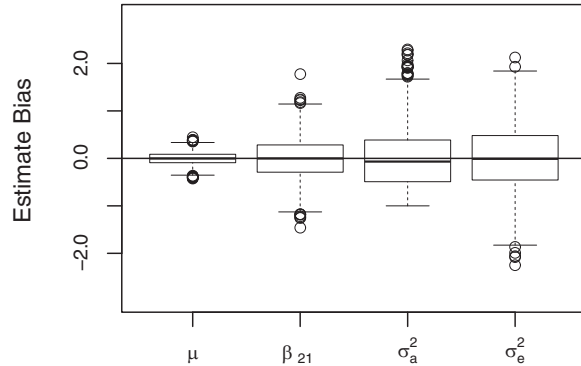

Figure 6 Parameter estimate bias when data are simulated under RHD incompatibility with a smaller sample size. Boxplots show bias of parameter estimates for the grand mean, MFG effects, additive genetic variance, and environmental variance using 100 three-generational families over 2000 replications using parameters from Simulation E with an RHD effect of 1.75 ($\mu = 40$, $\beta_{21} = 1.75$, $\sigma_a^2 = 1$, $\sigma_e^2 = 5$). A horizontal line is drawn at zero bias.

again see no bias for grand mean, NIMA, offspring, and environmental variance parameters (Fig. S9). Additive variance is again slightly underestimated over the 2000 repetitions. The proportion of variation explained by the NIMA and offspring effects in this scenario is 0.056 and the power for the three degrees of freedom test for NIMA or offspring effects is 0.82 (SE = 0.009).

Analysis When the MFG Mechanism Is Unknown

Our QMFG analyses thus far assume that our outcome is associated with two well-known mechanisms of MFG incompatibility. However, it may be that there is no *a priori* information about the underlying MFG model that influences a trait's value. Thus, we study the effects of using the general model, which imposes no constraints on the MFG parameters. First, we investigate the properties (type I error, power, and parameter estimates) of fitting such a model. Figure 3E shows the Q-Q plot for data simulated under the null hypothesis of no genetic effects (Simulation A) where all six MFG parameters β_{01}, β_{10}, β_{11}, β_{12}, β_{21}, β_{22} are tested for significance ($\lambda = 1.086$). This is a six degrees of freedom test where the full model (Model 8) is compared to the null model in which no MFG effects are estimated (Model 1) via the LRT. All the points lie within the confidence bounds; there is no bias in the type I error rate. The score test produces equivalent P-values (not shown). Parameter estimate bias is examined by simulating data given a count model (Simulation G). In
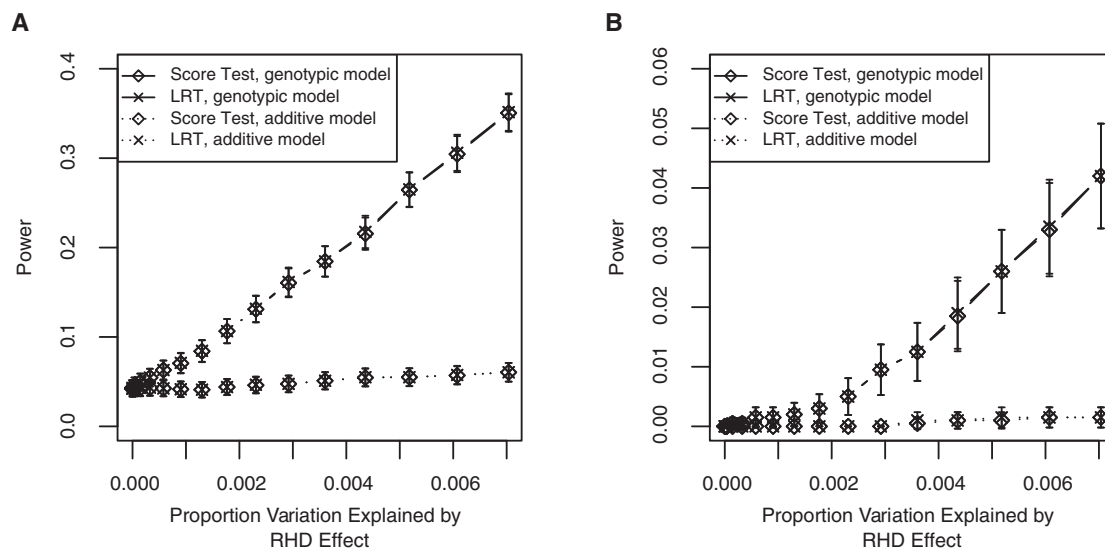
Figure 7 Power using an offspring effect only test for data simulated under RHD incompatibility. RHD effect sizes range from 0 to 0.7. Solid lines show the power for fitting the genotypic model, i.e., the two degrees of freedom test for offspring effects $(\beta_{.1}, \beta_{.2})$. Dotted lines show the power for fitting the additive model, i.e., the one degree of freedom test with the added constraint $2\beta_{.1} = \beta_{.2}$. Error bars represent approximate 95% confidence intervals. (A) Curves show power using the LRT and score test at significance level $\alpha = 0.05$. (B) Curves show power using the LRT and score test at significance level $\alpha = 0.001$.

Simulation G, each variant allele in the mother or offspring increases the offspring's phenotype by 0.17 and the effects are additive and independent. In this scenario, 0.007 of the variance is explained by the MFG effect. Figure 4C shows the boxplots of the parameter estimate bias over 2000 replicates when the general model is fit to the data. Again, unbiased parameter estimates are produced.

It is also of interest to examine the degree to which parameter estimate precision is reduced and power is lost when the underlying MFG model requires a less complex model, such as RHD or NIMA, but agnostically, the general model is fit. We consequently repeat the analysis of Simulations B and C, this time fitting the general six-parameter model (Model 8). As the boxplots displayed in Figures S10A and S10B illustrate, parameter estimates remain unbiased. The effect on power is visible in Figures 5A and 5B (dotted lines). As expected, the power curves follow a similar pattern to those from less complex models but are lower for both RHD incompatibility and NIMA examples. Under Simulation B conditions, the power of the LRT when $\alpha = 0.001$ reduces from 0.751 (SE = 0.010) to 0.467 (SE = 0.011) when the proportion of variation explained is 0.0044. Under Simulation C conditions, the power of the LRT at the same significance level reduces from 0.823 (SE = 0.009) to 0.702 (SE = 0.010) when the proportion of variation explained is 0.0058. These results demonstrate that in terms of power, the QMFG test performs well when there is no prior support for a restricted model,

thus avoiding possible model misspecification or misinterpretation. However, when there is prior support for a specific model (such as RHD incompatibility or NIMA), a restricted model can provide a substantial increase in power.

Power to Detect MFG Incompatibility in a Standard GWAS Analysis

Can a typical GWAS, which tests the effects of an offspring's genotype and ignores MFG interactions, be used as a first screen for MFG incompatibility? We address this question by using data simulated with RHD incompatibility effects ranging from 0 to 0.70 and comparing the power shown in Figure 5A with the power that results when testing for either offspring genotypic or additive effects. Figure 7 shows the two degrees of freedom test for an offspring genotypic effect model and the one degree of freedom test for an additive offspring allelic effect model with significance levels $\alpha = 0.05$ (Fig. 7A) and $\alpha = 0.001$ (Fig. 7B). For the LRT, the two degrees of freedom test involves fitting Models 1 and 4 and the one degree of freedom test involves fitting Models 1 and 9. Together these figures demonstrate that, compared to the correct test for an RHD effect, power is drastically reduced. In the case of a true underlying RHD effect of 0.70 (Simulation H), the parameters representing the effect of one variant allele in the offspring genotype (β_{01}, β_{11}) are biased upward when fitting the offspring only genotypic model
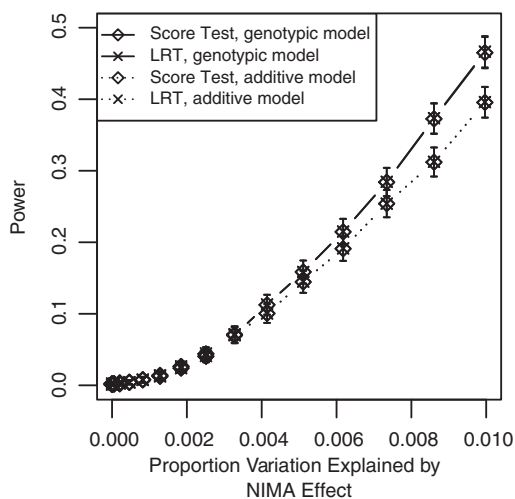
Figure 8 Power using an offspring effect only test for data simulated under a NIMA effect. Curves show power using the LRT and score test at significance level $\alpha = 0.001$. NIMA effect sizes range from 0 to 0.7. Solid lines show the power for fitting the genotypic model, i.e., the two degrees of freedom test for offspring effects $(\beta_{.1}, \beta_{.2})$. Dotted lines show the power for fitting the additive model, i.e., the one degree of freedom test with the added constraint $2\beta_{.1} = \beta_{.2}$. Error bars represent approximate 95% confidence intervals.

(Fig. S11A). The parameter representing the effect of being homozygous for the variant allele remains unbiased. The parameter for the offspring allelic effect is only slightly upwardly biased when fitting the additive model (Fig. S11B). We repeat both the two degrees of freedom test for offspring effects and one degree of freedom additive offspring effects analyses in data simulated with a NIMA effect ranging from 0 to 0.7 and no offspring effect. Figure 8 shows the resulting power curves for $\alpha = 0.001$. In this case, the power is not as severely reduced. However, when fitting the genotypic model to data with a true NIMA effect of 0.7 (Simulation I), we see a downward bias of offspring genotype parameter estimates (Fig. S12A). Figure S12B also shows there is parameter estimate bias when the additive model is fit to Simulation I data. Thus in the case of NIMA, a user may very well reject the null hypothesis but mistakenly attribute the effect to the offspring's genotype.

Model Misspecification

We find that an RHD effect is unlikely to be detected if an investigator uses a NIMA effect model. For instance, if we take data with a true RHD effect of 0.7 (Simulation H),

power drops from 0.97 (SE $= 0.004$) when the correct RHD effect model (Model 2) is fit to 0.008 (SE $= 0.002$) when an incorrect NIMA effect model is fit (Model 5). Although the parameter for a NIMA effect, β_{10}, is estimated on average to reduce the quantitative trait (Fig. S13A), the estimate would not likely be found significant. Thus, in the event that the model is misspecified, an RHD effect would not be misinterpreted as a NIMA effect but instead it would be missed. Likewise, misspecifying the QMFG model as an RHD effect model (Model 2) when data have a true underlying NIMA effect of 0.7 (Simulation I) would result in a missed effect. In this case, power drops from 0.999 (SE $= 0.001$) when the correct NIMA model is fit (Model 5) to 0.004 (SE $= 0.001$) when an incorrect RHD model is fit (Model 2). As shown in Figure S13B, the RHD parameter (β_{21}) is on average estimated to decrease the phenotype, though it would seldom be found significant given similar sample and effect sizes and therefore the NIMA effect would not be misinterpreted as an RHD effect. In both cases, although the true MFG effect would be missed, detecting false MFG effects is unlikely.

In Simulation J, a quantitative trait is simulated for offspring with an additive offspring allelic effect of 0.27. Mimicking a situation where the user incorrectly hypothesizes that there is an RHD effect on the quantitative trait, we fit Model 2 to the simulated data thus estimating the RHD parameter (β_{21}). Under these conditions where we would have power equal to 0.86 (SE $= 0.008$) if the correct model for offspring effects (Model 4) was used, the rejection rate when testing for an RHD effect is 0.002 (SE $= 0.001$). Figure 9A shows the degree of parameter estimate bias that follows from misspecifying the model. Taken together, these results show that it is possible that an offspring effect would be misinterpreted as an RHD-like effect that reduces the trait value, but that the null hypothesis of no RHD effect would rarely be rejected. If a user instead believes there may be a NIMA effect on phenotype, he or she may fit Model 3 (NIMA and offspring effects model) or Model 5 (NIMA effect only model) to the data. Because the NIMA and offspring effects model includes parameters for offspring allelic effects, we expect that the estimated parameters would not be biased although the power would be reduced. These expectations are confirmed by Figure 9B and the power for the three degrees of freedom test for NIMA or offspring effects is 0.81 (SE $= 0.009$). On the other hand, if the NIMA effect model (Model 5) is fit to the same data, the power when testing for a NIMA effect is 0.16 (SE $= 0.008$). The potential to reject the null hypothesis is higher than in the case of RHD, but a significant effect would probably not be detected. In Figure 9C it can be seen that if there was enough power, the offspring effect might be misinterpreted as a NIMA effect that decreases the quantitative phenotype. Finally, if the user has no *a priori* hypothesis and fits the most general model (Model 8), there is zero bias
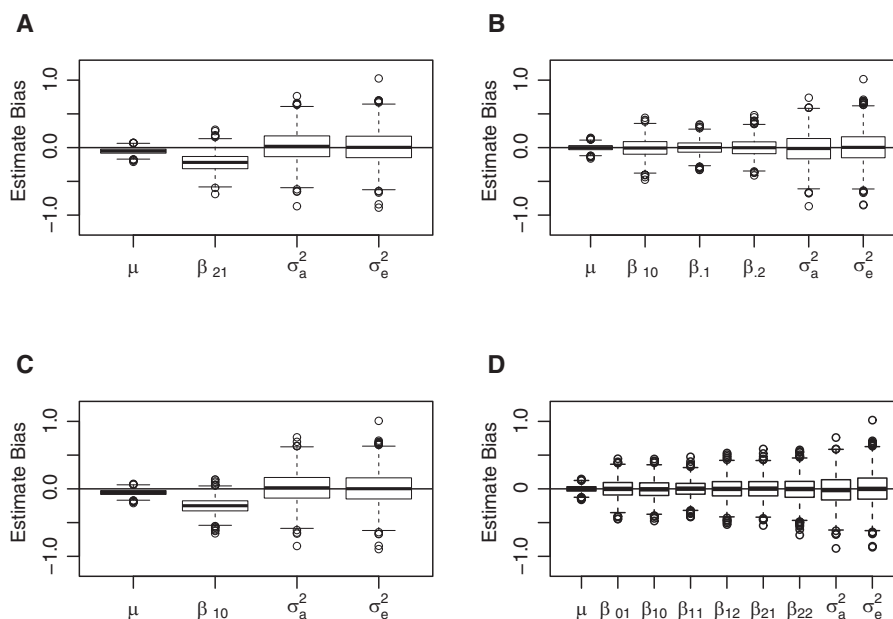
Figure 9 Parameter estimate bias due to model misspecification. Genotypes and quantitative traits for each replication were simulated for 1000 pedigrees using Simulation J data with an offspring allelic effect of 0.27 ($\mu = 40$, $\beta_{.1} = 0.27$, $\beta_{.2} = 0.54$, $\sigma_a^2 = 1$, $\sigma_e^2 = 5$). Boxplots show bias of parameter estimates, additive variance, and environmental variance over 2000 replications when the model is misspecified as (A) Model 2, the RHD effect model, (B) Model 3, the NIMA and offspring effects model, (C) Model 5, the NIMA effect model, and (D) Model 8, the general model. A horizontal line is drawn at zero bias.

(Fig. 9D) and the power to detect an effect is 0.69 (SE = 0.010). These results further demonstrate the advantage of using an MFG model that allows for offspring effects, which is a generalization of the standard GWAS analysis, when screening for MFG effects.

Effect of Missing Data on Type I Error

As is often the case with real data, missing data are an issue that must be considered. Assuming genotypes are missing at random, we compare type I error rates testing for various sets of MFG parameters given 0%, 5%, 10%, and 20% missing genotypes. Here, we use Simulation A data, randomly removing a percentage of genotypes with each repetition, and estimate type I error rates with a per test significance level of 0.05. With all the models tested, missing data did not significantly affect type I error rates (Table S2).

**Screening for MFG Incompatibility Using Pedigree-based GWAS Data**

The ability to quickly screen markers is demonstrated by running the QMFG test on data from the SAFHS. Missing geno-

types for the 944,565 SNPs across the genome were imputed using Mendel's imputation option. We removed 14,008 SNPs (1.48%) because there was at least one impossible maternal-offspring genotype combination observed at each of these SNPs in the imputation results. We omitted another 295,063 SNPs because they had minor allele counts less than 10 leaving a total of 635,494 SNPs for the analysis. Because we have no specific hypothesis regarding MFG interactions associated with HDL measures but aim to demonstrate the feasibility of using the QMFG test on pedigree-based GWAS data, we use an alternative hypothesis of NIMA or offspring effects, a straightforward generalization of the standard GWAS analysis. This model takes into account offspring effects, so in the case that there is an underlying offspring genotypic effect but no NIMA effect, we lose power by including the parameter for NIMA but avoid misspecifying the model.
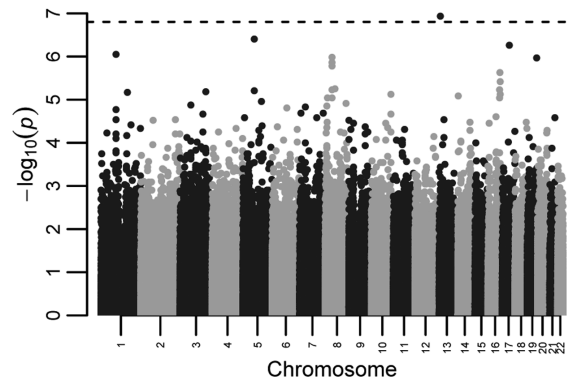
In these analyses we assume the minor allele of each SNP is the variant allele and include sex and age as fixed effects. Under the null hypothesis of no genetic effects, the estimate of the grand mean was 48.030 (SE = 2.094). The age effect estimate of 0.034 (SE = 0.069) was not significant. Women had significantly higher HDL levels than males (4.222 units

Table 4 Score test for NIMA or offspring effects on HDL from the San Antonio Family Heart Study.

Chromosome	Nearby gene	SNP	Score test statistic	P-value
13	USP12	rs1547189:G>A	35.1	1.16×10^{-7}
5	–	rs9293660:G>A	32.6	3.91×10^{-7}
17	NGFR	rs614455:T>C	31.9	5.49×10^{-7}
1	FAM69A	rs7521417:C>T	30.9	8.92×10^{-7}
8	LOC102723729	rs11987150:G>A	30.6	1.03×10^{-6}
19	ZNF888	rs10425203:G>A	30.5	1.08×10^{-6}
8	GINS4	rs13265966:T>C	30.0	1.38×10^{-6}
8	LOC102723729	rs11994079:G>T	29.6	1.68×10^{-6}
16	–	rs6564175:T>C	28.9	2.35×10^{-6}
16	WWOX	rs4267317:G>A	27.9	3.81×10^{-6}
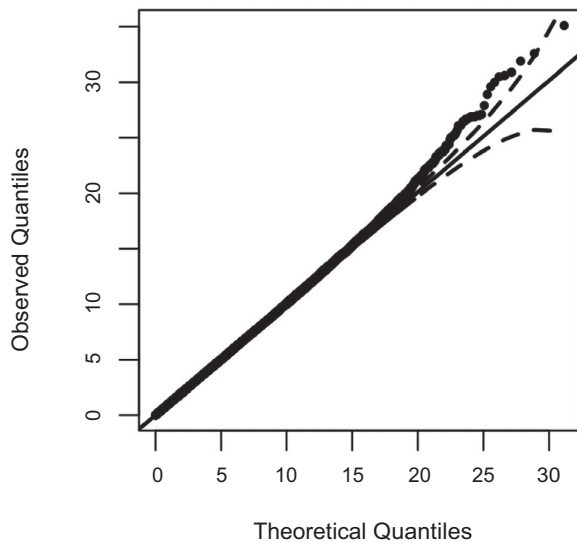
Figure 10 Q-Q plot for score test of the SAFHS data. Results from the three degrees of freedom test for NIMA or offspring effects ($\beta_{10}, \beta_{.1}, \beta_{.2}$) using the score test adjusting for age and sex ($\lambda = 1.012$). Data from the SAFHS consist of 635,494 SNPs from 419 offspring with HDL measurements in 43 multigenerational families.

Figure 11 Manhattan plot for score test of the SAFHS data. Results from the three degrees of freedom test for NIMA or offspring effects ($\beta_{10}, \beta_{.1}, \beta_{.2}$) using the score test adjusting for age and sex. Data from the SAFHS consists of 635,494 SNPs from 419 offspring with HDL measurements in 43 multigenerational families. A dashed horizontal line is drawn at the initial significance cutoff for an FDR of 10%.

higher, SE = 1.154). Figures 10 and 11 show the Q-Q plot ($\lambda = 1.012$) and Manhattan plot resulting from the three degrees of freedom QMFG score test for the alternative hypothesis of NIMA or offspring effects. The run time for Mendel to read in and perform the score tests for the 635,494 SNPs was 4 min and 28 s on a computer with 12 CPU cores (at 2.67 GHz). About 0.2% (1309) of sites fall outside the 95% confidence limits of the Q-Q plot (Figure 10). The 10 markers with the smallest P-values are shown in Table 4. Figure S14 shows the Q-Q plot of the same results with the top 10 hits removed ($\lambda = 1.012$). The lowest P-value, found with SNP

rs1547189, corresponds to an FDR (Benjamini & Hochberg, 1995) of 7.5% (Table 5). The parameter estimates for this SNP are shown in Table 6.

For rs1547189, we further refine our analysis. The two degrees of freedom test for offspring genetic effects for SNP rs1547189 on HDL has a P-value of 0.0007. For the one degree of freedom test for additive offspring effects ($2\beta_{.1} = \beta_{.2}$) using the same SNP, the P-value is 0.18. We also test for a NIMA effect in the presence of offspring effects adjusting for sex and age using an LRT. The P-value of 2.98×10^{-6} suggests that even when accounting for offspring genotypic effects, there may be an additional effect of NIMA. Taken together, these results suggest that there may be both an underlying recessive offspring effect and a NIMA effect on HDL for SNP rs1547189.

Table 5 False discovery rates for the San Antonio Family Heart Study analysis.

FDR	P-value threshold	Number of significant SNPs
5%	7.87×10^{-8}	0
7.5%	1.18×10^{-7}	1
10%	1.57×10^{-7}	1
15%	1.89×10^{-6}	8

Table 6 Parameter estimates for SNP rs1547189 from the SAFHS data.

Effect	Parameter	Estimate	Std Error
Grand mean	μ	46.10	2.72
NIMA	β_{10}	8.51	2.15
A/G* offspring	$\beta_{.1}$	−0.60	1.85
G/G offspring	$\beta_{.2}$	10.19	2.84
Female	β_{female}	4.32	0.55
Age	β_{age}	0.04	0.07
Additive variance	σ_a^2	80.66	19.64
Environmental variance	σ_e^2	63.60	14.60

*G is the minor allele for SNP rs1547189

Discussion

Our simulation studies show that the LRT version of the QMFG test leads to correct parameter estimates and inference and the score test version of the QMFG test provides equivalent inference to the LRT under both specific and general models of MFG incompatibility. The simulations under an RHD incompatibility scenario illustrate the QMFG test under a simple model, involving just one MFG parameter. We show that our approach has correct type I error rates for the LRT and score test, zero parameter estimation bias, and high power even when the proportion of variation explained by MFG incompatibility is small. The simulations under NIMA provide a more complicated, but still biologically pertinent scenario to evaluate the properties of the QMFG test, requiring an MFG effect and offspring effects to be tested jointly. In this setting, the effect on offspring with genotypes homozygous for the reference allele depends on their mother's genotype. With this scenario, we show the flexibility of the LRT to test effects jointly, marginally, and conditionally. When testing NIMA and offspring effects jointly and conditionally, again the LRT version of the QMFG test produces appropriate type I error rates, zero parameter estimation bias, and high power. This model can be extended by allowing for maternal genetic effects. We also demonstrate that it is possible to impose parameter restrictions to test for other situations such as a dominant offspring effects.

Additionally, we investigate MFG incompatibility testing in the case that there is no *a priori* information about the underlying MFG model. As expected, power is reduced when applying this general model, which imposes no constraints on the MFG parameters, to data simulated with a specific, more restricted MFG incompatibility. Our results indicate that even when the general model is fit to the data in place of the correct, simpler model, the QMFG LRT still produces unbiased parameter estimates. We also explore cases where the model is misspecified such that a model with only offspring effects is fit to data with true underlying MFG effects and find that power is greatly reduced. This is especially true in the case of testing for additive effects when the underlying MFG effect is RHD incompatibility, where power drops down to the type I error rate. This particularly low powered case with the additive model results because we have a SNP that, when viewed from the perspective of offspring effects, is displaying a weak amount of overdominance in which only a fraction of the offspring with heterozygous genotypes are expected to have different phenotypic values from offspring with either of the two homozygous genotypes. In the case of an underlying NIMA effect, there is more power to detect a genotypic effect when fitting a model with only offspring effects but in this case, the NIMA would be misinterpreted as a weak dominant effect on the quantitative trait. These results have implications for GWAS, which typically use additive models or Armitage trend tests, since, as we have shown, the genetic effects can be detected but misinterpreted, determined with lower power, or missed all together. The QMFG test is also subject to type I or type II error when the model is misspecified. In particular, when applying the NIMA effect model to data with an RHD effect, applying the RHD effect model to data with a NIMA effect, or applying an RHD or NIMA model to data with offspring effects, the null hypothesis is rejected at very low rates, indicating that in these cases the locus would be missed. Because most effects are very likely to be offspring genotype effects, we recommend using an MFG model that includes offspring genetic effects when screening large numbers of SNPs.

Unlike other methods that have been proposed to test for an association between a quantitative trait and MFG incompatibility, the QMFG test can handle small and large pedigrees simultaneously. With actual data from the SAFHS we verify that the QMFG score test is an effective and rapid screening tool for genome wide association studies. In this data set, family size varied greatly; the smallest family had eight members and spanned three generations; the largest family had 176 members and spanned five generations. We chose to analyze the data by jointly screening for NIMA and offspring genotypic effects using the score test. If the genetic effects only come from the offspring as typically assumed, our analysis would still be able to detect them, albeit with slightly lower power than in the typical GWAS.

Although none of the top 10 markers for NIMA or off-spring effects have been previously shown to be associated with HDL, the *WWOX* gene has been shown to be associated with HDL (Lee et al., 2008, Saez et al., 2010). Our result with the smallest *P*-value is for SNP rs1547189, which is an intron variant in the ubiquitin specific peptidase 12 (*USP12*) gene, with an FDR of 7.5%. Because it is possible that this marker's effect could be exclusively due to offspring effects, we use the LRT with the null hypothesis of only offspring effects and the alternative hypothesis of NIMA and offspring effects to test for NIMA effects in the possible presence of offspring effects with SNP rs1547189. Combined, our results are suggestive of a NIMA effect in the presence of offspring effects. To determine whether this association is due to a previously undetected NIMA effect on HDL or, perhaps what is more likely, is just a type I error, requires testing in other cohorts. However, we have clearly demonstrated the potential of the QMFG test to identify novel associations with quantitative traits that may not be detected in standard GWAS analysis models because the standard GWAS only considers offspring allelic effects. Additionally, our analysis demonstrates that the LRT is a useful tool to refine results following the rapid screening provided by the QMFG score test.

To date, no other studies have looked at the role of MFG incompatibility on quantitative traits in families larger than trios due to the lack of appropriate models and practical software. The QPL method (Kistner & Weinberg, 2005) as well as the QCPG method (Wheeler & Cordell, 2007) are both retrospective approaches in which the offspring genotype is modeled as a function of the quantitative trait and parental genotypes and are restricted to parent-offspring trios. These methods can be easily modified to test for maternal-offspring gene interactions (Wheeler & Cordell, 2007). However, Wheeler and Cordell's simulation results suggest that, compared to these two retrospective approaches to test for quantitative trait association in trios, a prospective, linear regression approach such as ours is likely to be more efficient but more sensitive to departures from normality. From our viewpoint, the main difficulty with retrospective approaches such as the QPL and QCPG is in generalizing them to work with a data set composed of dramatically different sized families. Another limitation is the interpretation of the estimated effects as they are scaled by the unknown trait variance. Furthermore, including covariates such as age and sex is not straightforward. With the QMFG test we have demonstrated the benefits of a prospective approach to rapidly test for MFG incompatibility in families of any size. It is a highly flexible and accurate method, which is also easy to execute with our user-friendly software.

It should be noted that in general, LMMs used for quantitative traits do not directly apply to binary phenotypes. This limitation results from the fact that the phenotypic variance of a dichotomous disease or trait depends on its incidence in the population (Falconer, 1965). Thus, estimates must be rescaled. Unlike with continuous traits, case–control studies are additionally susceptible to ascertainment bias. As a result, using LMMs on binary traits directly can result in loss of power as sample size increases, likely due to the amplification of inaccuracies caused by ascertainment bias (Yang et al., 2014). Methods to improve power for qualitative outcomes are based on a liability threshold principle, in which it is assumed that binary traits can be represented by an underlying normally distributed liability trait. If an individual's liability exceeds a threshold, then he or she has a phenotypic value of 1, otherwise 0, with the proportion of the normal distribution that exceeds the threshold being equal to trait incidence (Dempster & Lerner, 1950, Falconer, 1965). An approach by Hayeck et al. (2015) estimates the posterior mean liability (PML) of each individual conditional on the case-control status of all subjects, disease prevalence, and liability scale phenotypic covariance. The association between each SNP and PML is then tested. Accordingly, if the QMFG test is applied to dichotomous data, it is recommended that one adopt the Hayeck et al. (2015) approach.

We have implemented the QMFG test by modifying the statistical genetics software package Mendel. This option is scheduled for release in an upcoming version of Mendel. The validity of our method, together with the availability of convenient software, make the QMFG test a powerful tool for detecting undiscovered associations with complex diseases.

Acknowledgments

References

Ayers, K. L. & Lange, K. (2008) Penalized estimation of haplotype frequencies. *Bioinformatics* **24**, 1596–1602.

Benjamini, Y. & Hochberg, Y. (1995) Controlling the false discovery rate: A practical and powerful approach to multiple testing. *J R Stat Soc Ser B* **57**, 289–300.

Boerwinkle, E., Chakraborty, R. & Sing, C. F. (1986) The use of measured genotype information in the analysis of quantitative phenotypes in man. I. Models and analytical methods. *Ann Hum Genet* **50**, 181–194.

Cannon, M., Jones, P. B. & Murray, R. M. (2002) Obstetric complications and schizophrenia: Historical and meta-analytic review. *Am J Psychiat* **159**, 1080–1092.

Chen, W. M. & Abecasis, G. R. (2007) Family-based association tests for genomewide association scans. *Am J Hum Genet* **81**, 913–926.

Dahlquist, G. G., Patterson, C. & Soltesz, G. (1999) Perinatal risk factors for childhood type 1 diabetes in Europe. The EURODIAB Substudy 2 Study Group. *Diabetes Care* **22**, 1698–1702.

Dempster, E.R. & Lerner, I.M. (1950) Heritability of Threshold Characters. *Genetics* **35**, 212–236.

Devlin, B. & Roeder, K. (1999) Genomic control for association studies. *Biometrics* **55**, 997–1004.

Falconer, D. S. (1965) The inheritance of liability to certain diseases, estimated from the incidence among relatives. *Ann Hum Genet* **29**, 51–76.

Fox, J. (2008) *Applied regression analysis and generalized linear models*. Los Angeles: Sage.

Freedman, D., Deicken, R., Kegeles, L. S., Vinogradov, S., Bao, Y. & Brown, A. S. (2011) Maternal-fetal blood incompatibility and neuromorphologic anomalies in schizophrenia: Preliminary findings. *Prog Neuropsychopharmacol Biol Psychiat* **35**, 1525–1529.

Gregersen, P. K., Silver, J. & Winchester, R. J. (1987) The shared epitope hypothesis. An approach to understanding the molecular genetics of susceptibility to rheumatoid arthritis. *Arthritis Rheum* **30**, 1205–12013.

Harney, S., Newton, J., Milicic, A., Brown, M. A. & Wordsworth, B. P. (2003) Non-inherited maternal HLA alleles are associated with rheumatoid arthritis. *Rheumatology (Oxford)* **42**, 171–174.

Hayeck, T. J., Zaitlen, N. A., Loh, P. R., Vilhjalmsson, B., Pollack, S., Gusev, A., Yang, J., Chen, G. B., Goddard, M. E., Visscher, P. M., Patterson, N. & Price, A. L. (2015) Mixed model with correction for case-control ascertainment increases association power. *Am J Hum Genet* **96**, 720–730.

Hollister, J. M., Laing, P. & Mednick, S. A. (1996) Rhesus incompatibility as a risk factor for schizophrenia in male adults. *Arch Gen Psychiat* **53**, 19–24.

Insel, B. J., Brown, A. S., Bresnahan, M. A., Schaefer, C. A. & Susser, E. S. (2005) Maternal-fetal blood incompatibility and the risk of schizophrenia in offspring. *Schizophr Res* **80**, 331–342.

Jawaheer, D. & Gregersen, P. K. (2002) Rheumatoid arthritis. The genetic components. *Rheum Dis Clin North A.*, **28**, 1–15, v.

Juul-Dam, N., Townsend, J. & Courchesne, E. (2001) Prenatal, perinatal, and neonatal factors in autism, pervasive developmental disorder-not otherwise specified, and the general population. *Pediatrics* **107**, E63.

Kistner, E. O. & Weinberg, C. R. (2004) Method for using complete and incomplete trios to identify genes related to a quantitative trait. *Genet Epidemiol* **27**, 33–42.

Kistner, E. O. & Weinberg, C. R. (2005) A method for identifying genes related to a quantitative trait, incorporating multiple siblings and missing parents. *Genet Epidemiol* **29**, 155–165.

Kraft, P., Palmer, C. G., Woodward, A. J., Turunen, J. A., Minassian, S., Paunio, T., Lonnqvist, J., Peltonen, L. & Sinsheimer, J. S. (2004) RHD maternal-fetal genotype incompatibility and schizophrenia: Extending the MFG test to include multiple siblings and birth order. *Eur J Hum Genet* **12**, 192–198.

Lange, K. (2002) *Mathematical and statistical methods for genetic analysis*. New York: Springer.

Lange, K., Papp, J. C., Sinsheimer, J. S., Sripracha, R., Zhou, H. & Sobel, E. M. (2013) Mendel: The Swiss army knife of genetic analysis programs. *Bioinformatics* **29**, 1568–1570.

Lange, K. & Sinsheimer, J. S. (2004) The pedigree trimming problem. *Hum Hered* **58**, 108–111.

Lee, J. C., Weissglas-Volkov, D., Kyttälä, M., Dastani, Z., Cantor, R. M., Sobel, E. M., Plaisier, C. L., Engert, J. C., Van Greevenbroek, M. M. J., Kane, J. P., Malloy, M. J., Pullinger, C. R., Huertas-Vazquez, A., Aguilar-Salinas, C. A., Tusie-Luna, T., De Bruin, T. W. A., Aouizerat, B. E., Van Der Kallen, C. C. J., Croce, C. M., Aqeilan, R. I., Marcil, M., Viikari, J. S. A., Lehtimäki, T., Raitakari, O. T., Kuusisto, J., Laakso, M., Taskinen, M. -R., Genest, J. & Pajukanta, P. (2008) WW-domain-containing oxidoreductase is associated with low plasma HDL-C levels. *Am J Hum Genet* **83**, 180–192.

Lee, Y. K., Daito, Y., Katayama, Y., Minami, H. & Negishi, H. (2009) The significance of measurement of serum unbound bilirubin concentrations in high-risk infants. *Pediatr Int* **51**, 795–799.

Levine, P., Burnham, L., Katzin, E. M. & Vogel, P. (1941) The role of iso immunization in the pathogenesis of erythroblastosts fetalis. *Am J Obstet Gynecol* **42**, 925–937.

Li, M., Erickson, S. W., Hobbs, C. A., Li, J., Tang, X., Nick, T. G., Macleod, S. L., Cleves, M. A. & National Birth Defect Prevention Study. (2014) Detecting maternal-fetal genotype interactions associated with conotruncal heart defects: A haplotype-based analysis with penalized logistic regression. *Genet Epidemiol* **38**, 198–208.

Liang, M., Wang, X., Li, J., Yang, F., Fang, Z., Wang, L., Hu, Y. & Chen, D. (2010) Association of combined maternal-fetal TNF-alpha gene G308A genotypes with preterm delivery: A gene-gene interaction study. *J Biomed Biotechnol* 2010, 396184.

Lupo, P. J., Mitchell, L. E., Canfield, M. A., Shaw, G. M., Olshan, A. F., Finnell, R. H., Zhu, H. & National Birth Defects Prevention Study. (2014) Maternal-fetal metabolic gene-gene interactions and risk of neural tube defects. *Mol Genet Metab* **111**, 46–51.

Mitchell, B. D., Kammerer, C. M., Blangero, J., Mahaney, M. C., Rainwater, D. L., Dyke, B., Hixson, J. E., Henkel, R. D., Sharp, R. M., Comuzzie, A. G., Vandeberg, J. L., Stern, M. P. & Maccluer, J. W. (1996) Genetic and environmental contributions to cardiovascular risk factors in Mexican Americans. *The San Antonio Family Heart Study. Circulation* **94**, 2159–2170.

Newton, J. L., Harney, S. M., Wordsworth, B. P. & Brown, M. A. (2004) A review of the MHC genetics of rheumatoid arthritis. *Genes Immun* **5**, 151–157.

Palmer, C. G., Mallery, E., Turunen, J. A., Hsieh, H. J., Peltonen, L., Lonnqvist, J., Woodward, J. A. & Sinsheimer, J. S. (2008) Effect of Rhesus D incompatibility on schizophrenia depends on offspring sex. *Schizophr Res* **104**, 135–145.

Palmer, C. G., Turunen, J. A., Sinsheimer, J. S., Minassian, S., Paunio, T., Lonnqvist, J., Peltonen, L. & Woodward, J. A. (2002) RHD maternal-fetal genotype incompatibility increases schizophrenia susceptibility. *Am J Hum Genet* **71**, 1312–1319.

Procopciuc, L. M., Caracostea, G., Zaharie, G. & Stamatian, F. (2014) Maternal/newborn VEGF-C936T interaction and its influence on the risk, severity and prognosis of preeclampsia, as well as on the maternal angiogenic profile. *J Matern Fetal Neonatal Med* **27**, 1754–1760.

Saez, M. E., Gonzalez-Perez, A., Martinez-Larrad, M. T., Gayan, J., Real, L. M., Serrano-Rios, M. & Ruiz, A. (2010) WWOX gene is associated with HDL cholesterol and triglyceride levels. *BMC Med Genet* **11**, 148.

Silveira, I. G., Burlingame, R. W., Von Muhlen, C. A., Bender, A. L. & Staub, H. L. (2007) Anti-CCP antibodies have more diagnostic impact than rheumatoid factor (RF) in a population tested for RF. *Clin Rheumatol* **26**, 1883–1889.

Sinsheimer, J. S. & Creek, M. M. (2013) Statistical approaches for detecting transgenerational genetic effects in humans. In: *Epigenetics and complex traits* (eds. A. K. Naumova & C. M. T. Greenwood), New York: Springer, 247–264.

Stubbs, E.G., Ritvo, E.R. & Mason-Brothers, A. (1985) Autism and shared parental HLA antigens. *J Am Acad Child Psychiat* **24**, 182–185.

Van Der Horst-Bruinsma, I. E., Hazes, J. M., Schreuder, G. M., Radstake, T. R., Barrera, P., Van De Putte, L. B., Mustamu, D., Van Schaardenburg, D., Breedveld, F. C. & De Vries, R. R. (1998) Influence of non-inherited maternal HLA-DR antigens on susceptibility to rheumatoid arthritis. *Ann Rheum Dis* **57**, 672–675.

Visser, H., Le Cessie, S., Vos, K., Breedveld, F. C. & Hazes, J. M. W. (2002) How to diagnose rheumatoid arthritis early—A prediction model for persistent (erosive) arthritis. *Arthritis Rheum* **46**, 357–365.

Wheeler, E. & Cordell, H. J. (2007) Quantitative trait association in parent offspring trios: Extension of case/pseudocontrol method and comparison of prospective and retrospective approaches. *Genet Epidemiol* **31**, 813–833.

Yang, J., Zaitlen, N. A., Goddard, M. E., Visscher, P. M. & Price, A. L. (2014) Advantages and pitfalls in the application of mixed-model association methods. *Nat Genet* **46**, 100–106.

Zhou, H., Blangero, J., Dyer, T. D., Chan, K. K., Sobel, E. & Lange, K. (2015) Fast Genome-Wide QTL Association Mapping on Pedigree and Population Data. arXiv:1407.8253v2.

Supporting Information

Additional Supporting Information may be found in the online version of this article:

Table S1 San Antonio Family Heart Study subject counts.

Table S2 Estimated type I error rates when a proportion of genotypes are missing.

Figure S1 Comparison of LRT and score test p-values for Simulation A data.

Figure S2 Power to detect a NIMA effect in the presence of offspring effects.

Figure S3 Power to jointly detect NIMA and dominant offspring effects.

Figure S4 Parameter estimate bias for data simulated with NIMA, offspring, and maternal effects.

Figure S5 Effect of allele frequency on power.

Figure S6 Effect of additive genetic and environmental variance on power.

Figure S7 Parameter estimate bias when data are simulated for parent–offspring trios.

Figure S8 Q-Q plot for LRT when data are simulated with a smaller sample size.

Figure S9 Parameter estimate bias when data are simulated under NIMA and offspring effects with a smaller sample size.

Figure S10 Parameter estimate bias when the general model is fit.

Figure S11 Parameter estimate bias when a misspecified model is fit using Simulation H data.

Figure S12 Parameter estimate bias when a misspecified model is fit using Simulation I data.

Figure S13 Parameter estimate bias due to model misspecification.

Figure S14 Q-Q plot for the score test of the SAFHS data when top 10 hits are removed.

Chapter 4

Human Birth Weight and Reproductive Immunology: Testing for Interactions between Maternal and Offspring Genes.

Introduction

Complex familial disorders result from interactions between environmental and genetic factors. One such interaction, which can contribute to disease susceptibility and variation in quantitative traits, occurs when the fetal environment is modified by the interaction of proteins expressed from maternal and offspring genes. During pregnancy, the maternal and fetal semiallogenic cells come into direct contact, resulting in an intricate connection between the two individuals. Depending on the combination of maternal and offspring genotypes, maternal immune recognition of fetal cells is one possible form of maternal-fetal genotype (MFG) interaction. MFG interactions can alter the conditions in which the fetus develops and have the potential to impact offspring traits. MFG interactions have been shown to be involved in perinatal diseases [1-5] as well as those that do not manifest until later in life [6-15].

One example of an MFG interaction involves genes that are thought to regulate human birth weight [16-18]. Besides variation in birth weight being of intrinsic interest in human evolution, extremes in birth weight are strongly associated with obstetric complications and perinatal mortality. For instance, whereas high birth weight causes obstructed labor [19], pre-eclampsia and low birth weight are two consequences of poor placentation in early pregnancy [20]. Trophoblast invasion and spiral artery transformation are important processes that affect the maternal blood supply to the placenta and therefore impact fetal growth. Uterine NK (uNK) cells, a specific type of lymphocyte population produced by the maternal immune system,

accumulate around the invasive trophoblast cells. Human maternal uNK receptors, which are encoded by the killer immunoglobulin-like receptor (KIR) gene family, can bind to HLA molecules expressed by fetal trophoblasts thus forming an interaction between maternal and fetal cells [21]. This interaction between uNK KIR and trophoblast HLA can influence the balance between restricted and amplified fetal placental cell invasion, transformation of spiral arteries, and, in turn, fetal development.

Trophoblast cells express three HLA class I molecules: two non-classical (HLA-G and HLA-E) and one classical (HLA-C) [22]. Of these three, only HLA-C is polymorphic. Although there are 2,902 known *HLA-C* alleles, they can be placed into two groups, C1 and C2, when considering their effect on birth weight [23]. Allotypic recognition of C1 or C2 epitopes varies by *KIR* gene. It is important to note that beyond the influence of the offspring *HLA-C* genotype, the maternal *HLA-C* genotype may also play an important role in placental development. It is hypothesized that, during uNK cell development, the maternal KIR interacts with her own HLA-C molecules, thus "educating" or "licensing" her uNK cells and changing the way they interact with her offspring's HLA-C molecules during placentation [17,18,24,25]. This education hypothesis is supported by evidence that maternal MHC (major histocompatibility complex) class I antigens educate the uNK cells in murine models [26].

Fifteen *KIR* genes have been identified, mapping to chromosome 19q13.4 within the 1 Mb leukocyte receptor complex (LRC) [27]. As described in Marsh *et al*. [27], *KIR* genes are denoted by the number of extracellular immunoglobulin domains (2D or 3D) and the length of the cytoplasmic tail (L for long and S for short). There are two basic *KIR* haplotypes ('A' and 'B') that consist of a variable number of genes encoding activating and inhibitory receptors. 'A' haplotypes have a smaller number of genes that, with one exception, encode for inhibitory KIRs

51

whereas the 'B' haplotypes have additional *KIR* genes, most of which are activating [27]. Unique to the telomeric region of haplotype 'B' is *KIR2DS1*, which encodes a KIR that binds to HLA-C2 allotypes thus acting as an activating KIR for HLA-C [28].

Combinations of maternal and fetal *KIR* and *HLA-C* variants are associated with both birth weight extremes. Models involving *KIR* genes in both the telomeric and centromeric regions together with *HLA-C* have been found to be essential in explaining associations with pregnancy disorders including pre-eclampsia, fetal growth restriction, and recurrent miscarriage [17,29,30]. However Hiby *et al*. [16] found that a parsimonious model, which models the presence or absence of *KIR2DS1* in the maternal *KIR* telomeric region, was sufficient for modeling the interaction of *KIR* and *HLA-C* as a predictor of birth weight in normal pregnancies using subjects from the United Kingdom and Norway. Because effectively all Europeans with the *KIR* telomeric 'B' haplotype have a copy of the *KIR2DS1* gene, in this article treating the maternal *KIR* telomeric region as though it were a single gene with two alleles is comparable to modeling the presence or absence of *KIR2DS1*. Thus, we limit our modeling to two *KIR* telomeric haplotypes (which we refer to as *KIR-tA*, *KIR-tB*) in the mother, two maternal *HLA-C* allele groups (which we refer to as *HLA-C1*, *HLA-C2*), and the same two offspring *HLA-C* groups when considering normal variation in birth weight.

Until recently, association testing for MFG interactions with quantitative traits was limited to retrospective likelihood designs [31-33] as direct extensions of association testing for MFG interactions with qualitative traits [34,35]. In addition to potential difficulties in parameter interpretation, such approaches have typically been limited to case-parent trios and cannot easily account for the main effects of other covariates [36,37]. To address these modeling limitations, the Quantitative-MFG (QMFG) test was developed [37]. This linear mixed effects modeling

approach can test for various scenarios of joint maternal and offspring effects quickly and accurately and can handle pedigrees of any size. The QMFG test was originally developed to address interactions that occur at a single locus, but we extend the model to multiple loci in this article. Of course there are practical issues for the increased number of possible effects; with more than two loci with two alleles each, the sample sizes needed for accurate inference become large. Alternatively, constraints can be placed on the models *a priori* to decrease the number of parameters estimated, which reduces the required sample size. In this article, we demonstrate how the QMFG test can be extended for practical use with two loci and show that its statistical properties remain sound. The *KIR-HLA-C* interaction as a predictor of birth weight serves as an interesting and important example where our extension to the QMFG test can provide new insights. We apply the QMFG test to the United Kingdom cohort [16,17,29,30] used by Hiby *et al*. [16] to extend their findings and test the hypothesis of a *KIR-HLA-C* MFG interaction effect on human birth weight.

Methods

The Quantitative-MFG Test for Multiple Loci

For a single pedigree, the loglikelihood for the QMFG test is

$$L = -\frac{1}{2}\ln|\Omega| - \frac{1}{2}(y - v)^{t}\Omega^{-1}(y - v)$$

with mean vector v, covariance matrix Ω, and observed quantitative trait vector y. Maternal-offspring genotypes are treated as fixed effects and are consequently included in $v = A\beta$, where A is the QMFG design matrix and β is the column vector of regression coefficients. When modeling the joint effects of maternal and offspring genotypes, each QMFG parameter denotes a change in phenotype for the maternal-offspring genotype combinations of interest compared to

the grand mean. Both gene main effects and interactions can be modeled with this linear mixed model. The first column of the design matrix A consists of all ones (to allow the estimation of the grand mean) and the additional columns are indicator variables corresponding to each of the possible MFG combinations, the number depending on the number of maternal loci and alleles, the number of offspring loci and alleles, and the number of loci that are common between these two sets. Additional non-genetic covariates and their interactions can easily be incorporated as desired, for example in the case of birth weight, gestational age and sex can be included.

As mentioned previously for the example of birth weight, the model can be represented by one telomeric *KIR* locus with two haplotypes (*KIR-tA*, *KIR-tB*) and one *HLA-C* locus with two alleles (*HLA-C1*, *HLA-C2*). We include only the effects of the maternal haplotypes for the *KIR* telomeric region whereas we include the effects of both the maternal and offspring genotypes at the *HLA-C* locus to capture the association with birth weight [16] (Table 1). The vector of QMFG regression coefficients for these two loci therefore consists of parameters of the form β_{ijk} where i denotes the number of maternal *KIR-tB* haplotypes, j denotes the number of maternal *HLA-C2* alleles, and k denotes the number of offspring *HLA-C2* alleles. As with the single-SNP version of the QMFG test, one of the parameters for the MFG effects should be made the reference state or, equivalently, the sum of QMFG parameters should be set to some constant (typically zero) to avoid non-identifiability. In our analyses, parameter β_{000} is always set to zero, which denotes zero copies of the variant alleles (in this case, *KIR-tB* and *HLA-C2*) in the maternal and offspring genotypes. The general model therefore has 20 parameters to be estimated. As mentioned before, to make estimation practical we impose constraints on the QMFG parameters based on prior understanding of underlying immune response to reduce the number of parameters estimated and consequently reduce the degrees of freedom.

Familial correlations are treated as random effects by partitioning the residual variance. Often in genetic studies only two variance components are used to avoid over parameterization, one representing the additive genetic effects (σ_a^2) and one representing environmental random effects (σ_e^2) such that $\Omega = 2\sigma_a^2\Phi + \sigma_e^2\mathbf{I}$ [38]. Each element Φ_{st} of the global kinship coefficient matrix Φ is the probability that, at a randomly chosen autosomal locus, an allele chosen at random from subject s and an allele chosen at random from subject t match identically by descent. When s equals t the alleles are chosen with replacement. Because the environment is assumed to affect each subject independently, the environmental variance σ_e^2 is multiplied by the identity matrix $\mathbf{I}$. The environmental variance is always included even when there are thought to be no environmental factors to insure that the matrix Ω is positive definite. This model can be extended to include other variance components if desired [38].

Simulation Studies

To examine type I error, power, and parameter estimate accuracy for the QMFG test for multiple loci, we conduct simulation studies with 300 three-generation families, each having the structure shown in figure 1. Each family has five offspring for a total of 1500 phenotyped individuals. Haplotype frequencies for the *KIR* telomeric region and allele frequencies for *HLA-C* were simulated based on frequencies observed in white British populations [16]. Unless otherwise specified, the *KIR-tB* haplotype frequency is 20% and the *HLA-C2* allele frequency is 30%. Each simulation run consists of 2000 repetitions in which birth weight is simulated with a grand mean μ (3.5 kg) and a sex effect (β_{sex} = -0.2 kg if female). Variance components are simulated to allow for a high heritability of birth weight as found by Demerath and colleagues [39] (σ_a^2 = 0.2025 kg^2 and σ_e^2 = 0.0475 kg^2; residual heritability h^2 = 0.81). For the purpose of examining

55

the type I error of the QMFG test when using multiple loci, Scenario I data are simulated under

the null of no genetic effects. To investigate power and parameter estimate accuracy, we simulate

data with genetic effects inspired by the example of *KIR* and *HLA-C*. Scenario II data are

simulated under conditions consistent with the effect observed by Hiby *et al*. [16], that is, there is

a change in birth weight only when an offspring has more *HLA-C2* than his or her mother and the

mother has at least one copy of *KIR-tB*. Scenario III involves the same interaction effect in

Scenario II as well as a separate effect for the maternal *KIR-tB* that is independent of *HLA-C* and

a separate effect for more *HLA-C2* in the offspring than in the mother that is independent of *KIR*.

These simulation scenarios are summarized in Table 2.

Genomic control values (λ) are reported as an assessment of type I error accuracy [40].

The significance level used to estimate power is 0.001. Given that p is the proportion of rejected

tests and N is the number of simulation repetitions, the standard errors for the power estimates

are calculated as $SE = \sqrt{\frac{p(1-p)}{N}}$. Proportion of variation explained is estimated empirically by

dividing the difference in phenotypic variance under the null and alternative models by the

phenotypic variance under the null. All simulations and analyses are conducted using the

statistical genetics software package Mendel [41].

Model Descriptions

A variety of full and reduced models (Models 0-9) are fit to the simulated data (Table 3)

depending on the null and alternative hypotheses to be compared. In addition to genetic effects,

all models estimate parameters for a grand mean (μ), sex effect (β_{sex}), and variance components

(σ_a^2 and σ_e^2). As described above, the most general QMFG model has 20 parameters of the form

β_{ijk} where i denotes the maternal *KIR-tB* count, j denotes the maternal *HLA-C2* count, and k

denotes the offspring *HLA-C2* count. Because we have an *a priori* hypothesis about how *KIR* and

HLA-C affect birth weight, our most general *KIR-HLA-C* model, Model 4, reduces the 20

possible QMFG parameters to three parameters (Table 1). In this model, there is an effect for the

maternal *KIR* telomeric region, modeled as a *KIR-tB* dominant effect, denoted by β_{KIR} and there

is an effect for those offspring with more *HLA-C2* than their mother, denoted by β_{more}. Also in

this model, there is a *KIR-HLA-C* interaction β_{int}, an additional effect when the mother has at

least one *KIR-tB* allele and the offspring has more *HLA-C2* alleles than his/her mother. Thus

$\beta_{001} = \beta_{012} = \beta_{more}$, $\beta_{i00} = \beta_{i10} = \beta_{i11} = \beta_{i21} = \beta_{KIR}$ for $i = 1,2$, and $\beta_{i01} = \beta_{i12} = \beta_{more} +$

$\beta_{KIR} + \beta_{int}$ for $i = 1,2$. Otherwise, $\beta_{ijk} = 0$. As shown in Table 3, Models 1-3 & 5 include

additional constraints on the parameters in Model 4.

Because GWAS studies typically include only the genotypes of affected individuals

(offspring-only), Models 6-9 are used to explore the effects of misspecification when an

offspring-only model is incorrectly applied to data generated under an MFG interaction scenario.

For Models 6 and 7, offspring *HLA-C* effects are denoted by parameters $\beta_{C1/C2}$ and $\beta_{C2/C2}$.

Estimating an additive offspring *HLA-C* effect requires the additional constraint, $2\beta_{C1/C2} =$

$\beta_{C2/C2}$ (Model 7). For Models 8 and 9, offspring *KIR-tB* effects are denoted by parameters $\beta_{tA/tB}$

and $\beta_{tB/tB}$. Estimating an additive offspring *KIR-tB* effect requires the additional constraint,

$2\beta_{tA/tB} = \beta_{tB/tB}$ (Model 9).

United Kingdom Cohort

Details of the participants from the United Kingdom cohort study [16,17,29,30] were previously

published. The previous analysis of mother-offspring pairs from the United Kingdom and

Norway [16] found a significant association between the maternal *KIR* telomeric gene *KIR2DS1*

and increased offspring birth weight, especially in the presence of more *HLA-C2* in offspring

than their mothers but did not explicitly test for an interaction between these loci or determine if

there were additional independent effects of the loci. We use the mother-offspring pairs, *KIR*

telomeric haplotypes, and *HLA-C* genotypes from the United Kingdom cohort (one of the two

cohorts used by Hiby *et al*. [16]) in order to investigate MFG interactions using the QMFG test

in an effort to allow comparison to their analyses and to refine the characterization of the effects.

A detailed description of the genotyping are described elsewhere [17,30]. Both mothers and

offspring were genotyped for *HLA-C* and mothers' *KIR* haplotypes were determined. As

described previously, ethical approval was obtained from the Cambridge Research Ethics

Committee (reference nos. 01/197 and 05/Q0108/367; Cambridgeshire, U.K.). All subjects

provided informed written consent.

Pregnancies with pre-eclampsia or fetal growth restriction as well as normal pregnancies

were included in this cohort. Small babies (< 5th centile) were heavily oversampled and large

babies (≥ 90th centile) were slightly oversampled. As in Hiby *et al*. [16], we run our analyses only

on the offspring that had birth weight data > 5th centile, were firstborn singletons from full-term

births (38-42 weeks), and whose mothers were over the age of 18 and had no medical conditions

including gestational diabetes, hypertension, renal disease, and auto-immune disease. Of the 404

pregnancies that fit these criteria, one pregnancy was dropped due to a missing offspring *HLA-C*

genotype. These birth weights were approximately normally distributed.

Results

Modeling a Two-Locus Interaction Effect

To examine type I error for this extended QMFG test, data are simulated under the null of no

genetic effects (Scenario I). Based on the work of Hiby *et al.* [16], we examine the statistical properties of a model that requires that the mother have at least one copy of *KIR-tB* and the offspring have more *HLA-C2* alleles than his/her mother (Model 5) to affect a difference in birth weight. Using the a likelihood ratio test (LRT) to compare the null model of no genetic effects (Model 0) to the alternative model estimating a *KIR-HLA-C* interaction (Model 5) results in the Q-Q plot in figure 2A. All the points fall between the confidence bounds, showing that there is no bias in the type I error for the QMFG test in this scenario ($\lambda = 1.055$).

Under Scenario II, data are simulated with *KIR-HLA-C* interaction effects (β_{int}) ranging from 0.05 to 0.35 kg in increments of 0.02 kg. Fitting the same null (Model 0) and alternative (Model 5) models, the power to detect a *KIR-HLA-C* interaction effect is shown in figure 3A. When the significance level is 0.001, 80% power is reached when the *KIR-HLA-C* interaction effect is approximately 0.19 kg, that is, when the proportion of variation explained by the *KIR-HLA-C* interaction effect is approximately 0.011. Figure 4A shows the parameter estimate bias over 2000 simulation replicates when Model 5 is fit to data simulated with a *KIR-HLA-C* interaction effect of 0.19 kg. The parameter estimate biases are centered at zero. Together the type I error, power, and estimate bias demonstrate that the QMFG test has good statistical properties.

Modeling Two-Locus Main and Interaction Effects

Next, we check for bias in type I error rates for our most complex two-locus model using data simulated under the null of no genetic effects (Scenario I). Model 4, corresponding to our alternative hypothesis, estimates parameters for an effect of the dominant maternal *KIR-tB* haplotype (β_{KIR}), an effect of an offspring having more *HLA-C2* (β_{more}), and an effect of their

interaction (β_{int}). The LRT comparing Model 4 to the null model (Model 0), results in a three degrees of freedom test for main or interaction effects. Figure 2B together with the genomic control value, $\lambda = 0.991$, show there is no bias in the type I error rate. If instead Model 3 is used as the null model, the LRT tests for a *KIR-HLA-C* interaction effect (β_{int}) in the presence of their main effects ($\beta_{KIR}, \beta_{more}$). The Q-Q plot in figure 2C has the genomic control value $\lambda = 0.997$, again indicating that there is no bias in the type I error rate when using the QMFG test for this analysis.

We also conduct simulations under a scenario that involves a *KIR-HLA-C* interaction and main effects of a dominant-acting maternal *KIR-tB* haplotype and more *HLA-C2* in the offspring. To estimate power, birth weight data are simulated given a dominant acting maternal *KIR-tB* effect (β_{KIR}) of 0.05 kg, a more-*HLA-C2* effect (β_{more}) of -0.1 kg, and a *KIR-HLA-C* interaction effect (β_{int}) ranging from 0.05 to 0.35 kg (Scenario III). The power to detect either single-locus or interaction effects by comparing Model 4 to the null model (Model 0) is shown in figure 3B. The power to detect main or interaction effects is approximately 80% when the proportion of variation explained by the total of the three effects reaches 0.016, that is, when the effect size of the interaction is 0.19 kg. In this scenario, because there are effects of *KIR* and *HLA-C* even when there is no *KIR-HLA-C* interaction, the proportion of variation explained is not zero when the interaction effect size is zero. Note that as the interaction effect ranges from 0.05 to 0.09 kg (proportion of variation explained ~ 0.010), there is a small but reproducible discontinuity on the power curve. This is because the interaction effect size is similar in magnitude to the maternal-offspring *HLA-C* effect but in the opposite direction. It is therefore difficult to distinguish Model 4 from a model with only two effects ($\beta_{1ij} = \beta_{2ij}$ for all i, j and $\beta_{001} = \beta_{012}$; all other QMFG parameters are zero, $\beta_{010} = \beta_{011} = \beta_{021} = \beta_{022} = 0$) for this set of values.

All parameter estimates for these data are unbiased as shown by the boxplots of estimate bias in figure 4B. On average, the grand mean parameter is slightly underestimated, but the bias is not significantly different from zero. The power curve in figure 3C is the result of testing for a *KIR-HLA-C* interaction effect in the presence of main effects (Model 4 vs. Model 3). As expected, power for this one degree of freedom test increases as the effect of the interaction increases. When the interaction effect is 0.25 kg (proportion of variation explained equals 0.010), there is approximately 80% power to detect an interaction effect accounting for main effects.

Model Misspecification

We imagine that in some cases an MFG interaction may be poorly understood, so we examine how power is altered when fitting a model with more parameters than needed. Consider the situation where the true underlying effect is a *KIR-HLA-C* interaction effect with no main effects (Scenario II) and the user fits a model with parameters for *KIR* and *HLA-C* main effects in addition to the interaction (Model 4). Power to detect *KIR*, *HLA-C*, or two-locus interaction effects by fitting Model 4 to Scenario II data is shown in figure 1S. As expected, this three degrees of freedom test results in reduced power compared to the more appropriate one degree of freedom test shown in figure 3A. When the *KIR-HLA-C* interaction effect is equal to 0.19 kg, power drops from 0.82 to 0.62 when Model 4 is used as the alternative model rather than Model 5. The parameter estimate bias boxplots remain approximately centered at zero (fig. 2S). Thus as expected, when the model is misspecified to include main effects in addition to the true underlying interaction effect, power is decreased but parameter estimates are unbiased.

Effects of Using a Standard Offspring Model

An additional question of interest is whether a standard GWAS that uses only offspring genotypes at a single locus would be able to find evidence for an association to either of the two loci that make up the MFG interaction. For simulations in which an interaction between the maternal *KIR* telomeric region and maternal-offspring *HLA-C* provides the only genetic effect on birth weight (Scenario II), we investigate whether an offspring only analysis would detect an association at *HLA-C*. When data simulated under Scenario II ($\beta_{int} = \{0.05, 0.07, ..., 0.35\}$) are tested using an offspring *HLA-C* genotype model (Model 6, Table 3), power is drastically reduced (fig. 5A, solid line) compared to a QMFG model that includes an interaction effect (Models 4 or 5, Table 3). Here, Model 6 reflects the alternative hypothesis and Model 0 reflects the null hypothesis. The parameter estimate bias and boxplots for Model 6 analyses are displayed in figure 4C corresponding to data simulated with an interaction effect size of 0.35 kg (Scenario II). Over the 2000 simulations, the grand mean is slightly underestimated and the variance components are overestimated. Figure 5B (solid line) shows the power when testing Scenario II generated data using an additive offspring *HLA-C* model (Model 7 vs. Model 0). Power is slightly higher than for the offspring genotypic test, likely due to the reduced degrees of freedom, but remains severely attenuated compared to an analysis using Model 4 or 5.

Analogously, the offspring model might be used to test for an association of the *KIR* telomeric region with the trait. When using the data simulated under Scenario II and an offspring genotypic (Model 8) or additive (Model 9) model to detect an effect at the *KIR* locus, power is again drastically reduced (fig. 3SA and 3SB, respectively). When the *KIR-HLA-C* interaction effect size equals 0.19 kg (an effect large enough to yield 0.82 power when the correct model, Model 5, is used), the power to detect an offspring genotypic effect is ~0.001 when the

significance level is 0.001 and 0.0495 when the significance level is 0.05. These results are equivalent to the size of the test and are consistent with no genetic effect. Figure 4S shows the grand mean and genotypic parameters estimates are biased upwards for the genotypic model. Taken together, these results demonstrate that MFG multi-locus interactions would often be missed in standard GWAS analyses.

Effect of Population Frequency on the Standard Offspring Model

In the previous sections, *KIR-tB* had a frequency of 0.2 based on what has been observed in the white British population. It has previously been reported that in representative African populations, the frequency of the *KIR2DS1* gene is 0.07 [42]. Knowing that haplotype frequencies differ between populations, we evaluate how haplotype frequency for an intricate interaction such as this might change power when using standard offspring-effect models. We simulate samples with *KIR-tB* frequencies equal to $0.07, 0.2, 0.6,$ and 1.0 and use *KIR-HLA-C* interaction effect sizes ranging from 0.05 to 0.35 kg (Scenario II). As the *KIR-tB* frequency increases, power also increases for both the offspring *HLA-C* genotypic and additive models (fig. 5A and 5B, respectively). For the specific case where the *KIR-HLA-C* interaction effect size is 0.25 kg (fig. 5A, triangles), the estimated power to detect a genotypic offspring *HLA-C* effect ranges from 0.0035 (SE = 0.0013) when P(*KIR-tB*) = 0.07 to 0.78 (SE = 0.009) when P(*KIR-tB*) = 1.0. When testing for an additive offspring *HLA-C* effect the power ranges from 0.0035 (SE = 0.0013) when P(*KIR-tB*) = 0.07 to 0.86 (SE = 0.008) when P(*KIR-tB*) = 1.0 (fig. 5B, triangles). Thus, when using a standard offspring-only analysis, conclusions about the importance of variation in the *HLA-C* on the birth weight depend on the *KIR-tB* haplotype frequency in the mothers.

Effect of Population Frequency on the Single-Locus QMFG Test

Given the strong influence of the *KIR-tB* frequency on the standard offspring-only analysis, we next consider the effect of *KIR-tB* frequency when the *HLA-C* maternal-offspring effect is modeled properly but the effect of the maternal *KIR* is ignored. In other words, we are testing the more-offspring-*HLA-C2* effect using the single-locus QMFG test when a multi-locus QMFG test is appropriate. Setting *KIR-tB* frequencies equal to $0.07, 0.2, 0.6$, and 1.0, we simulate data with a *KIR-HLA-C* interaction effect (Scenario II) and test for an effect of more offspring *HLA-C2*. Here Model 2 corresponds to our alternative hypothesis and Model 0 corresponds to the null. As shown in figure 6, power to detect the *HLA-C* effect increases greatly as the *KIR-tB* frequency increases. When the *KIR-HLA-C* interaction effect size is 0.13 kg (proportion of variation explained ~0.006), power to detect a more-offspring-*HLA-C2* effect jumps from 0.0065 (SE = 0.0018) when P(*KIR-tB*) = 0.07 to 0.85 (SE = 0.008) when P(*KIR-tB*) = 1.0 (fig. 6, triangles). When using a single locus MFG test that explicitly models the HLA-C effects, the ability to find a significant effect of *HLA-C* on the trait depends on the population frequency of the *KIR-tB* haplotype.

Effect of Family Structure

Although one of the benefits of the QMFG test is the ability to use large families, genotyping extended families is not always feasible. To investigate the impact of family structure, we consider the minimal design that can be used and simulate 1500 mother-offspring pairs (750 female offspring and 750 male offspring). Note that when using mother-offspring pairs (or equivalently parent-offspring trios), the additive and environmental variances cannot be partitioned, so instead of estimating them separately, we estimate their sum (residual variance).

We consider how power and parameter estimates are affected when using mother-offspring pairs for data simulated under Scenario III. The three degrees of freedom LRT for *KIR*, *HLA-C*, or interaction effects (Model 4 vs Model 0) results in the power curve shown in figure 5S. Power decreases slightly compared to the power estimates for three-generation families (fig. 3B). For simulations with the *KIR-HLA-C* interaction effect equal to 0.19 kg, power drops from 0.793 when using three-generation families to 0.759 when mother-offspring pairs are used. As shown in figure 6S, parameter estimate bias is approximately centered at zero for the 2000 repetitions.

Application of the QMFG Test to UK Cohort Birth Weight Data

Previously published association analyses for birth weight using *KIR* and *HLA-C* looked at the effects of maternal *KIR* telomeric haplotypes in subgroups of offspring based on their maternal-fetal *HLA-C* genotypes [16]. These analyses did not test whether there were differences in the maternal *KIR* effect sizes between these subgroups. Using the QMFG test, we can extend these previous analyses by building a hierarchical model of *KIR*, *HLA-C*, and birth weight. With this linear model, we can directly determine if the effect of maternal *KIR-tB* on the offspring's birth weight varies depending on whether the offspring has more *HLA-C2* by testing the statistical significance of the *KIR-HLA-C* interaction parameter.

As shown in our simulation results, interactions between maternal and offspring loci can easily and accurately be modeled and tested using the QMFG analysis method. Using the same UK mother-offspring pairs as Hiby *et al*. [16], we fit various models (Models 1-5) that include single-locus and two-locus interaction effects between the *KIR* telomeric region and *HLA-C* genotypes for mothers and their offspring as well as the null model (Model 0). To compare how well each model fits these data, we use the Akaike Information Criterion (AIC). From the

candidate models, we can then select the model that minimizes the AIC as our best model. Table 4 shows the log-likelihoods and AIC values for each of our candidate models. For these data, the minimum AIC is provided by Model 4, which includes effects for the maternal *KIR-tB* haplotype and for more offspring *HLA-C2* than maternal *HLA-C2*, in addition to the *KIR-HLA-C* interaction.

The effect estimates for Model 4 are shown in Table 5 (second column). For this model, the reference group consists of offspring that have less or equal *HLA-C2* alleles than their mother and also have mothers with a *tA/tA* genotype at *KIR*. According to our analysis, offspring with more *HLA-C2* who also have mothers with *KIR* genotype *tA/tA* are born on average 0.1863 kg smaller compared to the reference group adjusting for sex (p = 0.005). Offspring with less or equal *HLA-C2* than their mother who also have a mother with at least one copy of the *KIR-tB* haplotype are estimated to be 0.0537 kg larger at birth compared to the reference group (p = 0.28). The interaction effect estimate indicates that having both more offspring *HLA-C2* and a mother with at least one *KIR-tB* confers the smallest change in birth weight compared to the reference group, increasing birth weight by 0.0119 kg adjusting for sex (p = 0.14). Examining Tables 4 and 5 and using likelihood ratio tests, we find little support for an MFG interaction, with (p = 0.14) or without (p = 0.75) accounting for the single locus effects. However there appear to be independent effects of KIR-tB and HLA-C2 (p = 0.006). Hiby and colleagues [16] do not adjust for gestational age in the subgroup analyses of their combined UK and Norwegian cohorts since they find that the effect of *KIR2DS1* on birth weight for is independent of gestational age. Since we are using only one of these cohorts, we also fit Model 4 adjusting for gestational age as a sensitivity analysis and discover that the effect estimates (Table 5, third column) and their p-values do not differ greatly.

Discussion

The link between maternal uNK cells and offspring trophoblast cells within the placenta presents a plausible setting for MFG interactions that affect fetal development. Motivated by a previous study that found a significant effect of the maternal *KIR2DS1* gene on human birth weight for offspring with more *HLA-C2* than their mother [16], we extend the QMFG analysis method to multiple loci and test the hypothesis of a *KIR-HLA-C* interaction effect on birth weight. In this article, we model the effect of having more *HLA-C2* present in the offspring's genotype than the mother's and a maternal dominant *KIR-tB* effect to allow a comparison to the analyses of Hiby *et al*. [16]. As discussed by these researchers [16], this model is a surrogate for a model that captures the effects of the maternal immune response to the fetus having an HLA-C antigen that is foreign to her antigens. That antigen would have to originate from the paternally derived gene. Although other models for the effects of the *KIR-tB* and *HLA-C* alleles are possible, we chose not to explore them in this article because of the limited sample size and constraints in the study design. In particular, Hiby's study design resulted in no fathers' *HLA-C* genotypes and therefore comparing the fit of the models presented in this article to a paternal parent of origin effect, a very plausible alternative model, is not possible (refer to Appendix C for further discussion).

Our simulation studies show the statistical validity of the QMFG test when extended to interactions between the maternal *KIR* telomeric haplotype, maternal *HLA-C* genotype, and offspring *HLA-C* genotype. In situations in which the model assumptions are consistent with the simulation scenario, the QMFG test has valid type I error rates, parameter estimate bias centered around zero, and high power even when the proportion of variation explained is low. To verify that these features are not greatly altered by family structure, we simulate data for mother-

offspring pairs, and find that parameter estimates remain unbiased and power is only slightly diminished.

We also investigate scenarios in which the true underlying MFG interaction is partially or entirely misspecified. As expected, power is reduced but parameter estimates are unbiased when a more general model is used in a case where a restricted model is sufficient. We also show that if a standard model, which considers only offspring genotypes, is fit to data generated under an underlying *KIR-HLA-C* interaction scenario, associations can be missed or identified incorrectly. The ability to detect an association at a locus involved in an MFG interaction using only offspring genotypes depends on the underlying nature of the MFG interaction, but in general power to detect the locus is greatly diminished. Similarly, if a single-locus QMFG test is used when two or more polymorphic loci are involved in the MFG interaction, power will likely be reduced. In particular using the offspring-only or single-locus QMFG analyses, if a researcher is unaware of the existence of the second polymorphic locus (*KIR*), then they might find that an *HLA-C* effect in a population where the *KIR-tB* haplotype is frequent, but fail to replicate the effect in a second population where the *KIR-tB* haplotype is infrequent even when *HLA-C* allele frequencies are the same in these two populations. The researcher would then be inclined to dismiss the first result as a false positive. Thus, like the case of offspring gene-gene interactions, models that incompletely capture the MFG inter-locus effects can easily lead to incorrect conclusions.

To date, no other study has directly looked for an association between *KIR-HLA-C* interactions and birth weight. We use a linear mixed modeling approach to analyze data from the United Kingdom cohort. Using minimum AIC to compare models, the model with effects for maternal *KIR-tB* and more *HLA-C2* alleles in the offspring's genotype than the mother's, and

their interaction, is determined to provide the best fit, however a model with just the independent locus effects (maternal *KIR-tB* and more *HLA-C2* alleles in offspring's genotype than the mother's) fits the data nearly as well. In the model with both loci and interaction effects (Model 4), *KIR* and *HLA-C* explain 3.0% of the phenotypic variation in human birth weight.

Using subgroup analyses, Hiby and colleagues [16] found that the effect of maternal *KIR* on birth weight was significant in offspring with more *HLA-C2* than their mothers and that this maternal effect was not significant in offspring with less or equal *HLA-C2* when they used mothers and offspring from both Norway [43] and the United Kingdom. From their analyses, it is difficult to determine whether the effect is exclusively an *KIR-HLA-C* interaction effect or whether than are also main effects of maternal *KIR* and more *HLA-C2* in the offspring's genotype than the mother's genotype. An advantage of using the QMFG test instead of conducting subgroup analyses is the ability to determine parameter estimates for maternal *KIR-tB*, more offspring *HLA-C2*, and *KIR-HLA-C* interaction effects. Furthermore, we can explicitly test whether the effect of more *HLA-C2* alleles in the presence of maternal *KIR-tB* (that is, the interaction) is statistically significant after accounting for the effects of the two loci acting independently (main effects). Using only the United Kingdom cohort we find that, in the presence of main effects, there is not a significant *KIR-HLA-C* interaction. There is a strong main effect of more offspring *HLA-C2*, whereas a *KIR* effect conditional on a *HLA-C* effect is less certain. Rather than speculate on differences in the previous study's modeling approaches to ours with regard to power and assumptions, we plan to analyze the same combined Norway [43] and United Kingdom cohort data as Hiby *et al.* [16] in the future. This reanalysis will provide a direct comparison of the two approaches and therefore will be more meaningful.

When the data consist of only mother-offspring pairs, general statistical packages that

include linear mixed model options can be used. However, for studies that have collected data from families of varying sizes, specialized software is needed. To perform our analyses, we extended the QMFG test to handle multiple loci in the statistical genetics software package Mendel. These extensions will be available in the next version of the freely available Mendel package. The power of our method to detect significant MFG interactions and our flexible software make the QMFG test an effective tool to consider when studying genetic factors associated with complex traits.

Table 1. Maternal-offspring genotype combinations and model parameters for the *KIR* telomeric region and *HLA-C*.

Maternal *KIR* telomeric region	Maternal *HLA-C*	Offspring *HLA-C*	General QMFG model	Model 4: Maternal *KIR*, more offspring *HLA-C2*, and interaction effects	Model 6: *HLA-C* offspring effects	Model 8: *KIR* offspring effects
tA/tA	*C1/C1*	*C1/C1*	0	0	0	0
tA/tA	*C1/C1*	*C1/C2*	β_{001}	β_{more}	$\beta_{C1/C2}$	0
tA/tA	*C1/C2*	*C1/C1*	β_{010}	0	0	0
tA/tA	*C1/C2*	*C1/C2*	β_{011}	0	$\beta_{C1/C2}$	0
tA/tA	*C1/C2*	*C2/C2*	β_{012}	β_{more}	$\beta_{C2/C2}$	0
tA/tA	*C2/C2*	*C1/C2*	β_{021}	0	$\beta_{C1/C2}$	0
tA/tA	*C2/C2*	*C2/C2*	β_{022}	0	$\beta_{C2/C2}$	0
tA/tB	*C1/C1*	*C1/C1*	β_{100}	β_{KIR}	0	$\beta_{tA/tB}$
tA/tB	*C1/C1*	*C1/C2*	β_{101}	$\beta_{KIR} + \beta_{more} + \beta_{int}$	$\beta_{C1/C2}$	$\beta_{tA/tB}$
tA/tB	*C1/C2*	*C1/C1*	β_{110}	β_{KIR}	0	$\beta_{tA/tB}$
tA/tB	*C1/C2*	*C1/C2*	β_{111}	β_{KIR}	$\beta_{C1/C2}$	$\beta_{tA/tB}$
tA/tB	*C1/C2*	*C2/C2*	β_{112}	$\beta_{KIR} + \beta_{more} + \beta_{int}$	$\beta_{C2/C2}$	$\beta_{tA/tB}$
tA/tB	*C2/C2*	*C1/C2*	β_{121}	β_{KIR}	$\beta_{C1/C2}$	$\beta_{tA/tB}$
tA/tB	*C2/C2*	*C2/C2*	β_{122}	β_{KIR}	$\beta_{C2/C2}$	$\beta_{tA/tB}$
tB/tB	*C1/C1*	*C1/C1*	β_{200}	β_{KIR}	0	$\beta_{tB/tB}$
tB/tB	*C1/C1*	*C1/C2*	β_{201}	$\beta_{KIR} + \beta_{more} + \beta_{int}$	$\beta_{C1/C2}$	$\beta_{tB/tB}$
tB/tB	*C1/C2*	*C1/C1*	β_{210}	β_{KIR}	0	$\beta_{tB/tB}$
tB/tB	*C1/C2*	*C1/C2*	β_{211}	β_{KIR}	$\beta_{C1/C2}$	$\beta_{tB/tB}$
tB/tB	*C1/C2*	*C2/C2*	β_{212}	$\beta_{KIR} + \beta_{more} + \beta_{int}$	$\beta_{C2/C2}$	$\beta_{tB/tB}$
tB/tB	*C2/C2*	*C1/C2*	β_{221}	β_{KIR}	$\beta_{C1/C2}$	$\beta_{tB/tB}$
tB/tB	*C2/C2*	*C2/C2*	β_{222}	β_{KIR}	$\beta_{C2/C2}$	$\beta_{tB/tB}$

Table 2. Simulation scenarios.

Scenario	Genetic effects	Simulated values*
I	None	$\beta_{KIR} = \beta_{more} = \beta_{int} = 0$
II	Interaction	$\beta_{KIR} = \beta_{more} = 0$ $\beta_{int} = \{0.05, 0.07, \ldots, 0.35\}$
III	Maternal *KIR-tB* More *HLA-C2* Interaction	$\beta_{KIR} = 0.05$ $\beta_{more} = -0.1$ $\beta_{int} = \{0.05, 0.07, \ldots, 0.35\}$

*All models additionally simulate a grand mean µ (3.5 kg), sex effect (β_{sex} = -0.2 kg if female), and variance components (σ_a^2 = 0.2025 kg² and σ_e^2 = 0.0475 kg²).

Table 3. Model descriptions.

Type	Model	Genetic effects modeled	Additional Constraints	Genetic parameters estimated*
Null model[1]	0	None	$\beta_{KIR} = \beta_{more} = \beta_{int} = 0$ or $\beta_{C1/C2} = \beta_{C2/C2} = 0$ or $\beta_{tA/tB} = \beta_{tB/tB} = 0$	0
QMFG models[2]	1	Maternal *KIR-tB*	$\beta_{more} = \beta_{int} = 0$	β_{KIR}
	2	More *HLA-C2*	$\beta_{KIR} = \beta_{int} = 0$	β_{more}
	3	Maternal *KIR-tB* More *HLA-C2*	$\beta_{int} = 0$	β_{KIR} β_{more}
	4	Maternal *KIR-tB* More *HLA-C2* Interaction	None	β_{KIR} β_{more} β_{int}
	5	Interaction	$\beta_{KIR} = \beta_{more} = 0$	β_{int}
Standard offspring models[3]	6	Genotypic offspring *HLA-C* effect	None	$\beta_{C1/C2}$ $\beta_{C2/C2}$
	7	Additive offspring *HLA-C* effect	$2\beta_{C1/C2} = \beta_{C2/C2}$	$\beta_{C1/C2}$
	8	Genotypic offspring *KIR* effect	None	$\beta_{tA/tB}$ $\beta_{tB/tB}$
	9	Additive offspring *KIR* effect	$2\beta_{tA/tB} = \beta_{tB/tB}$	$\beta_{tA/tB}$

*All models additionally estimate a grand mean (μ), sex effect (β_{sex}), and variance components (σ_a^2 and σ_e^2).
[1]Model 0 is a sub-model of all other models 1-9
[2]Models 1-3 & 5 are sub-models of Model 4 (parameterization shown in Table 1)
[3]Model 7 is a sub-model of Model 6 and Model 9 is a sub-model of Model 8 (parameterization of Models 6 and 8 shown in Table 1)

Table 4. Comparison of *KIR-HLA-C* models for real data (n = 403).

Model	Genetic effects	No. genetic parameters[1]	Log-likelihood	Proportion of variation explained[2]	AIC	ΔAIC from best model
0	None	0	137.45	N/A	-272.89	6.33
1	Maternal *KIR-tB*	1	139.46	0.010	-274.92	4.30
2	More *HLA-C2*	1	140.32	0.015	-276.64	2.58
3	Maternal *KIR-tB* More *HLA-C2*	2	142.53	0.025	-279.06	0.15
4	Maternal *KIR-tB* More *HLA-C2* Interaction	3	143.61	0.030	-279.22	0
5	Interaction	1	137.50	0.0005	-270.99	8.23

[1]All models additionally adjusted for sex
[2]Proportion of residual variation explained by *KIR* and *HLA-C* effects compared to Model 0.

Table 5. Effect estimates for Model 4 with and without adjusting for gestational age.

Covariates	Sex of fetus	Sex of fetus and gestational age
Number of subjects	403	403
Mean effect of maternal *KIR-tB* (SE)	0.0537 kg (0.0495)	0.0671 kg (0.0454)
Mean effect of more *HLA-C2* (SE)	-0.1863 kg (0.0656)	-0.1708 kg (0.0598)
Mean effect of interaction	0.1445 kg (0.0983)	0.1119 kg (0.0901)
Residual variance (SE)	0.1804 kg^2 (0.0127)	0.1494 kg^2 (0.0106)
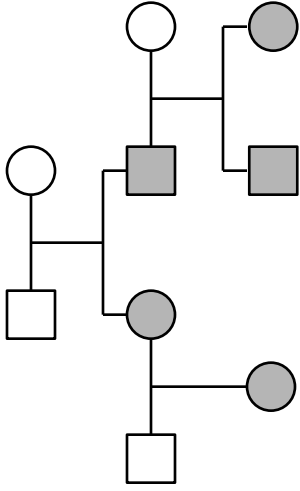
Fig. 1. Simulated family structure for Scenarios I, II, and III. Offspring used in simulation analyses are shaded gray.
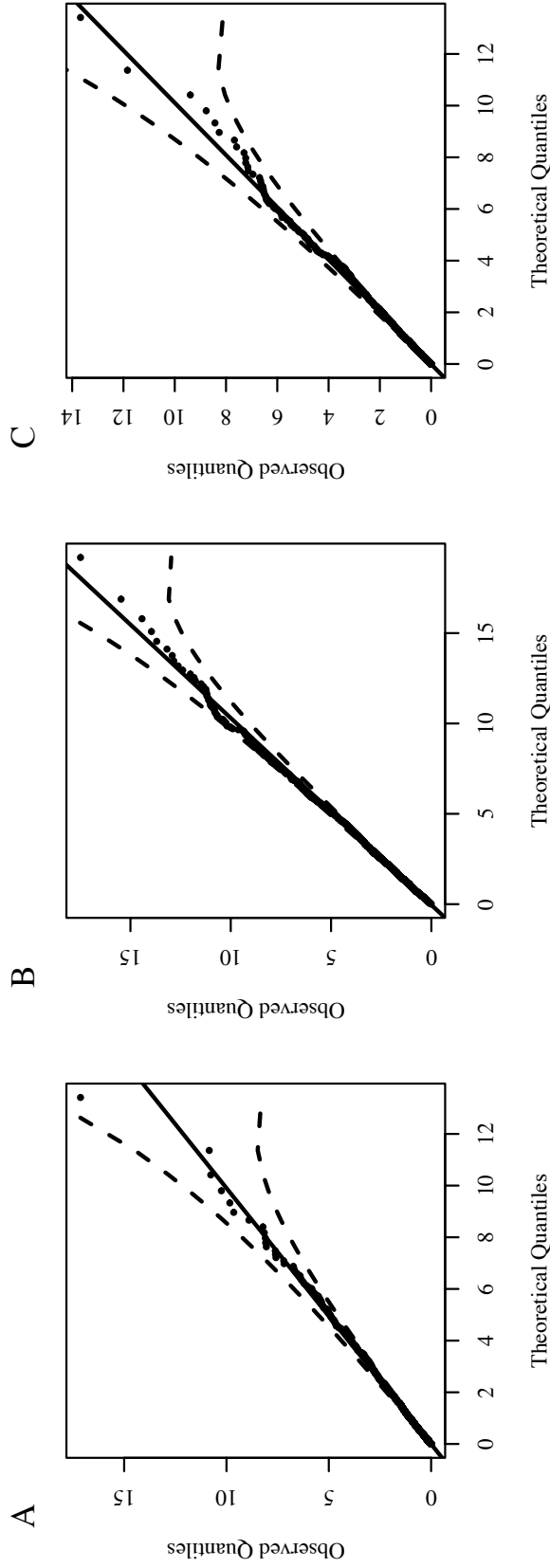
Fig. 2. Q-Q plots testing for *KIR* and *HLA-C* MFG effects. Genotypes and birth weight phenotypes were simulated for 300 pedigrees under the null of no genetic effects (Scenario I). Over 2000 repetitions the QMFG analysis was used to test an effect of (A) a *KIR-HLA-C* interaction (df = 1; $\lambda = 1.055$), (B) maternal *KIR-tB*, more offspring *HLA-C2*, or a *KIR-HLA-C* interaction (df = 2; $\lambda = 0.991$), and (C) a *KIR-HLA-C* interaction in the presence of maternal *KIR-tB* and more offspring *HLA-C2* main effects (df = 1; $\lambda = 0.997$).

74
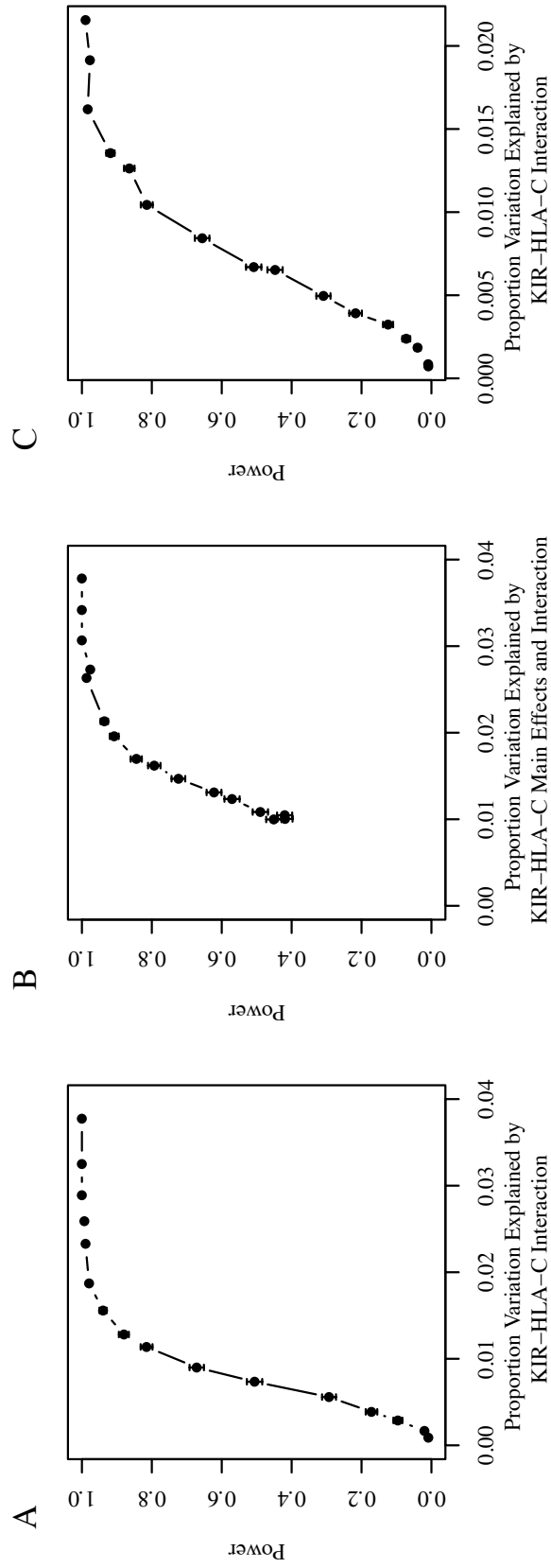
Fig. 3. Power to detect *KIR* and *HLA-C* MFG effects. Error bars represent approximate 95% confidence intervals. (A) Power to detect a *KIR-HLA-C* interaction effect when data are simulated under Scenario II (*KIR-HLA-C* interaction only). QMFG test for a *KIR-HLA-C* interaction effect (df = 1). (B) Power to detect *KIR*, *HLA-C*, or two-locus interaction effects when data are simulated under Scenario III (maternal *KIR*, more *HLA-C2* in offspring, and *KIR-HLA-C* interaction; df = 3). (C) Power to detect a *KIR-HLA-C* interaction effect in presence of main effects when data are simulated under Scenario III (df = 1).

75
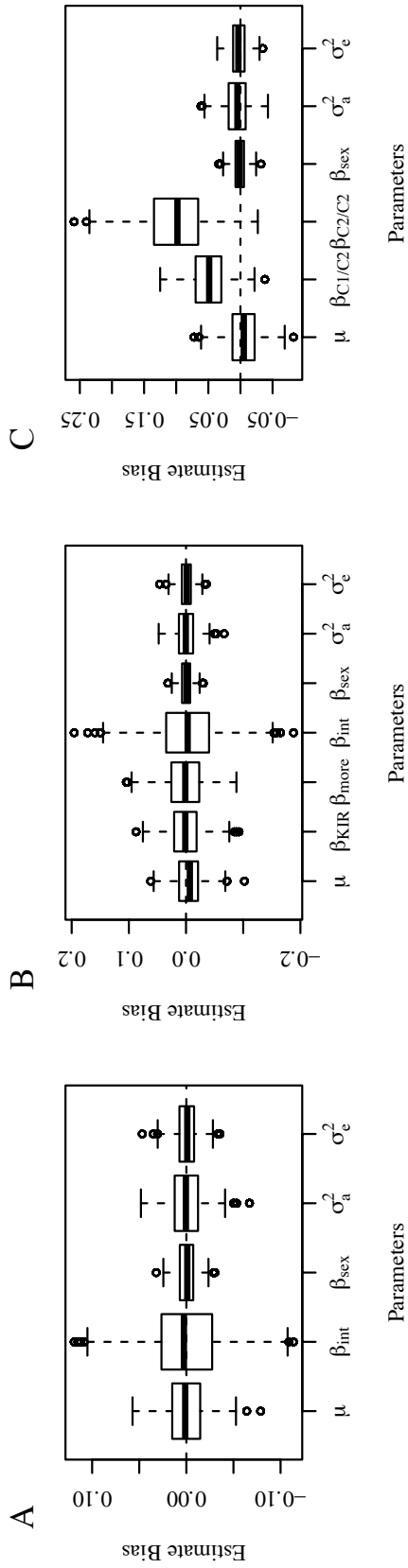
Fig. 4. Parameter estimate bias. Birth weight phenotypes simulated given $\mu = 3.5$ kg, $\beta_{sex} = -0.2$ kg, $\sigma_a^2 = 0.2025$ kg^2, and $\sigma_e^2 = 0.0475$ kg^2 as well as additional *KIR* and *HLA-C* effects. (A) A *KIR-HLA-C* interaction only model is fit to Scenario II (*KIR-HLA-C* interaction only) data with a *KIR-HLA-C* interaction effect of 0.19 kg ($\beta_{int} = 0.19$ kg). (B) A model with parameters for maternal *KIR*, more *HLA-C2* in offspring, and *KIR-HLA-C* interaction is fit to Scenario III (maternal *KIR*, more *HLA-C2* in offspring, and *KIR-HLA-C* interaction) data with one-locus and two-locus interaction effects ($\beta_{KIR} = 0.05$ kg, $\beta_{more} = -0.1$ kg, $\beta_{int} = 0.19$ kg). (C) Model is misspecified as a genotypic offspring HLA-C model and is fit to Scenario II data with a *KIR-HLA-C* interaction effect of 0.35 kg ($\beta_{int} = 0.35$ kg).

76
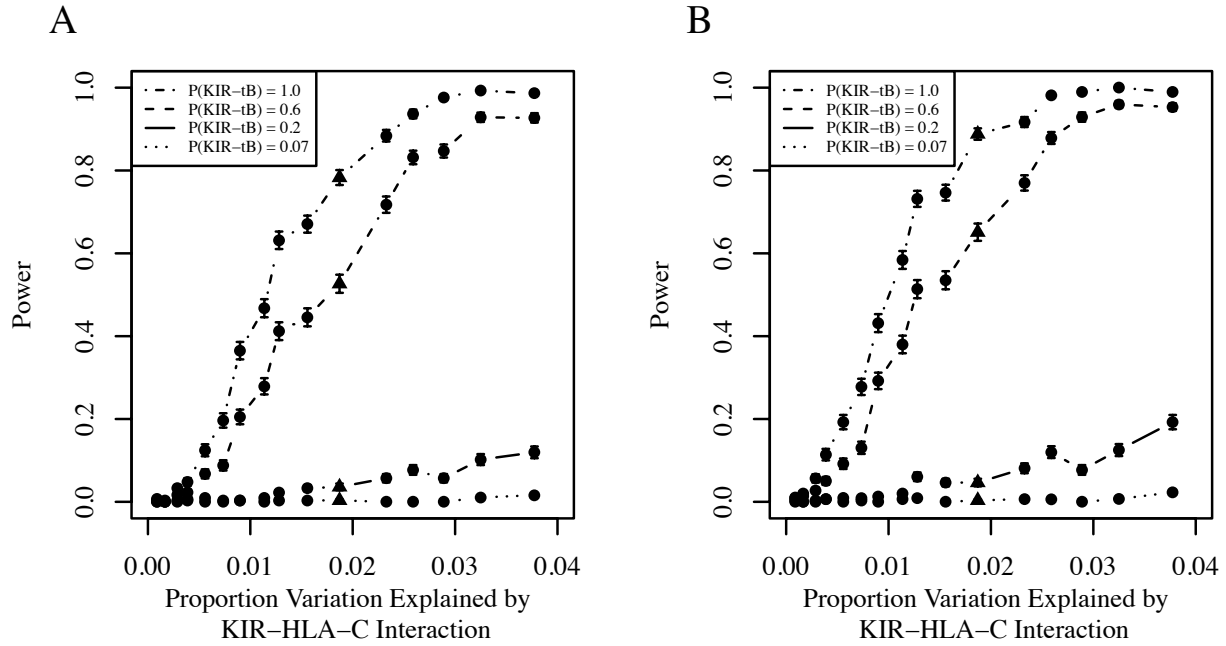
Fig. 5. Power to detect *HLA-C* offspring effects. Error bars represent approximate 95% confidence intervals. Data simulated under Scenario II (*KIR-HLA-C* interaction only) and the model is misspecified such that it tests for (A) genotypic offspring *HLA* effects (df = 2) and (B) additive offspring *HLA* effects (df = 1). Simulations in which the *KIR-HLA-C* interaction effect is 0.25 kg (the effect size when power is close to 80% for the genotypic model when P(*KIR-tB*) = 1.0) are represented with triangles.
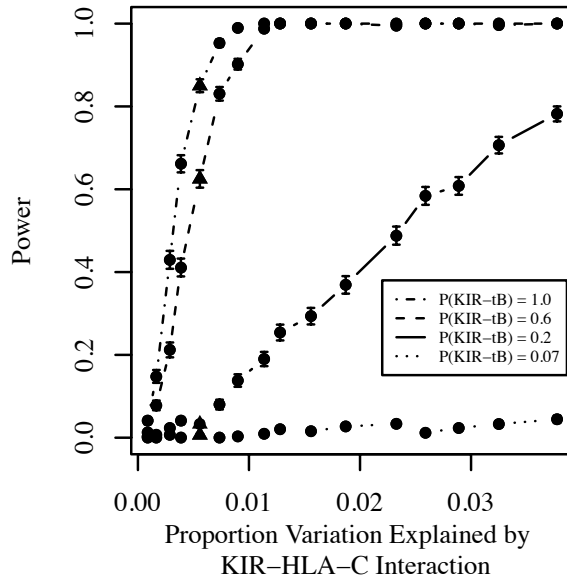
Fig. 6. Effect of *KIR-tB* frequency on the single-locus QMFG test. Data simulated under Scenario II (*KIR-HLA-C* interaction only) and the model is misspecified such that it tests for an effect of more offspring *HLA-C2* (df = 1). Simulations in which the *KIR-HLA-C* interaction effect is 0.13 kg (the effect size when power is 85% when P(*KIR-tB*) = 1.0) are represented with triangles. Error bars represent approximate 95% confidence intervals.

References

1   Liang M, Wang X, Li J, Yang F, Fang Z, Wang L, Hu Y, Chen D: Association of combined maternal-fetal TNF-alpha gene G308A genotypes with preterm delivery: a gene-gene interaction study. J Biomed Biotechnol 2010;2010:396184.

2   Li M, Erickson SW, Hobbs CA, Li J, Tang X, Nick TG, Macleod SL, Cleves MA, National Birth Defect Prevention Study: Detecting maternal-fetal genotype interactions associated with conotruncal heart defects: a haplotype-based analysis with penalized logistic regression. Genet Epidemiol 2014;38:198-208.

3   Lupo PJ, Mitchell LE, Canfield MA, Shaw GM, Olshan AF, Finnell RH, Zhu H, National Birth Defects Prevention Study: Maternal-fetal metabolic gene-gene interactions and risk of neural tube defects. Mol Genet Metab 2014;111:46-51.

4   Procopciuc LM, Caracostea G, Zaharie G, Stamatian F: Maternal/newborn VEGF-C936T interaction and its influence on the risk, severity and prognosis of preeclampsia, as well as on the maternal angiogenic profile. J Matern Fetal Neonatal Med 2014;27:1754-1760.

5   Li M, Li J, Wei C, Lu Q, Tang X, Erickson SW, MacLeod SL, Hobbs CA: A Three-Way Interaction among Maternal and Fetal Variants Contributing to Congenital Heart Defects. Ann Hum Genet 2016;80:20-31.

6   Stubbs EG, Ritvo ER, Mason-Brothers A: Autism and shared parental HLA antigens. J Am Acad Child Psychiatry 1985;24:182-185.

7   Hollister JM, Laing P, Mednick SA: Rhesus incompatibility as a risk factor for schizophrenia in male adults. Arch Gen Psychiatry 1996;53:19-24.

8   Dahlquist GG, Patterson C, Soltesz G: Perinatal risk factors for childhood type 1 diabetes in Europe. The EURODIAB Substudy 2 Study Group. Diabetes Care 1999;22:1698-1702.

9   Juul-Dam N, Townsend J, Courchesne E: Prenatal, perinatal, and neonatal factors in autism, pervasive developmental disorder-not otherwise specified, and the general population. Pediatrics 2001;107:E63.

10   Cannon M, Jones PB, Murray RM: Obstetric complications and schizophrenia: historical and meta-analytic review. Am J Psychiatry 2002;159:1080-1092.

11   Palmer CG, Turunen JA, Sinsheimer JS, Minassian S, Paunio T, Lonnqvist J, Peltonen L, Woodward JA: RHD maternal-fetal genotype incompatibility increases schizophrenia susceptibility. Am J Hum Genet 2002;71:1312-1319.

12   Newton JL, Harney SM, Wordsworth BP, Brown MA: A review of the MHC genetics of rheumatoid arthritis. Genes Immun 2004;5:151-157.

13   Insel BJ, Brown AS, Bresnahan MA, Schaefer CA, Susser ES: Maternal-fetal blood incompatibility and the risk of schizophrenia in offspring. Schizophr Res 2005;80:331-342.

14   Palmer CG, Mallery E, Turunen JA, Hsieh HJ, Peltonen L, Lonnqvist J, Woodward JA, Sinsheimer JS: Effect of Rhesus D incompatibility on schizophrenia depends on offspring sex. Schizophr Res 2008;104:135-145.

15   Freedman D, Deicken R, Kegeles LS, Vinogradov S, Bao Y, Brown AS: Maternal-fetal blood incompatibility and neuromorphologic anomalies in schizophrenia: Preliminary findings. Prog Neuropsychopharmacol Biol Psychiatry 2011;35:1525-1529.

16   Hiby SE, Apps R, Chazara O, Farrell LE, Magnus P, Trogstad L, Gjessing HK, Carrington M, Moffett A: Maternal KIR in combination with paternal HLA-C2 regulate human birth weight. J Immunol 2014;192:5069-5073.

17   Hiby SE, Apps R, Sharkey AM, Farrell LE, Gardner L, Mulder A, Claas FH, Walker JJ, Redman CW, Morgan L, Tower C, Regan L, Moore GE, Carrington M, Moffett A: Maternal

activating KIRs protect against human reproductive failure mediated by fetal HLA-C2. J Clin Invest 2010;120:4102-4110.

18   Moffett A, Colucci F: Co-evolution of NK receptors and HLA ligands in humans is driven by reproduction. Immunol Rev 2015;267:283-297.

19   Siggelkow W, Boehm D, Skala C, Grosslercher M, Schmidt M, Koelbl H: The influence of macrosomia on the duration of labor, the mode of delivery and intrapartum complications. Arch Gynecol Obstet 2008;278:547-553.

20   Odegard RA, Vatten LJ, Nilsen ST, Salvesen KA, Austgulen R: Preeclampsia and fetal growth. Obstet Gynecol 2000;96:950-955.

21   Moffett A, Loke C: Immunology of placentation in eutherian mammals. Nat Rev Immunol 2006;6:584-594.

22   Apps R, Murphy SP, Fernando R, Gardner L, Ahad T, Moffett A: Human leucocyte antigen (HLA) expression of primary trophoblast cells and placental cell lines, determined using single antigen beads to characterize allotype specificities of anti-HLA antibodies. Immunology 2009;127:26-39.

23   Robinson J, Halliwell JA, Hayhurst JD, Flicek P, Parham P, Marsh SG: The IPD and IMGT/HLA database: allele variant databases. Nucleic Acids Res 2015;43:D423-431.

24   Chazara O, Xiong S, Moffett A: Maternal KIR and fetal HLA-C: a fine balance. J Leukoc Biol 2011;90:703-716.

25   Kim S, Poursine-Laurent J, Truscott SM, Lybarger L, Song YJ, Yang L, French AR, Sunwoo JB, Lemieux S, Hansen TH, Yokoyama WM: Licensing of natural killer cells by host major histocompatibility complex class I molecules. Nature 2005;436:709-713.

26   Kieckbusch J, Gaynor LM, Moffett A, Colucci F: MHC-dependent inhibition of uterine NK cells impedes fetal growth and decidual vascular remodelling. Nat Commun 2014;5:3359.

27   Marsh SG, Parham P, Dupont B, Geraghty DE, Trowsdale J, Middleton D, Vilches C, Carrington M, Witt C, Guethlein LA, Shilling H, Garcia CA, Hsu KC, Wain H: Killer-cell immunoglobulin-like receptor (KIR) nomenclature report, 2002. Tissue Antigens 2003;62:79-86.

28   Stewart CA, Laugier-Anfossi F, Vely F, Saulquin X, Riedmuller J, Tisserant A, Gauthier L, Romagne F, Ferracci G, Arosa FA, Moretta A, Sun PD, Ugolini S, Vivier E: Recognition of peptide-MHC class I complexes by activating killer immunoglobulin-like receptors. Proc Natl Acad Sci U S A 2005;102:13224-13229.

29   Hiby SE, Regan L, Lo W, Farrell L, Carrington M, Moffett A: Association of maternal killer-cell immunoglobulin-like receptors and parental HLA-C genotypes with recurrent miscarriage. Hum Reprod 2008;23:972-976.

30   Hiby SE, Walker JJ, O'Shaughnessy K M, Redman CW, Carrington M, Trowsdale J, Moffett A: Combinations of maternal KIR and fetal HLA-C genes influence the risk of preeclampsia and reproductive success. J Exp Med 2004;200:957-965.

31   Kistner EO, Weinberg CR: Method for using complete and incomplete trios to identify genes related to a quantitative trait. Genet Epidemiol 2004;27:33-42.

32   Kistner EO, Weinberg CR: A method for identifying genes related to a quantitative trait, incorporating multiple siblings and missing parents. Genet Epidemiol 2005;29:155-165.

33   Wheeler E, Cordell HJ: Quantitative trait association in parent offspring trios: Extension of case/pseudocontrol method and comparison of prospective and retrospective approaches. Genet Epidemiol 2007;31:813-833.

34   Sinsheimer JS, Creek MM: Statistical Approaches for Detecting Transgenerational Genetic Effects in Humans; in Naumova AK, Greenwood CMT (eds): Epigenetics and Complex Traits, Springer New York, 2013, pp 247-264.

35   Sinsheimer JS, Palmer CG, Woodward JA: Detecting genotype combinations that increase risk for disease: maternal-fetal genotype incompatibility test. Genet Epidemiol 2003;24:1-13.

36   Ainsworth HE, Unwin J, Jamison DL, Cordell HJ: Investigation of Maternal Effects, Maternal-Fetal Interactions and Parent-of-Origin Effects (Imprinting), Using Mothers and Their Offspring. Genetic Epidemiology 2011;35:19-45.

37   Clark MM, Blangero J, Dyer TD, Sobel EM, Sinsheimer JS: The Quantitative-MFG Test: A Linear Mixed Effect Model to Detect Maternal-Offspring Gene Interactions. Ann Hum Genet 2016;80:63-80.

38   Lange K: Mathematical and statistical methods for genetic analysis, ed 2nd. New York, Springer, 2002.

39   Demerath EW, Choh AC, Czerwinski SA, Lee M, Sun SS, Chumlea WC, Duren D, Sherwood RJ, Blangero J, Towne B, Siervogel RM: Genetic and environmental influences on infant weight and weight change: the Fels Longitudinal Study. Am J Hum Biol 2007;19:692-702.

40   Devlin B, Roeder K: Genomic control for association studies. Biometrics 1999;55:997-1004.

41   Lange K, Papp JC, Sinsheimer JS, Sripracha R, Zhou H, Sobel EM: Mendel: the Swiss army knife of genetic analysis programs. Bioinformatics 2013;29:1568-1570.

42   Nakimuli A, Chazara O, Hiby SE, Farrell L, Tukwasibwe S, Jayaraman J, Traherne JA, Trowsdale J, Colucci F, Lougee E, Vaughan RW, Elliott AM, Byamugisha J, Kaleebu P,

Mirembe F, Nemat-Gorgani N, Parham P, Norman PJ, Moffett A: A KIR B centromeric region present in Africans but not Europeans protects pregnant women from pre-eclampsia. Proceedings of the National Academy of Sciences of the United States of America 2015;112:845-850.

43   Magnus P, Birke C, Vejrup K, Haugan A, Alsaker E, Daltveit AK, Handal M, Haugen M, Hoiseth G, Knudsen GP, Paltiel L, Schreuder P, Tambs K, Vold L, Stoltenberg C: Cohort Profile Update: The Norwegian Mother and Child Cohort Study (MoBa). Int J Epidemiol 2016

Chapter 5

Future Work in Statistical Methods for Transgenerational Effects

This dissertation has addressed the absence of methods for testing for a specific type of transgenerational effect, MFG interactions, on quantitative traits using arbitrary family structures. I have demonstrated that the QMFG analysis method is a powerful and accurate approach when investigating the effect of offspring and maternal genotypes jointly for a single locus or multiple bi-allelic loci. Here I outline possible extensions to the QMFG test and other methods for detecting transgenerational effects.

QMFG Extension for Missing Data

Often with real data, lack of complete phenotype and genotype data is an issue that must be considered. As discussed in Chapter 3, one way to handle missing genotypes is to use a standard genotype imputation program (Marchini & Howie, 2010) that requires dense genotyping and ignores family structure, inevitably producing impossible maternal-offspring genotype combinations. In such situations, either the entire SNP would have to be discarded from the QMFG analysis or the erroneous genotypes would have to be replaced with the most probable genotype. To address this issue with standard genotyping imputation programs, Wasiolek *et al*. (2015) have proposed an imputation approach informed by pedigree relationships. Utilizing the Majorize-Minimize (MM) algorithm, they reduce the number of Mendelian inconsistencies that result from genotype imputation. Though this method is an improvement, it still relies on dense genotype data and produces some impossible mother-offspring genotype combinations.

The most ideal solution would be an approach that imputes genotypes conditional on both

the family genotype and phenotype data. Such an approach would require the use of MCMC

methods that may be computationally intensive, especially when imputing many genotypes in

data consisting of large families with hundreds of thousands or millions of SNPs. A more

feasible approach to address the issue of generating Mendelian inconsistencies while

simultaneously requiring less computational power would fill in genotypes conditional on family

genotype data only. Given the dependence of maternal and offspring genotypes, I propose using

joint maternal and offspring genotype probabilities calculated based on family structure. Because

maternal and offspring genotypes depend on one another, probabilities for each maternal-fetal

genotype combination could be estimated given the other genotypes in the pedigree. By

including these estimated probabilities within the QMFG design matrix in the place of missing

data and fitting a mixture of normal distributions, I expect that QMFG parameters could be

accurately estimated. Though this approach imputes genotypes under the null, it would be a

worthwhile starting point for future work handling missing data. Once methods to allow for the

inclusion of incompletely genotyped pedigrees have been developed, the Mendel software's

current implementation of the QMFG test can be modified accordingly. Further simulations

should be performed to determine error rates, power, and the effects of model misspecification.

Given that the statistical properties are sound, this extension for missing data will be valuable for

researchers who have incomplete family genotypes and want to use the QMFG test on their data.

QMFG Score Test Extension for Multiple Loci

The aim of genome-wide association studies (GWAS) is to analyze DNA variants across the

human genome to identify genetic factors associated with common diseases and traits. As mentioned previously, GWAS have revealed numerous loci associated with human diseases and complex traits (Lander, 2011, Hirschhorn, 2009) and have the potential to shift traditional medical practices to personalized medicine (Katsios & Roukos, 2010). With the decreasing cost and increasing efficiency of genome-wide scans, it is more important than ever for statistical methods that can handle thousands or millions of markers quickly. I have shown that the QMFG score test is just such an effective and rapid screening tool when testing MFG interactions for one SNP at a time. A straightforward addition to the methods proposed in this dissertation would be to extend the QMFG score test to handle multiple loci (as I've done for the LRT in Chapter 3). This would require modifying the Mendel QMFG score test option for a single locus to allow users to specify multiple loci of interest and validating the extension via simulations studies.

The QMFG Test for Non-Normally Distributed Traits

The QMFG test is a linear mixed modeling approach used to study associations between MFG interactions and continuous traits assuming either that the trait is normally distributed or the central limit theorem is applicable. A useful extension to the QMFG test would be an option that allows for response variables from non-Gaussian distributions (binomial, Poisson, etc.). Generalized linear mixed models (GLMM) provide the framework to handle data from the exponential family of distributions (Tempelman, 1998) and accordingly, packages and procedures have been developed for various statistical programs to apply GLMM to family data in genetic association studies (Vazquez *et al.*, 2010, Wang *et al.*, 2015). Generalizing the QMFG test to fit GLMM to test for MFG interaction effects on non-normal traits is a viable area of future research.

One challenge of programing such a method involves the estimation of parameters. For linear mixed models, the likelihood has a closed form so maximum likelihood estimation can be directly performed. In most cases, the GLMM likelihood cannot be evaluated in closed form due to multi-dimensional integration and therefore must be approximated (Wang *et al.*, 2015). Methods for numerical integral approximation include Laplace approximation (Wolfinger, 1993), Monte Carlo integration (Pan & Thompson, 2007), and Bayesian methods via Markov chain Monte Carlo (Zeger & Karim, 1991). An alternative to approximation methods is linearization of the regression model via Taylor expansion (Wolfinger & Oconnell, 1993). The computational advantages and disadvantages of each method must be carefully considered prior to updating the Mendel software to fit GLMM. Once again, simulation studies would be essential in determining the statistical properties of this generalized-QMFG test.

In the case of many SNPs, it may be more practical to use the QMFG LMM to quickly screen markers prior to running a more computationally intensive GLMM. By using a loose threshold, the most significant SNPs could be selected and then results could be refined using a GLMM appropriate for the trait distribution. It has been shown that the LMM is robust when fit to data with a non-Gaussian or heteroscedastic error distribution (Jacqmin-Gadda *et al.*, 2007) suggesting it may function as an acceptable screening tool. An additional goal of this aim would be to investigate if and how error rates would be affected by misspecifying the model as linear when the trait truly comes from a non-normal distribution.

Methods to Assess Transgenerational Epigenetic Effects

Another field of genetics in which there is a need for statistical methods to study transgenerational effects is the area of epigenetics. Epigenetic effects are heritable factors that

determine expression of genes without altering the DNA sequence itself and are potential facilitators of transgenerational effects. The most studied epigenetic effect is the process of DNA methylation, in which a methyl group is added to DNA nucleotides and can affect the transcription of genes (Cortessis *et al.*, 2012). Methyl groups can be added to cytosine or adenine nucleotides and thus regions where cytosine and guanine nucleotides are found in high frequency (CpG sites) are targets for methylation. When these CpG sites are found within the promoters of genes, their methylation can lead to gene silencing. One interesting area of research would involve determining whether transgenerational effects, such as MFG interactions, act through methylation.

DNA methylation data are often expressed as the fraction of a specific CpG site that is methylated. In statistical analyses this fraction is called the beta value. One promising approach to investigating the role of methylation level as an intermediate phenotype for transgenerational effects is two-step Mendelian randomization (Relton & Davey Smith, 2012). Their strategy is to interrogate the role of epigenetic mediators of environmental exposures on disease risk. In the first step, a genetic proxy associated with the exposure is used to test for the conditional independence of the exposure and the beta value. In the second step, a different genetic proxy, independent of the first proxy and associated with the beta value, is used to test for the conditional independence of DNA methylation and the outcome or disease.

Of particular relevance to understanding transgenerational effects is that two-step Mendelian randomization can span generations. Relton and Davey Smith (2012) discuss the example of maternal alcohol use during pregnancy as the exposure, offspring methylation fraction at a particular CpG site as the intermediate phenotype and offspring cognition as the outcome. In this instance, an appropriate genetic proxy for alcohol consumption is the mother's

genotype at an associated locus and an appropriate genetic proxy for the beta value is the offspring's genotype at another locus, unlinked and independent of the first locus. One option may be to extend the two-step epigenetic Mendelian randomization method proposed by Relton and Davey Smith (2012) to allow for gene-gene interactions between generations and determine whether transgenerational effects influence offspring phenotypes by way of methylation.

References

Cortessis, V.K., Thomas, D.C., Levine, A.J., Breton, C.V., Mack, T.M., Siegmund, K.D., Haile, R.W. & Laird, P.W. (2012) Environmental epigenetics: prospects for studying epigenetic mediation of exposure-response relationships. *Hum Genet,* 131, 1565-89.

Hirschhorn, J.N. (2009) Genomewide association studies--illuminating biologic pathways. *N Engl J Med,* 360, 1699-701.

Jacqmin-Gadda, H., Sibillot, S., Proust, C., Molina, J.M. & Thiebaut, R. (2007) Robustness of the linear mixed model to misspecified error distribution. *Computational Statistics & Data Analysis,* 51, 5142-5154.

Katsios, C. & Roukos, D.H. (2010) Individual genomes and personalized medicine: life diversity and complexity. *Personalized Medicine,* 7, 347-350.

Lander, E.S. (2011) Initial impact of the sequencing of the human genome. *Nature,* 470, 187-97.

Marchini, J. & Howie, B. (2010) Genotype imputation for genome-wide association studies. *Nat Rev Genet,* 11, 499-511.

Pan, J.X. & Thompson, R. (2007) Quasi-Monte Carlo estimation in generalized linear mixed models. *Computational Statistics & Data Analysis,* 51, 5765-5775.

Relton, C.L. & Davey Smith, G. (2012) Two-step epigenetic Mendelian randomization: a strategy for establishing the causal role of epigenetic processes in pathways to disease. *Int J Epidemiol,* 41, 161-76.

Tempelman, R.J. (1998) Generalized linear mixed models in dairy cattle breeding. *J Dairy Sci,* 81, 1428-44.

Vazquez, A.I., Bates, D.M., Rosa, G.J., Gianola, D. & Weigel, K.A. (2010) Technical note: an R package for fitting generalized linear mixed models in animal breeding. *J Anim Sci,* 88**,** 497-504.

Wang, T., He, P., Ahn, K.W., Wang, X., Ghosh, S. & Laud, P. (2015) A re-formulation of generalized linear mixed models to fit family data in genetic association studies. *Front Genet,* 6**,** 120.

Wasiolek, R., Sinsheimer, J.S. & Lange, K. (2015) Genotype Imputation Informed by both Kinship and Linkage Disequilibrium [abstract]. ASHG; Baltimore, MD.

Wolfinger, R. (1993) Laplaces Approximation for Nonlinear Mixed Models. *Biometrika,* 80**,** 791-795.

Wolfinger, R. & Oconnell, M. (1993) Generalized Linear Mixed Models - a Pseudo-Likelihood Approach. *Journal of Statistical Computation and Simulation,* 48**,** 233-243.

Zeger, S.L. & Karim, M.R. (1991) Generalized Linear-Models with Random Effects - a Gibbs Sampling Approach. *Journal of the American Statistical Association,* 86**,** 79-86.

Appendix A

Supplemental Material for Chapter 3

Table S1 *San Antonio Family Heart Study subject counts*.

	Count
Total subjects	3637
Subjects with phenotype	1397
Subjects with phenotype and genotypes	1043
Offspring with phenotype and genotypes	855
Offspring with phenotype, genotypes, and genotyped mother	419

Table S2 *Estimated type I error rates when a proportion of genotypes are missing.*

Effects tested	Percent missing			
	0%	5%	10%	20%
RHD effect	0.053	0.045	0.054	0.046
NIMA or offspring effects	0.045	0.053	0.061	0.045
NIMA effects in the presence of offspring effects	0.050	0.053	0.060	0.049
Offspring effects in the presence of NIMA effects	0.042	0.052	0.055	0.044
Any MFG effects	0.048	0.048	0.048	0.048

Standard errors for all estimated type I error rates are approximately 0.005
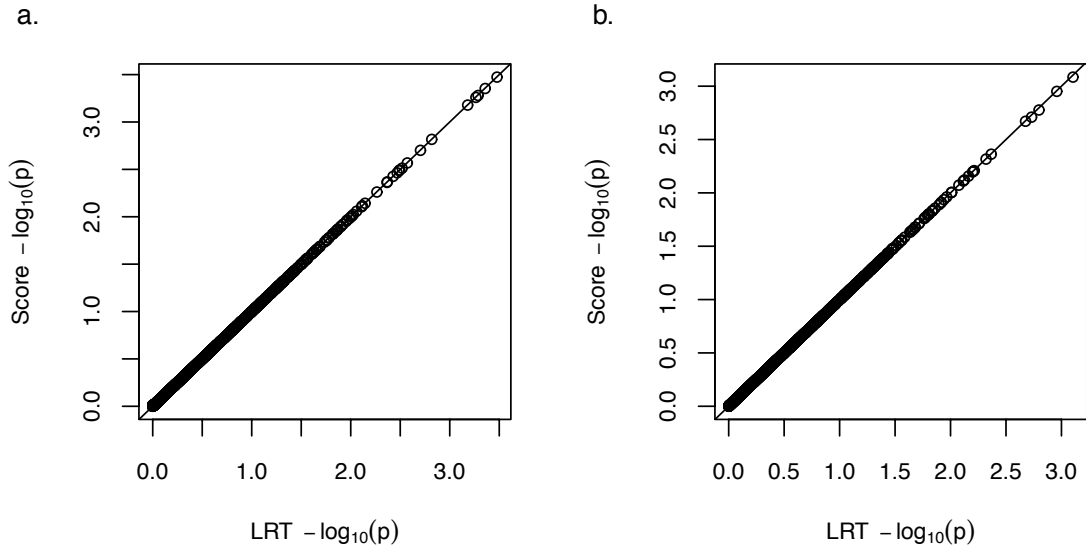
Figure S1 *Comparison of LRT and score test p-values for Simulation A data*. P-values resulting

from the LRT are compared to p-values resulting from a score test for the same data when testing

for (a) RHD effects (df = 1) and (b) NIMA or offspring effects (df = 3).
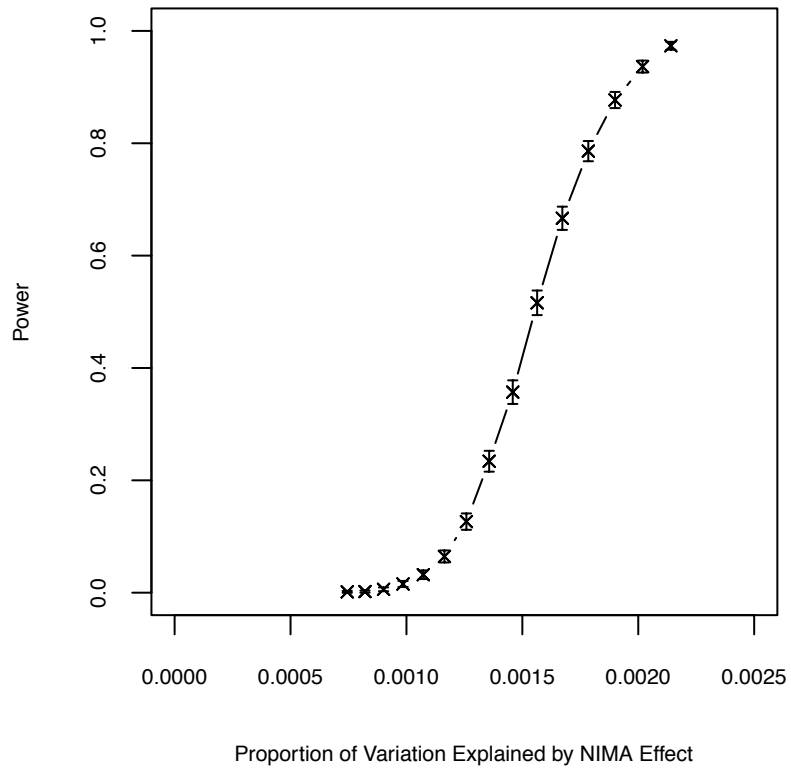
Figure S2 *Power to detect a NIMA effect in the presence of offspring effects.* Curves show power for the one degree of freedom test for a NIMA effect (β_{10}) using the LRT. NIMA effect sizes range from 0 to 0.7 and offspring effects are fixed ($\beta_{.1} = 0.18, \beta_{.2} = 0.36$). Error bars represent approximate 95% confidence intervals.
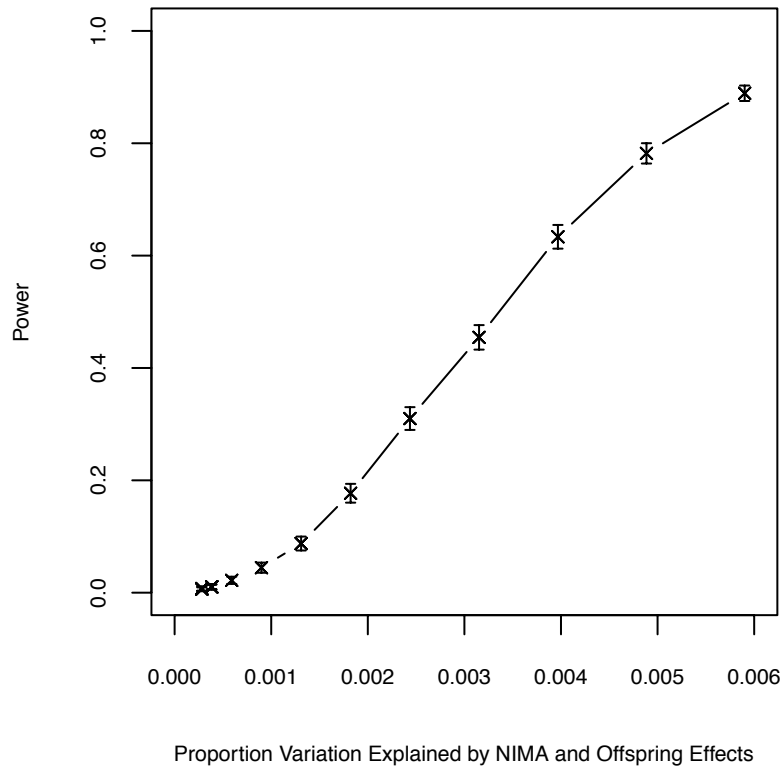
Figure S3 *Power to jointly detect NIMA and dominant offspring effects.* Curves show power for the two degrees of freedom test for a NIMA or dominant offspring effect $(\beta_{10}, \beta_{.1} = \beta_{.2})$ using the LRT and score test for a dominant offspring effect of 0.1 $(\beta_{.1} = \beta_{.2} = 0.1, \alpha = 0.001)$. NIMA effect sizes range from 0 to 0.6. Error bars represent approximate 95% confidence intervals.
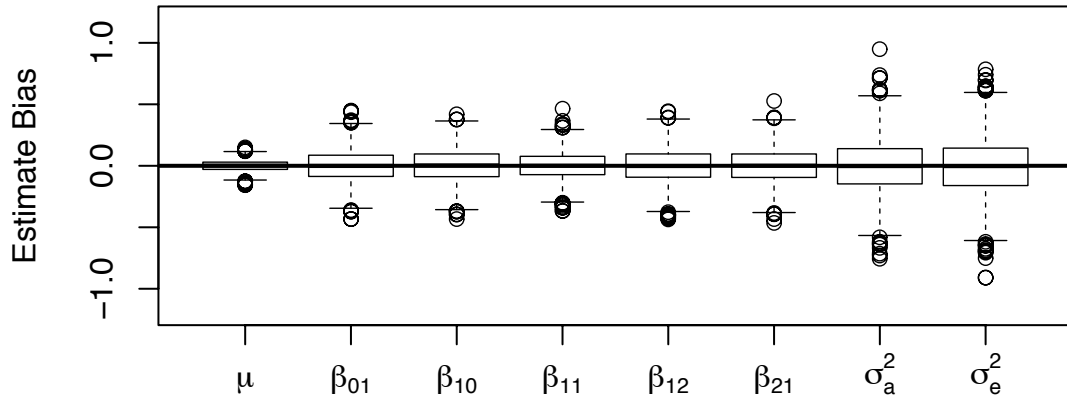
Figure S4 *Parameter estimate bias for data simulated with NIMA, offspring, and maternal effects*. Genotypes and quantitative traits for each replication were simulated for 1,000 three-generational pedigrees using Simulation D data with NIMA, offspring, and maternal effects ($\mu = 40, \beta_{01} = 0.18, \beta_{10} = 0.60, \beta_{11} = 0.36, \beta_{12} = 0.54, \beta_{21} = 0.54, \beta_{22} = 0.72, \sigma_a^2 = 1, \sigma_e^2 = 5$). Boxplot shows bias of parameter estimates and environmental variance over 2,000 replications. A horizontal line is drawn at zero bias.
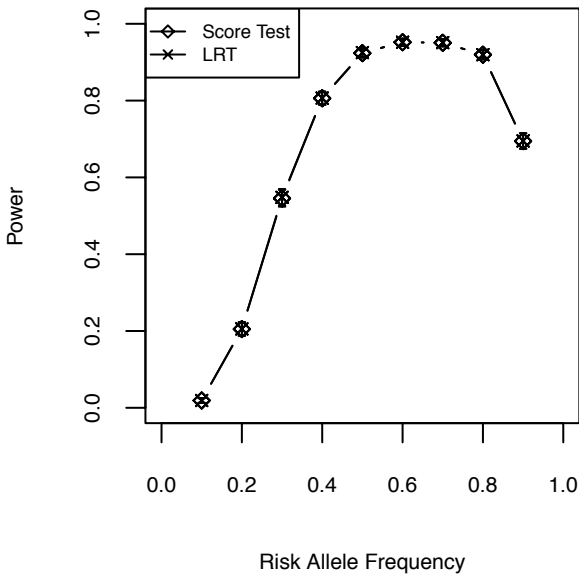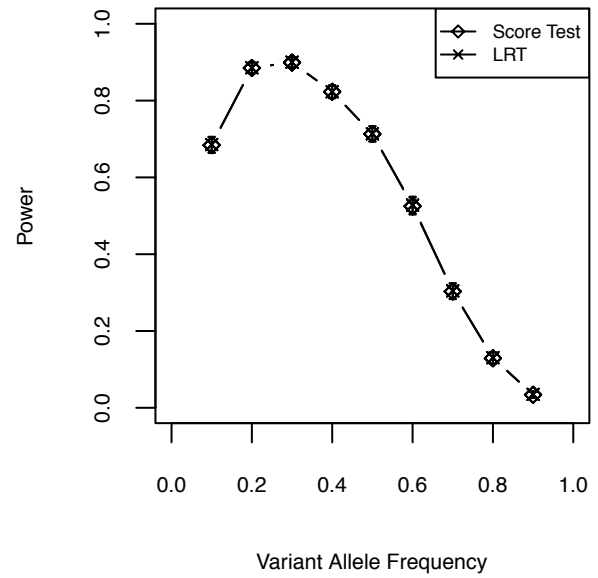
Figure S5 *Effect of allele frequency on power*. Variant allele frequency ranges from 0.1 to 0.9. Error bars represent approximate 95% confidence intervals. (a) Curves show power for the one degree of freedom test for an RHD effect (β_{21}) using Simulation B data for the LRT and score test. (b) Curves show power for the three degrees of freedom test for a NIMA or offspring effect ($\beta_{10}, \beta_{.1}, \beta_{.2}$) using Simulation D data for the LRT and score test.

a.

b.
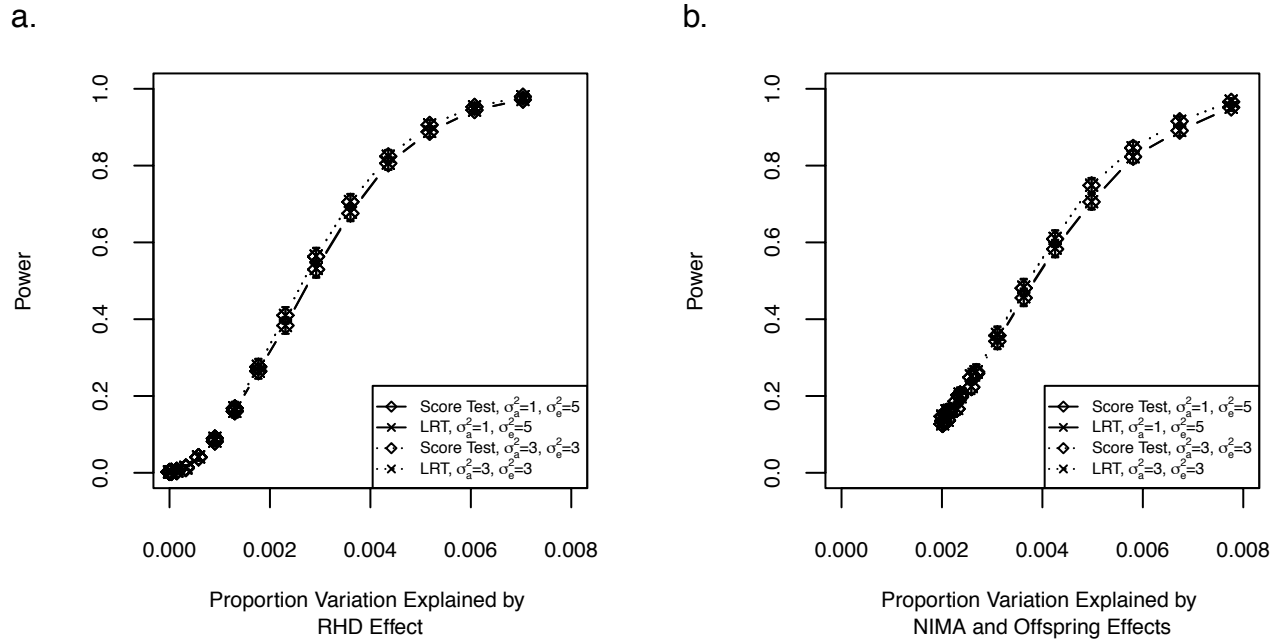
Figure S6 *Effect of additive genetic and environmental variance on power.* Solid lines show

power when $\sigma_a^2 = 1$ and $\sigma_e^2 = 5$, while dotted lines show power when additive genetic and

environmental variance simulation values are changed to $\sigma_a^2 = 3, \sigma_e^2 = 3$. Error bars represent

approximate 95% confidence intervals. (a) Curves show power for the one degree of freedom

test for an RHD effect (β_{21}) using the LRT and score test. RHD effect sizes range from 0 to 0.7.

(b) Curves show power for the three degrees of freedom test for NIMA or offspring effects

($\beta_{10}, \beta_{.1}, \beta_{.2}$) using the LRT and score test. NIMA effect sizes range from 0 to 0.7 and offspring

effects are fixed ($\beta_{.1} = 0.18, \beta_{.2} = 0.36$).
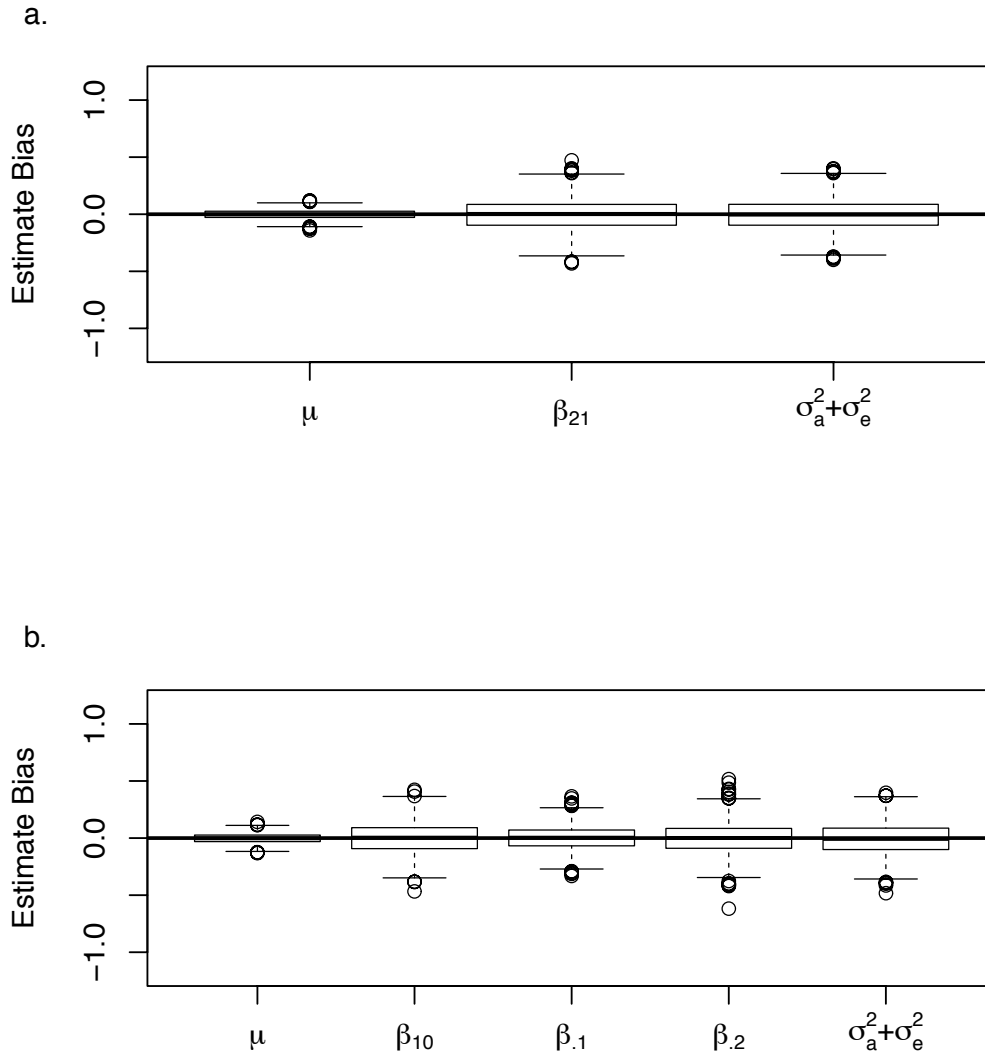
Figure S7 *Parameter estimate bias when data are simulated for parent-offspring trios.* Boxplots

show bias of parameter estimates for the grand mean, MFG effects, additive genetic variance,

and environmental variance using 1,000 parent-offspring trios over 2,000 replications using

parameters from (a) Simulation B with an RHD effect of 0.55 ($\mu = 40, \beta_{21} = 0.55, \beta_{01} = \beta_{10} =$

$\beta_{11} = \beta_{12} = \beta_{22} = 0, \sigma_a^2 = 1, \sigma_e^2 = 5$), and (b) Simulation C with a NIMA effect of 0.60 and an

additive offspring allelic effect of 0.18 ($\mu = 40, \beta_{10} = 0.60, \beta_{01} = \beta_{11} = \beta_{21} = 0.18, \beta_{12} =$

$\beta_{22} = 0.36, \sigma_a^2 = 1, \sigma_e^2 = 5$). A horizontal line is drawn at zero bias.
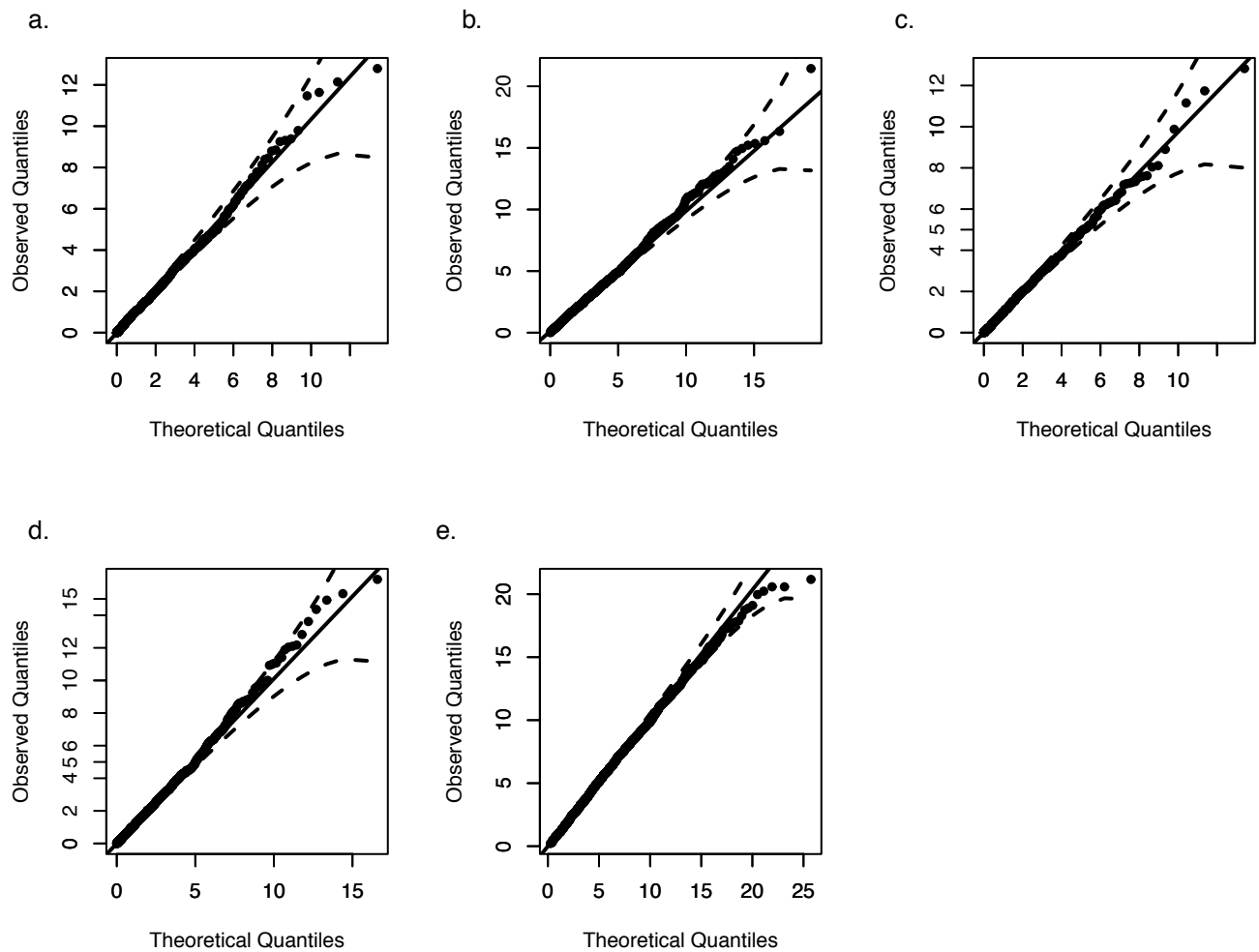
Figure S8 *Q-Q plot for LRT when data are simulated with a smaller sample size*. Genotypes and quantitative traits for each replication were simulated for 100 pedigrees under the null of no MFG effects and were tested with the LRT for (a) RHD effects (df = 1, λ = 1.094), (b) NIMA or offspring effects (df = 3, λ =1.000), (c) NIMA effects in the presence of offspring genotype effects (df = 1, λ = 0.972), (d) offspring effects in the presence of NIMA effects (df = 2, λ = 1.047), and (e) any MFG effects (df = 6, λ = 1.086).
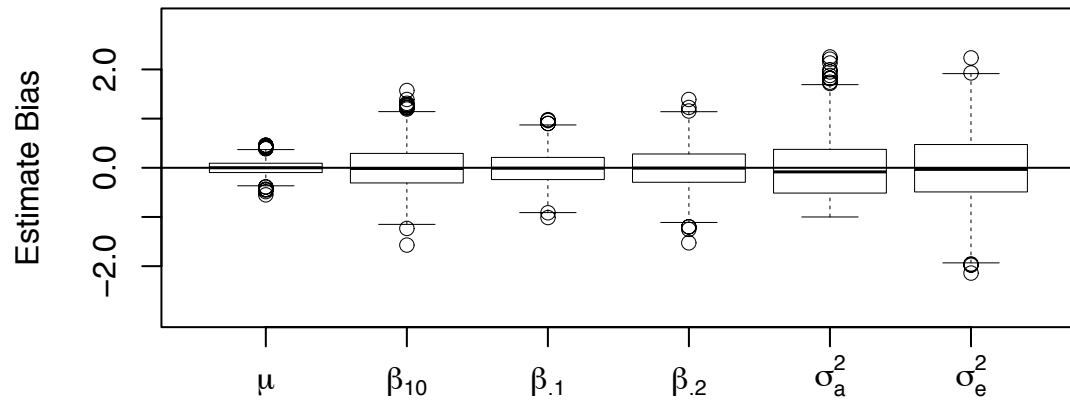
Figure S9 *Parameter estimate bias when data are simulated under NIMA and offspring effects with a smaller sample size*. Boxplots show bias of parameter estimates for the grand mean, MFG effects, additive genetic variance, and environmental variance using 100 three-generational families over 2,000 replications using parameters from Simulation F with a NIMA effect of 1.90 and an additive offspring allelic effect of 0.60 ($\mu = 40, \beta_{10} = 1.90, \beta_{.1} = 0.60, \beta_{.2} = 1.20, \sigma_a^2 = 1, \sigma_e^2 = 5$). A horizontal line is drawn at zero bias.
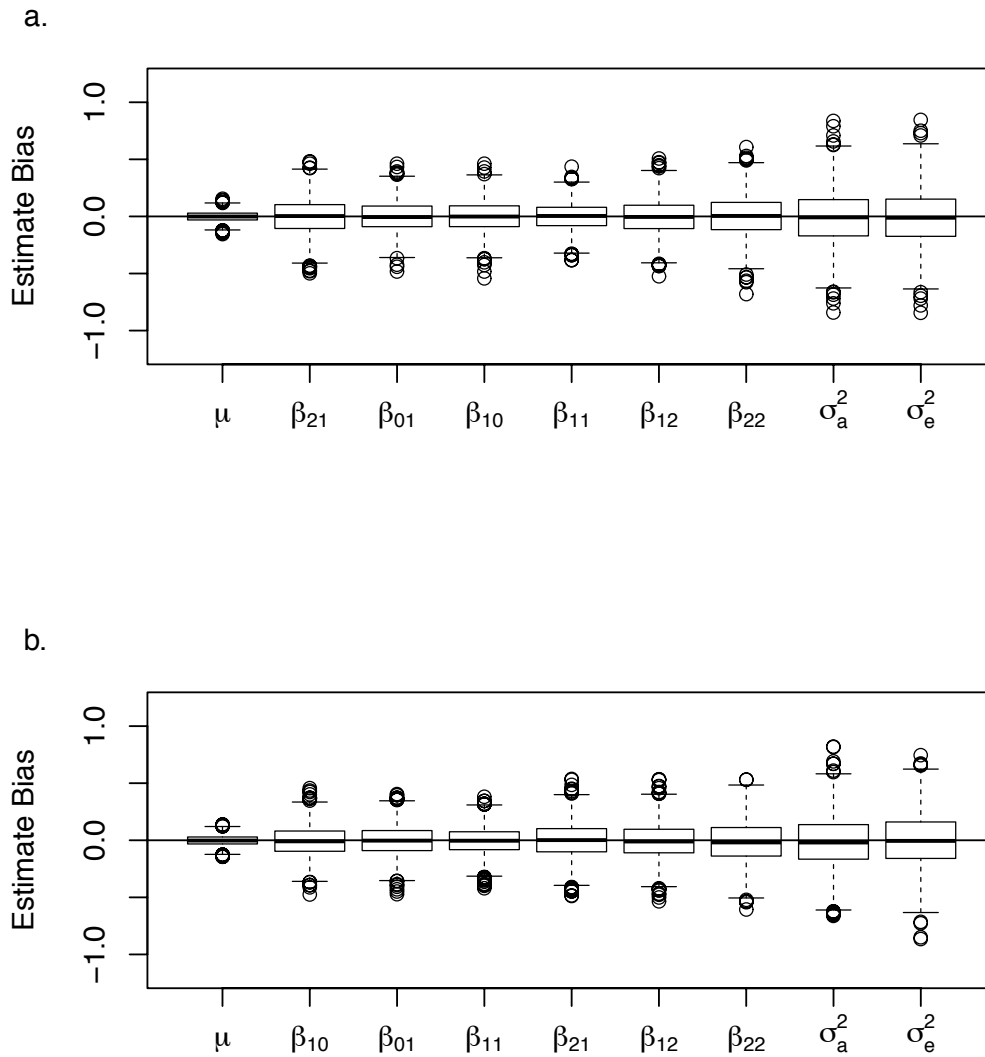
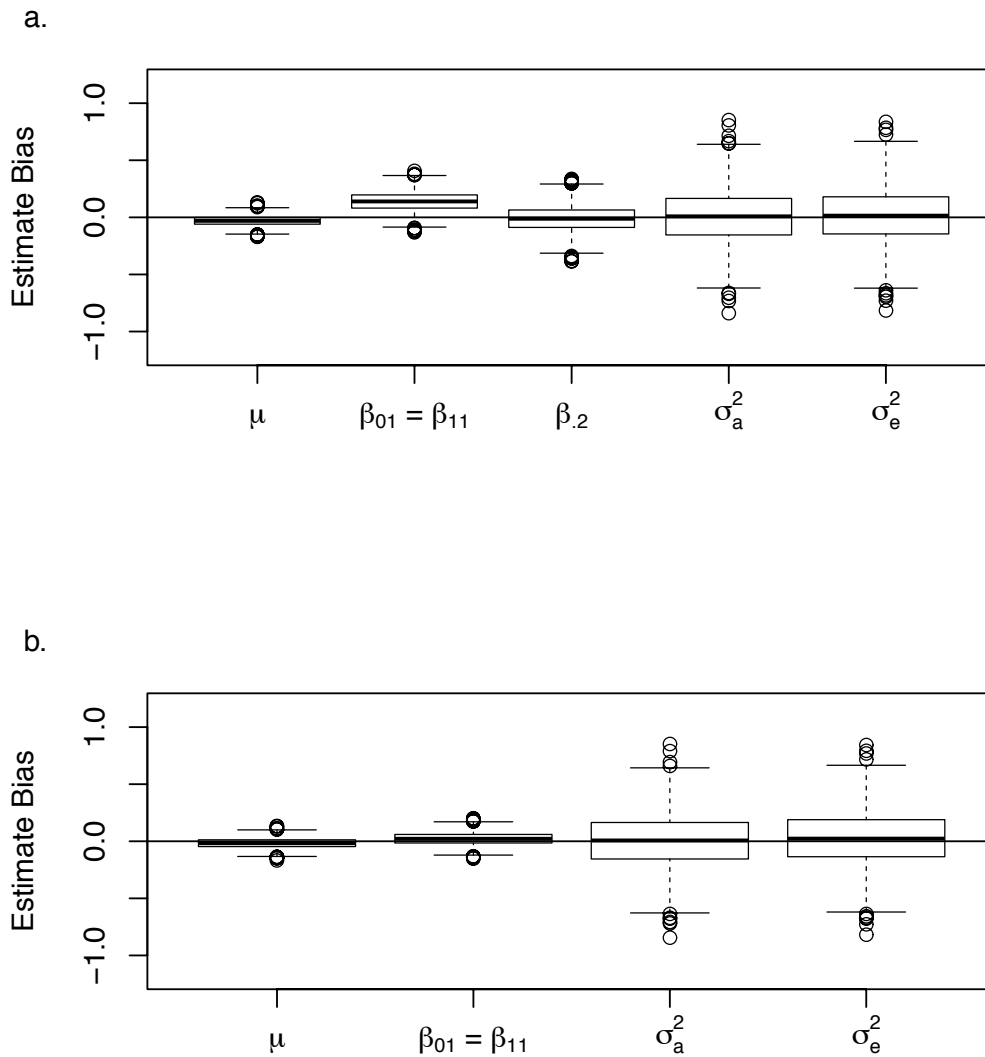**Figure S10** *Parameter estimate bias when the general model is fit.* Boxplots show bias of

parameter estimates for the grand mean, MFG effects, additive genetic variance, and

environmental variance over 2,000 replications when the general model is fit to (a) Simulation B

data with an RHD effect of 0.55 ($\mu = 40, \beta_{21} = 0.55, \beta_{01} = \beta_{10} = \beta_{11} = \beta_{12} = \beta_{22} = 0, \sigma_a^2 =$

$1, \sigma_e^2 = 5$) and (b) Simulation D data with a NIMA effect of 0.60 and an additive offspring

allelic effect of 0.18 ($\mu = 40, \beta_{10} = 0.60, \beta_{01} = \beta_{11} = \beta_{21} = 0.18, \beta_{12} = \beta_{22} = 0.36, \sigma_a^2 =$

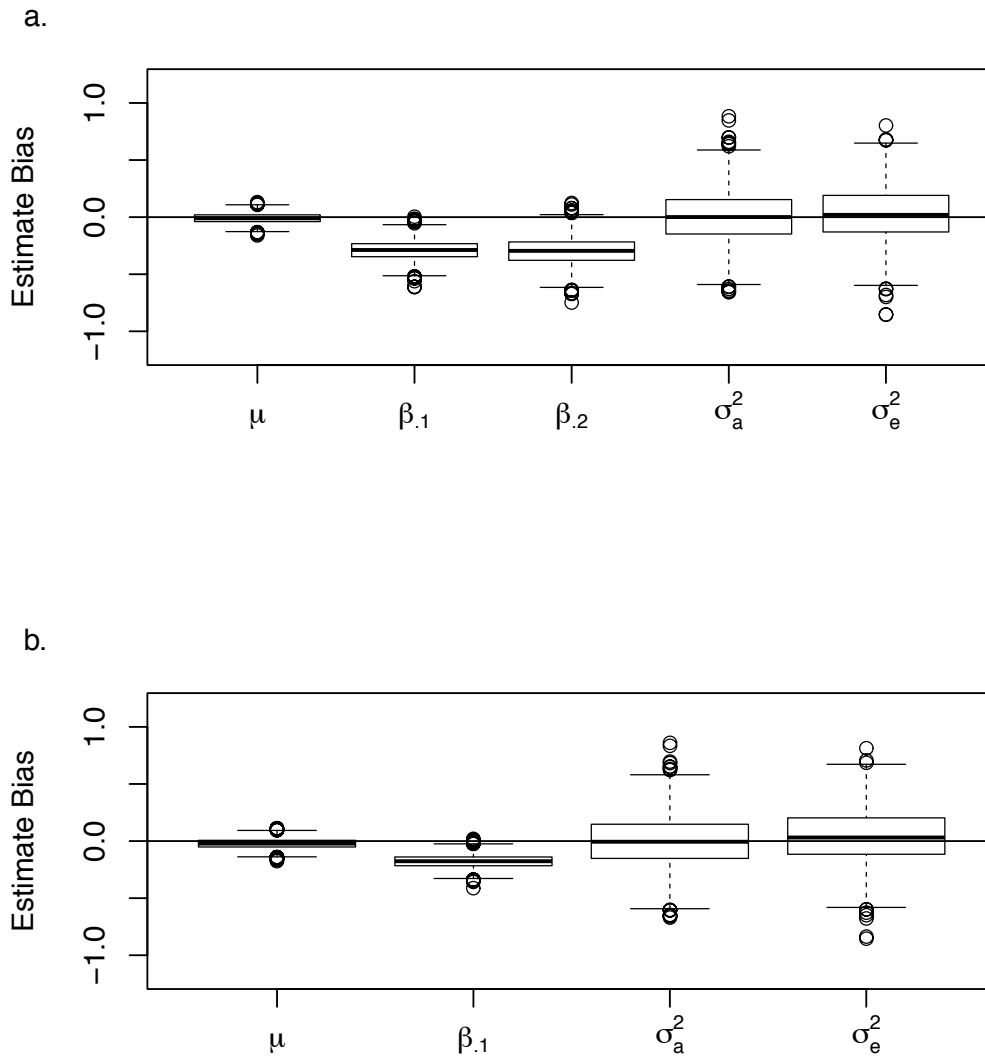$1, \sigma_e^2 = 5$). A horizontal line is drawn at zero bias.

Figure S11 *Parameter estimate bias when a misspecified model is fit using Simulation H data.*

Boxplots show bias of parameter estimates for the grand mean, offspring effects, additive genetic

variance, and environmental variance over 2,000 replications of data with a true underlying RHD

effect of 0.70 ($\mu = 40, \beta_{21} = 0.70, \beta_{01} = \beta_{10} = \beta_{11} = \beta_{12} = \beta_{22} = 0, \sigma_a^2 = 1, \sigma_e^2 = 5$) when

fitting (a) a genotypic model and (b) an additive model with the constraint $2\beta_{.1} = \beta_{.2}$. A

horizontal line is drawn at zero bias.

a.

b.

Figure S12 *Parameter estimate bias when a misspecified model is fit using Simulation I data.*
Boxplots show bias of parameter estimates for the grand mean, offspring effects, additive genetic
variance, and environmental variance over 2,000 replications of data with a true underlying
NIMA effect of 0.70 ($\mu = 40, \beta_{10} = 0.70, \beta_{01} = \beta_{11} = \beta_{21} = \beta_{12} = \beta_{22} = 0, \sigma_a^2 = 1, \sigma_e^2 = 5$)
when fitting (a) a genotypic model and (b) an additive model with the constraint $2\beta_{.1} = \beta_{.2}$. A
horizontal line is drawn at zero bias.
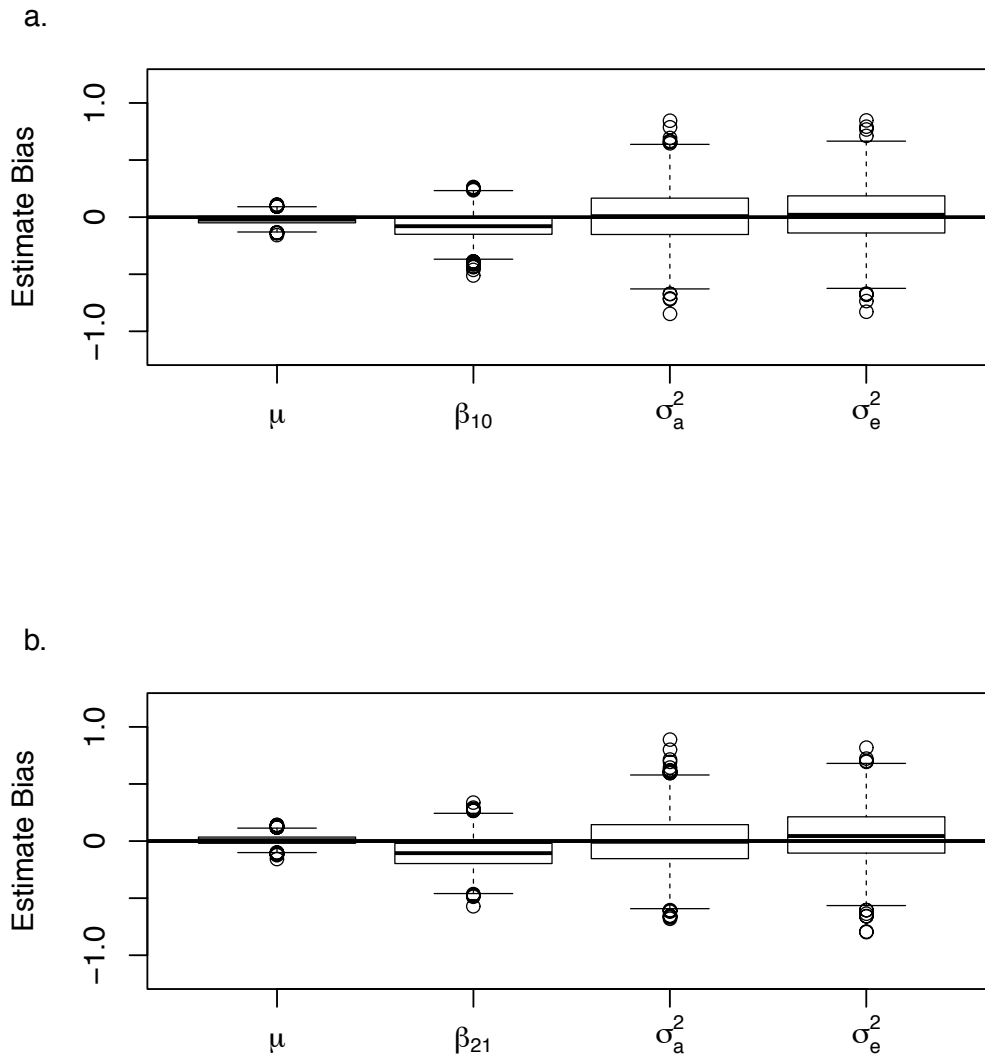
Figure S13 *Parameter estimate bias due to model misspecification*. Genotypes and quantitative traits for each replication were simulated for 1,000 pedigrees fitting (a) Model 5 (NIMA effect model) to Simulation B data with an RHD effect of 0.55 ($\mu = 40, \beta_{21} = 0.55, \beta_{01} = \beta_{10} = \beta_{11} = \beta_{12} = \beta_{22} = 0, \sigma_a^2 = 1, \sigma_e^2 = 5$), and (b) Model 2 (RHD effect model) to Simulation I with a NIMA effect of 0.70 ($\mu = 40, \beta_{10} = 0.70, \beta_{01} = \beta_{11} = \beta_{12} = \beta_{21} = \beta_{22} = 0, \sigma_a^2 = 1, \sigma_e^2 = 5$). A horizontal line is drawn at zero bias.

107
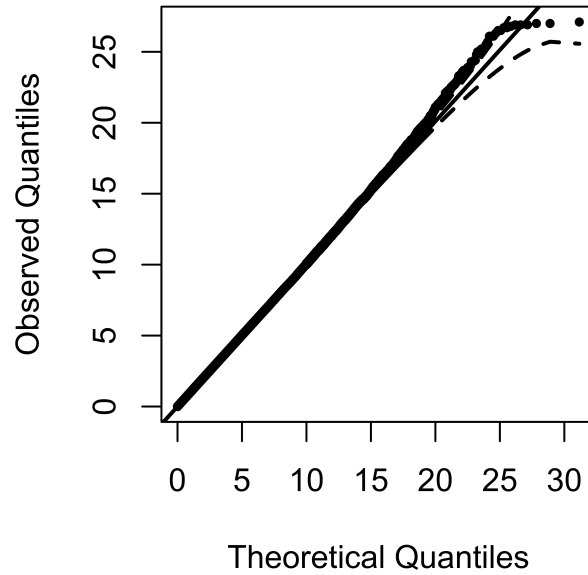
Figure S14 *Q-Q plot for score test of the SAFHS data when top 10 hits are removed*. Results from the 3 degrees of freedom test for NIMA or offspring effects $(\beta_{10}, \beta_{.1}, \beta_{.2})$ using the score test adjusting for age and sex with the top 10 hits removed $(\lambda = 1.012)$. Data from the SAFHS consist of 635,494 SNPs from 419 offspring with HDL measurements in 43 multi-generational families.
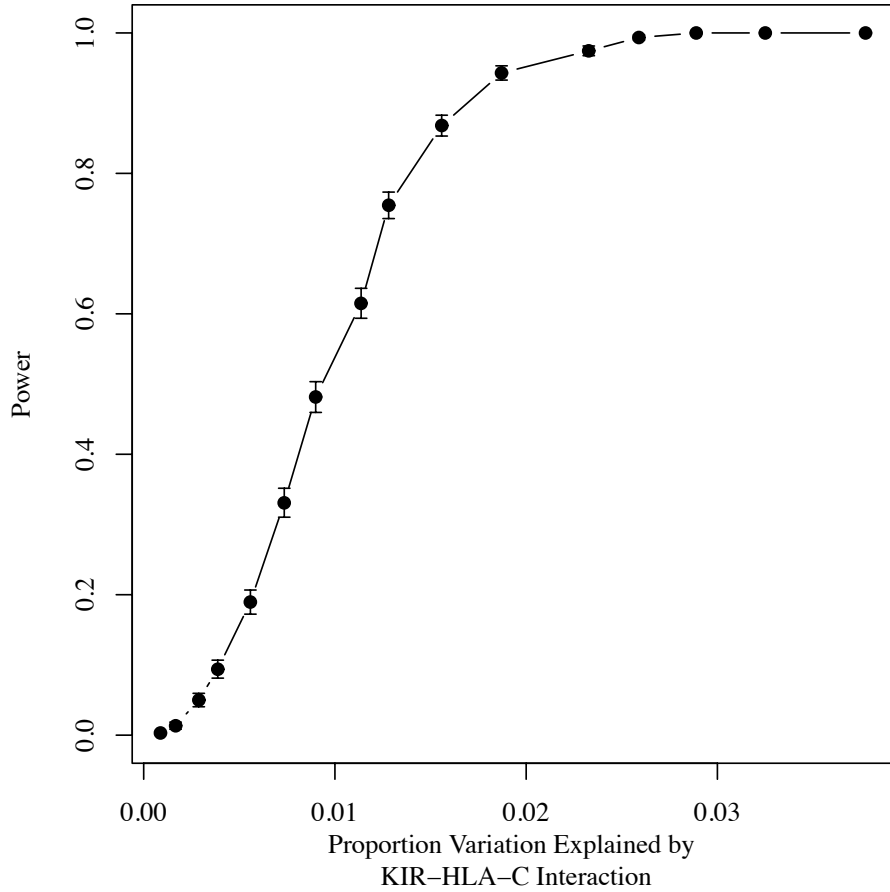
Appendix B

Supplemental Material for Chapter 4

Fig. 1S. *Power to detect a KIR, HLA-C, or KIR-HLA-C interaction effect when the model is misspecified as Model 4*. Data simulated under the Scenario II (*KIR-HLA-C* interaction only) conditions and Model 4 is fit (maternal *KIR*, more *HLA-C2* in offspring, and *KIR-HLA-C* interaction). QMFG compared Model 4 to Model 0 (no genetic effects), thus testing for *KIR*, *HLA-C*, or *KIR-HLA-C* interaction effects (df = 3). Error bars represent approximate 95% confidence intervals.
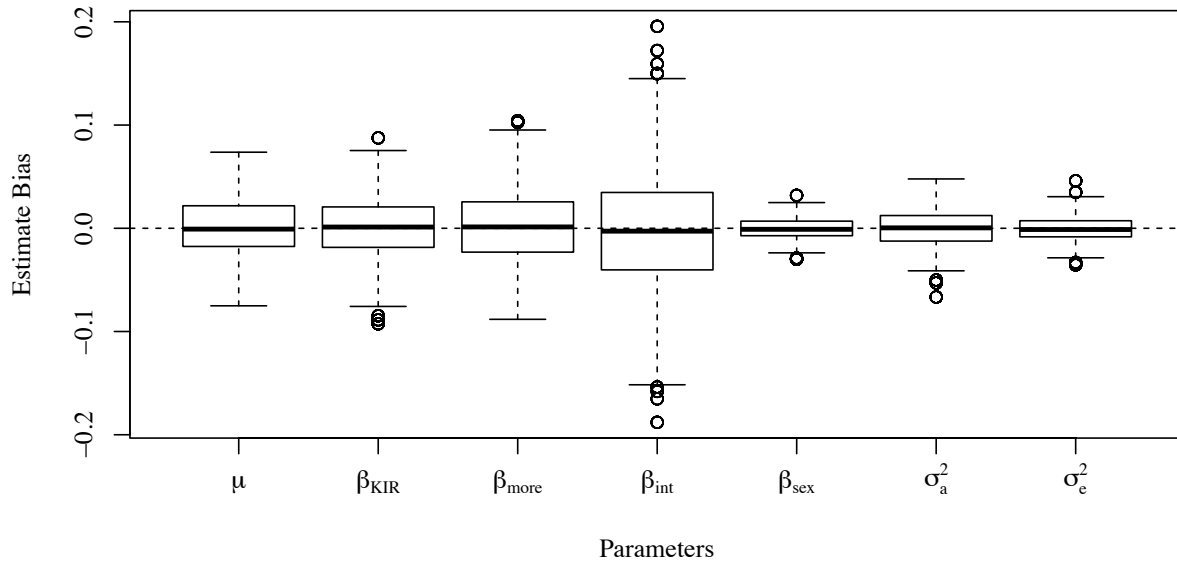
Fig. 2S. *Parameter estimate bias when the model is misspecified as Model 4*. Data simulated under the Scenario II (*KIR-HLA-C* interaction only) conditions and Model 4 is fit (maternal *KIR*, more *HLA-C2* in offspring, and *KIR-HLA-C* interaction). Birth weight phenotypes were simulated with a *KIR-HLA-C* interaction effect of 0.19 kg (μ = 3.5 kg, β_{KIR} = 0 kg, β_{more} = 0 kg, β_{int} = 0.19 kg, β_{sex} = -0.2 kg, σ_a^2 = 0.2025 kg^2, σ_e^2 = 0.0475 kg^2).
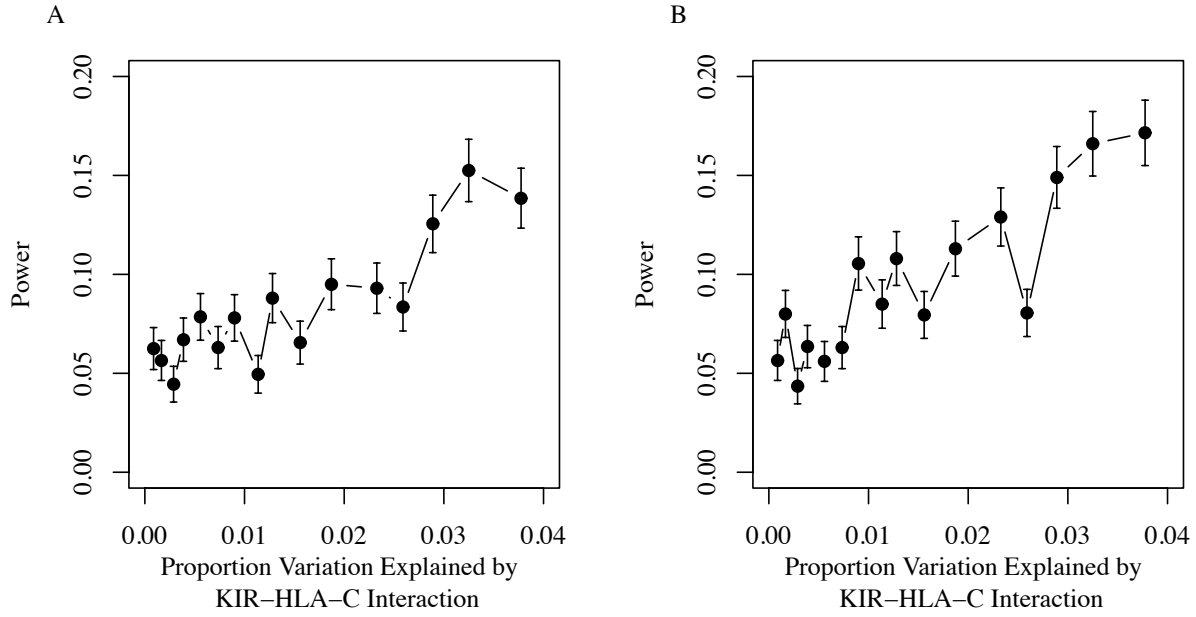
Fig. 3S. *Power for KIR offspring effects*. Significance level used to determine power is 0.05.

Data simulated under Scenario II (*KIR-HLA-C* interaction only) conditions and the model is

misspecified such that it tests for (A) genotypic offspring *KIR* effects (Model 8 vs. Model 0; df =

2) and (B) additive offspring *KIR* effects (Model 9 vs. Model 0; df = 1). Error bars represent
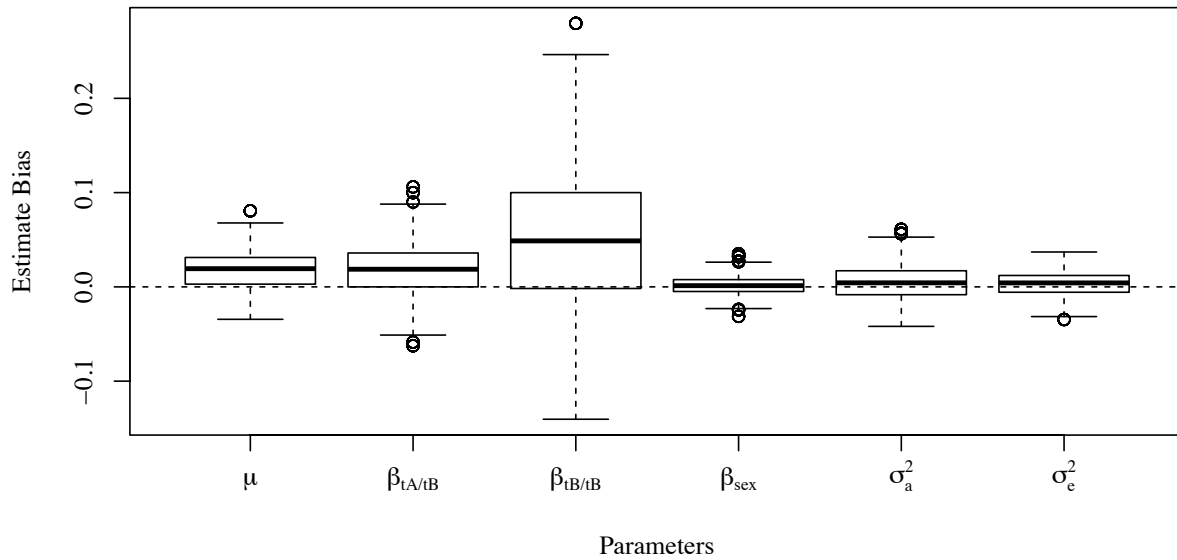
approximate 95% confidence intervals.

Fig. 4S. *Parameter estimate bias when the model estimates offspring KIR effects*. Model is misspecified as Model 8 (genotypic offspring *KIR*) and is fit to Scenario II (*KIR-HLA-C* interaction) data. Birth weight phenotypes were simulated with a *KIR-HLA-C* interaction effect of 0.35 kg (μ = 3.5 kg, β_{int} = 0.35 kg, β_{sex} = -0.2 kg, σ_a^2 = 0.2025 kg^2, σ_e^2 = 0.0475 kg^2).
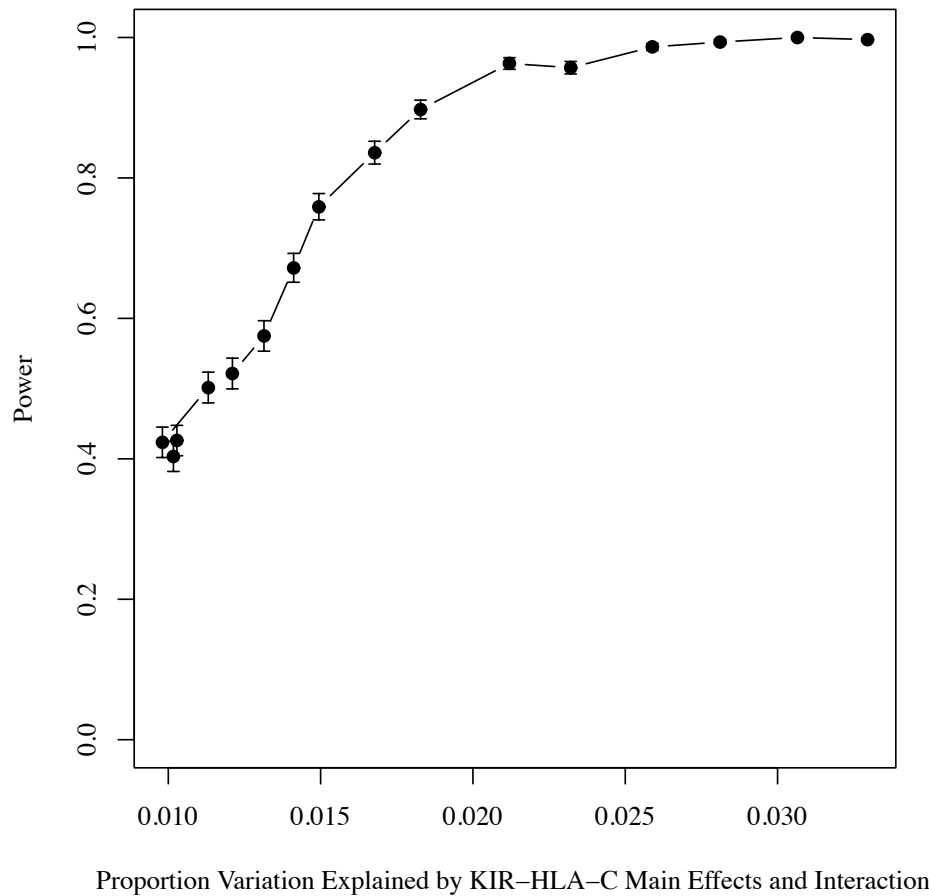
Fig. 5S. *Power to detect KIR, HLA-C, or two-locus interaction effects in mother-offspring pairs.* Data simulated under the Scenario III (maternal *KIR*, more *HLA-C2* in offspring, and *KIR-HLA-C* interaction) conditions. QMFG compared Model 4 (maternal *KIR*, more *HLA-C2* in offspring, and *KIR-HLA-C* interaction) to Model 0 (no genetic effects), thus testing for either a maternal *KIR-tB*, more *HLA-C2*, or *KIR-HLA-C* interaction effect (df = 3). Error bars represent approximate 95% confidence intervals. Note that, in this scenario because there are effects of *KIR* and *HLA-C* even when there is no *KIR-HLA-C* interaction, the proportion of variation explained is not zero when the interaction effect size is zero.
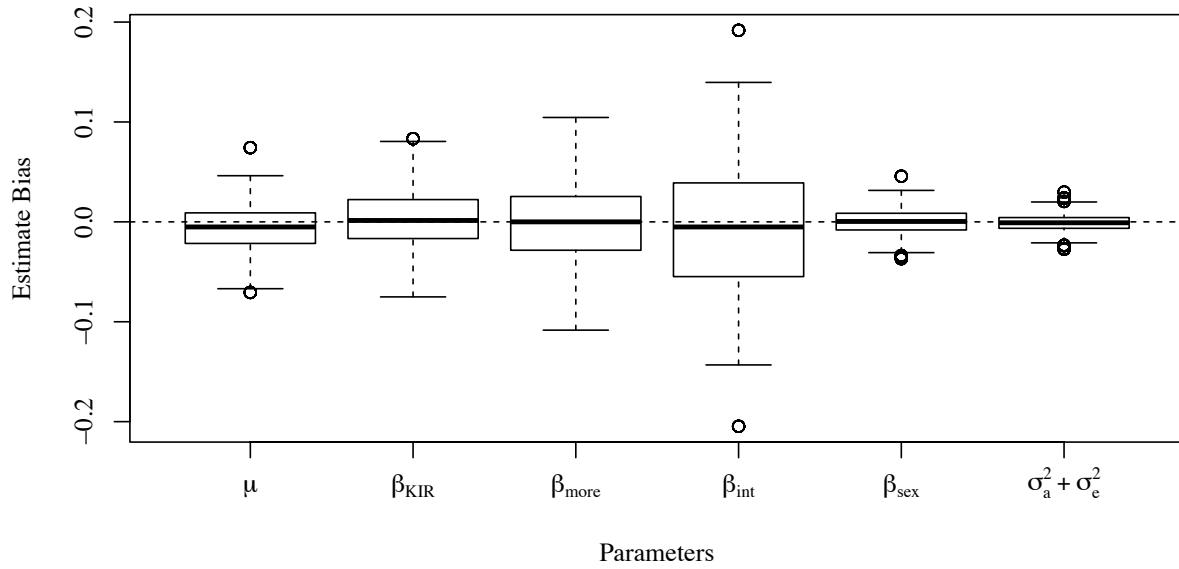
Fig. 6S. *Parameter estimate bias for mother-offspring pairs*. Model 4 (maternal *KIR*, more *HLA-C2* in offspring, and *KIR-HLA-C* interaction) is fit to data for mother-offspring pairs simulated under Scenario III (maternal *KIR*, more *HLA-C2* in offspring, and *KIR-HLA-C* interaction). Birth weight phenotypes were simulated with a *KIR-HLA-C* interaction effect of 0.19 kg ($\mu = 3.5$ kg, $\beta_{KIR} = 0.05$ kg, $\beta_{more} = -0.1$ kg, $\beta_{int} = 0.19$ kg, $\beta_{sex} = -0.2$ kg, $\sigma_a^2 + \sigma_e^2 = 0.25$ kg^2).

Appendix C

Alternative *KIR* and *HLA-C* Models

In Chapter 4, we model an effect of having more *HLA-C2* present in the offspring's genotype

than the mother's and an effect of the maternal *KIR-tB* haplotype. Modeling the effect of *HLA-C*

in this way is a proxy for modeling the effects of a maternal immune response to a fetus that has

an HLA-C antigen that is foreign to her own. That antigen would have to originate from the

paternally inherited gene. As discussed by Hiby and colleagues (2014), such a model limits the

ability to determine whether the impact of *HLA-C* is due to the extra dose of *C2* in the offspring

compared to his/her mother or a paternal parent of origin effect. No fathers were genotyped for

HLA-C in Hiby's United Kingdom cohort and thus testing for a parent of origin effect directly is

not possible. Here we explore a variety of alternative models (Table 1) in an attempt to elucidate

the mechanism by which the interaction between maternal and offspring *HLA-C* genotypes

impacts birth weight.

 The first alternative model we consider mimics the prototypical RHD example of MFG

incompatibility (refer to Chapters 3 and 4 for more detailed descriptions). This scenario is

motivated by evidence that the mother's KIR may be educated by her own *HLA-C* genotype. One

possibility is that a mother who has at least one *HLA-C2* allele in her genotype may be

completely desensitized to her offspring's *HLA-C2* alleles. Thus, we estimate an *HLA-C* effect

on birth weight only for *C1/C2* offspring with *C1/C1* mothers (Model C). We evaluate the fit of

this model compared to the null model of maternal *KIR* effects only (Model A) and to the more-

offspring-*HLA-C2* model (Model B). Comparing the AIC for Model C (-278.21) to Model B (-

279.06), we find no evidence that estimating the effect of *HLA-C* in this way provides a better fit

than the more-offspring-*HLA-C2* effect used in Chapter 4 and by Hiby and researchers (2014), but the fit of Model C is statistically significant compared to the maternal-*KIR*-only model (Model C versus Model A, p = 0.021) and we can't rule this model out as being a reasonable possibility.

Another hypothesis is that both mothers who don't have any *HLA-C2* alleles and mothers who do have *HLA-C2* in their genotypes react to the extra dose of offspring *HLA-C2* from the father (likely due to different *C2* allele type), but by varying degrees. Considering this possibility, the next model we fit (Model D) estimates separate parameters for 1) *C1/C2* offspring with *C1/C1* mothers and 2) *C2/C2* offspring with *C1/C2* mothers. This model is statistically significant compared to Model A (p = 0.038), but there is no evidence for a more nuanced effect of *C2* (comparing Model D to Model B, p = 0.517).

Next, we explore whether alternative models of a *HLA-C* paternal parent of origin effect provide an improvement in fit compared to the more-offspring-*HLA-C2* effect in Model B. The United Kingdom birth weight cohort doesn't include genotyped fathers. To model a paternal parent of origin effect therefore requires estimating an effect for maternal-offspring genotype combinations in which the offspring must have inherited an *HLA-C2* allele from the father (*C1/C2* offspring with *C1/C1* mothers, *C2/C2* offspring with *C2/C1* mothers, *C2/C2* offspring with *C2/C2* mothers). Note that in the case of a heterozygous offspring with a heterozygous mother, it is unclear which offspring allele has paternal origin without phased genotypes. Model E groups these *C1/C2* offspring with *C1/C2* mothers together with offspring who clearly have a paternally inherited *HLA-C2* allele. In contrast, Model F estimates separate parameters for 1) offspring with a paternally inherited *HLA-C2* and who have a mother that is not educated by *HLA-C2* alleles in her own genotype (*C1/C2* offspring with *C1/C1* mothers), 2) heterozygous

116

offspring with a heterozygous mother, and 3) offspring with a paternally inherited *HLA-C2* and who have a mother that is educated by the *HLA-C2* alleles in her own genotype (*C1/C2* offspring with *C2/C2* mothers or *C2/C2* offspring with *C2/C2* mothers). Neither Model E nor Model F is statistically significant compared to Model A (Model E versus Model A, p = 0.610; Model F versus Model A, p = 0.115). The AICs for these models are substantially larger than the AIC of Model B allowing us to dismiss these models.

Finally, we investigate the possibility that the *HLA-C* effect on birth weight is instead due to an offspring immune response to a foreign antigen in the mother. This maternal antigen-offspring antibody hypothesis is comparable to the NIMA effect described in Chapters 3 and 4. In Model G, we model an effect for offspring with a *C2/C2* genotype reacting to the *C1* allele in the genotype of their heterozygous mothers (*C2/C2* offspring with *C1/C2* mothers). The AIC for this model (-274.86) does not offer evidence that this model is greatly better than our original model (Model B). Moreover this model does not provide a significantly better fit to the data than the maternal-*KIR*-only model (Model G versus Model A. p = 0.585)

Through this exploration of alternative models, we cannot conclusively determine the mechanism by which the interaction between maternal and offspring *HLA-C* genotypes impacts birth weight. Of the seven models we can explore here, the model used in Chapter 4 and by Hiby and colleagues (2014) has the lowest AIC and thus can be considered the best model for this UK cohort, however Model C, which limits the immune effect to those mothers who are naïve with regards to the *C2* allele, fits nearly as well as Model B. To better dissect out the separate contributions of the inherited maternal and paternal *HLA-C* alleles in the future we would need the father's *HLA* genotypes.

Table 1. Comparison of alternative *KIR* and *HLA-C* models for UK cohort data (n = 403)

Model	Genetic effects	No. genetic parameters[1]	Log-likelihood	Proportion of variation explained[2]	AIC
A	Maternal *KIR-tB*	1	139.46	0.010	-274.92
B	Maternal *KIR-tB* More *HLA-C2*	2	142.53	0.025	-279.06
C	Maternal *KIR-tB* *C1/C2* offspring with a *C1/C1* mother	2	142.11	0.023	-278.21
D	Maternal *KIR-tB* *C1/C2* offspring with a *C1/C1* mother *C2/C2* offspring with a *C1/C2* mother	3	142.74	0.026	-277.49
E	Maternal *KIR-tB* *C1/C2* offspring with a *C1/C1* mother, *C1/C2* offspring with a *C1/C2* mother, *C2/C2* offspring with a *C1/C2* mother, or *C2/C2* offspring with a *C2/C2* mother	2	139.59	0.011	-273.18
F	Maternal *KIR-tB* *C1/C2* offspring with a *C1/C1* mother *C1/C2* offspring with a *C1/C2* mother *C2/C2* offspring with a *C1/C2* mother or *C2/C2* offspring with a *C2/C2* mother	4	142.43	0.025	-274.85
G	Maternal *KIR-tB* *C2/C2* offspring with a *C1/C2* mother	2	140.43	0.015	-274.86

[1]All models additionally adjusted for sex
[2]Proportion of variation explained by *KIR* and *HLA-C* effects.

References

Hiby, S.E., Apps, R., Chazara, O., Farrell, L.E., Magnus, P., Trogstad, L., Gjessing, H.K., Carrington, M. & Moffett, A. (2014) Maternal KIR in combination with paternal HLA-C2 regulate human birth weight. *J Immunol*, 192**,** 5069-73.