

Analysis of Genetic Interactions Involving Maternal and Offspring Genotypes at Different Loci: Power Simulation and Application to Testicular Cancer

Jérémie Nsengimana* and Jennifer H. Barrett

Section of Epidemiology and Biostatistics, Leeds Institute of Molecular Medicine, University of Leeds, Leeds, United Kingdom

The analyses of genetic interaction between maternal and offspring genotypes are usually conducted considering a single locus. Here, we propose testing maternal \times offspring ($M \times O$) and maternal \times maternal ($M \times M$) genotype interactions involving two unlinked loci. We reformulate the log-linear approach of analyzing cases and their parents (family trios) to accommodate two loci, fit fuller models to avoid confounding in a first analysis step and propose that the model be reduced to the most prominent effects in a second step. We conduct extensive simulations to assess the validity and power of this approach under various model assumptions. We show that the approach is valid and has good power to detect $M \times O$ and $M \times M$ interactions. For example, the power to detect a dominant interaction relative risk of 1.5 (both $M \times O$ and $M \times M$) is 70% with 300 trios and approaches 100% with 1,000 trios. Unlike the main effects, $M \times O$ and $M \times M$ interactions are conditionally independent of mating types, and consequently, their power is not affected by missing paternal genotypes. When applied to single-locus $M \times O$ interaction, our method is as powerful as other existing methods. Applying the method to testicular cancer, we found a nominally significant $M \times M$ interaction between single nucleotide polymorphisms from C-Kit Ligand (KITLG) and Sex Hormone Binding Globulin (SHBG) using 210 families (relative risk 2.2, $P = 0.03$). This finding supports a role of maternal hormones in offspring testicular cancer and warrants confirmation in a larger dataset. *Genet. Epidemiol.* 36:612–621, 2012. © 2012 Wiley Periodicals, Inc.

Key words: family trios; EM algorithm; two-locus log-linear model; type 1 error

Supporting Information is available in the online issue at wileyonlinelibrary.com.

*Correspondence to: Jérémie Nsengimana, Section of Epidemiology and Biostatistics, Leeds Institute of Molecular Medicine, Saint James Hospital, Cancer Genetics Building, Beckett Street, Leeds LS9 7TF, UK. E-mail: J.Nsengimana@leeds.ac.uk

Received 3 November 2011; Revised 10 May 2012; Accepted 17 May 2012

Published online 27 June 2012 in Wiley Online Library (wileyonlinelibrary.com/journal/gepi).

DOI: 10.1002/gepi.21655

INTRODUCTION

It has been suggested that maternal effects play an important role in species adaptation [Wolf, 2000], and there is a compelling body of evidence that complex adulthood diseases, such as diabetes, hypertension, schizophrenia, asthma, testicular, renal, and gastric cancers, are partially determined by antenatal factors [Drever et al., 2010; Eriksson et al., 2011; Maringhini et al., 2010; Winder et al., 2011]. Although the process appears to be dominated by in utero epigenetic changes, placental size, and foetal growth rate, all of which are directly related to maternal diet [Barker et al., 2010], maternal genes involved in placental formation, adhesion, vascularization, hormone secretion, or transport activity are also likely to be involved [Jansson et al., 2002].

Dissecting maternal effects is attracting renewed attention, and novel statistical methods and designs, such as those exploiting Assisted Reproduction Technology (ART), are emerging [Rice et al., 2009; Thapar et al., 2007; Zhou et al., 2011]. ART allows complete separation of the effects of the child's genotype from those of the maternal genotype when gestational and biological mothers are different. More classical family designs have also been used to jointly

analyze these effects [Weinberg et al., 1998], with the possibility to include maternal-foetal interaction at one locus [Ainsworth et al., 2011; Sinsheimer et al., 2003]. Effects that are seldom analyzed are the interaction between a maternal genotype at one locus and a child genotype at another locus, and the interaction of two maternal loci to raise the disease risk in their offspring. Maternal \times offspring ($M \times O$) interaction at different loci is ubiquitous in mammals [Wolf, 2000]. In a mouse model, it was shown that maternal thyroid hormones bind the embryonic thyroid receptors [Nucera et al., 2010] and a list of 81 foetal genes have been identified that are maternally regulated [SenthamaraiKannan et al., 2011]. $M \times O$ interactions may partially determine birth-weight, a trait associated with many adulthood diseases [Wolf, 2000]. Simulation studies have shown that when $M \times O$ interaction exists, the maternally expressed allele may persist longer at a higher frequency in the population than the zygotically expressed allele, and under certain genetic models, it escapes from the purifying selection, increasing the disease prevalence in the population [Priest and Wade, 2010]. An interaction between two maternal genes ($M \times M$ interaction) can also increase the child disease risk. For example, an association was found between craniofacial malformations and

an interaction between maternal single nucleotide polymorphism (SNP) genotypes in folate and choline metabolism genes [Mostowska et al., 2010a, b].

The risk of developing testicular germ cell tumor (TGCT) has long been associated with antenatal exposure to maternal hormones. When exposed to high levels of estrogens in vitro, primordial germ cells (PGCs, the embryonic precursors of gametes) start transcription of C-kit ligand (*KITLG*), and the accumulation of this growth factor eventually transforms PGCs into tumorigenic cells [Moe-Behrens et al., 2003]. We and others have found that SNPs from *KITLG* are the most associated with TGCT in genome-wide association studies (GWAS) [Kanetsky et al., 2009; Rapley et al., 2009] and it can be hypothesized that the disease-initiating events occur in the early embryonic development from an interplay between maternal estrogens and child or maternal *KITLG*. The purpose of this study is to investigate the application of log-linear methods to test the interactions between a maternal genotype at one locus and a child genotype at another unlinked locus and between two unlinked maternal loci. We use a method originally proposed to analyze gene-environment interactions [Umbach and Weinberg, 2000], considering as exposure a second locus maternal genotype. Our study extends and investigates this approach in several ways: first, because genetic effects from maternal and child genotypes can be mutually confounded, we fit a model including the main effects from child and maternal genotypes along with two-locus $M \times O$ and $M \times M$ interactions. We propose that a further analysis stage may be conducted, where the least significant effects are dropped from the model and we discuss the implications of this model selection process. Second, we simulated data to assess the relative merits (power and type 1 error) of different approaches to testing for $M \times O$ and $M \times M$ interactions: testing assuming mating symmetry, either specifying or estimating the second locus genotype frequency distribution, and testing assuming Hardy-Weinberg equilibrium (HWE). Third, we investigate the sensitivity of the $M \times O$ and $M \times M$ interactions test to the violation of the assumption of parental mating symmetry. Fourth, we compare the power to detect the interactions to the power to detect main effects. Finally, we use the Expectation-Maximization (EM) algorithm to analyze incomplete trios and show that $M \times O$ and $M \times M$ interaction tests are independent of missing paternal genotypes. Therefore, although the methods considered in this study have been used previously, we revisit them from a different angle, reformulating the log-linear framework to accommodate two-locus models and exploring various novel aspects before conducting the power simulations. We then apply the method to a TGCT dataset, examining the interaction between a maternal hormone-related gene and *KITLG*.

METHODS

TWO-LOCUS LOG-LINEAR MODEL

We assume that family trios have been genotyped at a first locus and that the mothers are additionally genotyped at a second locus, both unlinked and in linkage equilibrium with the first locus. Although some of the tests we propose rely only on the assumption of linkage equilibrium, $M \times O$ and $M \times M$ cannot be tested simultaneously without imposing the absence of both linkage and linkage disequi-

librium (LD). Therefore, throughout this article when we say “unlinked” we also imply absence of LD. Originally, the log-linear method was designed to analyze genetic effects from one locus in cases, their mothers, or both, and possibly subject to imprinting [Weinberg et al., 1998]. Under the assumption of mating symmetry, joint genotypes of parents define six mating types, and joint genotypes of a family trio define 15 categories (see Table I). If mothers of cases are additionally genotyped at a second locus (M_2), the 15 categories can be stratified on the second locus genotypes and the proportions of family trios falling in each of the new defined categories is given in Table I. Using the original notation of Weinberg et al. [1998], locus 1 has main maternal effects S_i and main offspring effects R_j , where i, j represent the number of susceptibility alleles carried by mother and child, respectively. S_i and R_j are the relative excess of disease risk in the offspring conferred by maternal and offspring disease alleles. Scaling factors μ_m are population frequencies of the mating types.

We introduce two new sets of parameters: ϑ_{ik} coefficient for $M \times M$ interaction between the two loci and Φ_{jk} for $M \times O$ interaction (index k is the number of maternal alleles at locus 2, assumed to be unlinked to locus 1). Stratifying the expected cell counts by the second locus implies that each of the new cells must be weighted by the population genotype frequencies p_k at this locus. Using Table I, a two-locus log-linear model of family trio counts can be written as:

$$\ln(n_{M=i,F=f,C=j,M_2=k}) \propto \ln(\mu_m) + \delta_k + \alpha_i + \beta_j + \varphi_{ik} + \phi_{jk} + \ln[p_k 2^{I_{i=j=f=1}}], \quad (1)$$

where $n_{M=i,F=f,C=j,M_2=k}$ is the expected number of family trios with specific maternal (M), paternal (F), and child (C) genotype at locus 1 and maternal (M_2) genotype at locus 2. Mating type index m on the right-hand side of the equation ranges from 1 to 6 as (M,F) on the left-hand side takes the values (2,2), (2,1), (2,0), (1,1), (1,0), and (0,0). Regression coefficients α_i , β_j , φ_{ik} , and ϕ_{jk} are, respectively, the natural logarithm of disease relative risks S_i , R_j , ϑ_{ik} , and Φ_{jk} of Table I. All these coefficients are set to zero for the baseline genotype, i.e., $\alpha_0 = \beta_0 = 0$, $\varphi_{ik} = 0$ if $i = 0$ or $k = 0$, and $\phi_{jk} = 0$ if $j = 0$ or $k = 0$. The last term of the equation represents a “weight” or “offset.” There are two different ways giving rise to a heterozygous offspring when both parents are heterozygous, and they are equally likely in the absence of imprinting. In this case, the cell count is weighted by a factor of 2, which is reflected in the “offset” (i.e., the indicator variable $I_{i=j=f=1}$ takes the value of 1). In addition, with a second locus, each cell count is proportional to the genotype frequencies p_k as shown in Table I, and this is also reflected in the “offset” term.

Parameter δ_k is the natural logarithm of τ_k shown in Table I, which fits the main effect of maternal genotype k at locus 2, assuming the absence of effect from the baseline genotype ($\delta_0 = 0$). Although we are not interested in this effect, its inclusion in the model is necessary to ensure that the interaction test does not depend on genotype frequencies. On closer inspection, it can be seen that each of the new cells created by stratifying on locus 2 is weighted by genotype frequencies independently of coefficients ϑ_{ik} and Φ_{jk} (Table I). It follows that if one were interested in testing the locus 2 main effect, knowing its frequency distribution with precision would be required. It would also be required if one fitted the interaction terms without this main effect in the

TABLE I. Theoretical proportions of family trio types in a two-locus model

Configuration	Mating type	M ^a	F ^a	C ^a	Expected cell proportions		
					M ₂ = 0 ^b	M ₂ = 1 ^b	M ₂ = 2 ^b
1	1	2	2	2	$p_0 S_2 R_2 \mu_1$	$p_1 \vartheta_{21} \Phi_{21} S_2 R_2 \mu_1 \tau_1$	$p_2 \vartheta_{22} \Phi_{22} S_2 R_2 \mu_1 \tau_2$
2	2	2	1	2	$p_0 S_2 R_2 \mu_2$	$p_1 \vartheta_{21} \Phi_{21} S_2 R_2 \mu_2 \tau_1$	$p_2 \vartheta_{22} \Phi_{22} S_2 R_2 \mu_2 \tau_2$
3	2	2	1	1	$p_0 S_2 R_1 \mu_2$	$p_1 \vartheta_{21} \Phi_{11} S_2 R_1 \mu_2 \tau_1$	$p_2 \vartheta_{22} \Phi_{12} S_2 R_1 \mu_2 \tau_2$
4	2	1	2	2	$p_0 S_1 R_2 \mu_2$	$p_1 \vartheta_{11} \Phi_{21} S_1 R_2 \mu_2 \tau_1$	$p_2 \vartheta_{12} \Phi_{22} S_1 R_2 \mu_2 \tau_2$
5	2	1	2	1	$p_0 S_1 R_1 \mu_2$	$p_1 \vartheta_{11} \Phi_{11} S_1 R_1 \mu_2 \tau_1$	$p_2 \vartheta_{12} \Phi_{12} S_1 R_1 \mu_2 \tau_2$
6	3	2	0	1	$p_0 S_2 R_1 \mu_3$	$p_1 \vartheta_{21} \Phi_{11} S_2 R_1 \mu_3 \tau_1$	$p_2 \vartheta_{22} \Phi_{12} S_2 R_1 \mu_3 \tau_2$
7	3	0	2	1	$p_0 R_1 \mu_3$	$p_1 \Phi_{11} R_1 \mu_3 \tau_1$	$p_2 \Phi_{12} R_1 \mu_3 \tau_2$
8	4	1	1	2	$p_0 S_1 R_2 \mu_4$	$p_1 \vartheta_{11} \Phi_{21} S_1 R_2 \mu_4 \tau_1$	$p_2 \vartheta_{12} \Phi_{22} S_1 R_2 \mu_4 \tau_2$
9	4	1	1	1	$2p_0 S_1 R_1 \mu_4$	$2p_1 \vartheta_{11} \Phi_{11} S_1 R_1 \mu_4 \tau_1$	$2p_2 \vartheta_{12} \Phi_{12} S_1 R_1 \mu_4 \tau_2$
10	4	1	1	0	$p_0 S_1 \mu_4$	$p_1 \vartheta_{11} S_1 \mu_4 \tau_1$	$p_2 \vartheta_{12} S_1 \mu_4 \tau_2$
11	5	1	0	1	$p_0 S_1 R_1 \mu_5$	$p_1 \vartheta_{11} \Phi_{11} S_1 R_1 \mu_5 \tau_1$	$p_2 \vartheta_{12} \Phi_{12} S_1 R_1 \mu_5 \tau_2$
12	5	1	0	0	$p_0 S_1 \mu_5$	$p_1 \vartheta_{11} S_1 \mu_5 \tau_1$	$p_2 \vartheta_{12} S_1 \mu_5 \tau_2$
13	5	0	1	1	$p_0 R_1 \mu_5$	$p_1 \Phi_{11} R_1 \mu_5 \tau_1$	$p_2 \Phi_{12} R_1 \mu_5 \tau_2$
14	5	0	1	0	$p_0 \mu_5$	$p_1 \mu_5 \tau_1$	$p_2 \mu_5 \tau_2$
15	6	0	0	0	$p_0 \mu_6$	$p_1 \mu_6 \tau_1$	$p_2 \mu_6 \tau_2$

^aM, F, C: maternal, paternal, and child genotypes at the first locus.

^bM₂ = maternal genotypes at the second locus with frequency $p_k, k = 0,1,2$. Genotypes are coded as the number of minor alleles (assumed to increase the risk).

model. However, as we will show in simulations, estimates of the interaction terms are independent of the genotype frequencies provided that the main effect is included. Furthermore, Equation (1) is valid for testing the main effect δ_k itself when the genotype frequencies are correctly specified in the offset term (see Results).

Equation (1) requires specifying the second locus genotype frequencies in a similar way to the environmental risk factor ratios used by Umbach and Weinberg [2000] (see their equation on p. 257). However, as we consider a three-level exposure (second locus genotype), we fit a two-level main effect ($\delta_k, k = 1,2$) instead of one level. Their term, which is the natural logarithm of the ratio of exposed to unexposed individuals in the population, would be written as $\ln(p_k/p_0)$ in our notation. The Umbach and Weinberg [2000] term μ_i is then equivalent to $\ln(\mu_m) + \ln(p_0)$ in our notation.

The locus 2 genotype frequencies could be estimated, instead of specifying them in an offset term as in Equation (1). However, now the second locus main effect cannot be estimated. Equation (2) describes this model:

$$\ln(n_{M=i,F=f,C=j,M_2=k}) \propto \ln(\mu_m) + \lambda_k + \alpha_i + \beta_j + \varphi_{ik} + \Phi_{jk} + \ln(2^{I_i=j=f=1}), \quad (2)$$

where λ_k equals $\ln(p_k/p_0), k = 1,2$. Using the ratio of frequencies in the model ensures that their sum is constrained to equal to 1.

An alternative approach to reduce the number of parameters is to use Equation (1) and assume HWE for both loci. In fact, μ_m can be expressed as a function of one allele frequency if HWE is assumed at locus 1 [Weinberg et al., 1998]. However, unlike the main effects S_i and R_j that depend directly on the mating types, estimates of the interaction terms ϑ_{ik} and Φ_{jk} are conditionally independent of these nuisance parameters when S_i and R_j are fitted in the model. To see this, note that in Table I the interaction coefficients ϑ_{ik} and Φ_{jk} are defined in columns where $M_2 = 1$ and $M_2 = 2$ and they are tested against the baseline column $M_2 = 0$. They are

independent from mating type strata because the same coefficient μ_m multiplies cell counts in these three columns on each row of Table I. Therefore, modeling all mating types or replacing them by one parameter does not affect the power to detect the interactions (see Results).

Using Equation (1) or Equation (2), M×M and M×O interactions can be tested in a likelihood ratio test of the full model including all the effects vs. a model excluding the effect investigated. Log-linear methods also allow the analysis of M×O interactions involving one locus (coefficient ρ_{ij} in Equation (3), defined for $i > 0$ and $j > 0$, set to zero if $I = 0$ or $j = 0$). A likelihood ratio test of the interaction term ρ_{ij} is equivalent to the multinomial test used in the EMIM software [Ainsworth et al., 2011] and to the method of Sinsheimer et al. [2003] implemented in the software package Mendel [Lange et al., 2001].

$$\ln(n_{M=i,F=f,C=j}) \propto \ln(\mu_m) + \alpha_i + \beta_j + \rho_{ij} + \ln(2^{I_i=j=f=1}). \quad (3)$$

INCORPORATING INCOMPLETELY GENOTYPED TRIOS

The EM algorithm is a powerful approach to jointly analyzing full trios and incompletely genotyped trios in the log-linear framework [Weinberg, 1999]. For the analysis of M×M and M×O interactions, tests from Equation (1) and (2) are completely immune to missing paternal genotypes. This is because, as said earlier, the interaction terms ϑ_{ik} and Φ_{jk} are conditionally independent of the mating type strata when S_i and R_j are included in the model. In other words, missing fathers cause an ambiguity of mating types that could have an impact on the power to detect main effects S_i and R_j , but the interactions ϑ_{ik} and Φ_{jk} are completely unaffected. However, when the missing parent is the mother, it is reasonable to assume that her genotypes will be unavailable for both loci, in which case the data contain little

TABLE II. Simulated dominant models to evaluate the interaction tests

Model	R	S	ρ	$\exp(\delta)$	ϑ	Φ	Purpose ^a
1	1	1	1	1	1	1	Type 1 error of Equation (1) and (2) assuming mating symmetry, 1 assuming HWE, full data and missing parents, robustness to frequency misspecification.
2	1	1	1	1	1.5	1	Power + type 1 error.
3	1	1	1	1	2	1	Power + type 1 error.
4	1	1	1	1	1	1.5	Power + type 1 error.
5	1	1	1	1	1	2	Power + type 1 error.
6	1.5	1.5	1	1	1	1	Type 1 error + power for main effects of child and maternal genotypes, estimation of genotype frequencies with Equation (2)
7	1	1	1	1	1.5	1.5	Power comparison Equation (1) vs. Equation (2), full data and missing parents, estimation of genotype frequencies with Equation (2)
8	1	1	1	1	2	2	Power comparison Equation (1) vs. Equation (2), full data and missing parents, estimation of genotype frequencies with Equation (2)
9	1	1	1.5	1	1	1	Power comparison Equation (3) vs. EMIM.
10	1	1	2	1	1	1	Power comparison Equation (3) vs. EMIM.

^a R , S , Φ , ϑ , and δ represent, respectively, the main effect of child locus 1, main effect of maternal locus 1, two-locus $M \times O$ interactions, two-locus $M \times M$ interactions, and maternal second locus main effect. The purpose of simulations was to evaluate the power and/or type 1 error to detect $M \times M$ and $M \times O$ interactions in 300 and 1,000 full trios using Equation (1), assuming mating symmetry and correctly specifying locus 2 genotype frequencies, unless otherwise indicated. Dominant models are simulated for all relevant effects, i.e., $R = R_1 = R_2$, $S = S_1 = S_2$, $\rho = \rho_{11} = \rho_{12} = \rho_{21} = \rho_{22}$, $\vartheta = \vartheta_{11} = \vartheta_{12} = \vartheta_{21} = \vartheta_{22}$, and $\Phi = \Phi_{11} = \Phi_{12} = \Phi_{21} = \Phi_{22}$ and locus 2 has no main maternal effect ($\delta_1 = \delta_2 = 0$). The data simulated in these models were analyzed assuming dominant inheritance of all the effects tested and two nuisance parameters δ_1 and δ_2 were fitted for the main maternal effect of the second locus.

information about locus 2. We investigate the method's efficiency for $M \times M$ and $M \times O$ interactions when the EM algorithm is used to analyze trios with missing parental genotypes. More information on the application of the EM algorithm to incomplete family trios can be found in Weinberg [1999].

SIMULATIONS

We designed simulations to evaluate the type 1 error and power of Equation (1) and (2) to test two-locus $M \times M$ and $M \times O$ interactions as well as 3 for single-locus $M \times O$ interaction. In all simulations, we considered one locus genotyped in family trios (locus 1) and another, unlinked and in linkage equilibrium in the general population, genotyped only in the mothers (locus 2). In all simulated models, 1,000 replicates were generated where the baseline disease incidence rate was 0.1, for both loci HWE was assumed and the risk allele frequency (RAF) was 0.3. Under dominant inheritance (Table II), we simulated one example of a small dataset (300 family trios) and one moderate sample size (1,000 family trios). Only the sample size of 1,000 trios was considered for codominant genetic inheritance (Table III). Some of the simulated models were intended to assess type 1 error or power for both $M \times M$ and $M \times O$ interactions, while other models were designed to assess the type 1 error for one type of interaction and the power for the other. These models therefore allow possible confounding between the two effects to be assessed. They also allow an evaluation of the power gain in models restricted to only the genuine effects compared to a more saturated model. Models 11a–14a (Table III) were simulated under parental mating asymmetry to assess the sensitivity of the interaction tests to viola-

tion of the mating symmetry assumption. In these models, locus 1 RAF was 0.3 in females, but two equally sized strata of males were considered: one with RAF 0.1 and another with RAF 0.5, with HWE within each stratum. This corresponds to an overall RAF of 0.3 for males in the combined population, but gives rise to asymmetry in the mating patterns, since, for example, a female is now more likely to be heterozygote than is a male. For locus 2, the RAF remained 0.30 in females and in each of the two male strata, and the two loci were unlinked as before.

To demonstrate the independence of interaction tests from genotype frequency specification, datasets simulated under model 1 (see Table II) were analyzed assuming a broad range of the locus 2 genotype frequency distributions. To illustrate the immunity of the interaction tests to missing paternal data, we generated missing paternal genotypes in 10–80% of families in model 1 simulations. To assess the ability of the EM algorithm to handle missing data for two loci, we simulated missing maternal genotypes in 5–30% of families. Using data simulated under parental mating symmetry, we further compare: (1) the power of detecting two-locus $M \times M$ and $M \times O$ interactions using Equation (1) and (2), both assuming mating symmetry and Equation (1) assuming HWE (Equation (1).hwe); (2) the power to detect these interactions vs. the power to detect the main effects S_i and R_j of similar magnitude; and (3) the power to detect single-locus $M \times O$ interaction using Equation (3) vs. EMIM. The analysis including the main child and maternal effects from locus 1, a main effect from locus 2 and the two-locus $M \times O$ and $M \times M$ interactions under dominance or codominance will be referred to as a full model, while the analysis restricted to a subset of these effects will be referred to as a reduced or restricted model. Although our primary focus is on the validity and power of the interactions tests, we also report as supplementary data the risk parameter

TABLE III. Simulated codominant models^a

Model	R_1	R_2	S_1	S_2	Φ_{11}	Φ_{12}	Φ_{21}	Φ_{22}	ϑ_{11}	ϑ_{12}	ϑ_{21}	ϑ_{22}	δ_1	δ_2
11 + 11a	1	1	1	1	1.3	1.69	1.69	2.197	1	1	1	1	1	1
12 + 12a	1	1	1	1	1	1	1	1	1.4	1.96	1.96	2.744	1	1
13 + 13a	1	1	1	1	1.3	1.69	1.69	2.197	1.4	1.96	1.96	2.744	1	1
14 + 14a	1.3	1.69	1.4	1.96	1	1	1	1	1	1	1	1	1	1

^a R_j , S_i , Φ_{jk} , and ϑ_{ik} and δ_k represent, respectively, the main effect of child locus 1, main effect of maternal locus 1, two-locus $M \times O$ interactions, two-locus $M \times M$ interactions, and maternal second locus main effect. Models 11–14 were simulated under parental mating symmetry while 11a–14a were simulated under parental mating asymmetry with similar relative risks. Data simulated under these models were analyzed using Equation (1) assuming an unconstrained genetic model for all parameters of interest (S_i , R_j , Φ_{jk} , and ϑ_{ik}) along with nuisance parameters δ_1 and δ_2 . Robustness to genetic model misspecification was assessed by repeating the analysis assuming dominant effects in the analysis.

estimates for the simulated models. All data were simulated with a purpose-built C program and analyzed in STATA using Poisson regression. These programs are available from the authors on request.

APPLICATION TO TGCT

We applied two-locus log-linear models to 210 families (147 full trios and 63 case-mother pairs), testing the interactions between *KITLG* SNPs (rs1508595 and rs995030, each separately) and rs6259, a missense SNP from Sex Hormone Binding Globulin (*SHBG*; it binds estrogens and other sex hormones, regulating their availability). The two *KITLG* SNPs were the most significant in a TGCT GWAS and showed independent signals in multivariable analysis (Rapley et al., 2009). The LD between *KITLG* SNPs was estimated using 640 subjects (370 case parents and 270 unrelated controls) with STATA add-on command *pwld*.

RESULTS

VALIDITY AND POWER OF TWO-LOCUS INTERACTION TESTS IN FULL TRIOS

Table IV summarizes the type 1 error and power to detect $M \times M$ and $M \times O$ interactions at the nominal significance level of 0.05 in different simulated datasets with 300 or 1,000 trios. There are five sections in the table: the top section shows for simulated models 1 and 6 the type 1 error of the interactions using Equations (1) or (2), both assuming mating symmetry. For model 1, this section further depicts the robustness of 1 to misspecifying the second locus genotype frequencies and assesses the type 1 error of 1 under the assumption of HWE at locus 1 (Equation 1_hwe). All the results in this section show that the interaction tests have a non-inflated type 1 error using all the methods. The highest error observed is 5.8%, which is not significantly higher than 5% (one-sided P -value for the difference is 0.14). Results also confirm that the interaction tests based on Equation (1) are robust to genotype frequency misspecification: data analysis assuming a frequency distribution of 0.01, 0.18, 0.81 as well as 0.25, 0.50, 0.25 gave exactly the same interaction error rate as when the correct genotype frequencies of 0.09, 0.42, and 0.49 were used. As explained earlier, this is because the uncertainty is compensated for by the inclusion of the second locus main effect. However, it is important to note that this main effect can only be accurately estimated if the genotype frequency distribution is correctly speci-

fied (see next paragraph). Model 6 results show that main and interaction effects are not confounded if they are all fitted.

The second section of Table IV shows models 2–5 where one type of interaction was simulated under the null while the other was under the alternative hypothesis. The results in this section show that the interaction tests are always valid and have reasonable power. For example, in simulation model 3 with 300 trios, the $M \times O$ test has a type 1 error rate of 4.5% and the $M \times M$ test has a power of 73.6% at a nominal 5% significance level. The third section of Table IV shows the power for both $M \times M$ and $M \times O$ interactions (models 7 and 8) using Equations (1), 1_hwe, and (2). In all simulated models and sample sizes, Equations (1) and (2) have the same power to detect the interactions, since the only difference between them is in the treatment of the confounded main effects and genotype frequencies, and Equation 1_hwe shows similar power. For example, in model 8 with 300 families, the powers are 65.0% and 66.6% for the two interactions with 1 and 2, while they are 65.9% and 67.5% with 1_hwe. Results in the fourth section of Table IV (models 11–14) were obtained assuming an unconstrained genotypic model estimating 12 different effects and eight nuisance parameters. They show that $M \times O$ and $M \times M$ tests remain valid and retain good power with 1,000 full trios. Results in the last section of Table IV show that two-locus interaction tests remain valid and do not lose power when true parental mating is asymmetric but symmetry is assumed in the analysis.

Supplementary Material Table S1 shows that the frequencies of the second locus genotypes are accurately estimated with Equation (2), although this is based on the assumption that the second locus has no main effect. Risk parameter estimates using Equation (1) are unbiased and generally close to their true values (Tables S2–S5 in Supplementary Material). These risk estimates remain accurate when parental mating asymmetry is simulated but symmetry is assumed in the analysis.

POWER FOR MAIN EFFECTS AND SINGLE-LOCUS $M \times O$ INTERACTION

Although assuming HWE does not improve the power to detect $M \times M$ and $M \times O$ interactions (see Table IV), it does improve the power for the main effects as shown in Supplementary Material Table S6 (model 6). Overall, the power to detect the main effects is substantially higher than the power to detect interactions. As has been reported by others [Sinsheimer et al., 2003], we found that assuming mating

TABLE IV. Power and type 1 error (%) of the interactions using full trios ($\alpha = 0.05$)

Simulated Model ^a	No. families	Method ^b	Assumed p_k if Equation (1) ^c	ϑ^d	Φ^d
1	300	Equation (1)	True	4.8	4.6
1	300	Equation (1)	False 1	4.8	4.6
1	300	Equation (1)	False 2	4.8	4.6
1	300	Equation 1_hwe	True	4.5	4.7
1	300	Equation (2)	-	4.8	4.6
1	1,000	Equation (1)	True	4.4	3.8
1	1,000	Equation (1)	False 1	4.4	3.8
1	1,000	Equation (1)	False 2	4.4	3.8
1	1,000	Equation 1_hwe	True	4.2	3.8
1	1,000	Equation (2)	-	4.4	3.8
6	300	Equation (1)	True	5.2	5.7
6	300	Equation (2)	-	5.2	5.7
6	1,000	Equation (1)	True	5.8	5.4
6	1,000	Equation (1)	-	5.8	5.4
2	300	Equation (1)	True	32.7	4.3
3	300	Equation (1)	True	73.6	4.5
4	300	Equation (1)	True	5.1	34.6
5	300	Equation (1)	True	5.2	73.9
2	1,000	Equation (1)	True	81.9	5.5
3	1,000	Equation (1)	True	99.5	5.3
4	1,000	Equation (1)	True	5.7	82.2
5	1,000	Equation (1)	True	5.4	99.6
7	300	Equation (1)	True	31.9	34.4
7	300	Equation 1_hwe	True	32.1	34.4
7	300	Equation (2)	-	31.9	34.4
8	300	Equation (1)	True	65.0	66.6
8	300	Equation 1_hwe	True	65.9	67.5
8	300	Equation (2)	-	65.0	66.6
7	1,000	Equation (1)	True	78.2	79.2
7	1,000	Equation 1_hwe	True	78.7	79.3
7	1,000	Equation (2)	-	78.2	79.2
8	1,000	Equation (1)	True	99.0	99.0
8	1,000	Equation 1_hwe	True	99.3	99.0
8	1,000	Equation (2)	-	99.0	99.0
11	1,000	Equation (1)	True	4.8	60.6
12	1,000	Equation (1)	True	84.9	5.0
13	1,000	Equation (1)	True	85.5	61.0
14	1,000	Equation (1)	True	4.9	5.1
11a	1,000	Equation (1)	True	4.5	63.5
12a	1,000	Equation (1)	True	84.5	4.3
13a	1,000	Equation (1)	True	85.4	62.3
14a	1,000	Equation (1)	True	4.0	4.6

^aEach model was simulated in one run with 300 trios and in a separate run with 1,000 trios, each repeated 1,000 times, and the data were analyzed using the approaches in column 3 in a full model assuming dominant (models 1–8) or codominant inheritance (models 11–14 and 11a–14a).

^bEquation (1) and (2) assume mating symmetry, Equation (1)_hwe assumes HWE at locus 1. Genotype frequencies estimated with Equation (2) are supplied in Supplementary Material Table S1.

^cThe true genotype frequency distribution simulated was 0.09/0.42/0.49 for the minor homozygous/heterozygous/major homozygous. False 1 and False 2 distributions assumed in some analyses were, respectively, 0.01/0.18/0.81 and 0.25/0.50/0.25.

^dPower or type 1 error in same column; see simulated models in Tables II and III. Φ and ϑ are $M \times O$ and $M \times M$ interactions, respectively.

type symmetry in testing the main child effect remains valid when the mating is asymmetric while testing the main maternal effect is invalid (Table S6 in Supplementary Material, model 13a). We stated in the Methods section that Equation (1) could be used to test the second locus main effect when its genotype frequency distribution is known, and Supplementary Material Table S7 gives a confirmation of

this statement. However, a small deviation from the correct frequency can cause considerable effect inflation, invalidating the test of the main effect. Using the data simulated in model 1, we fitted Equation (3) to assess its type 1 error for single-locus $M \times O$ interaction. The type 1 error found at level 0.05 was 5.2% with 1,000 trios and 5.5% with 300 trios. With sample sizes 300 and 1,000 trios, we found that

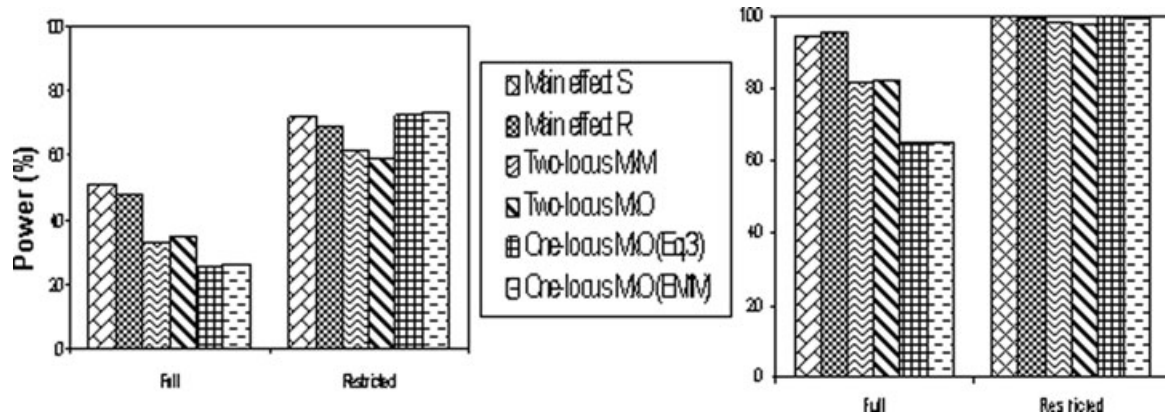


Fig. 1. Power ($\alpha = 0.05$) to detect different types of genetic effects with similar strength (relative risk 1.5) using full and restricted models. The restricted model fits the truly simulated effects plus the second locus main effect when there is an interaction involving this locus. Data from simulated model 6 for R and S, model 2 for M×M interaction, model 4 for two-locus M×O interaction and model 9 for single-locus M×O interaction. Left panel: 300 full trios; right panel: 1,000 full trios. Analyses conducted using Equation (1) unless otherwise indicated. The analysis assumed a dominant genetic model.

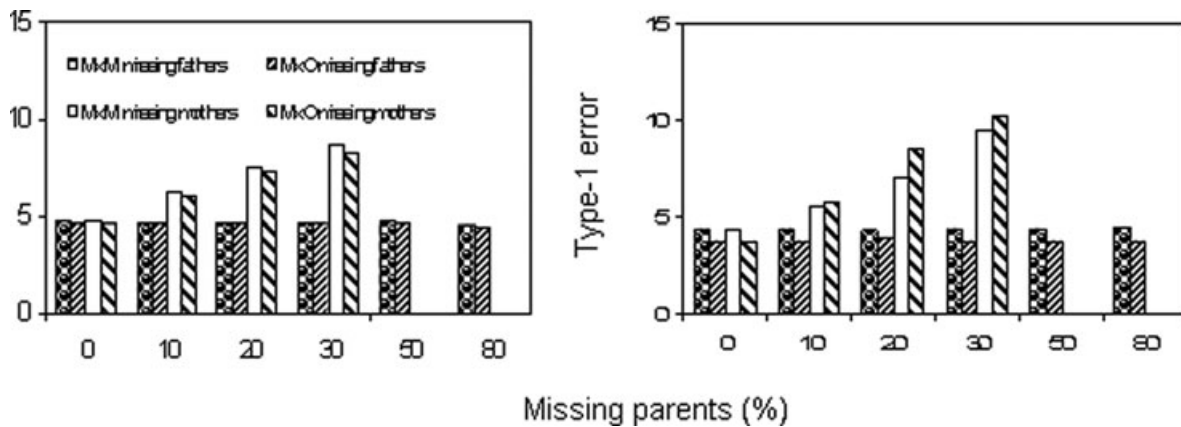


Fig. 2. Type 1 error rate for M×M and M×O interactions at nominal level 0.05 with missing parents. Left panel: total sample of 300 families; right panel: total sample of 1,000 families. Data from simulated model 1 and analyzed using the full model 1. The analysis assumed a dominant genetic model.

Equation (3) and EMIM have similar power to detect single-locus M×O interactions (Fig. 1).

MODEL REDUCTION

All the results presented so far were obtained when fitting a full model. While this full model is necessary to detect all the existing effects, the large number of terms impacts on the power to detect each of them. Restricting the analysis to the truly simulated effects (in addition to locus 2 main effect terms when there is an interaction involving this locus), the power is substantially increased, as can be seen in Figure 1.

TYPE 1 ERROR AND POWER FOR THE INTERACTIONS WITH MISSING PARENTS

The validity, power, and estimation accuracy of the M×M and M×O interaction tests using Equation (1) are not affected by missing fathers up to 80% (Figs. 2 and 3; Table S2 in Supplementary Material). By contrast, when genotypes are missing in mothers of cases at the two loci considered,

both interaction tests have an inflated type 1 error. In an initial sample size of 300 trios, there is a significant inflation with 10% missing mothers or more, while this occurs with 15% or more missing mothers in an initial sample size of 1,000 trios (Fig. 2). With a total sample size of 1,000 families, power to detect M×M and M×O interactions is unaffected by up to 10% missing mothers (Table S8 in Supplementary Material), while their risk parameter estimates remain accurate with up to 20% missing mothers (Table S2 in Supplementary Material). When parental mating asymmetry is simulated but the symmetry is assumed in the analysis, as much as 80% missing fathers do not affect the validity, the power, or the estimation accuracy of the two-locus interactions test (Tables S5 and S9 in Supplementary Material).

ROBUSTNESS TO MODEL MISSPECIFICATION

When the correct model for the interactions is codominant, assuming a dominant model in the analysis does not invalidate the test, nor does it cause a power loss (Table S10 in Supplementary Material). However, as expected, the

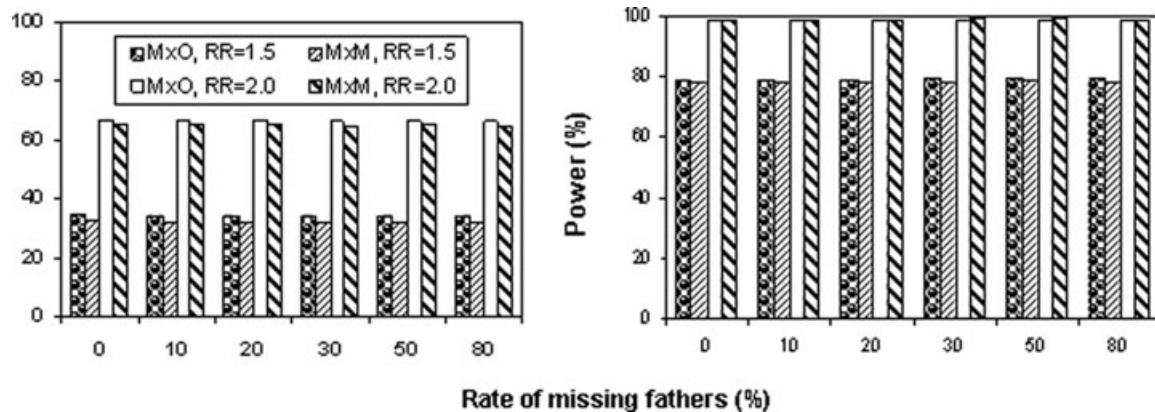


Fig. 3. Power for M×M and M×O interactions at nominal level 0.05 with 10–80% missing fathers. Left panel: total sample of 300 families; right panel: total sample of 1,000 families. Data from simulated models 7 and 8 and analyzed using the full model 1. The analysis assumed a dominant genetic model.

parameters estimation becomes inaccurate when the effects exist (Table S11 in Supplementary Material).

TGCT DATASET ANALYSIS

Despite showing independent effects on TGCT in GWAS, we found that *KITLG* rs1508595 and rs995030 are in strong LD ($D' = 0.91$, $r^2 = 0.65$, MAF = 0.17 and 0.14, respectively). In the original results from the GWAS [Kanetsky et al., 2009; Rapley et al., 2009], where major rather than minor alleles were tested, allelic relative risks were estimated to be 3.1 and 2.6, respectively, and the pattern of risk was consistent with a multiplicative genetic model. In view of the small sample size and rarity of the minor homozygote genotype, we here carry out a primary analysis of the 147 full trios and 63 case-mother pairs assuming dominant main effects and interactions using Equation (1) in full and reduced models. Case-father pairs were not used since simulations showed that they could inflate type 1 error. The reduced model included the two most prominent effects in the full model. Table V reports the results: in the full model, the most significant effect was the inherited *KITLG* genotypes. Whichever SNP analyzed showed a protective dominant minor allele relative risk $R \approx 0.40$ ($P \leq 4 \times 10^{-4}$), confirming the GWAS results. The second most significant effect was the M×M interaction involving rs1508595 and rs6259 (relative risk 2.2, $P = 0.07$, with a similar result when rs995030 was analyzed). In a model restricted to the two most prominent effects, the effect of inherited genotypes in *KITLG* is maintained, and the M×M interaction is also nominally significant between rs1508595 and rs6259 ($P = 0.03$). In an analysis assuming a multiplicative model for the main effects and dominant interaction terms, the results were very similar, reflecting the low number of minor homozygotes (results not shown).

DISCUSSION

Maternal risk factors for child disease include environmental exposures and genetic effects (i.e., untransmitted alleles) and it has been suggested that maternal genetic effects might contribute to the so-called “missing heritability” [Nadeau, 2009; Zhou et al., 2011]. Different designs and methods have been used to analyze maternal exposures

and their interactions with offspring genotypes, including at genomewide scale [Beatty et al., 2011]. Most often, the designs and/or analysis methods do not allow for a full investigation of all the effects that can potentially be involved. Here, we revisit the log-linear approach for family trio analysis to investigate its efficiency to detect intergenerational epistasis. We propose fitting a full model in stage 1 to avoid confounding and to drop the least prominent effects in a second stage to increase the power. As with any regression analysis, if the same data are used for model selection and testing of parameters, there will be some inflation of type 1 error rate [Hurvich and Tsai, 1990], so that results must be interpreted bearing this in mind, or, preferably but often unrealistically, an independent dataset should be used. The approach can be applied to genome-wide studies but requires prioritizing the analysis by choosing which locus will be tested for both maternal and inherited genotype effects (locus 1) or only for its interactions (locus 2). We have only examined the performance of the methods under the assumption that the two interacting loci are unlinked and not in LD.

Overall findings of this study are that: (1) two-locus log-linear modeling is effective in detecting M×O and M×M interactions; (2) similar power can be achieved to detect M×O and M×M interactions of similar strength; this power can be substantial, although lower than the power to detect the main effects; (3) M×O and M×M interactions are conditionally independent from paternal genotypes but are inflated by missing maternal genotypes; (4) in scenarios where HWE applies, its assumption increases the power to detect main effects but does not affect the interaction tests; (5) M×O and M×M tests are robust to the violation of mating symmetry assumption but, as has already been observed in single-locus models, the main maternal effect needs to be included in the model, although it is not itself robust to the violation of this assumption [Sinsheimer et al., 2003]; (6) in testing for interaction effects, the method is insensitive to any inaccuracy in the second locus genotype distribution, and specifying this distribution is as powerful as estimating it in the model; (7) if the second locus genotype distribution is well specified, the method is valid to test the main effect of the second locus; (8) the risk to testicular cancer may be increased by an interaction between maternal variants in *KITLG* and *SHBG*.

TABLE V. TGCT family data analysis with Equation (1)

Analysis model	Locus 1 ^a	Locus 2 ^a	R (P-value) ^b	S (P-value) ^b	Φ (P-value) ^b	‡ (P-value) ^b
Full	<i>KITLG</i> /rs995030	<i>SHBG</i> /rs6259	0.41 (3×10^{-4})	1.05 (0.84)	1.34 (0.56)	1.63 (0.26)
	<i>KITLG</i> /rs1508595		0.37 (4×10^{-4})	0.85 (0.55)	1.32 (0.59)	2.19 (0.07)
Restricted to two effects	<i>KITLG</i> /rs995030	<i>SHBG</i> /rs6259	0.43 (2×10^{-4})	-	-	1.84 (0.12)
	<i>KITLG</i> /rs1508595		0.40 (2×10^{-4})	-	-	2.19 (0.03)

^aMinor allele frequencies are 0.14 for rs995030, 0.17 for rs1508595, and 0.20 for rs6259.

^bEstimated relative risks under a dominant model for locus 1 inherited genotype (R), locus 1 maternal genotype (S), two-locus M×O (Φ) and M×M (‡) interactions.

Significant results at $p < 0.05$ are highlighted in boldface.

The main limitation of the approach is sample size requirements, which is a common feature of all interaction analyses. However, the power to detect intergenerational interactions is reasonable and comparable to the power to detect intragenerational interactions or single-locus M×O interaction. With 300 trios, a dominant relative risk of 1.5 is detectable with 70% power in the restricted model, while power approaches 100% in samples of 1,000 trios (Fig. 1). The full model can be subject to multicollinearity in small datasets or when the analyzed allele is too rare. This step should be used to avoid mutual confusion of different effects if there is no collinearity, and a second analysis step modeling only the most prominent effects can increase the power, although this is subject to the above caveats.

The application of the method to TGCT confirmed the finding of two genome-wide studies: inherited *KITLG* variants (minor alleles) have a strong protective effect against the disease (relative risk 0.4, $P = 2 \times 10^{-4}$, assuming a dominant model). The method further detected a nominally significant M×M interaction between *KITLG* rs1508595 and *SHBG* rs6259 (reduced model relative risk = 2.2, $P = 0.03$). A simulation of similar conditions to this real dataset shows that we had 60% and 77% power to detect this interaction in full and reduced models, respectively. There was no significant main maternal effect, and a model including imprinting showed that it was not significant. TGCT has an early age of onset and a good response to chemotherapy [Rapley and Nathanson, 2010]. Its familial risk is higher than for other cancers, yet linkage studies were not successful in identifying susceptibility loci. The relative risk conferred by *KITLG* is one of the largest published so far and it has been replicated within and across studies. These unusual characteristics may derive from complex biological mechanisms such as the one reported here, although a larger sample is needed for confirmation. Both *KITLG* and *SHBG* have been previously implicated in TGCT, and if their interaction could be verified it would consolidate the observation made in *in vitro* studies that *KITLG* is activated by estrogen.

It has been demonstrated for single-locus analysis that power and the ability to check assumptions can be enhanced by supplementing the case-parent trios with unrelated controls and one or both parents [see e.g., Vermeulen et al., 2009], and some of these benefits are also likely to apply when two loci are under consideration. This investigation was motivated by the analysis of an existing dataset of testicular cancer cases and their parent(s). Study design is often largely driven by practical considerations; since testicular cancer primarily affects young men, recruiting testicular cancer cases attending a clinic, often accompanied by a parent, was an efficient recruitment strategy, whereas recruiting unaffected young men and their parent(s) may be more dif-

ficult. Thus, the focus of our investigation was on methods of analysis rather than design.

ACKNOWLEDGMENTS

This study was funded by Cancer Research UK grant C588/A4994. We thank Helen Snowden, Mark Harland, and Juliette Randerson-Moor for genotyping SNPs in candidate genes to support our method evaluation. We are grateful to the manuscript reviewers for their insightful comments and suggestions. We declare no conflict of interest.

REFERENCES

- Ainsworth HE, 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(1):19–45.
- Barker DJP, Gelow J, Thornburg K, Osmond C, Kajantie E, Eriksson JG. 2010. The early origins of chronic heart failure: impaired placental growth and initiation of insulin resistance in childhood. *Eur J Heart Fail* 12(8):819–825.
- Beaty TH, Ruczinski I, Murray JC, Marazita ML, Munger RG, Hetmanski JB, Murray T, Redett RJ, Fallin MD, Liang KY, Wu T, Patel PJ, Jin SC, Zhang TX, Schwender H, Wu-Chou YH, Chen PK, Chong SS, Cheah F, Yeow V, Ye X, Wang H, Huang S, Jabs EW, Shi B, Wilcox AJ, Lie RT, Jee SH, Christensen K, Doheny KE, Pugh EW, Ling H, Scott AF. 2011. Evidence for gene-environment interaction in a genome wide study of nonsyndromic cleft palate. *Genet Epidemiol* 35:469–478.
- Drever N, Saade GR, Bytautiene E. 2010. Fetal programming: early-life modulations that affect adult outcomes. *Curr Allergy Asthm R* 10(6):453–459.
- Eriksson JG, Kajantie E, Thornburg KL, Osmond C, Barker DJ. 2011. Mother's body size and placental size predict coronary heart disease in men. *Eur Heart J* 32(18):2297–2303.
- Hurvich CM, Tsai CL. 1990. The impact of model selection on inference in linear-regression. *Am Stat* 44(3):214–217.
- Jansson T, Ekstrand Y, Bjorn C, Wennergren M, Powell TL. 2002. Alterations in the activity of placental amino acid transporters in pregnancies complicated by diabetes. *Diabetes* 51(7):2214–2219.
- Kanetsky PA, Mitra N, Vardhanabhuti S, Li MY, Vaughn DJ, Letrero R, Ciosek SL, Doody DR, Smith LM, Weaver J and others. 2009. Common variation in *KITLG* and at 5q31.3 predisposes to testicular germ cell cancer. *Nat Genet* 41(7):811–815.
- Lange K, Cantor R, Horvath S, Perola M, Sabatti C, Sinsheimer J, Sobel E. 2001. Mendel version 4.0: a complete package for the exact genetic analysis of discrete traits in pedigree and population data sets. *Am J Hum Genet* 69(4):504.
- Maringhini S, Corrado C, Maringhini G, Cusumano R, Azzolina V, Leone F. 2010. Early origin of adult renal disease. *J Matern Fetal Neonatal Med* 23:84–86.

- Moe-Behrens GHG, Klinger FG, Eskild W, Grotmol T, Haugen TB, De Felici M. 2003. Akt/PTEN signaling mediates estrogen-dependent proliferation of primordial germ cells in vitro. *Mol Endocrinol* 17(12):2630–2638.
- Mostowska A, Hozyasz KK, Biedziak B, Misiak J, Jagodzinski PP. 2010a. Polymorphisms located in the region containing BHMT and BHMT2 genes as maternal protective factors for orofacial clefts. *Eur J Oral Sci* 118(4):325–332.
- Mostowska A, Hozyasz KK, Wojcicki P, Dziegielewska M, Jagodzinski PP. 2010b. Associations of folate and choline metabolism gene polymorphisms with orofacial clefts. *J Med Genet* 47(12):809–815.
- Nadeau JH. 2009. Transgenerational genetic effects on phenotypic variation and disease risk. *Hum Mol Genet* 18(R2):R202–R210.
- Nucera C, Muzzi P, Tiveron C, Farsetti A, La Regina F, Foglio B, Shih S-C, Moretti F, Della Pietra L, Mancini F and others. 2010. Maternal thyroid hormones are transcriptionally active during embryo-foetal development: results from a novel transgenic mouse model. *J Cell Mol Med* 14(10):2417–2435.
- Priest NK, Wade MJ. 2010. Maternal-zygotic epistasis and the evolution of genetic diseases. *J Biomed Biotechnol* 2010:478732.
- Rapley EA, Nathanson KL. 2010. Predisposition alleles for testicular germ cell tumour. *Curr Opin Genet Dev* 20(3):225–230.
- Rapley EA, Turnbull C, Al Olama AA, Dermizakis ET, Linger R, Huddart RA, Renwick A, Hughes D, Hines S, Seal S and others. 2009. A genome-wide association study of testicular germ cell tumor. *Nat Genet* 41(7):807–811.
- Rice F, Harold GT, Boivin J, Hay DF, van den Bree M, Thapar A. 2009. Disentangling prenatal and inherited influences in humans with an experimental design. *Proc Natl Acad Sci USA* 106(7):2464–2467.
- SenthamaraiKannan P, Sartor MA, O'Connor KT, Neumann JC, Klyza Jr JP, Succop PA, Wagner BD, Karyala S, Medvedovic M, Menon AG. 2011. Identification of maternally regulated fetal gene networks in the placenta with a novel embryo transfer system in mice. *Physiol Genomics* 43(7):317–324.
- Sinsheimer JS, Palmer CGS, Woodward JA. 2003. Detecting genotype combinations that increase risk for disease: the maternal-fetal genotype incompatibility test. *Genet Epidemiol* 24:1–13.
- 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.
- Umbach DM, Weinberg CR. 2000. The use of case-parent triads to study joint effects of genotype and exposure. *Am J Hum Genet* 66(1):251–261.
- Vermeulen SH, Shi M, Weinberg CR, Umbach DM. 2009. A hybrid design: case-parent triads supplemented by control-mother dyads. *Genet Epidemiol* 33(2):136–144.
- Weinberg CR. 1999. Allowing for missing parents in genetic studies of case-parent triads. *Am J Hum Genet* 64(4):1186–1193.
- 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(4):969–978.
- Winder NR, Krishnaveni GV, Hill JC, Karat CLS, Fall CHD, Veena SR, Barker DJP. 2011. Placental programming of blood pressure in Indian children. *Acta Paediatrica* 100:653–660.
- Wolf JB. 2000. Gene interactions from maternal effects. *Evolution* 54(6):1882–1898.
- Zhou JJ, Pelka S, Lange K, Palmer CGS, Sinsheimer JS. 2011. Dissecting prenatal, postnatal, and inherited effects: ART and design. *Genet Epidemiol* 35(6):437–446.