Detection of Parent-of-Origin Effects Based on Complete and Incomplete Nuclear Families with Multiple Affected Children

Ji-Yuan Zhou^{a, b} Yue-Qing Hu^{a, c} Shili Lin^d Wing K. Fung^a

^aDepartment of Statistics and Actuarial Science, The University of Hong Kong, Hong Kong, ^bDepartment of Mathematics, Huaihua University, Huaihua, Hunan, and ^cDepartment of Mathematics, Southeast University, Nanjing, China; ^dDepartment of Statistics, The Ohio State University, Columbus, Ohio, USA

Key Words

Parent-of-origin effects • Genomic imprinting • Missing parent • Incomplete nuclear family • Complete nuclear family • Genotypic relative risk • Population stratification demographic model • Assortative mating demographic model

Abstract

Parent-of-origin effects are important in studying genetic traits. More than 1% of all mammalian genes are believed to show parent-of-origin effects. Some statistical methods may be ineffective or fail to detect linkage or association for a gene with parent-of-origin effects. Based on case-parents trios, the parental-asymmetry test (PAT) is simple and powerful in detecting parent-of-origin effects. However, it is common in practice to collect nuclear families with both parents as well as nuclear families with only one parent. In this paper, when only one parent is available for each family with an arbitrary number of affected children, we firstly develop a new test statistic 1-PAT to test for parent-of-origin effects in the presence of association between an allele at the marker locus under study and a disease gene. Then we extend the PAT to accommodate complete nuclear families each with one or more affected children. Combining families with both parents and families with only one parent, the C-PAT is proposed to detect parent-of-origin effects. The validity of the test statistics is verified by simulation in various scenarios of parameter values. A power study shows that using the ad-

ditional information from incomplete nuclear families in the analysis greatly improves the power of the tests, compared to that based on only complete nuclear families. Also, utilizing all affected children in each family, the proposed tests have a higher power than when only one affected child from each family is selected. Additional power comparison also demonstrates that the C-PAT is more powerful than a number of other tests for detecting parent-of-origin effects.

Copyright © 2008 S. Karger AG, Basel

Introduction

Parent-of-origin effects, also known as ‘genomic imprinting’, are important in studying genetic traits. More and more genes are found to show parent-of-origin effects. Morison et al. [1] constructed an imprinted-gene database and more than 1% of all mammalian genes are believed to be imprinted (<http://igc.otago.ac.nz>). For some complex diseases such as Beckwith-Wiedemann, Prader-Willi, and Angleman syndromes, parent-of-origin effects have been demonstrated [2, 3]. For other complex diseases such as diabetes, hereditary paragangliomas, schizophrenia, intrauterine growth retardation, autism, neural tube defects, and obesity, parent-of-origin effects are suspected or hypothesized to play an important role, although specific genes have not been identified conclusively as yet [4–10].

Recently, there has been considerable interest in the detection of parent-of-origin effects. For case-parents trios and single markers, Weinberg et al. [11] reported a versatile log-linear model for candidate genes to test/estimate linkage disequilibrium, maternal effects, and parent-of-origin effects. Based on case-parents trios, Weinberg [12] developed the parental-asymmetric test (PAT) in the case of no maternal genotype effects and the parent-of-origin likelihood ratio test (PO-LRT) based on the logistic model to test for parent-of-origin effects, and found that the PAT was much more powerful than the PO-LRT. Based on the PAT, Becker et al. [13] proposed the HAP-PAT to test for parent-of-origin effects in the case of multiple affected children and multiple tightly linked markers. On the other hand, the log-linear methods were applicable to a disease gene [11, 12, 14]. If the candidate gene under study is not a marker allele per se, the recombination fraction between the marker locus (ML) and the disease susceptibility locus (DSL) should be taken into account in the analysis. In this situation, the log-linear model is not strictly correct for a marker under alternatives to the null hypothesis [12].

In practical studies, it is often the case to collect genotyping information from complete nuclear families (families with both parents) as well as incomplete nuclear families (families in which not both parents are available). Information from these incomplete nuclear families could be generally incorporated to improve the statistical power of a test. However, Rampersaud et al. [10] commented that one cannot study parent-of-origin effects when both parents are missing. The incomplete nuclear families hereafter simply refer to those with only one parent. Rampersaud et al. [10] suggested the combined likelihood ratio test (Combined LRT) for parent-of-origin effects in the presence of missing parental genotypes by incorporating additional information from the genotypes of unaffected siblings to improve inference of missing parental data. As shown in [13], it is more reasonable to consider all affected children of a nuclear family than to randomly choose one affected child from each family in conducting parent-of-origin tests. However, it is difficult to extend the approaches in the framework of log-linear model to general nuclear families with multiple affected children [10, 13].

In this paper, we firstly develop a new test statistic 1-PAT to test for parent-of-origin effects in the case of incomplete nuclear families with an arbitrary number of affected children, and show that it is more efficient to use all affected offspring in a family in conducting parent-of-origin effects tests. Then, we extend the PAT to accom-

modate complete nuclear families each with one or more affected children and propose the combined statistic C-PAT of the PAT and 1-PAT by using both the complete nuclear families and incomplete nuclear families in the presence of association between an allele at the ML and a disease gene. The validity of the proposed parent-of-origin tests is verified through simulation in various scenarios of parameter values. A power study shows that the C-PAT by capturing the information from incomplete nuclear families has a higher power than the PAT which only uses the observed complete nuclear families. The power comparison also illustrates that the C-PAT is more powerful than a number of other tests for detecting parent-of-origin effects. In summary, the C-PAT incorporating incomplete families is simple and powerful in testing for parent-of-origin effects in the case of no maternal genotype effects when association between an allele at the ML and a disease gene is present.

Methods

Background and Notation

We consider a DSL with mutant allele D and normal allele d , and the ML of interest with two alleles M_1 and M_2 . The four ordered genotypes at the DSL are D/D , D/d , d/D , and d/d , respectively. The left allele of the slash denotes the allele transmitted from the father and the right one denotes the allele transmitted from the mother. Since parent-of-origin effects are taken into account, the risks of two heterozygous genotypes D/d and d/D at the DSL may be different. We denote the four associated risks of genotypes D/D , D/d , d/D and d/d by $\phi_{D/D}$, $\phi_{D/d}$, $\phi_{d/D}$, and $\phi_{d/d}$, respectively. We assume that the risk with only one mutant D is between the risk with no mutant and the risk with two mutants, i.e., $\phi_{d/d} \leq \phi_{D/d}$, $\phi_{d/D} \leq \phi_{D/D}$. Thus the degree of imprinting $I = (\phi_{D/d} - \phi_{d/D})/2$ ranges from $(\phi_{d/d} - \phi_{D/D})/2$ to $(\phi_{D/D} - \phi_{d/d})/2$ and is used to measure the parent-of-origin effects [15]. Specifically, $I > 0$ indicates a maternal parent-of-origin effect or equivalently paternal expression, $I < 0$ represents the reversal, and $I = 0$ denotes either no parent-of-origin effects or no effect of the gene on risk. There are two extreme cases, one is complete paternal parent-of-origin effect, i.e., $\phi_{D/d} = \phi_{d/d}$ and $\phi_{d/D} = \phi_{D/D}$, and the other is complete maternal parent-of-origin effect, i.e., $\phi_{D/d} = \phi_{D/D}$ and $\phi_{d/D} = \phi_{d/d}$. Note that there are four haplotypes M_1D , M_1d , M_2D , and M_2d at the ML and DSL. Let θ be the recombination fraction between the ML and DSL. In order to test for parent-of-origin effects, we assume that the ML and DSL are in linkage disequilibrium (LD), which is measured by D' [16].

We begin by describing the notation for case-parents trio data. For convenience, let 0, 1 and 2 represent the marker genotypes M_2M_2 , M_1M_2 and M_1M_1 , respectively, and F , M and C represent the genotypes of the father, mother and child, respectively, and so F , M and C take possible values of 0, 1 or 2. Throughout this paper, the mating symmetry is assumed as in [12], i.e., $P(F = f, M = m) = P(F = m, M = f)$ for all $f, m = 0, 1, 2$.

When there are missing parental genotypes, we follow Hu et al. [17] by assuming that there is nondifferential availability or missingness of parental genotype data as presented by Allen et al. [18]. That is, we assume that whether the genotype of a parent is missing is independent of his/her underlying genotype. Moreover, the conditional distribution of the trio marker genotype configuration FMC given that the child is a case are the same irrespective of whether the father's or the mother's genotype is missing.

We describe the existing statistic PAT for detecting parent-of-origin effects based on case-parents trios as below. Suppose we have n independent case-parents trios, each with known marker genotypes FMC for the father, mother, and affected child. Theoretically, the total number of possible FMC combinations is $3^3 = 27$, but only 15 of these types are genetically possible. All the fifteen types of family and the corresponding conditional probabilities $s_1, s_2, \dots, s_{15}$ are listed in table 1. For example, $s_1 = P(FMC = 212 | \text{child is affected})$ is the conditional probability that a family falls into the first category $F = 2, M = 1, C = 2$, given the child is a case. Detailed expressions of the s_i 's for a homogenous population can be found in [19]. Denote $N_{F>M, C=1}$ and $N_{F<M, C=1}$ as the numbers of case-parents trios with heterozygous child in which the father carries more and fewer copies of marker allele M_1 than the mother, respectively. Under the condition that there are no maternally-mediated genetic effects, the PAT statistic can be expressed as

$$\text{PAT} = \frac{N_{F>M, C=1} - N_{F<M, C=1}}{\sqrt{N_{F>M, C=1} + N_{F<M, C=1}}} \quad (1)$$

Note that the PAT proposed by Weinberg [12] is the square of the right side of equation (1). The PAT is valid for testing parent-of-origin effects without assuming Hardy-Weinberg equilibrium (HWE) or the ML being a DSL per se.

Methods when Only One Parent Is Available

Now suppose we have n_M case-mother pairs each with known marker genotype pair MC for the mother and affected child, and n_P case-father pairs each with known marker genotype pair FC for the father and affected child. Theoretically, the total number of possible MC combinations is $3^2 = 9$, but only 7 of these types are genetically possible, i.e., $MC = 22, 21, 12, 11, 10, 01, 00$. Similarly, the combination FC takes these 7 possible types too. Note that testing for parent-of-origin effects is just to test the equality between $\phi_{D|d}$ and $\phi_{d|D}$, i.e., the heterozygous risks when the mutant allele D is paternally and maternally inherited, respectively. In the presence of association between the marker allele M_1 and mutant allele D , we therefore consider the difference between the numbers of heterozygous children who inherit the marker allele M_1 from the father and mother, respectively. For case-parents trios, that difference is just $N_{F>M, C=1} - N_{F<M, C=1}$ as in the numerator of the PAT in equation (1). For case-mother pairs, that difference can be expressed as $N_{M<C, C=1} - N_{M>C, C=1}$, where $N_{M<C, C=1}$ is the number of heterozygous children with homozygous mother $M = 0$, which indicates that the children's marker allele M_1 is inherited from the father, and $N_{M>C, C=1}$ is the number of heterozygous children with homozygous mother $M = 2$, which signifies that the children's marker allele M_1 is inherited from the mother. Similarly, for case-father pairs, we are interested in the difference $N_{F>C, C=1} - N_{F<C, C=1}$, where $N_{F>C, C=1}$ and $N_{F<C, C=1}$ are defined by analogy. Jointly using case-mother pairs and case-father pairs, we should investigate the following weight-

Table 1. Classification of all 15 family types for nuclear families each with a single affected child, together with the notation for the corresponding conditional probabilities of each trio, given that the child is a case

FMC^a	Conditional probability	FMC^a	Conditional probability	FMC^a	Conditional probability
212	s_1	111 ^b	s_6	010	s_{11}
122	s_2	110	s_7	222	s_{12}
211	s_3	101	s_8	201	s_{13}
121	s_4	011	s_9	021	s_{14}
112	s_5	100	s_{10}	000	s_{15}

^a F, M , and C denote the genotypes of the father, mother, and child, respectively.

^b For this cell, the child could have inherited the marker allele M_1 from either parent.

ed summation of the above two differences $w(N_{M<C, C=1} - N_{M>C, C=1}) + (1 - w)(N_{F>C, C=1} - N_{F<C, C=1})$, where $w = n_P/(n_M + n_P)$. Furthermore, from table 1, we can have

$$\begin{aligned} E(N_{M<C, C=1}) &= n_M(s_8 + s_{13}), \\ E(N_{M>C, C=1}) &= n_M(s_4 + s_{14}), \\ E(N_{F>C, C=1}) &= n_P(s_3 + s_{13}), \\ E(N_{F<C, C=1}) &= n_P(s_9 + s_{14}), \end{aligned}$$

and so $E[w(N_{M<C, C=1} - N_{M>C, C=1}) + (1 - w)(N_{F>C, C=1} - N_{F<C, C=1})] = (n_M n_P)/(n_M + n_P)[(s_3 - s_4) + (s_8 - s_9) + 2(s_{13} - s_{14})]$. Under the null hypothesis of no parent-of-origin effects, we have $s_3 = s_4, s_8 = s_9, s_{13} = s_{14}$ [17]. Therefore, $E[w(N_{M<C, C=1} - N_{M>C, C=1}) + (1 - w)(N_{F>C, C=1} - N_{F<C, C=1})] = 0$ under the null hypothesis. Further, $w^2 N_{M \neq C, C=1} + (1 - w)^2 N_{F \neq C, C=1} + (n_M + n_P)^{-1}(N_{M<C, C=1} - N_{M>C, C=1})(N_{F>C, C=1} - N_{F<C, C=1})$ is an unbiased estimator of the variance of $w(N_{M<C, C=1} - N_{M>C, C=1}) + (1 - w)(N_{F>C, C=1} - N_{F<C, C=1})$ under the null hypothesis (see Appendix). So we propose the following 'parental-asymmetric test' when only one parent is available to test for parent-of-origin effects:

1 - PAT =

$$\frac{w(N_{M<C, C=1} - N_{M>C, C=1}) + (1 - w)(N_{F>C, C=1} - N_{F<C, C=1})}{\sqrt{w^2 N_{M \neq C, C=1} + (1 - w)^2 N_{F \neq C, C=1} + (n_M + n_P)^{-1}(N_{M<C, C=1} - N_{M>C, C=1})(N_{F>C, C=1} - N_{F<C, C=1})}} \quad (2)$$

The 1-PAT in equation (2) is asymptotically normally distributed. The region of rejection for testing for parent-of-origin effects is as follows: $|1\text{-PAT}| > z_{\alpha/2}$, where $z_{\alpha/2}$ is the upper $\alpha/2$ point of a standard normal distribution and α is the significance level.

Now we consider the situation where multiple affected children's genotypes are available. Suppose we have n_M incomplete nuclear families in which the fathers are missing and n_P incomplete nuclear families in which the mothers are missing. So we have a total of $n_I = n_M + n_P$ incomplete nuclear families. For each incomplete nuclear family, let every child be paired with the parent and the resulting pair is naturally termed as a case-parent pair.

If there are n_i affected children in the i -th incomplete family, $1 \leq i \leq n_I$, then we have

$$n_{CMP} = \sum_{i=1}^{n_I} n_i$$

case-mother pairs and

$$n_{CFP} = \sum_{i=n_I+1}^{n_M} n_i$$

case-father pairs. Based on these case-parent pairs, the 1-PAT can be extended as follows:

$$1\text{-PAT} = \frac{T}{\sqrt{\text{Var}_0(T)}}, \quad (3)$$

where

$$T = w \sum_M \sum_C (I_{M<C,C=1} - I_{M>C,C=1}) + (1-w) \sum_F \sum_C (I_{F>C,C=1} - I_{F<C,C=1})$$

and

$$\begin{aligned} \text{Var}_0(T) &= w^2 \left[\sum_M \sum_C I_{M \neq C, C=1} + 2 \sum_M \sum_{j < k} (I_{M < C_j, C_j=1} - I_{M > C_j, C_j=1}) \right. \\ &\quad \left. (I_{M < C_k, C_k=1} - I_{M > C_k, C_k=1}) \right] \\ &+ (1-w)^2 \left[\sum_F \sum_C I_{F \neq C, C=1} + 2 \sum_F \sum_{j < k} (I_{F > C_j, C_j=1} - I_{F < C_j, C_j=1}) \right. \\ &\quad \left. (I_{F > C_k, C_k=1} - I_{F < C_k, C_k=1}) \right] \\ &+ \frac{n_{CFP}^2 \sum_{i=1}^{n_I} n_i^2 + n_{CMP}^2 \sum_{i=n_I+1}^{n_M} n_i^2}{n_{CMP} n_{CFP} (n_{CMP} + n_{CFP})^2} \sum_M \sum_C (I_{M < C, C=1} - I_{M > C, C=1}) \\ &\quad \sum_F \sum_C (I_{F > C, C=1} - I_{F < C, C=1}) \end{aligned}$$

is an unbiased estimator of the variance of T under the null hypothesis (see Appendix). The weight is $w = n_{CFP}/(n_{CMP} + n_{CFP})$, and $I_{\text{comparison statement}} = 1$ when the comparison statement holds and is 0 otherwise. In the term $\sum_M \sum_C$, the first summation sums over all mothers and the second summation is for all children with the same mother, and in the term $\sum_M \sum_{j < k}$, the first summation sums over all mothers and the second summation is for all combinations of children C_j and C_k with the same mother. The summations in the terms $\sum_F \sum_C$ and $\sum_F \sum_{j < k}$ are similarly defined. Note that the terms $\sum_M \sum_{j < k}$ and $\sum_F \sum_{j < k}$ on the right side of the above equation take into account the dependency amongst sibship within each family.

Methods Combining Complete and Incomplete Nuclear Families

We first extend the PAT to be suited for the case of the families each with multiple affected children. Suppose F_i and M_i are the

father's and mother's marker genotypes in the i -th family, respectively, $1 \leq i \leq n$. Let n_i denote the number of affected children in the i -th family and C_{ij} be the genotype of the j -th child in the i -th family, $i = 1, 2, \dots, n; j = 1, 2, \dots, n_i$. Under the null hypothesis of no parent-of-origin effects,

$$E \left[\sum_{i=1}^n \sum_{j=1}^{n_i} (I_{F_i > M_i, C_{ij}=1} - I_{F_i < M_i, C_{ij}=1}) \right] = 0.$$

Further,

$$\sum_{i=1}^n \left(\sum_{j=1}^{n_i} I_{F_i \neq M_i, C_{ij}=1} + 2 \sum_{j < k} I_{F_i \neq M_i, C_{ij}=1, C_{ik}=1} \right)$$

is an unbiased estimator of the variance of

$$\sum_{i=1}^n \sum_{j=1}^{n_i} (I_{F_i > M_i, C_{ij}=1} - I_{F_i < M_i, C_{ij}=1})$$

under the null hypothesis (see Appendix), where the summation $\sum_{j < k} I_{F_i \neq M_i, C_{ij}=1, C_{ik}=1}$ denotes the total number of paired children who are heterozygous in the i -th family in which the parents have different marker genotypes. So the corresponding test statistic incorporating multiple affected children can be expressed as follows:

$$\text{PAT} = \frac{\sum_{i=1}^n \sum_{j=1}^{n_i} (I_{F_i > M_i, C_{ij}=1} - I_{F_i < M_i, C_{ij}=1})}{\sqrt{\sum_{i=1}^n \left(\sum_{j=1}^{n_i} I_{F_i \neq M_i, C_{ij}=1} + 2 \sum_{j < k} I_{F_i \neq M_i, C_{ij}=1, C_{ik}=1} \right)}}. \quad (4)$$

It is common in practice to collect complete nuclear families as well as incomplete nuclear families. Suppose we have n complete nuclear families, n_M incomplete nuclear families with missing father and n_P incomplete nuclear families with missing mother, then the following combined test statistic C-PAT of the PAT and 1-PAT is proposed to test for parent-of-origin effects

$$\text{C-PAT} = \frac{\sum_{i=1}^n \sum_{j=1}^{n_i} (I_{F_i > M_i, C_{ij}=1} - I_{F_i < M_i, C_{ij}=1}) + T}{\sqrt{\sum_{i=1}^n \left(\sum_{j=1}^{n_i} I_{F_i \neq M_i, C_{ij}=1} + 2 \sum_{j < k} I_{F_i \neq M_i, C_{ij}=1, C_{ik}=1} \right)} + \text{Var}_0(T)}, \quad (5)$$

where T and $\text{Var}_0(T)$ are defined below equation (3).

Notice that the power of the C-PAT remains unchanged when the two marker alleles M_1 and M_2 are interchanged. As a result, we only need to consider the case where the marker allele frequency is not larger than 0.5 in the simulation study.

Simulation Study

Settings

In this section, we study the size and power of 1-PAT and C-PAT by simulation. The recombination fraction between the DSL and ML is set to be $\theta = 0.001$ in all simulations. We consider the following four parent-of-origin

effects models that represent various scenarios of imprinting degree: complete paternal parent-of-origin effect (PEM1: $\phi_{D/D} = \phi_{d/D} = 0.6$, $\phi_{D/d} = \phi_{d/d} = 0.2$), incomplete paternal parent-of-origin effect (PEM2: $\phi_{D/D} = 0.6$, $\phi_{D/d} = 0.225$, $\phi_{d/D} = 0.575$, $\phi_{d/d} = 0.2$), incomplete maternal parent-of-origin effect (PEM3: $\phi_{D/D} = 0.6$, $\phi_{D/d} = 0.575$, $\phi_{d/D} = 0.225$, $\phi_{d/d} = 0.2$), and complete maternal parent-of-origin effect (PEM4: $\phi_{D/D} = \phi_{D/d} = 0.6$, $\phi_{d/D} = \phi_{d/d} = 0.2$). To gauge the size of a test, we also consider simulations under the null model, a model without parent-of-origin effects (PEM0: $\phi_{D/D} = 0.6$, $\phi_{D/d} = \phi_{d/D} = 0.4$, $\phi_{d/d} = 0.2$).

We define a parameter $\beta = P$ (the missing parent is father|one parent is missing in a family) for convenience, which means that the father is missing with probability β and the mother with $1 - \beta$ for each incomplete family. For ease of reference, we call it missing father rate (among incomplete families) hereafter. In the first simulation, we consider the situation where β ranges from 0.2 to 0.8 in increments of 0.1. The remaining simulations in this section are all performed with $\beta = 0.5$. The incomplete family rate τ is defined as the ratio of the number of families with missing father or mother to that of total families. The incomplete family rate varies within the range 0–1 and in increment of 0.1, unless noted otherwise. Note that C-PAT is reduced to PAT when the incomplete family rate τ is 0 and to 1-PAT when the incomplete family rate τ is 1. We utilize three types of family samples, each with a total of 200 affected children, FS1: 200 families each with one affected offspring, FS2: 100 families each with two affected offspring, and FS3: 100 families each with one affected offspring and 50 families each with two affected offspring.

The population stratification demographic model (PSM) and the assortative mating demographic model (AMM) [20, 21] are used to assess the performance of the C-PAT. In the population stratification model, we consider two subpopulations in the population under study. Let the frequencies of haplotypes M_1D , M_1d , M_2D , and M_2d in the first (second) population be 0.2, 0.15, 0.05, and 0.6 (0.4, 0.1, 0.1, and 0.4), respectively, where $D' = 0.7$ (0.6). A family belongs to the first population with probability 0.7 and to the second population with probability 0.3. For simplicity, the HWE is assumed to hold in each of the two subpopulations. Firstly, we generate the father's and mother's haplotypes at the ML and a DSL for each subpopulation based on those four haplotype frequencies. Then, the haplotypes of the child are generated from the parents' haplotypes with the recombination fraction θ . We assign the affection status of the parents and child

Table 2. Type I error rates (in 100%) of 1-PAT under the null parent-of-origin effects model PEM0

β	Population stratification model		Assortative mating model	
	FS1	FS2	FS1	FS2
0.2	4.72	4.66	3.90	4.11
0.3	4.73	4.61	4.38	4.84
0.4	4.95	4.81	4.44	4.94
0.5	4.83	4.35	4.93	5.23
0.6	5.05	4.83	4.42	4.95
0.7	4.97	4.66	4.66	4.34
0.8	4.70	4.69	4.04	4.29

according to their genotypes at the DSL and the corresponding four risks $\phi_{D/D}$, $\phi_{D/d}$, $\phi_{d/D}$ and $\phi_{d/d}$.

In the assortative mating model, we suppose that a sample is taken from the population where the frequencies of haplotypes M_1D , M_1d , M_2D , and M_2d are 0.2, 0.15, 0.05, 0.6, respectively, which leads to $D' = 0.7$. In this population, 80% of the families were generated through random mating and 20% of the families were generated through assortative mating where the father and mother have the same affection status. We use the same idea as described in the population stratification demographic model to generate family data.

For each set of parameter values, we evaluate the actual size and power by simulation with 10,000 replicates. The actual sizes/powers are estimated as the proportions of replicates of rejecting the null hypothesis at significance level $\alpha = 5\%$ when the null/alternative hypothesis holds.

Sizes and Powers of 1-PAT

Firstly, we investigate the effect of different missing father rates, or the β values, on the performance of the proposed 1-PAT when the sum $n_M + n_P$ is fixed. Simulations under the null hypothesis of no parent-of-origin effects (PEM0) yield the simulated type I error rates of 1-PAT. Table 2 shows the actual sizes of the 1-PAT for two types of samples, FS1 and FS2, under both the population stratification demographic model and assortative mating demographic model. As can be seen from the table, the actual sizes are generally quite close to the nominal 5% level, regardless of the missing father rates, the types of family samples, or the population models.

Figure 1 shows the powers of 1-PAT against β under four parent-of-origin effects models: PEM1–PEM4, two types of family samples: FS1 and FS2, and two population models: PSM and AMM. As expected, the maximum power occurs when the missing rates are the same for fa-

Fig. 1. Powers of 1-PAT against missing father rate β under four parent-of-origin effects models, PEM1–PEM4. **a** FS1 and population stratification model (PSM). **b** FS2 and PSM. **c** FS1 and assortative mating model (AMM). **d** FS2 and AMM. In each plot, the results are based on 10,000 replicates and are assessed at the 5% level.

ther and mother (i.e., $\beta = 0.5$), and the effect of β on the power of 1-PAT can be significant. These observations are consistent across the various parent-of-origin effects models, types of family samples, and population models. Also as expected, there is a greater power for detecting parent-of-origin effects when the parent-of-origin effect is complete (PEM1 and PEM4) than when it is incomplete (PEM2 and PEM3). Furthermore, when the missing father rate β is less than 0.5, power for detecting parent-of-origin effects when the underlying parent-of-origin effects model is paternal is generally greater (albeit only slightly) than when it is maternal with the corresponding degree of imprinting (i.e., PEM1 vs. PEM4 or PEM2 vs.

PEM3). On the other hand, when β is greater than 0.5, the reverse is true. For simplicity, we fix $\beta = 0.5$ for the remaining of the simulations. Although the total numbers of affected offspring are the same in FS1 and FS2, FS2 is much more powerful, indicating synergistic effect of multiple affected children within a family for testing for parent-of-origin effects.

Sizes and Powers of C-PAT

Table 3 gives the actual sizes of C-PAT under the null parent-of-origin effects model PEM0 for family sample types FS1 and FS3 under both the population stratification demographic model and assortative mating demo-

Fig. 2. Powers of C-PAT and PAT under PEM1 and PEM2 plotted against incomplete family rate τ . **a** FS1 and PSM. **b** FS3 and PSM. The results are based on 10,000 replicates and are assessed at the 5% level.

graphic model. The incomplete family rate τ takes values from the interval $[0, 1]$ in increments of 0.1, while the missing father rate within the incomplete families is taken to be fixed at $\beta = 0.5$ as noted above. As can be seen from table 3, the actual type I error rates are all close to the nominal 5% level, signifying the validity of C-PAT as a test for parent-of-origin effects.

For power assessment of the proposed tests and for comparison with the original and extended PAT, we conduct simulations under complete paternal parent-of-origin effect (PEM1) and incomplete paternal parent-of-origin effect (PEM2). The results under the other two models with maternal parent-of-origin effect (PEM3 and PEM4) are similar and are omitted here. In addition to C-PAT that takes the entire sample (complete and incomplete families) into consideration, we also consider the power of the original PAT by using only the complete families (i.e., by removing the incomplete families before the analysis). We consider family sample types FS1 and FS3 under the population stratification demographic model. The results under the assortative mating demographic model are similar and are omitted here. Figure 2 plots the actual power against the incomplete family rate τ . As can be seen from the figure, the power for detecting parent-of-origin effects under PEM1 is higher than under PEM2 for each corresponding setting, which is consistent with the observations made in figure 1. In general, C-PAT

Table 3. Type I error rates (in 100%) of C-PAT under the null parent-of-origin effects model PEM0

Incomplete family rate (in 100%)	Stratification population model		Assortative mating model	
	FS1	FS3	FS1	FS3
0	4.83	5.16	4.79	5.25
10	4.92	5.22	4.53	5.02
20	5.11	5.06	5.16	5.02
30	5.04	4.97	4.85	5.02
40	5.06	5.00	4.88	5.09
50	4.80	5.06	4.84	4.94
60	4.55	4.94	4.91	4.71
70	4.86	5.01	4.67	4.77
80	4.97	5.30	4.83	4.77
90	5.23	5.28	4.91	4.95
100	4.57	4.72	4.49	4.46

by including the incomplete families is more powerful than its counter part PAT based only on complete families. For example, under the FS1 population stratification model (fig. 2a), when the incomplete family rate τ is 30%, the powers of C-PAT and PAT are 68.6 and 59.5%, respectively, under PEM1. The gain in power of C-PAT is about 10% compared to PAT without including the incomplete families. Compared to the setting when all the families

Fig. 3. Powers of C-PAT under PEM2 and the population stratification demographic model. The results are based on 10,000 replicates and are assessed at the 5% level. The powers based on FS1 (200 families with one affected child) are represented by $\square$, the powers based on FS2.1 (100 families with one randomly selected affected child out of two affected children) are represented by $\square$ with dots, and the powers based on FS2 (100 families with two affected offspring) are represented by $\blacksquare$.

are fully genotyped ($\tau = 0$), C-PAT with $\tau = 30\%$ is only 7% less powerful. One peculiar feature exhibited in the two plots of figure 2 (and also in the similar plots under other parameter settings that are omitted here as mentioned above) is that the power of C-PAT is somewhat higher when the incomplete family rate is 100% compared to when it is 90%. This could be because the heterogeneous information based on the combination of both complete and incomplete families ($\tau = 90\%$) may result in a higher variability than the homogeneous information based on only incomplete families ($\tau = 100\%$).

By comparing figure 2a with b, it can be seen that the power of C-PAT (or PAT) based on FS3 is larger than based on FS1. This is consistent with our observation from figure 1 (comparing FS1 vs. FS2) that families with multiple affected children provide more information for detecting parent-of-origin effects when the total numbers of affected children across all families are the same. To further substantiate this observation, we conduct a power study based on three sample types, FS1, FS2, and FS2.1 in which we only randomly select one of the two children from each of the FS2 families in our analysis. Figure 3 displays the powers of C-PAT against the incomplete family rate in increments of 0.5 under the popula-

tion stratification demographic model, the three sample types, FS1, FS2.1, and FS2, and parent-of-origin effects model PEM2. These results show a 17–23% gain in power for C-PAT with FS2 when compared to FS1, confirming the results in figure 1 when incomplete family rate $\tau = 100\%$. The power of C-PAT is also higher based on FS2 than on FS2.1, as anticipated. The results with the assortative mating model are similar, and are omitted.

Discussion

In this paper, we extend the parental-asymmetry test (PAT) to nuclear families with both parents and an arbitrary number of affected children to test for parent-of-origin effects in the presence of association. Our simulation study shows that families with multiple affected children provide more information for detecting parent-of-origin effects compared to families including only one affected child, even when the total numbers of affected children are the same in both samples. This indicates that our extension of PAT to arbitrary nuclear families is a worthwhile exercise. We further develop a PAT-like test (1-PAT) for nuclear families with only one parent genotyped (incomplete families). For a dataset comprised of both complete and incomplete families, we propose a combined PAT, or C-PAT test, that utilizes both the extended PAT and 1-PAT for accommodating both family types. Our simulation results show that C-PAT indeed controls the size well under the null hypothesis of no parent-of-origin effects. We assess the power of C-PAT under four parent-of-origin effects models, three types of family samples, various incomplete family rates and missing father rates, and two population models. We also compare C-PAT with PAT that omits incomplete families in its analysis. Our simulation results show that, even when 100% of the families are incomplete, the power for detecting parent-of-origin effects can still reach 80% of that when all the families have complete data. On the other hand, deleting incomplete families can lead to a tremendous loss of power.

There have been a handful of other tests developed or suggested in the literature, and as such it would be of interest to compare their performances. Based on trio data, in addition to PAT, Weinberg [12] also proposed another test, PO-LRT, to test for parent-of-origin effects. However, PO-LRT is less powerful than PAT when there is no maternal effects and it has also been pointed out that it is difficult to extend PO-LRT to accommodate nuclear families with multiple affected children [10, 13]. On the

Table 4. Powers (in 100%) of PO-LRT and 1-PAT at the nominal $\alpha = 5\%$ level for simulation with 1,000 replicates under the population stratification demographic model of [12] having $\theta = 0$

γ_2	γ_{1p}	γ_{1m}	Sample size	PO-LRT ^a	1-PAT ^b	1-PAT ^c
4	1	4	100	55.2	87.9	61.9
			200	83.1	99.5	95.7
			400	98.6	100	100
5	2	4	100	19.0	44.0	21.0
			200	33.5	68.5	45.8
			400	58.4	94.0	78.4
6	2	6	100	34.8	75.4	62.0
			200	59.6	97.4	87.1
			400	88.9	99.2	99.7

^a Cited from table 5 of [12] based on trio data with both parents.

^b Based on pair data with one parent when $\beta = 0.5$.

^c Based on pair data with one parent when $\beta = 0.8$.

other hand, the Combined LRT is applicable to the case of missing data and could use the unaffected children's information to infer the missing parental genotypes [10]. It extended the PO-LRT by incorporating the estimated missing parental genotypes into analysis. Like PO-LRT, the Combined LRT is also valid in the presence of maternal genotype effects, but it can only accommodate families with one affected child. In contrast, our proposed C-PAT is applicable to a combination of complete and incomplete nuclear families with an arbitrary number of affected children, but it is not valid as a test for parent-of-origin effects when there are maternal effects.

Here we carry out a power comparison of the 1-PAT with PO-LRT, using the population stratification demographic model of [12]: a 20% subpopulation with an allele frequency of 0.3 and $\phi_{d/d} = 0.05$, and the remaining 80% subpopulation with an allele frequency of 0.1 and $\phi_{d/d} = 0.01$, each in HWE. For convenience, let $\gamma_2 = \phi_{D/D}/\phi_{d/d}$, $\gamma_{1p} = \phi_{D/d}/\phi_{d/d}$, and $\gamma_{1m} = \phi_{d/D}/\phi_{d/d}$ denote the genotype relative risks [22], respectively. The relationship between γ_2 , γ_{1p} , γ_{1m} and R_2 , R_1 , I_m of [12] is $\gamma_2 = R_2$, $\gamma_{1p} = R_1$, $\gamma_{1m} = R_1 I_m$. We consider a given sample of families each with both parents and compare the power of the PO-LRT based on this sample with that of the 1-PAT when either father or mother in each family is missing. Table 4 displays the power of the PO-LRT for all trio data with both parents (0% missing rate) and the power of the 1-PAT for all pair data each with only one parent and a single affected child (100% missing rate) when $\beta = 0.5$ or 0.8 based on the sim-

ulation with 1,000 replicates. The power of the PO-LRT in table 4 is cited from table 5 of [12]. Table 4 shows that the power of the 1-PAT based on pair data can be substantially higher than that of the PO-LRT based on full trio data. For example, when $\gamma_2 = 6$, $\gamma_{1p} = 2$, $\gamma_{1m} = 6$, and the sample size is 100, the powers of the 1-PAT are respectively 75.4% when $\beta = 0.5$ and 62.0% when $\beta = 0.8$, which are much larger than the power of the PO-LRT at 34.8%.

In general, the power of a test statistic based on a sample comprising completely genotyped nuclear families and incomplete nuclear families will not exceed that based on a sample of the same size but comprising the fully genotyped nuclear families (i.e., assuming no missing data). Thus, the Combined LRT is less powerful than the PO-LRT. Consequently, the Combined LRT would have less power than the C-PAT using both complete and incomplete nuclear families in testing for parent-of-origin effects.

There is another comparison that we have performed with published methods in the literature. Hu et al. [17] developed the POET and 1-POET tests for complete and incomplete families, respectively, and suggested a combined C-POET test that can utilize both complete and incomplete families in the same sample. To this end, we implemented the C-POET test as suggested. Comparing 1-PAT to 1-POET, and C-PAT to C-POET under a number of settings show that the former is more powerful (results not shown).

Our results for the sizes and the powers of the tests are based on the asymptotic normality assumption, which might be questionable for sample sizes smaller than those presented thus far. To assess the validity of the normality assumption for smaller data sets, we have also carried out additional simulations with sample sizes half of those presented in tables 2 and 3, and we did not find any evidence of notable deviation from the nominal α level of 5% (results not shown). Since most practical studies would have sample sizes at least commensurate with the smallest of those in our simulations, we conclude that using the asymptotic normality assumption is generally reasonable and effective.

From tables 2 and 3, we find that PAT, 1-PAT and C-PAT are all valid when the missingness of parental genotype data is independent of his/her affection status, even though the probability of missingness may be sex dependent. However, affected offspring get the mutant allele more likely from an affected parent. To investigate how sex and affection status dependent willingness of parents to participate in the study may affect the performance of these three tests, we consider the following two sampling scenarios:

Table 5. Type I error rates (in 100%) of PAT, 1-PAT and C-PAT at the nominal $\alpha = 5\%$ level for simulation with 5,000 replicates under the null parent-of-origin effects model PEM0 and the population stratification demographic model as described in the Simulation section, having 400 families with one affected child

Scenario a		Scenario b			
(p_1, p_2, p_3)	PAT	(q_1, q_2, q_3, q_4)	PAT ^a	1-PAT ^b	C-PAT ^c
(0.90, 0.70, 0.95)	5.42	(0.90, 0.70, 1.00, 1.00)	5.48	4.00	5.54
(0.90, 0.70, 0.50)	5.26	(0.90, 0.50, 0.95, 0.80)	5.82	4.08	5.70
(0.90, 0.50, 0.95)	6.74	(0.90, 0.30, 0.95, 0.80)	8.94	4.42	8.58
(0.90, 0.50, 0.45)	7.82	(0.70, 0.50, 0.85, 0.70)	4.86	4.58	5.16
(0.70, 0.50, 0.95)	5.22	(0.70, 0.30, 0.85, 0.70)	6.30	5.02	6.66
(0.70, 0.50, 0.30)	5.66	(0.50, 0.30, 0.90, 0.70)	5.06	4.82	5.66
(0.70, 0.30, 0.95)	7.94	(0.50, 0.30, 0.70, 0.50)	5.40	4.68	5.66
(0.70, 0.30, 0.10)	8.78	(0.30, 0.10, 0.90, 0.70)	5.94	5.48	5.88
(0.50, 0.30, 0.95)	5.70	(0.30, 0.10, 0.70, 0.50)	5.84	5.22	6.22
(0.50, 0.30, 0.10)	5.80	(0.30, 0.10, 0.50, 0.30)	5.82	4.84	6.06

^a Using only the complete families. ^b Using only the incomplete families. ^c Using both the complete and incomplete families.

a) The whole trio tends to be missing when the willingness of an affected parent to participate in the study is sex dependent. Specifically, we use three probabilities to determine whether the whole trio is missing or not: $p_1 = P$ (whole trio participates|father is unaffected and mother is affected), $p_2 = P$ (whole trio participates|father is affected and mother is unaffected) and $p_3 = P$ (whole trio participates|father is affected and mother is affected). We assume that the whole trio will always participate in the study when both parents are unaffected. Note that $p_1 \neq p_2$ signifies sex-dependent missingness.

b) Only parents less willing to participate in the study tend to be missing and such missingness is sex and affection status dependent. Specifically, we use the following four probabilities to determine whether the parent is missing or not: $q_1 = P$ (mother participates|mother is affected), $q_2 = P$ (father participates|father is affected), $q_3 = P$ (mother participates|mother is unaffected), $q_4 = P$ (father participates|father is unaffected). We then utilized these four probabilities to determine whether a family is complete or incomplete and whether the father or mother is missing in an incomplete family, where father's willingness to participate in the study is independent of that of the mother in each family.

Under the population stratification demographic model described in the Simulation section, we evaluate the actual sizes of PAT (only using the complete families), 1-PAT (only using the incomplete families) and C-PAT (using both the complete and incomplete families) under

both scenarios a) and b). Table 5 shows the respective actual type I error rates of PAT, 1-PAT and C-PAT at the nominal 5% level under the null parent-of-origin effects model PEM0. These results are based on 5,000 replicates with 400 families with one affected child in each sample. It is shown in table 5 that 1-PAT controls the size well for different affection status and sex-dependent preference under scenario b). It appears that PAT and C-PAT are also robust to modest departure from random missingness under both scenarios a) and b). However, when the departure from random missingness is extreme, the type I error rates of PAT and C-PAT can be inflated, under both a) and b). For example, when (q_1, q_2, q_3, q_4) takes the value (0.90, 0.30, 0.95, 0.80), the type I error rates of PAT and C-PAT are respectively 8.94 and 8.58%, which deviate greatly from the nominal 5%. On the other hand, the type I error rate of 4.42% for 1-PAT is close to the nominal 5% level.

The above results are not surprising. We know that PAT measures the difference between $N_{F > M, C = 1}$ and $N_{F < M, C = 1}$ (see equation (1)) and tests whether this difference is significantly close to zero. A potential source of bias for PAT is sex-dependent willingness of affected parents to participate in a study, as exemplified in scenarios a) and b), which indeed may lead to spurious parent-of-origin effects, as we see in table 5. Nevertheless, when affection status and sex-dependent preferences are only modest, PAT and C-PAT may still be safely used. On the other hand, for case-parent pairs, 1-PAT compares the

numbers of the two homozygous genotypes within the same sex parents, leading to its robustness against differential participation in the two sexes (difference between q_1 and q_2 and/or difference between q_3 and q_4).

The 1-PAT is developed to test for parent-of-origin effects when association between the marker allele M_1 and the mutant allele D is present. Nevertheless, following the argument in HAP-PAT [13], 1-PAT may also be explored for its potential as a test for association in the presence of parent-of-origin effects. To this end, we evaluate the size of 1-PAT under the null hypothesis of no association under both the population stratification demographic model and assortative mating demographic model. The corresponding type I error rates of the 1-PAT are found to be close to the nominal 5% level (results not shown), which provides preliminary evidence indicating the validity of 1-PAT as a test of association in the presence of parent-of-origin effects. Further investigation is therefore warranted to study its power for such a purpose.

When there are missing genotypes in a pedigree, Ding et al. [23] extended the pedigree disequilibrium test (PDT) [24] to X-chromosomal markers and developed the X-chromosomal Monte Carlo PDT to test for LD. Borrowing the idea from [23], the C-PAT might be extended to the case that genotypes of some individuals in a pedigree are missing, which is our further work of interest.

Acknowledgments

We would like to thank two reviewers for their insightful and helpful suggestions which greatly improved our presentation. This work was partially supported by a Hong Kong RGC CERG Research Grant (HKU 702207P), the National Natural Science Foundation of China (10561008), the Science Foundation of Southeast University (9207011430), the National Institute of Health grant 5R01HG002657, and the Scientific Research Fund of Huaihua University.

Software

A software for the C-PAT is available. Please contact Wing K. Fung via e-mail at wingfung@hku.hk.

Appendix

Unbiased Estimator of Variance of 1-PAT

We consider the situation where each incomplete nuclear family has one affected offspring. We know that

$$\begin{aligned} E(I_{M<C,C=1} - I_{M>C,C=1}) &= s_8 + s_{13} - s_4 - s_{14}, \\ E(I_{F>C,C=1} - I_{F<C,C=1}) &= s_3 + s_{13} - s_9 - s_{14}, \end{aligned}$$

and their sum is zero under the null hypothesis of no parent-of-origin effects. Let

$$\begin{aligned} B &= E(I_{M<C,C=1} - I_{M>C,C=1} \mid \text{no parent-of-origin effects}) \\ &= -E(I_{F>C,C=1} - I_{F<C,C=1} \mid \text{no parent-of-origin effects}), \end{aligned}$$

then under the null hypothesis

$$\begin{aligned} &E[w(N_{M<C,C=1} - N_{M>C,C=1}) + (1-w)(N_{F>C,C=1} - N_{F<C,C=1})]^2 \\ &= E[w^2 N_{M \neq C,C=1} + (1-w)^2 N_{F \neq C,C=1}] \\ &\quad + w^2 n_M(n_M - 1)B^2 + (1-w)^2 n_F(n_F - 1)B^2 - 2w(1-w)n_M n_F B^2 \\ &= E[w^2 N_{M \neq C,C=1} + (1-w)^2 N_{F \neq C,C=1}] - n_M n_F (n_M + n_F)^{-1} B^2 \\ &= E[w^2 N_{M \neq C,C=1} + (1-w)^2 N_{F \neq C,C=1}] \\ &\quad + (n_M + n_F)^{-1} (N_{M<C,C=1} - N_{M>C,C=1})(N_{F>C,C=1} - N_{F<C,C=1}). \end{aligned}$$

So, $w^2 N_{M \neq C,C=1} + (1-w)^2 N_{F \neq C,C=1} + (n_M + n_F)^{-1} (N_{M<C,C=1} - N_{M>C,C=1})(N_{F>C,C=1} - N_{F<C,C=1})$ is an unbiased estimator of the variance of $w(N_{M<C,C=1} - N_{M>C,C=1}) + (1-w)(N_{F>C,C=1} - N_{F<C,C=1})$ under the null hypothesis of no parent-of-origin effects.

When there are multiple affected offspring in a family, the unbiased estimator of the variance under the null hypothesis of no parent-of-origin effects can be similarly obtained.

Unbiased Estimator of Variance of PAT when Multiple Affected Children Are Available

Since, under the null hypothesis of no parent-of-origin effects,

$$\begin{aligned} &E \left[\sum_{i=1}^n \sum_{j=1}^{n_i} (I_{F_i>M_i,C_{ij}=1} - I_{F_i<M_i,C_{ij}=1}) \right]^2 \\ &= \sum_{i=1}^n E \left[\sum_{j=1}^{n_i} (I_{F_i>M_i,C_{ij}=1} - I_{F_i<M_i,C_{ij}=1}) \right]^2 \\ &= \sum_{i=1}^n E \left[\sum_{j=1}^{n_i} (I_{F_i>M_i,C_{ij}=1} - I_{F_i<M_i,C_{ij}=1})^2 \right. \\ &\quad \left. + 2 \sum_{j<k} (I_{F_i>M_i,C_{ij}=1} - I_{F_i<M_i,C_{ij}=1})(I_{F_i>M_i,C_{ik}=1} - I_{F_i<M_i,C_{ik}=1}) \right] \\ &= E \left[\sum_{i=1}^n \sum_{j=1}^{n_i} (I_{F_i \neq M_i,C_{ij}=1} + 2 \sum_{j<k} I_{F_i \neq M_i,C_{ij}=1,C_{ik}=1}) \right], \end{aligned}$$

so

$$\sum_{i=1}^n \left(\sum_{j=1}^{n_i} I_{F_i \neq M_i,C_{ij}=1} + 2 \sum_{j<k} I_{F_i \neq M_i,C_{ij}=1,C_{ik}=1} \right)$$

is an unbiased estimator of the variance of

$$\sum_{i=1}^n \sum_{j=1}^{n_i} (I_{F_i>M_i,C_{ij}=1} - I_{F_i<M_i,C_{ij}=1}).$$

References

- Morison IM, Paton CJ, Cleverley SD: The imprinted gene and parent-of-origin effect database. *Nucleic Acids Res* 2001;29:275–276.
- Falls JG, Pulford DJ, Wylie AA, Jirtle RL: Genomic imprinting: implications for human disease. *Am J Pathol* 1999;154:635–647.
- Ziegler A, König IR: A statistical approach to genetic epidemiology: concepts and applications. Weinheim, Wiley-VCh, 2006.
- Chatkupt S, Lucek PR, Koenigsberger MR, Johnson WG: Parental sex effect in spina bifida: a role for genomic imprinting? *Am J Med Genet* 1992;44:508–512.
- Temple IK, James RS, Crolla JA, Sitch FL, Jacobs PA, Howell WM, Betts P, Baum JD, Shield, JP: An imprinted gene(s) for diabetes? *Nat Genet* 1995;9:110–112.
- van Schothorst EM, Jansen JC, Bardoel AF, van der Mey AG, James MJ, Sobol H, Weissenbach J, Van Ommen GJ, Cornelisse CJ, Devilee P: Confinement of PGL, an imprinted gene causing hereditary paragangliomas, to a 2-cM interval on 11q22-q23 and exclusion of DRD2 and NCAM as candidate genes. *Eur J Hum Genet* 1996;4:267–273.
- Abel KM: Fetal origins of schizophrenia: testable hypotheses of genetic and environmental influences. *Br J Psychiatry* 2004;184:383–385.
- Dong CH, Li WD, Geller F, Lei L, Li D, Grolla OY, Hebebrand J, Amos CI, Nicholls RD, Price RA: Possible genomic imprinting of three human obesity-related genetic loci. *Am J Hum Genet* 2005;76:421–437.
- Samaco RC, Hogart A, LaSalle JM: Epigenetic overlap in autism-spectrum neurodevelopmental disorders: MECP2 deficiency causes reduced expression of UBE3A and GABRB3. *Hum Mol Genet* 2005;14:483–492.
- Rampersaud E, Morris RW, Weinberg CR, Speer MC, Martin ER: Power calculations for likelihood ratio tests for offspring genotype risks, maternal effects, and parent-of-origin (POO) effects in the presence of missing parental genotypes when unaffected siblings are available. *Genet Epidemiol* 2007;31:18–30.
- Weinberg CR, Wilcox AJ, Lie RT: 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* 1998;62:969–978.
- Weinberg CR: Methods for detection of parent-of-origin effects in genetic studies of case-parents triads. *Am J Hum Genet* 1999;65:229–235.
- Becker T, Baur MP, Knapp M: Detection of parent-of-origin effects in nuclear families using haplotype analysis. *Hum Hered* 2006;62:64–76.
- Weinberg CR: Allowing for missing parents in genetic studies of case-parent triads. *Am J Hum Genet* 1999;64:1186–1193.
- Strauch K, Fimmers R, Kurz T, Deichmann KA, Wienker TF, Baur MP: Parametric and nonparametric multipoint linkage analysis with imprinting and two-locus-trait models: Application to mite sensitization. *Am J Hum Genet* 2000;66:1945–1957.
- Lewontin RC: On measures of gametic disequilibrium. *Genetics* 1988;120:849–852.
- Hu YQ, Zhou JY, Sun F, Fung WK: The transmission disequilibrium test and imprinting effects test based on case-parent pairs. *Genet Epidemiol* 2007;31:273–287.
- Allen AS, Rathouz PJ, Satten GA: Informative missingness in genetic association studies: case-parent designs. *Am J Hum Genet* 2003;72:671–680.
- Zhou JY, Hu YQ, Fung WK: A simple method for detection of imprinting effects based on case-parents trios. *Heredity* 2007;98:85–91.
- Sun FZ, Flanders WD, Yang QH, Khoury MJ: Transmission disequilibrium test (TDT) when only one parent is available: the 1-TDT. *Am J Epidemiol* 1999;150:97–104.
- Sun FZ, Flanders WD, Yang QH, Zhao HY: Transmission/disequilibrium tests for quantitative traits. *Ann Hum Genet* 2000;64:555–565.
- Risch N, Merikangas K: The future of genetic studies of complex human diseases. *Science* 1996;273:1516–1517.
- Ding J, Lin S, Liu Y: A Monte Carlo pedigree disequilibrium test for X-chromosome markers. *Am J Hum Genet* 2006;79:567–573.
- Martin ER, Monks SA, Warren LL, Kaplan NL: A test for linkage and association in general pedigrees: The pedigree disequilibrium test. *Am J Hum Genet* 2000;67:146–154.