

Research Article

Detection of Fetomaternal Genotype Associations in Early-Onset Disorders: Evaluation of Different Methods and Their Application to Childhood Leukemia

Jasmine Healy,¹ Mathieu Bourgey,¹ Chantal Richer,¹ Daniel Sinnett,^{1,2}
and Marie-Helene Roy-Gagnon^{1,3}

¹ Sainte-Justine Hospital Research Center, University of Montreal, 3175 Chemin de la Côte-Sainte-Catherine, Room B-467, Montreal, QC, Canada H3T 1C5

² Department of Pediatrics, Faculty of Medicine, University of Montreal, 3175 Chemin de la Côte-Sainte-Catherine, Room 7955, Montreal, QC, Canada H3T 1C5

³ Department of Social and Preventive Medicine, Faculty of Medicine, University of Montreal, P. O. Box 6128, Station Centre-Ville, Montreal, QC, Canada H3C 3J7

Correspondence should be addressed to Daniel Sinnett, daniel.sinnett@umontreal.ca

Received 31 August 2009; Revised 11 November 2009; Accepted 15 March 2010

Academic Editor: Wenjiang J. Fu

Copyright © 2010 Jasmine Healy et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Several designs and analytical approaches have been proposed to dissect offspring from maternal genetic contributions to early-onset diseases. However, lack of parental controls halts the direct verification of the assumption of mating symmetry (MS) required to assess maternally-mediated effects. In this study, we used simulations to investigate the performance of existing methods under mating asymmetry (MA) when parents of controls are missing. Our results show that the log-linear, likelihood-based framework using a case-triad/case-control hybrid design provides valid tests for maternal genetic effects even under MA. Using this approach, we examined fetomaternal associations between 29 SNPs in 12 cell-cycle genes and childhood pre-B acute lymphoblastic leukemia (ALL). We identified putative fetomaternal effects at loci *CDKN2A* rs36228834 ($P = .017$) and *CDKN2B* rs36229158 ($P = .022$) that modulate the risk of childhood ALL. These data further corroborate the importance of the mother's genotype on the susceptibility to early-onset diseases.

1. Introduction

The risk for early-onset disorders can be influenced both by the inherited genotype of the child as well as by parentally-mediated genetic effects [1]. The mother has a crucial role in early-onset disease predisposition as she provides the prenatal environment [2, 3] and can influence her offspring's risk of disease not only as a genetic donor but also through the effects of her genes acting directly on the intrauterine milieu or indirectly through fetomaternal gene-gene interactions [1, 4–6]. Given the important role the mother can play in shaping disease susceptibility in her offspring, focusing solely on the genotype of the child in association testing could, in certain instances, be misleading.

Several family-based tests have been proposed to dissect offspring and maternal genetic contributions to early-onset disorders including the case-parent designs of Wilcox et al. [1] and Weinberg et al. [3] using a log-linear framework, and of Cordell et al. [7, 8], which uses a conditional logistic regression framework. Designs using alternative family structures have also been suggested: the hybrid design based on augmenting a set of case-parent trios with a set of parents of unrelated controls [9], the “pent” design consisting of the affected child, mother, father, and maternal grandparents [2], the case-mother/control-mother dyad design [10] or the design consisting of case-parent triads supplemented by control-mother dyads [11]. However, the use of these alternative designs can be limited by the difficulty of obtaining grandparental data or sufficiently

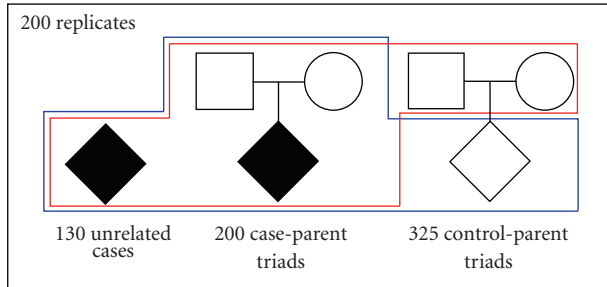


FIGURE 1: Cohort data used for simulations. 200 replicate datasets were generated for each model. For all of our calculations we used the same overall number of affected individuals and case-parents however the methods differed in how the case and control data were utilized in each individual test (see Materials and Methods). The same number of affected individuals ($n = 330$) and case-parents ($n = 400$) was used under all three methods. For the combined case-control and conditional logistic regression analysis (CC+CLR) and the log-linear analysis using unrelated controls rather than their parents (HD-NPC), we used the genotypes of unrelated controls ($n = 325$, in blue), whereas for the hybrid design using parents of controls in a log-linear framework (HD), we incorporated the genotypic information of the parents of these controls rather than the controls themselves ($n = 650$, in red).

large samples of parents of unaffected children, and by the increase in costs incurred for genotyping these additional sets of individuals. On the other hand, in many ongoing genetic association studies it often occurs that parents of cases and unrelated control individuals are collected in parallel, in which case designs based solely on case-triad data could lead to considerable losses in power if unrelated case-control data are disregarded. In the event that both case-parent triads as well as case-control data are ascertained, additional designs/analytical strategies are needed for fetomaternal genotype association testing.

The first aim of this study was to evaluate the adaptability of existing methods to deal with mixed data sets consisting of both case-parent triads and case-control data. We used simulations to investigate the validity and power of (1) Weinberg and Umbach's hybrid design [9] treating parents of controls as missing (HD-NPC) and (2) a classic case-control test in conjunction with Cordell et al.'s conditional logistic regression method [7] (CC+CLR), to distinguish between offspring and maternal genetic contributions to disease. Given that Cordell et al.'s approach relies solely on the use of case-parent triads, combining it with a case-control test will allow us to maximize the use of available genotype information. Finally, we compared both these approaches to a third hypothetical, ideal situation, in which genotype data from parents of controls would be available and Weinberg and Umbach's hybrid design could be used (HD) as described in [9].

However lack of parents of controls in HD-NPC and CC+CLR precludes the direct verification of the assumption of mating symmetry (MS) that is required to assess maternally-mediated effects. Mating symmetry (MS) refers

to the hypothesis that for a parental genotype pair, the frequency in the population for a given mother-father genotype assignment is the same as for the reverse father-mother assignment [1, 12]. Departures from this symmetry could lead to genotype frequency differences among males and females mating in the population which, in the context of fetomaternal association testing, could lead to confounding and spurious maternal associations. As such, the HD approach, using an auxiliary sample of parents of controls to obtain direct information on mating frequencies [9], is the only method that allows for the assumption of symmetry to be directly tested and readily accommodated. It is unclear however how this method performs when parents of controls are missing and MA is present. And since parents of unaffected controls are not available in most ongoing studies, it is important to assess the robustness of these fetomaternal association tests under such circumstances. Therefore, we evaluated type I error rates and power of the three methods (HD, HD-NPC, and CC+CLR) under varying degrees of MA, and genotypic risk models involving child, mother or both child and mother jointly, in order to identify the analytical approach that is most reliable for dissecting child and maternal genetic contributions to early-onset diseases in the absence of parents of controls.

Another aim of this study was to use these methods to investigate fetomaternal associations in a real mixed dataset of childhood pre-B acute lymphoblastic leukemia (ALL) patients. ALL is a hematological malignancy resulting from chromosomal alterations and mutations that affect molecular pathways that disrupt lymphoid progenitor cell differentiation [13, 14]. There is well-established evidence for prenatal initiation of the leukemogenesis process in children [15–18]. Moreover, parental exposures to environmental carcinogens or use of medication have been identified as potential risk factors for childhood leukemia [19–23] and transplacental carcinogen exposure has been involved in the development of certain subtypes of ALL [24]. Although the risk of leukemia from environmental exposures in utero or in early childhood is likely to be influenced by genetic variation at both the level of the child and the mother, the role of maternally-mediated genetic effects in childhood leukemia susceptibility remains undefined. Here we performed a candidate gene association study using both ALL case-parent triads and unrelated controls to assess the impact of 29 SNPs from 12 cell-cycle genes in both mother and child on childhood pre-B ALL risk.

2. Materials and Methods

2.1. Methods Used to Test for Fetomaternal Genotype Associations. We compared three analytical approaches for the detection of early-onset disease associations. In the event that both case-parent triads and unrelated case-control data are available, we tested (1) a combined method in which a case-control genotypic test was carried out in conjunction with the conditional logistic regression test of Cordell et al. to detect associations at the level of the child and mother, respectively (CC+CLR); and (2) the log-linear, likelihood-based approach of Weinberg and Umbach [9]

TABLE 1: Forward stepwise likelihood-ratio testing procedure used to dissect child and maternal genotype associations.

Weinberg and Umbach's log-linear approach using case-triads, unrelated cases and unrelated controls (HD-NPC) or parents of controls (HD)	Genotypic case-control test combined with the conditional logistic regression approach of Cordell et al. using case-triads (CC+CLR)
<i>Step 1.</i> Child genotypic effect Null versus CG LR Chi-square (2df) And Maternal genotypic effect Null versus MG LR Chi-square (2df)	<i>Step 1.</i> Child genotypic association test (CC) Chi-square or Fisher's exact test (2df) And Maternal genotypic effect (CLR) Null versus MG LR Chi-square (2df)
<i>Step 2. (depending on Step 1)</i> Maternal effect given child effect CG versus CG+MG LR Chi-square (2df) Or Child effect given maternal effect MG versus CG+MG LR Chi-square (2df)	<i>Step 2. (depending on Step 1)</i> Maternal effect given child effect (CLR) CG versus CG+MG LR Chi-square (2df) Or Child effect given maternal effect (CLR) MG versus CG+MG LR Chi-square (2df)

Likelihood-ratio tests were performed in a forward stepwise fashion. The most significant single-step test (Child versus Null or Mother versus Null) was tested against a joint effects model in a 2 degree-of-freedom likelihood-ratio test (Child + Mother versus Child or Child + Mother versus Mother). LR Chi2 indicates likelihood-ratio chi-square test; df, degrees of freedom; GC Child genotype relative risk; GM, Mother genotype relative risk.

TABLE 2: The eight simulation models used for evaluation of the fetomaternal association tests.

Model	Child effect	Mother effect	MAF	ΔC
1	—	—	0.05 to 0.25	MS
2	GC ₁₁ = 1; GC ₁₂ = 2; GC ₂₂ = 3	—	0.05 to 0.25	MS
3	—	GM ₁₁ = 1; GM ₁₂ = 2; GM ₂₂ = 3	0.05 to 0.25	MS
4	GC ₁₁ = 1; GC ₁₂ = 2; GC ₂₂ = 3	GM ₁₁ = 1; GM ₁₂ = 2; GM ₂₂ = 3	0.05 to 0.25	MS
5	—	—	0.3	0 to 1
6	GC ₁₁ = 1; GC ₁₂ = 2; GC ₂₂ = 3	—	0.3	0 to 1
7	—	GM ₁₁ = 1; GM ₁₂ = 2; GM ₂₂ = 3	0.3	0 to 1
8	GC ₁₁ = 1; GC ₁₂ = 2; GC ₂₂ = 3	GM ₁₁ = 1; GM ₁₂ = 2; GM ₂₂ = 3	0.3	0 to 1

(–): indicates a null risk model where the genotype relative risks (GRRs) are $GRR_{11} = GRR_{12} = GRR_{22} = 1$. GC indicates Child genotype relative risk; GM, Mother genotype relative risk; MAF: Minor allele frequency; ΔC : mating-pair disequilibrium. MS indicates mating symmetry where $\Delta C = 0$.

using an additional set of unrelated cases and unrelated controls as proxies for parental control genotype information (HD-NPC). We also compared these two approaches to (3) one in which parents of controls are also available and therefore the hybrid design (case-parent/parents of controls) can be used through log-linear, likelihood-based analysis (HD).

It should be noted that the combined CC+CLR approach is not a modification of the conditional logistic regression approach of Cordell et al. but rather an adaptation in its use to detect fetomaternal associations. Cordell et al.'s approach relies solely on the use of case-parent triads. Since disregarding any available unrelated case-control genotype data reduces power, we used Pearson's chi-square tests or Fisher's exact tests (CC), as appropriate, and conditional logistic regression (CLR) in parallel on partially overlapping data. The former were used on all available cases to identify genotype associations at the level of the child whereas the latter was used on case-parent triads to identify maternally-contributed effects. Results from the two tests were not combined, rather if a significant association was found in the

child CC test then the CLR test was used (albeit on a reduced case set) in order to dissect offspring and maternal effects. Similarly, if a significant result was found for the mother test using CLR, then CLR was further used to distinguish a main effect of the mother from a joint fetomaternal effect (see likelihood-ratio testing below).

2.2. Likelihood-Ratio Testing to Dissect Child from Maternal Genetic Effects. Given that the offspring will be enriched for the risk allele by simple Mendelian inheritance [25], it is important to discriminate between direct effects of a maternal genotype or of a child genotype from a joint fetomaternal effect. To do so, we used a forward stepwise likelihood-ratio testing procedure. In the first step, we performed two single-step tests to investigate associations at the level of the child and mother separately (Table 1). For CC+CLR, Pearson's chi-square or Fisher's exact tests were performed in R (version 2.6.2), to compare genotype distributions in cases (unrelated and triad cases) versus controls. In parallel, we used case-parent triads and logistic

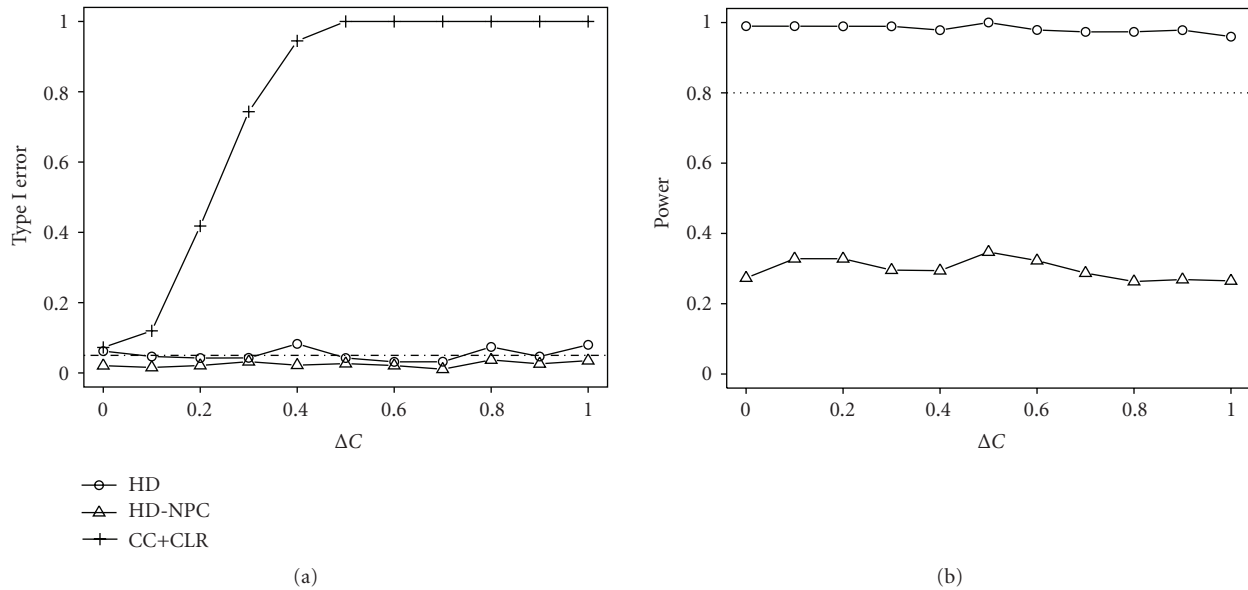


FIGURE 2: Type I error rates and power for the maternal association test under mating asymmetry. (a) Type I error rates are given for the Mother versus Null test as a function of departure from mating symmetry, as measured by ΔC , under the null model where $GC_{11} = GC_{12} = GC_{22} = GM_{11} = GM_{12} = GM_{22} = 1$ (Table 2, Model 5). Allele prevalence is set at $q = 0.3$. (b) Power to detect a maternal effect is shown as a function of departure from mating symmetry, as measured by ΔC , for a scenario with multiplicative effects of the mother (Table 2, Model 7). Allele prevalence is set at $q = 0.3$. MA models (9 mating-type parameters) were used for log-linear regression under the assumption of mating asymmetry. Horizontal reference lines indicate type I error rate of $\alpha = 0.05$ (a) and power = 0.8 (b). HD: hybrid design using parents of controls in a log-linear framework; HD-NPC: log-linear analysis using unrelated controls rather than their parents; CC+CLR: combined case-control and conditional logistic regression analysis.

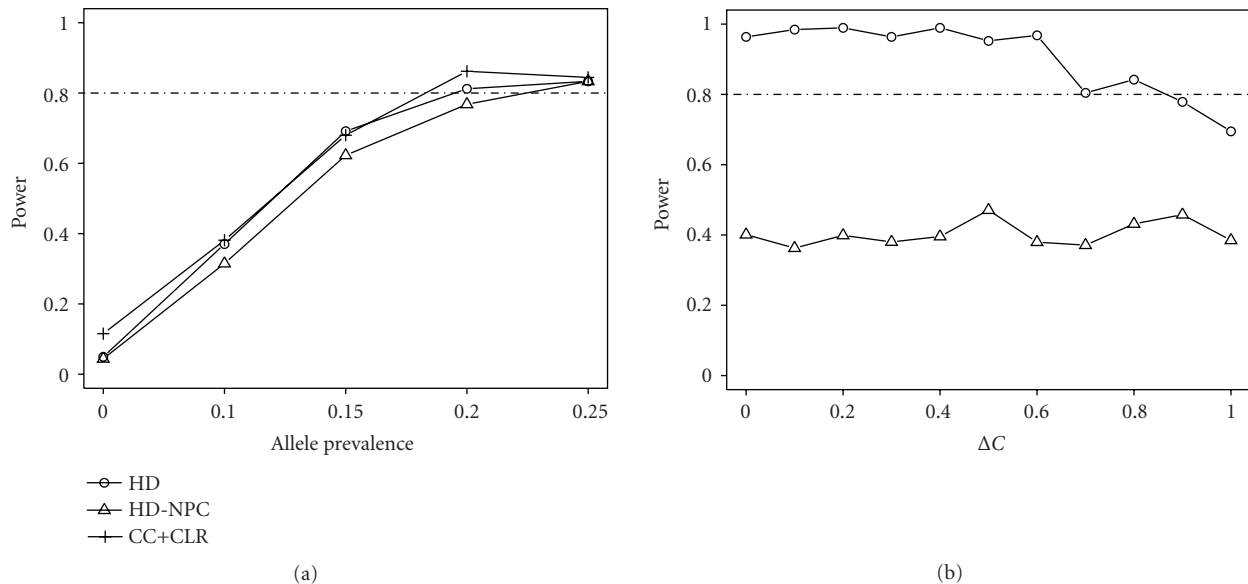


FIGURE 3: Power of the forward stepwise procedure to detect joint fetomaternal associations. (a) Power for the HD, HD-NPC and CC+CLR methods is shown as a function of allele prevalence for a scenario with mating symmetry and multiplicative effects of both Child and Mother (Table 2, Model 4). (b) Power for HD and HD-NPC is shown as a function of departure from mating symmetry, as measured by ΔC , for a scenario with mating asymmetry and multiplicative effects of both Child and Mother (Table 2, Model 8). Allele prevalence is set at $q = 0.3$. MA models (9 mating-type parameters) were used for log-linear regression under the assumption of mating asymmetry. Horizontal reference lines indicate power = 0.8. HD, hybrid design using parents of controls in a log-linear framework; HD-NPC: log-linear analysis using unrelated controls rather than their parents; CC+CLR: combined case-control and conditional logistic regression analysis.

TABLE 3: Parental genotype distributions under mating symmetry and mating asymmetry.

Mating type	Parental genotypes (Mother-Father)	Expected distribution	
		Mating symmetry	Mating asymmetry
0	11-11	$N * \mu_0$	$N * \mu_0$
1	11-12	$N * \mu_1$	$(2 - C_1) * N * \mu_1$
	12-11	$N * \mu_1$	$C_1 * N * \mu_1$
2	11-22	$N * \mu_2$	$(2 - C_2) * N * \mu_2$
	22-11	$N * \mu_2$	$C_2 * N * \mu_2$
3	12-12	$N * \mu_3$	$N * \mu_3$
4	12-22	$N * \mu_4$	$(2 - C_4) * N * \mu_4$
	22-12	$N * \mu_4$	$C_4 * N * \mu_4$
5	22-22	$N * \mu_5$	$N * \mu_5$

Alleles for a biallelic locus are denoted 1 and 2 and the corresponding genotypes 11, 12 or 22. N indicates the number of individuals in the sample; μ_i the i th mating type probability under the assumption of random mating; C_i the mating-pair disequilibrium for the i th parental couple. C is a multiplicative factor between 0 and 2 that describes the over representation (>1), under representation (<1) or symmetry ($=1$) of a mate-pair combination in the corresponding i th mating type.

regression conditioning on exchangeable parental genotypes (CEPG) [7, 8], using the DGCgenetics package for R developed by D. Clayton (available at http://www-gene.cimr.cam.ac.uk/clayton/software/DGCgenetics_1.0.tar.gz) to test for maternally-mediated associations through a two degree-of-freedom likelihood-ratio test (Mother versus Null). For HD-NPC and HD, log-linear regression analysis was performed using the LEM software (Log-linear and event history analysis with missing data using the EM algorithm) [26] and LEM script files provided by Weinberg et al. (available at <http://www.niehs.nih.gov/research/resources/software/hybrid/index.cfm>) [9, 26]. Two degree-of-freedom likelihood-ratio tests were performed to independently test for offspring (Child versus Null) and maternal (Mother versus Null) associations in the log-linear framework.

In the second step, if a significant effect was found in the first step (based on a Bonferroni correction of $P < .025$ for the 2 tests performed), then the most significant model was tested against a joint effects model (Child + Mother versus Child or Child + Mother versus Mother) in a two degree-of-freedom likelihood-ratio test (Table 1). For CC+CLR this meant that conditional logistic regression was used on the reduced case-triad data set to test either Child + Mother versus Child or Child + Mother versus Mother depending on the results from step one. Whereas log-linear regression analysis was performed on case-triads and unrelated cases and controls in HD-NPC and on case-triads, unrelated cases and parents of controls in HD.

2.3. Simulated Data. We simulated cohort data under several conditions to model different genotypic risk effects involving either child, mother or both child and mother (joint fetomaternal effects) (Table 2). Under each model 200 replicate datasets were simulated using the R software. Replicates

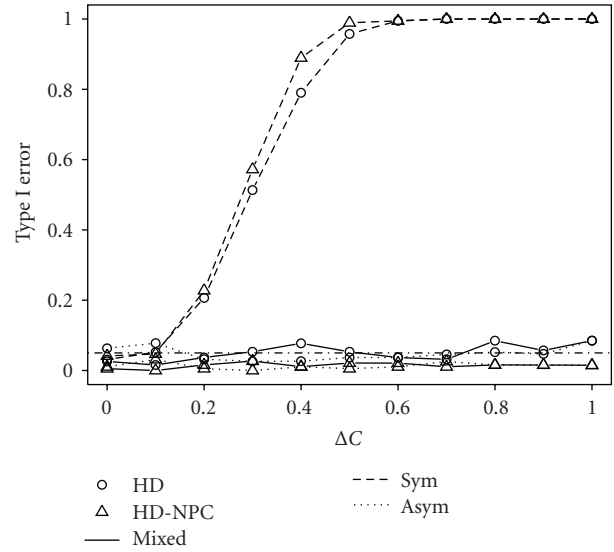


FIGURE 4: Type I error rate for the HD and HD-NPC approaches for the maternal association test assuming either mating symmetry or asymmetry. Type I error rates are shown as a function of departure from mating symmetry, as measured by ΔC , under the null model where $GC_{11} = GC_{12} = GC_{22} = GM_{11} = GM_{12} = GM_{22} = 1$ (Table 2, Model 5). Allele prevalence is set at $q = 0.3$. Log-linear, likelihood-based testing (Mother versus Null) was performed assuming either mating asymmetry estimating nine mating-type parameters in the likelihood-ratio tests (Asym), or mating symmetry using only six mating-type parameters in the likelihood-ratio tests (Sym), or with no a priori hypothesis, using a 3 degree-of-freedom likelihood-ratio test to first test the assumption of mating symmetry (Mating Asymmetry (9df) versus Mating Symmetry (6df)) and then using either the mating asymmetry models when asymmetry was detected or the mating symmetry models when it was not (Mixed). Horizontal reference line indicates type I error rate of $\alpha = 0.05$. HD, hybrid design using parents of controls in a log-linear framework; HD-NPC, log-linear analysis using unrelated controls rather than their parents.

that did not fit the simulated models were not included in the analysis. To imitate our childhood leukemia cohort, we simulated genotypic data for 200 case-parent triads, 130 unrelated cases and 325 unrelated control-parent triads for each replicate. For all of our calculations we used the same overall number of affected individuals however the methods differed in how the case and control data were utilized in each individual test (Figure 1). For all of our child-based tests we used the same number of affected individuals and unrelated controls (Figure 1): 330 cases (200 case-triads and 130 unrelated cases) were compared to 325 unrelated controls. However the number of cases and population-based controls used for maternal and fetomaternal association testing varied depending on the method: CC+CLR used the genotypes from 200 case-triads only; HD-NPC used the genotypes of 330 cases (200 case-triads and 130 unrelated cases) and of 325 unrelated controls; and HD incorporated the genotypic information of the parents ($n = 650$) rather than the unrelated controls themselves (Figure 1).

In HD and HD-NPC, unrelated cases and controls were used by treating their parental genotypes as missing and the

EM algorithm implemented in the LEM software was used to infer missing genotype information [26]. Unrelated cases and cases belonging to triads were considered to have similar penetrance and thus similar genotypic relative risks (GRR). We assumed Hardy-Weinberg equilibrium and the absence of population stratification in the form of admixture. We evaluated the different approaches in terms of type I error rate and power to detect associations by counting the number of replicates found to be significantly associated over the total number of replicates that fit the specified model.

2.4. Simulated Scenarios. We first assessed the behaviour of the three methods in scenarios in which MS was assumed across parents. Under MS, HD-NPC and HD were performed by forcing six mating-type variables in the log-linear model [9].

A second set of simulations was performed in which we assessed the performance of the tests assuming differences in genotype frequencies between males and females mating in the population, that is, assuming varying levels of mating asymmetry. MA was evaluated in terms of the degree of departure from the expected mate-pair probability under symmetry. We used C_i to denote the disequilibrium for the i th parental mating type. C_i is a multiplicative factor between 0 and 2 that describes the over representation ($C > 1$), under representation ($C < 1$) or symmetry ($C = 1$) of a mate-pair combination in the corresponding i th mating type. The level of departure from MS is denoted as ΔC , a numerical value ranging from -1 to 1 , with $C = 1 + \Delta C$. The expected parental genotype distributions under the assumption of MS and MA are shown in Table 3. By varying the departure from mating symmetry, ΔC , we introduced varying levels of asymmetry into our simulations (Table 2). In this study we assumed a model for MA where $C_1 = C_2$ and $C_4 = 1$. Supplementary Figure 1, mentioned in Supplementary Material that is available online at doi: 10.1155/2010/369534, shows how departures from symmetry, as measured by ΔC , translate into differences in reciprocal mating types and overall genotype frequencies between males and females mating in the population. Under MA, HD-NPC and HD were performed by forcing nine mating-type variables in the log-linear model [9] and we tested for the presence of asymmetry by comparing the nine mating-type model to the six mating-type model with a three degree-of-freedom likelihood-ratio test using a cut-off of $P < .05$ to declare deviation from symmetry. A less conservative P value $< .10$ was also used (data not shown) and yielded similar results.

2.5. Childhood Acute Lymphoblastic Leukemia Data

2.5.1. Study Subjects. We investigated fetomaternal associations in a pre-B acute lymphoblastic leukemia (ALL) cohort. The study population has been previously described [27, 28]. Briefly, incident cases of childhood pre-B ALL ($n = 321$) were diagnosed in the Division of Hematology-Oncology of the Sainte-Justine Hospital in Montreal, Canada, between October 1985 and November 2006. Our cohort includes 189 boys and 132 girls with a median age of 4.7 years,

all French-Canadian from the province of Quebec, Canada. Parental DNA was available for 203 of the probands. Healthy controls ($n = 329$) consisted of French-Canadian individuals recruited while using clinical departments other than Hematology-Oncology of the Sainte-Justine Hospital.

2.5.2. SNPs, Genotyping and Quality Control Checks. We selected 29 SNPs from 12 candidate cell-cycle genes for the analysis (Table 4). Genes were selected based on their function in regulating the G1/S cell-cycle checkpoint. Regulatory SNPs (found to lie within the proximal promoter region) were chosen based on the hypothesis that variation in gene dosage of such critical cell-cycle genes due to functional regulatory polymorphisms could influence cancer susceptibility by altering cell homeostasis [28]. For the purpose of this study using a French-Canadian cohort, we considered European-specific SNPs previously identified in [29]. DNA was isolated from buccal epithelial cells, peripheral blood or bone marrow in remission as previously described [30]. SNPs were genotyped using the Luminex xMAP/Autoplex Analyser CS1000 system (Perkin Elmer, Waltham, MA). Genetic variants were amplified using allele-specific primer extension in multiplexed assays and hybridized to Luminex MicroPlex TM -xTAG Microspheres as per Koo et al. [31]. Primer sequences, amplification conditions, and reaction conditions are available upon request. Genotypes were called using the Automatic Luminex Genotyping (ALG) software [32]. Three negative controls and three sample duplicates were used on each 96-well DNA plate. The average genotype call rate was 99.8% and rates of discordance were below 3.3%. In addition, Hardy-Weinberg equilibrium was tested using the X^2 goodness of fit test and PedCheck (Version 1.1) was used to identify genotype incompatibilities using the familial data [33]; inconsistent case-parent trios were removed from the analyses. Multiple testing corrections were performed on the single-step association tests using the Benjamini-Hochberg false discovery rate (FDR) method with a type I error rate of 10%; nominal P values are shown.

3. Results and Discussion

Although there are currently no data to document the frequency of events that lead to mating distortions in human populations, it is biologically plausible that MA might commonly occur. It is known that assortative mating (selection of a mate on the basis of phenotype leading to correlation between phenotypes of mated individuals with respect to a given trait) can lead to genotype frequency differences between males and females [34–36]. Other mechanisms leading to mating asymmetry however are unclear and their evolutionary consequences much less understood. If, for a specific marker, MA results in a departure from Hardy-Weinberg equilibrium, this genetic marker would be excluded from an association study following quality control. However MA could also arise through mating selection but with discrimination acting oppositely in each sex, or through differential individual mating success for the genotypes of each sex. Both these processes could lead to genotype

TABLE 4: Genes and DNA variants genotyped in the pre-B ALL association study.

Gene (Chromosome)	DNA variant	Position	MAF
CCND1 (11q13)	rs1944129	69,163,116	0.4876
	rs36225395	69,163,517	0.4523
CDC25a (3p21)	rs1903061	48,206,923	0.1028
CDKN1A (6p21)	rs733590	36,753,181	0.3616
	rs762624	36,753,566	0.2714
	rs2395655	36,753,674	0.3968
CDKN1B (12p13)	rs3759217	12,759,719	0.1159
	rs35756741	12,759,968	0.0865
	rs36228499	12,761,203	0.4342
CDKN2A (9p21)	rs36228834	21,965,319	0.0512
CDKN2B (9p21)	rs36229158	22,000,681	0.0282
	rs2069416	22,000,004	0.3742/0.0271
	rs2069418	21,999,698	0.4272
E2F1 (20q11)	rs3213141	31,738,041	0.2405
HDAC1 (1p35)	rs1741981	32,529,026	0.3302
	rs36212121	32,529,102	0.0031
	rs36212119	32,529,840	0.0846
MADH3 (15q22)	rs36221701	65,143,543	0.1199
	rs36222034	65,144,732	0.1111
	rs11633026	65,144,812	0.1235
MDM2 (12q15)	rs1144944	67,486,752	0.4954
	rs3730485	67,487,073–67,487,112	0.4052
	rs937282	67,488,064	0.483
	rs2279744	67,488,847	0.3662
RB1 (13q14)	rs1573601	47,774,358	0.2484
TGFB1 (19q13)	rs2317130	46,553,514	0.3141
	rs4803457	46,553,199	0.3937
	rs11466313	46,553,177–46,553,178	0.3096
	rs1800469	46,552,136	0.3127

DNA variant positions relative to dbSNP build 130. MAF indicates minor allele frequency and was calculated on a control cohort consisting of 329 healthy individuals of European descent.

frequency differences between sexes that would not lead to detectable deviations from Hardy-Weinberg equilibrium but that may incur important biases in fetomaternal association testing. In addition to these biological causes, low levels of MA could arise in a study sample simply due to the sampling process. Although parents of controls allow direct testing

for bias due to MA in fetomaternal genotype association testing [9], these samples are difficult to collect and a method that can combine case-control and family-based data and provide a valid analytical framework for fetomaternal association testing in the presence of MA is currently not available.

3.1. Simulation Study. In this study we used simulations to investigate the ability of three fetomaternal genotype association tests: (1) the log-linear, likelihood-based method of Weinberg and Umbach [9] using a case-parent/case-control design (HD-NPC), (2) the conditional logistic regression approach of Cordell et al. [7] combined with a case-control test (CC+CLR), and (3) Weinberg and Umbach's hybrid design using log-linear regression analysis, to distinguish between offspring and maternal genetic contributions to disease in the presence of MA. It should be noted that since both the log-linear and CLR frameworks are equivalent in terms of power and type I error for the detection of maternal genotype effects [7], our conclusions regarding the CC+CLR approach would also apply to a case-control combined with the log-linear linear framework of Wilcox et al. [1] and Weinberg et al. [3] using case-parent triads. We evaluated type I error rates and power of the methods under varying degrees of MA, and genotypic risk models involving child, mother or joint effects of both child and mother (Table 2). For clarity reasons, we present the results for multiplicative genotype effect models only. However similar results were obtained under dominant and recessive models, with recessive models yielding expected decreases in power across all methods, particularly at low allele frequencies.

As expected, all three methods showed similar low type I error rates, around 5% and similar power, above 80% for the detection of child effects ($GC_{11} = 1$; $GC_{12} = 2$; $GC_{22} = 3$) under MS as well as under MA (data not shown). For a maternal main effect ($GM_{11} = 1$; $GM_{12} = 2$; $GM_{22} = 3$), type I error and power under MS were also within the acceptable ranges (data not shown). By contrast, under MA the CC+CLR method yielded unacceptably high type I error rates for the Mother test (Figure 2(a)). Although we expected that the method developed by Cordell et al. would be susceptible to the confounding incurred by MA, we found that CLR does not withstand even low levels of asymmetry ($\Delta C \sim 0.1$) so that even weak assumptions concerning population distributions of parental genotypes could lead to important bias.

The validity of the Mother tests for HD and HD-NPC were unaffected by MA, with type I error rates below the 5% threshold (Figure 2(a)). Power of the HD design was maintained at 100% and was unaffected by MA under the simulation conditions considered here, whereas power of HD-NPC considerably dropped, averaging around 30% (Figure 2(b)). When HD-NPC is used under asymmetry, genotypes for the parents of controls are inferred based on mating-type frequencies estimated from the parents of the cases and based on the assumptions that the control offspring genotypes follow Mendelian proportions in relation to their parents [9]. Hence, the maternal effect present in the case

triads is partially captured in the inference of the mating-type frequencies for the parents of controls, resulting in a loss of power to detect this maternal effect as it becomes confounded with the estimated asymmetry.

The stepwise procedure allows maternal and case effects to be distinguished by estimating maternal effects independently of offspring effects and provides a valid test for joint fetomaternal associations. Under the null model and MS, the stepwise likelihood-ratio testing procedure yielded type I error rates close to 0% (data not shown), most likely due to the over-conservative Bonferroni correction that was applied. Since the Mother and Child tests are not completely independent a permutation test procedure would provide less-conservative type I error estimates. When we modelled multiplicative effects of both offspring and maternal effects ($GC_{11} = GM_{11} = 2$ and $GC_{12} = GM_{12} = 3$), the power to detect fetomaternal associations using the forward stepwise procedure was comparable for all three methods and increased with increasing allele prevalence for each method, reaching 80% for MAFs > 0.20 (Figure 3(a)).

We then evaluated type I error rates and power for the stepwise procedure in the presence of MA. The performance of both HD and HD-NPC was unaffected by MA with type I error rates close to zero even in the presence of high levels of asymmetry (Supplementary Figure 2). The type I error rate of the CC + CLR approach was close to 5% (Supplementary Figure 2) given that both child and maternal effects had to be falsely detected in order for the replicate to be counted as a false-positive and the case-control component of the test was robust against spurious child associations. Because the CLR maternal test is not valid under even low levels of MA, we assessed the power of the other two approaches to detect fetomaternal associations for varying levels of MA (Figure 3(b)). HD performed significantly better than HD-NPC: sensitivity of the HD design averaged around 100% whereas power of HD-NPC was again significantly lower, averaging around 40%.

Based on the above findings, no method seemed to provide a net advantage under MS for these simulation conditions; nor did we observe any significant loss of power or robustness when the conditional logistic regression or the log-linear, likelihood-based approaches were used without parents of controls. Although power was significantly reduced, the log-linear, likelihood-based approach using controls rather than parents had little effect on the specificity of the association tests in the presence of MA. However, for the analyses performed on MA simulated datasets, we forced the estimation of nine mating-type parameters. In practice, no a priori assumptions regarding MS could be made. To verify the robustness of both methods for the detection of maternal (and fetomaternal) genotype effects in the presence of asymmetry, we measured type I error rates for the Mother test for scenarios in which either MA or MS models are assumed, and for a scenario in which no a priori hypothesis is made but rather MS is first evaluated in a three degree-of-freedom likelihood ratio test and the appropriate models (MA or MS) are subsequently used for association testing. These results show that if one assumes MS and this assumption is violated, type I error rates for

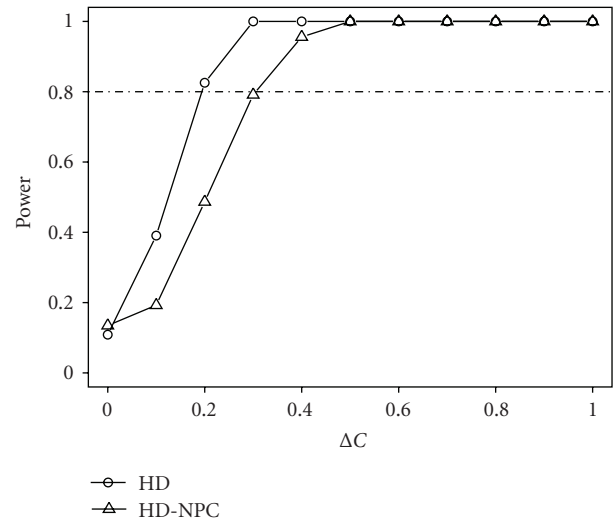


FIGURE 5: Power of HD and HD-NPC to detect mating asymmetry. Power is shown as a function of departure from mating symmetry, as measured by ΔC , under the null model where $GC_{11} = GC_{12} = GC_{22} = GM_{11} = GM_{12} = GM_{22} = 1$ (Table 2, Model 5). Allele prevalence is set at $q = 0.3$. Mating asymmetry was evaluated in the log-linear framework using a 3 degree-of-freedom likelihood-ratio test comparing the 9 mating-type parameter model under MA to a 6 mating-type parameter model under MS. Horizontal reference line indicates power = 0.8. HD: hybrid design using parents of controls in a log-linear framework; HD-NPC: log-linear analysis using unrelated controls rather than their parents.

the maternal test for both HD-NPC and HD are significantly inflated (Figure 4). However, first testing for asymmetry and then adjusting the association analyses accordingly provides accurate type I error rates for both methods. Similar results were obtained for the forward stepwise procedure (data not shown).

The specificity of the HD-NPC test therefore relies on its ability to detect MA and then use mating-type models accounting for asymmetry to test for association. Given that there are no biological references for the amount of MA that occurs in human populations, we evaluated the capacity of the HD and HD-NPC methods to detect various levels of MA (Figure 5). Our simulation results showed that the power of HD-NPC to detect asymmetry above $\Delta C = 0.4$ was comparable to that of HD. For a risk allele frequency of $q = 0.3$, HD and HD-NPC reached 80% power at $\Delta C = 0.25$ and $\Delta C = 0.35$, respectively. Under low levels (ΔC ranging from 0 to 0.2) the sensitivity of both methods to detect asymmetry was threatened, especially for the HD-NPC approach (Figure 5). On the other hand, the lack of power of HD-NPC and HD to detect low levels of asymmetry is compensated by the fact that, without any a priori assumptions regarding mating symmetry, both methods maintained low type I error rates, at least under the simulation conditions presented here (Figure 4).

Therefore if asymmetry is not strong enough to be detected by the MA test it should not be falsely interpreted

TABLE 5: Distribution of *CDKN2A* rs36228834 and *CDKN2B* rs36229158 genotypes and associated risk estimates for pre-B ALL susceptibility among children.

Gene, DNA variant, and genotype	No. (%)			Controls	Model	Log-linear regression analysis		
	ALL patients	ALL mothers	ALL fathers			Genotype	Child OR (95% CI)	<i>P</i>
<i>CDKN2A</i>								
rs36228834					Child versus Null	TA versus TT	2.48 (1.45–4.15)	.001
TT	266 (86.6)	160 (93.0)	149 (86.6)	298 (93.7)		AA versus TT	9.87 (0.89–109.69)	
TA	39 (12.7)	12 (7.0)	22 (12.8)	19 (6.0)		TA/AA versus TT	2.56 (1.54–4.26)	<.0005
AA	2 (0.7)	0	1 (0.6)	1 (0.3)	Child + Mother versus Null	TA versus TT	3.13 (1.81–5.40)	<.0005
						AA versus TT	—	
						TA/AA versus TT	2.56 (1.54–4.26)	<.0005
<i>CDKN2B</i>								
rs36229158					Child versus Null	CT versus CC	1.77 (0.98–3.21)	.054
CC	277 (91.4)	164 (95.4)	155 (90.1)	302 (94.7)		TT versus CC	8.25 (0.75–91.3)	
CT	24 (7.9)	8 (4.6)	16 (9.3)	16 (5.0)		CT/TT versus CC	1.86 (1.04–3.34)	.037
TT	2 (0.7)	0	1 (0.6)	1 (0.3)	Child + Mother versus Null	CT versus CC	2.32 (1.23–4.35)	.033
						TT versus CC	—	
						CT/TT versus CC	2.44 (1.29–4.60)	.006

Percentages indicate number of individuals with a given genotype/total number of genotyped individuals. Risk estimation was performed using log-linear regression analysis as implemented in the LEM software. Child odd ratios were measured using regression models consisting of the child genotype effect only (Child versus Null) or both child and mother genotypes (Child + Mother versus Null). Mating symmetry (i.e., six mating-type parameters) was assumed at both loci. *P* values of the Wald test provided by LEM are shown for either the 2 degree-of-freedom (2 child genotype effects) or 1 degree-of-freedom (1 child genotype effect resulting from the collapsed heterozygous/homozygous rare genotypes) tests. OR indicates odds ratio; CI: confidence interval.

as a maternally-mediated effect. By contrast if a maternal effect is present and HD-NPC is used to test for mating asymmetry without parents of controls, type I error rates (of false detection of MA) are high (ranging from 0.8 to 0.9 for MAFs of 0.10 to 0.25, resp.) leading to a subsequent loss of power to detect maternal effects due to the over-parameterisation under the MA models (data not shown). Together these results show that the log-linear, likelihood-based stepwise procedure using unaffected offspring provides a valid framework to evaluate MS without leading to spurious maternal associations. And when parents of controls cannot be ascertained but an additional set of unrelated controls is available, one can safely use this approach to test for fetomaternal associations if willing to accept that certain confounded maternally-mediated effects may be missed when parental mating is asymmetric.

3.2. Fetomaternal Association Study of Childhood Acute Lymphoblastic Leukemia. Guided by our simulation results we went on to test for fetomaternal associations between 29 SNPs in the proximal promoter regions of 12 cell-cycle genes [29] and the susceptibility to childhood pre-B ALL. SNP frequencies were in agreement with those previously reported in other populations of European descent and all distributions were in Hardy-Weinberg equilibrium. Our dataset consisted of 118 pre-B ALL patients, 203 ALL case-parent triads, and 329 unrelated controls. The lack of parents

of controls prevents us from excluding MA in the source population. Based on the results from our simulation study, we used the log-linear framework to perform likelihood-based testing in a stepwise fashion. For each SNP, we performed a three degree-of-freedom likelihood-ratio test for asymmetry implemented in the LEM software using a slightly less stringent *P* value <.10 to reject symmetry in order to reduce false-positives in the tests for maternal effects. Under this threshold, we identified MA at variants rs1144944 (*P* = .086) and rs3730485 (*P* = .095) of the *MDM2* gene, as well as at *CDKN2B* variant rs2069416 (*P* = .076); we did not detect asymmetry at any of the remaining loci tested (*P* values >.10; data not shown). Consequently, MA models (nine mating-type parameters) were used to test for association at these three SNPs whereas MS models (six mating-type parameters) were used for the remaining 26 SNPs (see Figure 6, for the Child and Mother single-step test results and Supplementary Table 1 for complete likelihood-ratio chi-square test results). Nominally significant genotype associations at the level of the child were identified for *CDKN2A* rs36228834 (Child versus Null; *P* = .0007), *CDKN1B* rs35756741 (Child versus Null; *P* = .0235) and *CDKN2B* rs2069416 (Child versus Null; *P* = 0.0063); however only *CDKN2A* rs36228834 and *CDKN2B* rs2069416 remained significant after multiple testing corrections (Supplementary Table 1). None of the other 26 loci revealed any significant child-mediated genetic associations with ALL and no significant maternal

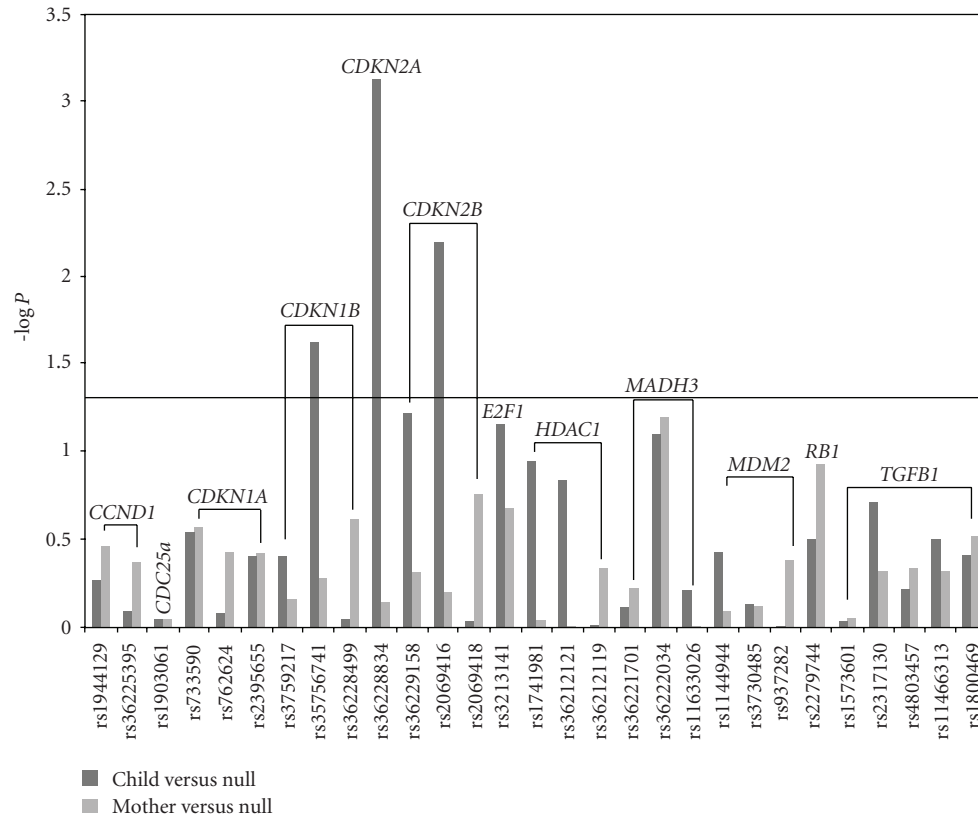


FIGURE 6: Log-linear, likelihood-ratio association analysis between 29 regulatory SNPs from 12 cell-cycle genes and childhood pre-B acute lymphoblastic leukemia (ALL). Log-linear regression analysis was performed in LEM using 203 case-triads, 118 unrelated ALL patients and 329 controls. Results of the likelihood-ratio chi-square tests ($-\log P$) are shown for the single-step Child (blue) and Mother (red) tests. Mating symmetry (i.e., six mating-type parameters) was assumed at all loci but variants rs1144944 and rs3730485 of the *MDM2* gene, as well as *CDKN2B* rs2069416 for which MA models (nine mating-type parameters) were used to test for association. Horizontal reference line indicates P value of .05. Note that for the tri-allelic variant *CDKN2B* rs2069416 ($A > T, C$) individuals were grouped according to their T allele such that $** > *T > TT$. See Supplementary Table 1 for complete likelihood-ratio chi-square test results.

genotype effects were identified through the Mother versus Null log-linear test (Figure 6 and Supplementary Table 1).

Nominally significant SNPs were further analyzed to detect putative joint fetomaternal effects. After accounting for the effect of the child's genotype, we found a significant maternal association at *CDKN2A* rs36228834 (Child + Mother versus Child; $P = .0168$) only (Supplementary Table 1). The *CDKN2A* rs36228834 A allele was overrepresented in patients when compared with controls (genotype Fisher's exact $P = .005$) and carriers of the A risk allele were 2.5-fold more susceptible to ALL (Child versus Null; TA versus TT: OR = 2.48; 95%CI (1.45–4.15); TA/AA versus TT: OR = 2.56; 95%CI (1.54–4.26)). This risk was further increased when the mother's genotype was included in the model (Child + Mother versus Null; TA versus TT: OR = 3.13; 95%CI (1.81–5.40); TA/AA versus TT: OR = 3.20; 95%CI (1.85–5.53)) (Table 5). No further maternal association was detected for *CDKN1B* rs35756741 (Child + Mother versus Child; $P = .59$) or *CDKN2B* rs2069416 (Child + Mother versus Child; $P = .33$) (Supplementary Table 1).

These results provide evidence of a novel fetomaternal effect at the *CDKN2A* rs36228834 locus that may influence pre-B ALL susceptibility among children, and a possible joint effect of both mother and child genotypes without main effects at *CDKN2B* rs36229158. Interestingly, although variant *CDKN2B* rs36229158 presented only a suggestive effect at the level of the child (Child versus Null; $P = .06$), a significant association was found for the mother after we accounted for the genotype of the child (Child + Mother versus Child; $P = .0217$) (Supplementary Table 1). Including the mother's genotype in the regression model significantly increased the risk 2.3-fold for carriers of a C allele (Child + Mother versus Null; CT versus CC: OR = 2.32; 95%CI (1.23–4.35); CT versus CC: OR = 2.44; 95%CI (1.29–4.60)) (Table 5).

Independent replication is required in order to confirm the significance of these associations given that some of these variants did not withstand multiple testing correction (*CDKN1B* rs35756741 and *CDKN2B* rs36229158); and caution is warranted in the interpretation of the risk estimates as risk allele frequencies at loci *CDKN2A* rs36228834 and

CDKN2B rs36229158 were low (MAFs ≤ 0.05), yielding large confidence intervals, particularly for the rare homozygous genotype classes. We also recognize that the interpretation of our real-data results relies in part on our simulation results and there are certain limitations to our simulation study. These include the restricted number of models used in the method evaluations and the important assumptions of absence of population substructure and Hardy-Weinberg equilibrium, which necessitate further investigation. Our conclusions on the validity of the HD-NPC approach under MA should hold for other models. As we noted, the detection of asymmetry and maternal effects are partially confounded with HD-NPC and it is difficult to imagine a model of asymmetry for which differences in mating type frequencies would not be detected by the asymmetry test but would be captured by the Mother versus Null test, thus leading to increased type I error. However, the reduction in power of HD-NPC compared to HD will likely be affected by the underlying genetic and asymmetry models and should be assessed under a wider range of models. Other important genetic effects should also be investigated, such as mother-gene child-gene interactions and parent-of-origin effects which are not addressed here but can also be involved in early-onset disorder risk.

Nonetheless, our results provide evidence that genes that regulate the cell cycle could play an important role during fetal development when the rate of cell growth and division is high both in child and mother. *In silico* analysis using the Match software [37] revealed that all three variant loci lead to the disruption of putative transcription factor binding sites, including the loss of binding sites for FOS and MYB at *CDKN2A* rs36228834. The FOS oncoprotein stimulates transcription of genes containing AP-1 regulatory elements and may transform cells through alterations in DNA methylation and in histone deacetylation [38]. Expression of FOS is 100-fold greater in human fetal membranes than in other normal human tissues and cells [39]. The MYB transcription factor is essential for hematopoiesis and controls the proliferation and differentiation of hematopoietic stem and progenitor cells [40]. MYB is frequently involved in hematopoietic disorders including ALL [41]. Although the biological relevance of our findings remains to be elucidated, our study suggests that promoter variation in the cell-cycle inhibitor gene *CDKN2A*, and possibly *CDKN2B*, could disrupt transcription-factor binding and influence gene expression during gestation. Disregulated cell division caused by aberrant cell-cycle inhibitor gene expression in both mother and child could disrupt the maternal-fetal interface and affect important physiological processes such as the growth of the fetus and/or normal haematopoiesis and potentially lead to increased susceptibility to ALL.

4. Conclusions

Given the unique nature of childhood disorders, the investigation of parental genetics and maternally-contributed effects is a prerequisite not only for understanding disease etiology but also to pave the way toward new opportunities in preventive medicine. Although the most powerful approach

is ideally the most desirable, in practice the best approach might be one that combines both valid detection of the possible underlying genetic associations involved in early-onset disorders and a feasible design in terms of ascertainment and genotyping costs. We have shown that the log-linear, likelihood-based framework using a case-triad/case-control design retains the ability to control for bias due to MA and can provide valid tests for maternally-contributed genotype effects even when the assumption of symmetry fails. Despite a modest sample size, we successfully used this approach to identify putative fetomaternal genotype effects in cell-cycle inhibitor genes *CDKN2A* and *CDKN2B* that are associated with modified risks of childhood pre-B ALL. Although these genes have been previously associated with ALL [28], we have shown for the first time that their influence on ALL risk might be driven, in part, by the maternal genotype. This study provides the first indication that maternal genotype effects can influence the risk of developing pediatric ALL, providing useful insights into the genetic mechanisms underlying this early-onset disease.

Acknowledgments

The authors are indebted to all the patients and their parents who consented to participate in this study. They thank Ekaterini Kritikou for critically reviewing and editing the manuscript. This study was supported by research funds provided by the Canadian Institutes for Health Research. J. Healy is the recipient of a NSERC Canada Graduate's scholarship. M. Bourgey is the recipient of a fellowship from the Cole Foundation. D. Sinnett holds the François-Karl Viau Chair in Pediatric Oncogenomics and is a scholar of the Fonds de la Recherche en Santé du Québec. J. Healy and M. Bourgey contributed equally to the work presented in this paper.

References

- [1] A. J. Wilcox, C. R. Weinberg, and R. T. Lie, "Distinguishing the effects of maternal and offspring genes through studies of 'case-parent triads,'" *American Journal of Epidemiology*, vol. 148, no. 9, pp. 893–901, 1998.
- [2] L. E. Mitchell and C. R. Weinberg, "Evaluation of offspring and maternal genetic effects on disease risk using a family-based approach: the "pent" design," *American Journal of Epidemiology*, vol. 162, no. 7, pp. 676–685, 2005.
- [3] C. R. Weinberg, A. J. Wilcox, and R. T. Lie, "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," *American Journal of Human Genetics*, vol. 62, no. 4, pp. 969–978, 1998.
- [4] J. S. Sinsheimer, C. G. S. Palmer, and J. A. Woodward, "Detecting genotype combinations that increase risk for disease: the maternal-fetal genotype incompatibility test," *Genetic Epidemiology*, vol. 24, no. 1, pp. 1–13, 2003.
- [5] C. R. Weinberg, "Studying parents and grandparents to assess genetic contributions to early-onset disease," *American Journal of Human Genetics*, vol. 72, no. 2, pp. 438–447, 2003.
- [6] C. G. S. Palmer, J. A. Turunen, J. S. Sinsheimer, et al., "RHD maternal-fetal genotype incompatibility increases

- schizophrenia susceptibility,” *American Journal of Human Genetics*, vol. 71, no. 6, pp. 1312–1319, 2002.
- [7] H. J. Cordell, B. J. Barratt, and D. G. Clayton, “Case/pseudocontrol analysis in genetic association studies: a unified framework for detection of genotype and haplotype associations, gene-gene and gene-environment interactions, and parent-of-origin effects,” *Genetic Epidemiology*, vol. 26, no. 3, pp. 167–185, 2004.
- [8] H. J. Cordell and D. G. Clayton, “A unified stepwise regression procedure for evaluating the relative effects of polymorphisms within a gene using case/control or family data: application to HLA in type 1 diabetes,” *American Journal of Human Genetics*, vol. 70, no. 1, pp. 124–141, 2002.
- [9] C. R. Weinberg and D. M. Umbach, “A hybrid design for studying genetic influences on risk of diseases with onset early in life,” *American Journal of Human Genetics*, vol. 77, no. 4, pp. 627–636, 2005.
- [10] M. Shi, D. M. Umbach, S. H. Vermeulen, and C. R. Weinberg, “Making the most of case-mother/control-mother studies,” *American Journal of Epidemiology*, vol. 168, no. 5, pp. 541–547, 2008.
- [11] S. H. Vermeulen, M. Shi, C. R. Weinberg, and D. M. Umbach, “A hybrid design: case-parent triads supplemented by control-mother dyads,” *Genetic Epidemiology*, vol. 33, no. 2, pp. 136–144, 2009.
- [12] D. J. Schaid and S. S. Sommer, “Genotype relative risks: methods for design and analysis of candidate-gene association studies,” *American Journal of Human Genetics*, vol. 53, no. 5, pp. 1114–1126, 1993.
- [13] M. Greaves, “Molecular genetics, natural history and the demise of childhood leukaemia,” *European Journal of Cancer*, vol. 35, no. 14, pp. 1941–1953, 1999.
- [14] J. D. Rowley, “Molecular genetics in acute leukemia,” *Leukemia*, vol. 14, no. 3, pp. 513–517, 2000.
- [15] J. L. Wiemels, G. Cazzaniga, M. Daniotti, et al., “Prenatal origin of acute lymphoblastic leukaemia in children,” *The Lancet*, vol. 354, no. 9189, pp. 1499–1503, 1999.
- [16] C. M. McHale, J. L. Wiemels, L. Zhang, et al., “Prenatal origin of TEL-AML1-positive acute lymphoblastic leukemia in children born in California,” *Genes Chromosomes and Cancer*, vol. 37, no. 1, pp. 36–43, 2003.
- [17] M. Greaves, “Infection, immune responses and the aetiology of childhood leukaemia,” *Nature Reviews Cancer*, vol. 6, no. 3, pp. 193–203, 2006.
- [18] M. F. Greaves, “Aetiology of acute leukaemia,” *The Lancet*, vol. 349, no. 9048, pp. 344–349, 1997.
- [19] J. S. Chang, S. Selvin, C. Metayer, V. Crouse, A. Golembesky, and P. A. Buffler, “Parental smoking and the risk of childhood leukemia,” *American Journal of Epidemiology*, vol. 163, no. 12, pp. 1091–1100, 2006.
- [20] W. Wen, X. O. Shu, J. D. Potter, et al., “Parental medication use and risk of childhood acute lymphoblastic leukemia,” *Cancer*, vol. 95, no. 8, pp. 1786–1794, 2002.
- [21] C. Infante-Rivard, J. Siemiatycki, R. Lakhani, and L. Nadon, “Maternal exposure to occupational solvents and childhood leukemia,” *Environmental Health Perspectives*, vol. 113, no. 6, pp. 787–792, 2005.
- [22] C. Infante-Rivard and D. Sinnett, “Preconceptional paternal exposure to pesticides and increased risk of childhood leukaemia,” *The Lancet*, vol. 354, no. 9192, p. 1819, 1999.
- [23] M. L. Kwan, C. Metayer, V. Crouse, and P. A. Buffler, “Maternal illness and drug/medication use during the period surrounding pregnancy and risk of childhood leukemia among offspring,” *American Journal of Epidemiology*, vol. 165, no. 1, pp. 27–35, 2007.
- [24] M. S. Pombo-de-Oliveira, S. Koifman, P. I. C. Araújo, et al., “Infant acute leukemia and maternal exposures during pregnancy,” *Cancer Epidemiology Biomarkers and Prevention*, vol. 15, no. 12, pp. 2336–2341, 2006.
- [25] S. Buyske, “Maternal genotype effects can alias case genotype effects in case-control studies,” *European Journal of Human Genetics*, vol. 16, no. 7, pp. 783–785, 2008.
- [26] E. J. C. G. van den Oord and J. K. Vermunt, “Testing for linkage disequilibrium, maternal effects, and imprinting with (in)complete case-parent triads, by use of the computer program LEM,” *American Journal of Human Genetics*, vol. 66, no. 1, pp. 335–338, 2000.
- [27] M. Krajcinovic, D. Labuda, C. Richer, S. Karimi, and D. Sinnett, “Susceptibility to childhood acute lymphoblastic leukemia: influence of CYP1A1, CYP2D6, GSTM1, and GSTT1 genetic polymorphisms,” *Blood*, vol. 93, no. 5, pp. 1496–1501, 1999.
- [28] J. Healy, H. Belanger, P. Beaulieu, M. Lariviere, D. Labuda, and D. Sinnett, “Promoter SNPs in G1/S checkpoint regulators and their impact on the susceptibility to childhood leukemia,” *Blood*, vol. 109, no. 2, pp. 683–692, 2007.
- [29] D. Sinnett, P. Beaulieu, H. Belanger, et al., “Detection and characterization of DNA variants in the promoter regions of hundreds of human disease candidate genes,” *Genomics*, vol. 87, no. 6, pp. 704–710, 2006.
- [30] A. Baccichet, S. K. Qualman, and D. Sinnett, “Allelic loss in childhood acute lymphoblastic leukemia,” *Leukemia Research*, vol. 21, no. 9, pp. 817–823, 1997.
- [31] S. H. Koo, T. C. Ong, K. T. Chong, C. G. L. Lee, F. T. Chew, and E. J. D. Lee, “Multiplexed genotyping of ABC transporter polymorphisms with the Bioplex suspension array,” *Biological Procedures Online*, vol. 9, no. 1, pp. 27–42, 2007.
- [32] M. Bourgey, M. Larivière, C. Richer, and D. Sinnett, “Genotype calling of luminex arrays,” submitted.
- [33] J. R. O’Connell and D. E. Weeks, “PedCheck: a program for identification of genotype incompatibilities in linkage analysis,” *American Journal of Human Genetics*, vol. 63, no. 1, pp. 259–266, 1998.
- [34] A. S. Kondrashov and M. Shpak, “On the origin of species by means of assortative mating,” *Proceedings of the Royal Society B*, vol. 265, no. 1412, pp. 2273–2278, 1998.
- [35] M. Kirkpatrick and V. Ravigné, “Speciation by natural and sexual selection: models and experiments,” *American Naturalist*, vol. 159, supplement 3, pp. S22–S35, 2002.
- [36] M. A. R. de Cara, N. H. Barton, and M. Kirkpatrick, “A model for the evolution of assortative mating,” *American Naturalist*, vol. 171, no. 5, pp. 580–596, 2008.
- [37] A. E. Kel, E. Gossling, I. Reuter, E. Cheremushkin, O. V. Kel-Margoulis, and E. Wingender, “MATCH: a tool for searching transcription factor binding sites in DNA sequences,” *Nucleic Acids Research*, vol. 31, no. 13, pp. 3576–3579, 2003.
- [38] A. V. Bakin and T. Curran, “Role of DNA 5-methylcytosine transferase in cell transformation by fos,” *Science*, vol. 283, no. 5400, pp. 387–390, 1999.
- [39] R. Muller, I. M. Verma, and E. D. Adamson, “Expression of c-onc genes: c-fos transcripts accumulate to high levels during development of mouse placenta, yolk sac and amnion,” *The EMBO Journal*, vol. 2, no. 5, pp. 679–684, 1983.

- [40] M. L. Sandberg, S. E. Sutton, M. T. Pletcher, et al., “c-Myb and p300 regulate hematopoietic stem cell proliferation and differentiation,” *Developmental Cell*, vol. 8, no. 2, pp. 153–166, 2005.
- [41] E. Clappier, W. Cuccuini, A. Kalota, et al., “The C-MYB locus is involved in chromosomal translocation and genomic duplications in human T-cell acute leukemia (T-ALL), the translocation defining a new T-ALL subtype in very young children,” *Blood*, vol. 110, no. 4, pp. 1251–1261, 2007.