

Report

Parental Genotypes in the Risk of a Complex Disease

Damian Labuda,^{1,2} Maja Krajinovic,^{1,2} Audrey Sabbagh,¹ Claire Infante-Rivard,^{1,3} and Daniel Sinnett^{1,2}

¹Centre de recherche, Hôpital Sainte Justine, ²Département de pédiatrie, Université de Montréal, and ³Department of Epidemiology and Biostatistics and Occupational Health, McGill University, Montreal

Our understanding of the genetic etiology of complex disorders is still elusive. According to the common-variant/common-disease hypothesis, frequent functional polymorphisms are the best candidates for disease-susceptibility alleles. Implicitly, we also assume that disease-susceptibility alleles are preferentially transmitted from parents to the affected offspring and that this effect can be captured by the transmission/disequilibrium test (TDT). However, our study of genetic predisposition to childhood acute lymphoblastic leukemia suggests that a focus on the patient's genotype might, in certain instances, be misleading. Our results indicate that, at least at some loci, parental genetics might be of primary importance in predicting the risk of cancer in this pediatric model of a complex disease. Consequently, in addition to TDT, other complementary strategies will need to be simultaneously applied to dissect genetic predisposition to complex disorders.

The great progress made toward an understanding of Mendelian inheritance in humans has, to a large extent, been due to the characterization of molecular origins of monogenic diseases. Mutations underlying these conditions, dominant or recessive, are typically rare and usually highly penetrant. In contrast, our understanding of the genetic etiology of complex disorders, such as diabetes, cancer, cardiovascular diseases, or asthma, is limited. According to the common-disease/common-variant hypothesis (Cargill et al. 1999; Chakravarti 1999; Reich and Lander 2001), frequent functional polymorphisms, rather than rare mutations, appear to be the best candidates for disease-susceptibility alleles. This is consistent with the role of APOE $\epsilon 4$ allele in Alzheimer disease (Corder et al. 1993), factor V^{Leiden} in deep venous thrombosis (Bertina et al. 1994), or PPAR γ Pro12Ala polymorphism in type II diabetes (Altshuler et al. 2000), to name a few examples. It is also generally assumed that the environmental context influences the phenotypic expression of these genetic variants. The resultant dis-

ease would thus originate through the collective contribution of different genes, environmental factors, and their interactions. Studies of linkage and association between disease and marker alleles are intended to identify loci affecting the disease. The transmission/disequilibrium test (TDT) (Spielman et al. 1993) is considered to be a powerful test in the mapping of complex diseases (Risch and Merikangas 1996). However, additional strategies should be considered when one is tracing the genetic etiology of complex diseases (Bell and Taylor 1997; Long et al. 1997; Muller-Myhsok and Abel 1997; Scott et al. 1997; Haines and Pericak-Vance 1998; Pericak-Vance 1998; Speer 1998; Terwilliger and Goring 2000).

It is plausible that, in the genetic association studies of common genetic diseases, one should consider, in addition to the patients' genetics, the parental genetic context. For example, given the correlation between birth weight, influenced by the mother's genetics, and susceptibility to a number of common diseases occurring during adult life (Ferguson et al. 2000; Godfrey and Barker 2001; Morley and Dwyer 2001), one can hypothesize that the mother's genotype at the particular loci contributes to and even predetermines the offspring's health. In the case of cancer, maternal exposure during pregnancy may influence disease risk during childhood (Infante-Rivard et al. 1999), as well as during adulthood (Ekbom 1998), especially given that, com-

Received March 1, 2002; accepted for publication April 19, 2002; electronically published June 7, 2002.

Address for correspondence and reprints: Dr. Damian Labuda, Centre de recherche, Hôpital Sainte-Justine, 3175 Côte-Sainte-Catherine, Montréal, Québec, H3T 1C5 Canada. E-mail: damian.labuda@umontreal.ca

© 2002 by The American Society of Human Genetics. All rights reserved. 0002-9297/2002/7101-0023\$15.00

pared to adults, the fetus appears to be more vulnerable to genetic damage by carcinogens (Perera 1997). Fetal effects resulting from maternal exposure might be affected by the efficiency of the mother's detoxification systems and, thus, by her genotype. The consequences can remain "imprinted," as acquired susceptibility, even if a child has not inherited the disease-susceptibility allele itself. By extending this concept to a prezygotic stage, at the level of gametes, one could hypothesize that the paternal genetics may play a similar role (Infante-Rivard and Sinnott 1999). In males, for example, there are a greater number of cell divisions during spermatogenesis than during oogenesis. This leads to a relative excess of mutations resulting naturally from DNA replication errors, which explains the higher risk of mutations that is associated with increased paternal age (Crow 2000). The number of mutations is increased because of carcinogenic exposure. This effect might be enhanced further by the presence of genetic variants in components of DNA repair that control both replication errors and the level of "environmental" DNA damage (Miller et al. 2001).

We hypothesized, therefore, that, in certain instances—and especially in the presence of exposure and genetic polymorphisms—susceptibility to a disease could be a consequence of a patient's mother's and/or father's genotypes, rather than of the patient's own combination of alleles; in other words, if the parental genotypes, rather than the offspring's genotypes, were responsible for genetic susceptibility to a disease, analysis such as TDT (Spielman et al. 1993) would fail to detect the disease-susceptibility allele(s). Figure 1 illustrates the possible outcomes of TDT and case-control analysis, given that either the patient's or the parents' genotypes are associated with susceptibility to a disease. In TDT, a significant deviation from random transmission of alleles from heterozygous parents to their offspring can be interpreted as association and linkage (fig. 1A). At the same time, an appropriate case-control study with sufficient power would detect an association, because the frequency of the tested allele significantly differs between patients and control individuals. On the other hand, let us consider a situation in which only the parents' (either mother's or father's or both) genotypes account for the child's susceptibility. The population of parents is then expected to differ from the control population. Parent-to-child transmissions could be random, leading to a negative TDT result (fig. 1B); however, in spite of a random parent-to-patient transmission, the patients and control individuals would differ as well. When parents are themselves enriched in disease-susceptibility alleles, even if they transmit randomly, their offspring also will have more disease-susceptibility alleles than are found in control individuals. If TDT analysis is not included, then, when the scenario of figure 1B applies, the

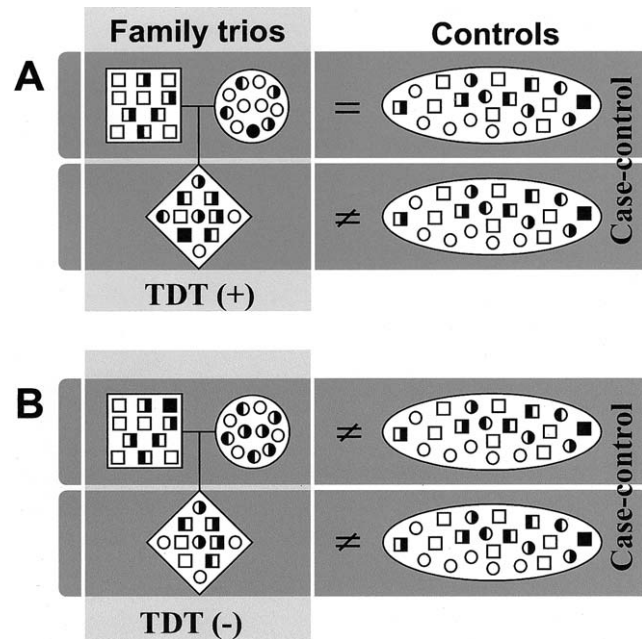


Figure 1 Types of study designs and their outcomes, depending on the underlying genetics. *A*, Situation expected when TDT is used and when patient's genetics counts. Parents are like control individuals, and there is a bias in the transmission of disease-susceptibility alleles (*half-blackened symbols*) to the affected children. This results in higher frequency of at-risk or protective alleles in patients, which also is apparent in a comparison with control individuals. *B*, Frequency of disease-susceptibility alleles. The frequency is higher in patients than in control individuals; however, a random transmission of parental chromosomes to the affected children results in a negative TDT. Case-control analysis reveals differences between patients and control individuals, simply reflecting what already had been present in an earlier generation, seen as differences between control individuals and parents (either only the mother or only the father or both). Plus (+) and minus (-) signs denote significant and nonsignificant TDT results, respectively, whereas equality (=) and inequality (≠) signs refer to susceptibility-allele frequency in cases compared with that in control individuals. The study outcome depicted in panel *A* corresponds to the results obtained with *GSTP1**C, and that in panel *B* corresponds to the results obtained with *CYP2E1**5, as reported in table 1.

patients' susceptibility resulting from the case-control analysis would be erroneously interpreted as a result of their own genotypes, whereas it reflects only the effect of the parental genotypes.

Indeed, our study of two genetic loci predisposing to childhood acute lymphoblastic leukemia (ALL) suggests that, at certain loci, the parental genotypes, rather than the genotype of the patient, are the predicting risk factor in this pediatric model of a complex disease. Consequently, we propose that a number of complementary mapping strategies will be required, such as the combination of case-control and TDT tests, to dissect genetic predisposition to complex disorders.

In the analysis of genetic susceptibility to ALL we

performed both case-control (Krajinovic et al. 1999) and case-only studies (Infante-Rivard et al. 1999), in a group of patients with ALL who were of French-Canadian origin and who were treated at the Sainte-Justine Hospital in Montreal. Several genes encoding carcinogen-metabolizing enzymes appeared to play a role in susceptibility to this disease (Sinnott et al. 2000). The parents of the participating patients with ALL were also recruited, to assess the effect of their genetics, in addition to that of their affected children. Genotyping of patients' parents was performed as described elsewhere for patients and control individuals (Krajinovic et al. 1999, 2002; Sinnott et al. 2000; M. Krajinovic, D. Labuda, and D. Sinnott, unpublished data). Genotype-frequency differences between cases and control individuals were assessed by χ^2 statistics. The level of significance was calculated by Fisher's exact test, and the strength of association was expressed by the odds ratio (OR) and its 95% CI. The TDT was run by the TDT/S-TDT program (version 1.1) (Spielman et al. 1993). Risk (with 95% CI) associated with transmitted variants was calculated according to the association measures for matched pairs (Lachin 2000). In table 1, we report the results obtained from >80 case-parents trios, applying both the TDT and the case-control analysis and using, as cases, either patients or their parents and the control individuals reported elsewhere.

The analysis of two genetic loci, those for glutathione S-transferase P1 (GSTP1 [MIM 134660]) and cytochrome P450 2E1 (CYP2E1 [MIM 124040]), revealed

results corresponding to the scenarios illustrated in figure 1A and B, respectively. The effect of the GSTP1*C variant is related to its presence/absence in the affected individual. The TDT analysis documents a reduced transmission of this variant, to the patients with ALL, from their heterozygous parents, implying that GSTP1*C is possibly protective against ALL. In a case-control study (table 1; M. Krajinovic, D. Labuda, and D. Sinnott, unpublished data), a similar association was suggested. In contrast, the variant CYP2E1*5, which, in a case-control study, appeared to confer an increased risk of ALL (Krajinovic et al. 2002), was negative in the TDT analysis. This indicates that the purported disease-susceptibility allele CYP2E1*5 is randomly transmitted from parents to the affected children. If the observed association is not spurious (see below), it suggests that the model, assuming that the genetic susceptibility is a function of patient's genotype, might not be adequate. Furthermore, the case-control results are replicated when the parents are substituted for the cases. Thus, an increased frequency of the disease-susceptibility allele among patients appears to reflect its overabundance in an earlier generation. The variant CYP2E1*5, conferring elevated transcription levels, has been associated with higher cancer risk (Watanabe et al. 1994; Shields et al. 1996). This effect includes childhood ALL (Krajinovic et al. 2002), the risk of which has been found to be particularly enhanced in relation to an increased alcohol consumption (Infante-Rivard et al. 2002). Therefore, it is plausible that the parents' genetics combined with al-

Table 1
Case-Control and TDT Analysis

ANALYSIS ^a	NO. (%) OF INDIVIDUALS					
	Patients vs. Control Individuals		Mothers vs. Control Individuals		Fathers vs. Control Individuals	
	Patients	Controls	Mothers	Controls	Fathers	Controls
CYP2E1*5:						
Absent	87 (90.6)	292 (97.0)	110 (90.2)	144 (98.0)	92 (88.5)	148 (96.1)
Present	9 (9.4)	9 (3.0)	12 (9.8)	3 (2.0)	12 (11.5)	6 (3.9)
OR _{cc}	3.4 (1.3–8.7)		5.2 (1.4–19.0)		3.2 (1.2–8.9)	
P _{cc}	.02		.007		.02	
OR _{TDT}	1.0 (.4–2.5)					
P _{TDT}	1.0					
GSTP1*C:						
Absent	77 (92.8)	264 (87.4)	99 (84.6)	129 (86.6)	82 (85.4)	135 (88.2)
Present	6 (7.2)	38 (12.6)	18 (15.4)	20 (13.4)	14 (14.6)	18 (11.8)
OR _{cc}	.5 (.2–1.3)		1.2 (.6–2.3)		1.3 (.6–2.7)	
P _{cc}	.2		.7		.6	
OR _{TDT}	.4 (.2–.8)					
P _{TDT}	.01					

NOTE.—Because only patients whose parents' DNA was available were included here, their number is smaller than that in our previous reports of case-control analysis, in which the following results were obtained: (i) OR_{CYP2E1*5} = 2.8 (range 1.2–6.5), P = .02, n = 174; and (ii) OR_{GSTP1*C} = 0.7 (range 0.4–1.1), P = .1, n = 278 (Krajinovic et al. 2002; M. Krajinovic, D. Labuda, and D. Sinnott, unpublished data).

^a cc = Case control.

cohol intake may increase cancer risk in their offspring. The risk, originally due to CYP2E1*5, is, however, transmitted to the future potential patients independently of this genetic variant (see fig. 1B).

A major concern in case-control studies is the possibility of a false-positive association due to nonhomogeneity of the population structure. Indeed, one of the great advantages of the TDT analysis is that it eliminates the possibility of spurious association due to population heterogeneity (Falk and Rubinstein 1987; Spielman et al. 1993; Risch and Merikangas 1996; Pericak-Vance 1998; Speer 1998). Negative TDT results observed in the case of the CYP2E1 gene could therefore simply indicate such spurious association obtained from the case-control analysis. There are several reasons to believe that this is not the case. First, both patients and control individuals originate from the same French-Canadian population of Quebec, where no regional clustering of childhood ALL was observed. Both patients and control individuals are expected to share, to the same extent, the ancestry among ~5,000 immigrants who founded “Nouvelle-France” during the 17th century (Courville 1996). Second, the analysis of our control population by the STRUCTURE program (Pritchard et al. 2000) clusters all individuals into a single population (proportion of individual’s genome originating in a single population [Q] is 0.994), thus suggesting population homogeneity and supporting our inclusion criteria (Krajinovic et al. 1999). The STRUCTURE program (Pritchard et al. 2000) was used, with the assumption that there are three population clusters, on the population of control individuals that were genotyped at 11 multiallelic and 7 biallelic loci distributed over the genome (Krajinovic et al. 1999, 2000, 2002; M. Krajinovic, D. Labuda, and D. Sinnett, unpublished data). Third, the case-control results reported in the present study (i.e., those which use only patients within familial trios) are consistent with those obtained with a larger cohort of our patients, where chances of a type I error are smaller (see the general “Note” to table 1). The effect of GSTP1*C was suggested by case-control analysis, and, in the present study, this result was reinforced by TDT. Fourth, the observed associations between CYP2E1*5 and cancer susceptibility are not limited to childhood ALL as noted by others (Shields et al. 1996; Hung et al. 1997). The effect of CYP2E1*5—via parental contribution in the context of gene-environment interaction, such as that due to alcohol exposure—is biologically plausible (Infante-Rivard et al. 2002). Finally, TDT works only for cases in which the patient’s genotypes count and thus, by default, is not suitable to detect the observed effect that parental genetics has in the present study.

Although the importance of parents’ exposure has often been studied, the influence that their disease-suscep-

tibility alleles have on the development of diseases in their offspring has been surprisingly neglected. The results described in the present study could represent a new paradigm in our comprehension of the mapping of complex disorders, since they may be extended to at least some clinical entities of other disorders, beyond cancer or childhood diseases. This possibility is suggested by recent epidemiological studies linking poor intrauterine growth, due to maternal factors, with health problems in adult life (Nyirenda and Seckl 1998). Our results also remind us that computational approaches, despite their power, may be misleading if they are based on inappropriate assumptions (see Terwilliger and Goring 2000). Instead, relying on a single test that assesses only limited genetic effect, we propose that, whenever possible, a wide spectrum of study designs should be considered; this may also mean, for example, that the analysis of gene-environment interaction should be included, in addition to analysis of the immediate effects that genetics has at the level of the patient and the parents.

Acknowledgments

We are grateful to the families of the patients with ALL, for their collaboration. M.K. and D.S. are scholars of the Fonds de la Recherche en Santé du Québec, and C.I.-R. holds a Canada Research Chair from McGill University. This study was supported by the Fondation Charles Bruneau, the Power Corp, Inc., and the Fondation de l’Hôpital Sainte-Justine, as well as by grants from the Canadian Institutes of Health Research.

Electronic-Database Information

Accession numbers and the URL for data presented herein are as follows:

Online Mendelian Inheritance in Man (OMIM), <http://www.ncbi.nlm.nih.gov/omim> (for GSTP1 [MIM 134660] and CYP2E1 [MIM 124040])

References

- Altshuler D, Hirschhorn JN, Klannemark M, Lindgren CM, Vohl MC, Nemesh J, Lane CR, Schaffner SF, Bolck S, Brewer C, Tuomi T, Gaudet D, Hudson TJ, Daly M, Groop L, Lander ES (2000) The common PPARgamma Pro12Ala polymorphism is associated with decreased risk of type 2 diabetes. *Nat Genet* 26:76-80
- Bell DA, Taylor JA (1997) Genetic analysis of complex disease. *Science* 275:1327-1328 (discussion 1329-1330)
- Bertina RM, Koeleman BP, Koster T, Rosendaal FR, Dirven RJ, de Ronde H, van der Velden PA, Reitsma PH (1994) Mutation in blood coagulation factor V associated with resistance to activated protein C. *Nature* 369:64-67
- Cargill M, Altshuler D, Ireland J, Sklar P, Ardlie K, Patil N, Shaw N, Lane CR, Lim EP, Kalyanaraman N, Nemesh J, Ziaugra L, Friedland L, Rolfe A, Warrington J, Lipshutz R, Daley GQ, Lander ES (1999) Characterization of single-

- nucleotide polymorphisms in coding regions of human genes. *Nat Genet* 22:231–238
- Chakravarti A (1999) Population genetics—making sense out of sequence. *Nat Genet* 21:56–60
- Corder EH, Saunders AM, Strittmatter WJ, Schmechel DE, Gaskell PC, Small GW, Roses AD, Haines JL, Pericak-Vance MA (1993) Gene dose of apolipoprotein E type 4 allele and the risk of Alzheimer's disease in late onset families. *Science* 261:921–923
- Courville S (1996) Atlas historique du Québec: population et territoire. Les Presses de l'Université Laval, Sainte-Foy, Québec, Canada
- Crow JF (2000) The origins, patterns and implications of human spontaneous mutation. *Nat Rev Genet* 1:40–47
- Ekbom A (1998) Growing evidence that several human cancers may originate in utero. *Semin Cancer Biol* 8:237–244
- Falk CT, Rubinstein P (1987) Haplotype relative risks: an easy reliable way to construct a proper control sample for risk calculations. *Ann Hum Genet* 51:227–233
- Ferguson P, Mohr L, Hoel D, Lipsitz S, Lackland D (2000) Possible relationship between birth weight and cancer incidence among young adults. *Ann Epidemiol* 10:471
- Godfrey KM, Barker DJ (2001) Fetal programming and adult health. *Public Health Nutr* 4:611–624
- Haines JL, Pericak-Vance MA (1998) Approaches to gene mapping in complex human diseases. John Wiley & Sons, New York
- Hung HC, Chuang J, Chien YC, Chern HD, Chiang CP, Kuo YS, Hildesheim A, Chen CJ (1997) Genetic polymorphisms of CYP2E1, GSTM1, and GSTT1: environmental factors and risk of oral cancer. *Cancer Epidemiol Biomarkers Prev* 6:901–905
- Infante-Rivard C, Krajinovic M, Labuda D, Sinnett D (2002) Childhood acute lymphoblastic leukemia associated with parental alcohol consumption and carcinogen-metabolizing genetic polymorphisms. *Epidemiology* 13:277–281
- Infante-Rivard C, Labuda D, Krajinovic M, Sinnett D (1999) Risk of childhood leukemia associated with exposure to pesticides and with gene polymorphisms. *Epidemiology* 10:481–487
- Infante-Rivard C, Sinnett D (1999) Preconceptional paternal exposure to pesticides and increased risk of childhood leukemia. *Lancet* 354:1819
- Krajinovic M, Labuda D, Richer C, Karimi S, Sinnett D (1999) Susceptibility to childhood acute lymphoblastic leukemia: influence of CYP1A1, CYP2D6, GSTM1, and GSTT1 genetic polymorphisms. *Blood* 93:1496–1501
- Krajinovic M, Richer C, Sinnett H, Labuda D, Sinnett D (2000) Genetic polymorphisms of N-acetyltransferases 1 and 2 and gene-gene interaction in the susceptibility to childhood acute lymphoblastic leukemia. *Cancer Epidemiol Biomarkers Prev* 9:557–562
- Krajinovic M, Sinnett H, Richer C, Labuda D, Sinnett D (2002) Role of NQO1, MPO and CYP2E1 genetic polymorphisms in the susceptibility to childhood acute lymphoblastic leukemia. *Int J Cancer* 97:230–236
- Lachin JM (2000) Biostatistical methods: the assessment of relative risks. John Wiley & Sons, New York
- Long AD, Grote MN, Langley CH (1997) Genetic analysis of complex diseases. *Science* 275:1328 (discussion 1329–1330)
- Miller MC III, Mohrenweiser HW, Bell DA (2001) Genetic variability in susceptibility and response to toxicants. *Toxicol Lett* 120:269–280
- Morley R, Dwyer T (2001) Fetal origins of adult disease? *Clin Exp Pharmacol Physiol* 28:962–966
- Muller-Myhsok B, Abel L (1997) Genetic analysis of complex diseases. *Science* 275:1328–1329 (discussion 1329–1330)
- Nyirenda MJ, Seckl JR (1998) Intrauterine events and the programming of adulthood disease: the role of fetal glucocorticoid exposure. *Int J Mol Med* 2:607–614
- Perera FP (1997) Environment and cancer: who are susceptible? *Science* 278:1068–1073
- Pericak-Vance MA (1998) Linkage disequilibrium and allelic association. In: Haines JL, Pericak-Vance MA (eds) Approaches to gene mapping in complex human diseases. John Wiley & Sons, New York, pp 323–333
- Pritchard JK, Stephens M, Rosenberg NA, Donnelly P (2000) Association mapping in structured populations. *Am J Hum Genet* 67:170–181
- Reich DE, Lander ES (2001) On the allelic spectrum of human disease. *Trends Genet* 17:502–510
- Risch N, Merikangas K (1996) The future of genetic studies of complex human diseases. *Science* 273:1516–1517
- Scott WK, Pericak-Vance MA, Haines JL (1997) Genetic analysis of complex diseases. *Science* 275:1327 (discussion 1329–1330)
- Shields PG, Ambrosone CB, Graham S, Bowman ED, Harrington AM, Gillenwater KA, Marshall JR, Vena JE, Laughlin R, Nemoto T, Freudenheim JL (1996) A cytochrome P4502E1 genetic polymorphism and tobacco smoking in breast cancer. *Mol Carcinog* 17:144–150
- Sinnett D, Krajinovic M, Labuda D (2000) Genetic susceptibility to childhood acute lymphoblastic leukemia. *Leuk Lymphoma* 38:447–462
- Speer MC (1998) Sample size and power. In: Haines JL, Pericak-Vance MA (eds) Approaches to gene mapping in complex diseases. John Wiley & Sons, New York, pp 161–200
- Spielman RS, McGinnis RE, Ewens WJ (1993) Transmission test for linkage disequilibrium: the insulin gene region and insulin-dependent diabetes mellitus (IDDM). *Am J Hum Genet* 52:506–516
- Terwilliger JD, Goring HH (2000) Gene mapping in the 20th and 21st centuries: statistical methods, data analysis, and experimental design. *Hum Biol* 72:63–132
- Watanabe J, Hayashi S, Kawajiri K (1994) Different regulation and expression of the human CYP2E1 gene due to the RsaI polymorphism in the 5'-flanking region. *J Biochem (Tokyo)* 116:321–326