

Proceedings

Open Access

Power of the 2-locus TDT for testing the interaction of two susceptibility genes

Salma Kotti*^{1,2}, Mathieu Bourgey^{1,2} and Françoise Clerget-Darpoux^{1,2}

Address: ¹Université Paris-Sud, UMR-S535, 94817 Villejuif, France and ²Inserm U535, Hôpital Paul Brousse, BP1000, 94817 Villejuif, France

Email: Salma Kotti* - kotti@vjf.inserm.fr; Mathieu Bourgey - bourgey@vjf.inserm.fr; Françoise Clerget-Darpoux - clerget@vjf.inserm.fr

* Corresponding author

from Genetic Analysis Workshop 15
St. Pete Beach, Florida, USA. 11–15 November 2006

Published: 18 December 2007

BMC Proceedings 2007, 1(Suppl 1):S65

This article is available from: <http://www.biomedcentral.com/1753-6561/1/S1/S65>

© 2007 Kotti et al; licensee BioMed Central Ltd.

This is an open access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/2.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Abstract

We recently proposed a new strategy: 2-locus TDT for detecting two susceptibility genes through their interaction in trio families. We apply our method to two candidate genes, A and C, on the Genetic Analysis Workshop 15 (GAW15) simulated rheumatoid arthritis data and study the power to identify an interactive effect of these genes.

This study was performed with full knowledge of the answers.

Introduction

Most multifactorial diseases result from complex interactions among multiple genes with modest effects and various environmental factors. Current strategies usually aim to detect the individual effect of each factor.

We have recently proposed a new strategy: the 2-locus TDT (transmission-disequilibrium test), which tests for interaction between two candidate genes [1]. Like the one-locus TDT [2], it applies to trio families including the patient and his two parents. Our method is used in the context of studying good candidate pathways. Indeed, we assume prior information that the disease under study is probably caused by particular genes recognized as key ele-

ments of the metabolic pathway. We apply the 2-locus TDT to Problem 3 simulated Genetic Analysis Workshop 15 (GAW15) rheumatoid arthritis (RA) data and compute its power for detecting the interaction of two susceptibility genes.

Data

The GAW15 Problem 3 consisted of 100 replicates of 1500 nuclear families with two affected sibs. From each replicate, we extracted 1500 trio families: one affected (the index case) and his two parents.

We focus the analyses on two candidate genes in the RA susceptibility: Locus A on chromosome 16 and Locus C

on chromosome 6. Locus A has an indirect effect on RA hazard. Its implication on the susceptibility to the disease depends on the HLA-DRB1 locus located on chromosome 6. Locus C is in strong linkage disequilibrium (LD) with HLA-DRB1. Both genes are involved directly or indirectly in the RA outcome through the HLA-DRB1 gene and thus do not act independently. We apply the 2-locus TDT to the SNPs indicated as functional in the two candidate genes in each replicate of 1500 trios.

Notation

We denote A_1, A_2 and C_1, C_2 the SNP alleles of genes A and C, respectively. For each patient, there are nine possible genotypes that we denote $A_iA_jC_kC_l$, where A_i, A_j and C_k, C_l are the alleles inherited from the two parents at the loci A and C, respectively. Subscripts i, j, k and l can be either 1 or 2. A_2 and C_2 are considered to be the high risk alleles and $A_2A_2C_2C_2$ are the reference genotype.

We denote the nine penetrances $P(\text{affected} | A_iA_jC_kC_l)$ by f_{ijkl} and by ε_{ijkl} the eight penetrances relatively to the reference genotype $A_2A_2C_2C_2$. Thus, $\varepsilon_{2222} = 1$. Let M_G be the general model in which ε_{ijkl} may take any value between 0 and 1. When the two genes A and B act independently, the lines and the columns of the relative penetrance matrix are proportional (multiplicative matrix) and may be written with four parameters: μ, ν, ϕ and π (see Table 1). The corresponding restricted model is denoted M_R .

Method

The 2-locus TDT is achieved in two steps. In the first step, the parameters (μ, ν, ϕ , and π) of the restricted model and ε_{ijkl} of the general model are estimated by their maximum likelihood estimator (MLE). In the second one, the restricted model is compared to the general one by a maximum likelihood ratio test.

The estimation of these parameters requires specifying the allele frequencies at each locus. This was achieved by the affected family-based controls (AFBAC) method, using the parental alleles untransmitted to the patients from which the family was ascertained [3]. Here, the families were not ascertained through an affected patient but through a patient with an affected sib. Our parameter estimation must also take into account this specific mode of

Table 1: Penetrance estimation under the restricted model

Gene C	Gene A		
	A_1A_1	A_1A_2	A_2A_2
C_1C_1	$\phi\mu$	$\pi\mu$	μ
C_1C_2	$\phi\nu$	$\pi\nu$	ν
C_2C_2	ϕ	π	1

ascertainment and consequently was achieved by using the marker association segregation chi-square (MASC) program [4].

We denote by $\max_G L_G$ and $\max_R L_R$ the likelihood values for the estimated parameters of the general and restricted models, respectively. The likelihood ratio test (LRT) was calculated as follows:

$$LRT = -2 \ln \frac{\max_R L_R}{\max_G L_G}$$

This ratio follows, under the null hypothesis, a chi-square distribution with four degrees of freedom (df) corresponding to the difference in the number of parameters between the two models.

For each replicate of 1500 trios, the LRT was computed and the significance of the test was assessed in order to evaluate the power of our test in the present situation.

Results

Because patients with genotypes $A_1A_1C_1C_1, A_1A_2C_1C_1$, and $A_2A_2C_1C_1$ are very rare in all replicates, we fixed the relative penetrances of these genotypes at 10^{-3} . Consequently, under the general model, we estimate only five parameters and, under the restricted model, only three parameters ν, ϕ , and π . Thus, in the present situation, we have 2 degrees of freedom for the LRT. Figures 1 and 2 show the distributions of the LRT and p -values in 98 replicates (the program failed in 2 replicates), respectively. There is a good percentage of replicates (44/98) for which

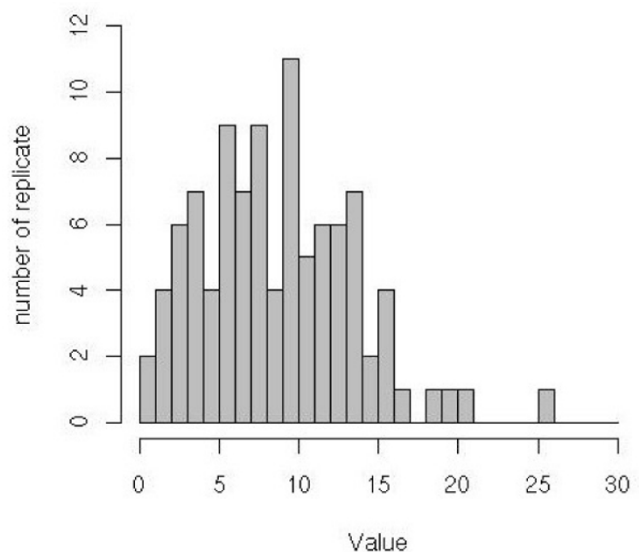


Figure 1
Distribution of the LRT values.

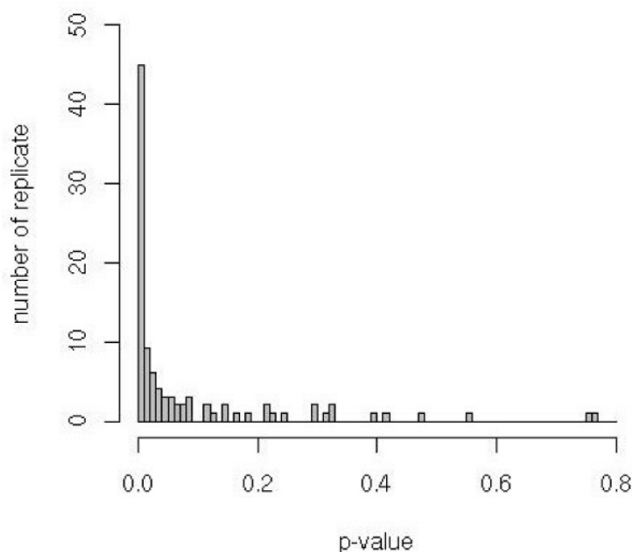


Figure 2
Distribution of the LRT *p*-values.

the test was significant at 1% level. The power of 2-locus TDT test (at a 5% level) is 66%.

We illustrate the variation of the LRT and *p*-values on two contrasting replicates: 9 and 27. In the Replicate 27, there is a strong evidence for interaction: LRT = 14.35 and $p = 7 \times 10^{-4}$. In contrast, we did not find interaction in Replicate 9 because the LRT = 2.49 and $p = 0.29$.

Tables 2 and 3 show the parameter estimates under the general model for Replicates 9 and 27, respectively. The parameter estimates under the restricted model are $\nu = 0.135$, $\phi = 0.688$, and $\pi = 0.886$ for Replicate 9 and $\nu = 0.168$, $\phi = 0.682$, and $\pi = 0.909$ for replicate 27. It is clear that the departure from multiplicativity in the relative penetrance matrix of the general model is stronger for Replicate 27 than for Replicate 9. In particular, for Replicate 27, the relative penetrance for the genotype $A_1A_2C_2C_2$ is half of the one expected under the restricted model.

Table 2: Replicate 9: Matrix of relative penetrances for the general model

Gene C	Gene A		
	A_1A_1	A_1A_2	A_2A_2
C_1C_1	0.001	0.001	0.001
C_1C_2	0.105^a	0.109	0.113
C_2C_2	0.668	0.879	1

^a Bold indicates parameter estimates.

Discussion

In this paper, we studied the performance of the 2-locus TDT for detecting the interaction of the genes A and C in the susceptibility to RA simulated data and show that interaction may be detected with a power of 66% at a 5% level.

The 2-locus TDT searches for a departure from multiplicativity in the two-locus penetrance matrix. This departure cannot be observed without an effect of the two genes. Therefore, the 2-locus TDT may also be used for detecting the effect of two candidate genes through their interaction.

More generally, the 2-locus TDT may be particularly interesting in the situation of weak effect of one or both genes and strong interaction. In such a situation, traditional methods will fail because they will stop in the absence of noticeable main effect. In fact, they ignore the possibility that the interactive effect of multilocus functional genetic variants may play a larger role than the single-locus effect in the determinism of one trait [1].

Our test applies to candidate genes already recognized as key elements of a metabolic pathway thought to be involved in the disease pathogenesis. There are difficulties in good candidate choice and the use of tagSNPs as markers for most genetic variations in the human genome has become a cornerstone of genetic research. There are, however, reasons to think that the expectations they have raised are somewhat inflated and may lead to unwarranted neglect of other strategies. In particular, we believe in particular that identification of biologically relevant pathways followed by candidate gene strategies may be quite effective at identifying new variants and their interaction in disease susceptibility [5].

Competing interests

The author(s) declare that they have no competing interests.

Acknowledgements

This work was supported by a grant from the Fondation pour la Recherche Médicale.

Table 3: Replicate 27: Matrix of relative penetrances for the general model

Gene C	Gene A		
	A_1A_1	A_1A_2	A_2A_2
C_1C_1	0.001	0.001	0.001
C_1C_2	0.137	0.133	0.083
C_2C_2	0.623	0.868	1

^a Bold indicates parameter estimates.

This article has been published as part of *BMC Proceedings* Volume 1 Supplement 1, 2007: Genetic Analysis Workshop 15: Gene Expression Analysis and Approaches to Detecting Multiple Functional Loci. The full contents of the supplement are available online at <http://www.biomedcentral.com/1753-6561/1?issue=S1>.

References

1. Kotti S, Bickeboller H, Clerget-Darpoux F: **Strategy for detecting susceptibility genes with weak or no marginal effect.** *Hum Hered* 2007, **63**:85-92.
2. Spielman RS, McGinnis RE, Ewens WJ: **Transmission test for linkage disequilibrium: the insulin gene region and insulin-dependent diabetes mellitus (IDDM).** *Am J Hum Genet* 1993, **52**:506-516.
3. Thomson G: **Mapping disease genes: family-based association studies.** *Am J Hum Genet* 1995, **57**:487-498.
4. Clerget-Darpoux F, Babron MC, Prum B, Lathrop GM, Deschamps I, Hors J: **A new method to test genetic models in HLA associated diseases: the MASC method.** *Ann Hum Genet* 1988, **52**:247-258.
5. Bourgain C, Genin E, Cox N, Clerget-Darpoux F: **Are genome-wide association studies all that we need to dissect the genetic component of complex human diseases?** *Eur J Hum Genet* 2007, **15**:260-263.

Publish with **BioMed Central** and every scientist can read your work free of charge

"BioMed Central will be the most significant development for disseminating the results of biomedical research in our lifetime."

Sir Paul Nurse, Cancer Research UK

Your research papers will be:

- available free of charge to the entire biomedical community
- peer reviewed and published immediately upon acceptance
- cited in PubMed and archived on PubMed Central
- yours — you keep the copyright

Submit your manuscript here:
http://www.biomedcentral.com/info/publishing_adv.asp

