ARTICLE

Joint effects of HLA, *INS*, *PTPN22* and *CTLA4* genes on the risk of type 1 diabetes

M. Bjørnvold · D. E. Undlien · G. Joner ·
K. Dahl-Jørgensen · P. R. Njølstad · H. E. Akselsen ·
K. Gervin · K. S. Rønningen · L. C. Stene

Abstract

Background/hypothesis HLA, *INS*, *PTPN22* and *CTLA4* are considered to be confirmed type 1 diabetes susceptibility genes. HLA, *PTPN22* and *CTLA4* are known to be involved in immune regulation. Few studies have systematically investigated the joint effect of multiple genetic variants. We evaluated joint effects of the four established genes on the risk of childhood-onset type 1 diabetes.

Methods We genotyped 421 nuclear families, 1,331 patients and 1,625 controls for polymorphisms of *HLA-DRB1*, −*DQA1* and −*DQB1*, the insulin gene (*INS*, −23 *Hph*I), *CTLA4* (*JO27_1*) and *PTPN22* (*Arg620Trp*).

Results The joint effect of HLA and *PTPN22* on type 1 diabetes risk was significantly less than multiplicative in the case-control data, but a multiplicative model could not be rejected in the trio data. All other two-way gene–gene interactions fitted multiplicative models. The high-risk HLA genotype conferred a very high risk of type 1 diabetes (OR 20.6, using the neutral-risk HLA genotype as reference). When including also intermediate-risk HLA genotypes together with risk genotypes at the three non-HLA loci, the joint odds ratio was 61 (using non-risk genotypes at all loci as reference).

Conclusion Most established susceptibility genes seem to act approximately multiplicatively with other loci on the risk of disease except for the joint effect of HLA and *PTPN22*. The joint effect of multiple susceptibility loci conferred a very high risk of type 1 diabetes, but applies to a very small proportion of the general population. Using multiple susceptibility genotypes compared with HLA genotype alone seemed to influence the prediction of disease only marginally.

Keywords *CTLA4* · Genes · HLA · *INS* · Interaction · *PTPN22* · Type 1 diabetes

M. Bjørnvold (✉) · D. E. Undlien
Institute of Medical Genetics,
Faculty Division Ullevål University Hospital, University of Oslo,
P.O. Box 1036, Blindern,
NO-0315 Oslo, Norway
e-mail: marit.bjornvold@medisin.uio.no

D. E. Undlien · H. E. Akselsen · K. Gervin
Department of Medical Genetics, Ullevål University Hospital,
Oslo, Norway

G. Joner · K. Dahl-Jørgensen
Department of Paediatrics, Ullevål University Hospital,
Oslo, Norway

G. Joner · K. Dahl-Jørgensen
Faculty of Medicine, University of Oslo,
Oslo, Norway

P. R. Njølstad
Section for Paediatrics, Department of Clinical Medicine,
University of Bergen,
Bergen, Norway

P. R. Njølstad
Department of Paediatrics, Haukeland University Hospital,
Bergen, Norway

K. S. Rønningen · L. C. Stene
Division of Epidemiology, Norwegian Institute of Public Health,
Oslo, Norway

Abbreviations

ROC receiver operating characteristic
SNP single nucleotide polymorphism

Introduction

The risk of complex diseases such as type 1 diabetes is generally thought to be influenced by multiple genetic and non-genetic factors, and it has been hypothesised that interactions between genes, or epistasis, are very common for such diseases [1]. The presence of interactions could be one of the reasons why searching for susceptibility loci for many diseases has been less successful than expected [2]. When moving from monogenic diseases to complex diseases, it seems reasonable to assess more than one locus at a time, although models become increasingly complex as the number of loci increases [3].

Whereas few or no common genetic variants have been firmly established for most common diseases [4], there are now at least four genetic loci that are established as causally involved in the aetiology of type 1 diabetes. They give us a unique possibility to evaluate gene–gene interactions among established susceptibility genes. Specific allelic combinations of *DRB1*, *DQA1* and *DQB1* in the human leucocyte antigen (HLA) complex, variants in the insulin gene (*INS*), the cytotoxic T lymphocyte antigen-4 gene (*CTLA4*) and the protein tyrosine phosphatase, non-receptor type 22 gene (*PTPN22*) have been repeatedly associated with type 1 diabetes susceptibility [5–8] using different approaches. All established loci are thought to be involved somehow in immune regulation, but details of the mechanisms relating the polymorphisms to risk of type 1 diabetes are in most cases poorly understood.

Evaluating the joint effects of genes contributes important information for risk prediction, and is also thought to provide information about biological interactions, although the latter is controversial and more complex than commonly thought [2, 3, 9].

Previous studies have assessed interaction between HLA and *INS* and reported divergent results [10–18]. The reported results of the joint effect of HLA and *INS* are confusing not only because they have shown diverging results but also because the definitions and terminology of interactions are not consistent [19]. The interpretation of statistical interaction depends on the choice of scale used to measure the effects [2]. Although additivity of risks is often taken as independence [20], multiplicativity of risk is sometimes also taken as independence [21]. The joint action of HLA and *INS* has variously been described as being multiplicative [14, 15], additive [13], providing evidence of interaction [15], and non-interacting [13, 14, 17].

Few studies have investigated the more recently established susceptibility loci *PTPN22* and *CTLA4* in the context of joint effects on the risk of type 1 diabetes. The studies that have been done have mainly concluded that there is a non-interaction [22, 23] effect, but here also the results have diverged [24–27].

The aim of our study was to assess the joint effects of the four established susceptibility loci HLA, *INS*, *CTLA4* and *PTPN22* in type 1 diabetes, using a consistent approach with both population-based case-control and family trio designs and with large sample sizes.

Methods

Participants We analysed two independent type 1 diabetes data sets. One nuclear family set consisted of 421 trios with mother, father and one child diagnosed in Norway with type 1 diabetes before age 15 years (225 [53.4%] of the affected children were boys). The families were collected between 1993 and 1997. In families with more than one affected sibling, only the proband was included in the analyses. The case-control data set consisted of 1,331 type 1 diabetes patients (51.9% boys and 48.1% girls) and 1,625 (51.4% boys and 48.6% girls) control participants aged <15 years. In analyses involving age of disease onset, we divided the data sets according to age of disease onset of the affected child into three groups (0–4.9, 5–9.9 and ≥10 years). The controls were randomly selected from the official population registry among children born between 1985 and 1999 and recruited in 2001, as previously described [28]. The patients in the case-control material were from the Norwegian Childhood Diabetes Registry consecutively recruited between 1997 and 2000 [29] and between 2002 and 2005. The type 1 diabetes patients and their family members were recruited by the Norwegian Childhood Diabetes Study Group, including all paediatric departments in Norway. All type 1 diabetes patients were diagnosed according to EURODIAB criteria [30]. The study was approved by the local ethics committee, and informed consent was obtained from all participants or their parents.

Genomic DNA extraction and genotyping In the majority of the type 1 diabetes case-control samples we used DNA extracted from buccal cells [31]. In all remaining samples, DNA was extracted from peripheral whole blood using a salting-out protocol. Genotyping of *HLA-DRB1*, −*DQA1* and −*DQB1* was performed using PCR-SSOP (sequence-specific oligonucleotide probing) mainly following published methods [32], or PCR-SSP (sequence-specific primer) [33, 34], or using time-resolved fluorescence technology in the Delfia assay (Perkin-Elmer Life Sciences, Turku, Finland). HLA genotypes were grouped into four risk categories based on *DQB1*, *DQA1* and *DRB1* genotypes, including *DRB1*04* subtyping. The majority of *DRB1*04-DQA1*0301-DQB1*0302* haplotypes in Norway are *DRB1*0401* or −*0404* (constituting >94% of the haplotypes) [35]. In the present study the other rare

subtypes are referred to as *DRB1*04XX*. Because of the almost complete linkage disequilibrium between *INS-VNTR* allele classes and the *−23 HphI* polymorphism, we genotyped −23 *HphI* (rs689) as a marker for the *INS-VNTR*. The −23 *HphI* A allele corresponds to the *VNTR* class I and the −23 *HphI* T allele correspond to the *VNTR* class III. In *PTPN22* we genotyped the single nucleotide polymorphism (SNP) *Arg620Trp* (rs2476601). The SNP *JO27_1* (rs11571297) was genotyped in *CTLA4*. SNP genotyping was performed by TaqMan allelic discrimination assays on an ABI 7900HT DNA Analyzer (Applied Biosystems, Foster City, CA USA). Primer and probe sequences are shown in Electronic Supplementary Material (ESM) Table 1. The PCR conditions are available on request.

Data analysis The HLA haplotypes were grouped as high risk, intermediate risk, neutral risk and low risk according to the following criteria: high-risk category, *DRB1*0401/ 04XX-DQA1*03-DQB1*0302/DRB1*03-DQA1*05-DQB1*0201* (*DR4-DQ8/DR3-DQ2*); low-risk category, all genotypes with at least one *DQB1*0602* allele; intermediate-risk category, *DRB1*0404-DQ8/DR3-DQ2, DR3-DQ2/ DR3-DQ2, DR4-DQ8/DR4-DQ8* (with the exception of *DRB1 0404-DQ8* homozygotes, which were grouped as neutral), *DRB1*0401* or *040XX-DQ8/X* (X≠*DQB1*0602* or *DR3-DQ2*). The remaining haplotypes were grouped in the neutral-risk category. For assessment of two-locus joint effects, we pooled genotypes of *INS*, *PTPN22* and *CTLA4* as follows: *INS* class I/I genotypes were compared with I/III together with III/III genotypes. The *PTPN22* TT and CT genotypes were compared with CC. *CTLA4* (*JO27_1*) TT genotypes were compared with TC and CC genotypes. Not all individuals were genotyped for non-HLA polymorphisms because of lack of DNA. We did not exclude individuals with some missing genotypes to prevent the loss of important information when looking at joint effects between polymorphisms. The available numbers of individuals for each analysis are seen in the tables. Data were presented using stratified 2×2 tables and analysed using logistic regression models including interaction (product) terms, in SPSS for Windows (version 14.0; SPSS, Chicago, IL, USA). In addition to formal analyses treating HLA categories as categorical in the logistic regression, we also tested for interactions when treating HLA category as a continuous variable coded 1, 2, 3, 4, thus maximising the power under alternative models where the effect of a non-HLA locus (as measured by the odds ratio [OR]) was assumed to decrease (or increase) (logit) linearly over the four HLA risk categories (test for interaction with one degree of freedom). Case-only analyses were used to estimate interaction parameters and to test for deviation from multiplicative effects using logistic regression [36]. In addition to the increased power obtained by utilising all

cases (from case-control and trio materials) simultaneously, case-only analyses have increased power by making the implicit assumption that there is no association between the two loci in the population; i.e. the OR for their association is 1.0 in the population. Thus, under reasonable assumptions the case-only analysis makes the most efficient use of data to assess deviation from multiplicative models. In the case-control analysis, likelihood ratio tests comparing nested logistic regression models were used as global tests for interaction. The transmission disequilibrium test [37] was performed using the UNPHASED application implemented in the UNPHASED software version 2.4 [38]. For the trio data, 95% confidence intervals for the relative risk were estimated using conditional logistic regression in UNPHASED. Receiver operating curve (ROC) and confidence bounds for the area under the curve were estimated assuming a non-parametric distribution and analyses were done using SPSS version 14.0. Genotypes were added sequentially in order of likelihood ratio (or equivalently by the absolute risk conferred by a given genotype combination estimated using Bayes' formula). Two four-locus genotype combinations were absent among cases in our material, and a very low value for the estimated absolute risk was imputed for these to allow inclusion in the ROC curve estimation with all four loci simultaneously. A *p* value <0.05 was considered to be statistically significant.

Results

The single-locus main effects are shown in ESM Table 2 (case-control data) and ESM Table 3 (trios). Compared with the neutral HLA risk category, the high-risk category showed a strong association with type 1 diabetes, with OR 20.6; for the intermediate-risk category the OR was 5.7 and for the low-risk category it was 0.09. *INS*, *PTPN22* and *CTLA4* also showed an association with type 1 diabetes, as expected. The transmission of the risk allele in the nuclear families confirmed the associations in *INS*, *PTPN22* and *CTLA4* (*JO27_1*), although with borderline significance for *JO27_1* (ESM Table 3).

Joint effect of HLA and PTPN22 ORs for the effect of *PTPN22* varied across the HLA risk categories and were significant in some of the subgroups. The ORs were smaller for the risk-conferring HLA genotypes, indicating negative deviation from a multiplicative model. A global test of interaction (with 3 *df*) between HLA and *PTPN22* in the logistic regression model confirmed a significant interaction (*p*=0.024). In the trio data, the relative risk conferred by the *PTPN22* T allele was similar in the strata defined by HLA group, with no evidence for deviation from a multiplicative model (ESM Table 4). A case-only analysis among all cases

from the case-control and family materials (ESM Table 5) supported a significant negative deviation from multiplicative effects, with weaker ORs conferred by *PTPN22* in the HLA risk categories (3 *df* test for interaction; *p*=0.028). When treating HLA-encoded risk as a continuous variable in the analysis (1 *df*), the interaction was even more statistically significant (*p*=0.003). There was no association between HLA and *PTPN22* among the controls (3 *df* test; *p*=0.19). We tried to fit the case-control data to an additive odds model using generalised linear models in STATA (version 9), as described by Skrondal [39]. However, convergence was not obtained, suggesting that the data did not fit well to an additive model.

Joint effect of HLA and INS The 3 *df* test for interaction between *INS* and HLA was not statistically significant (*p*= 0.67). There was also no statistically significant deviation from a multiplicative model in the trio data (test for interaction, *p*=0.5) (ESM Table 4) or in the case-only analysis (3 *df* test; *p*=0.49) (ESM Table 5); even when treating HLA-encoded risk as a continuous variable in the analysis there was no significance (1 *df* test, *p*=0.12). There was also no association between *INS* and HLA among controls, as expected (3 *df* test; *p*=0.41).

Joint effect of HLA and CTLA4 The ORs for *CTLA4* in the different HLA categories (Table 1) indicated no deviation

Table 1 Interaction between *HLA-INS* (-23HphI), *HLA-PTPN22* (*Arg620Trp*) and *HLA-CTLA4* (*JO27_1*) in the case–control data set using logistic regression

HLA category[a]	Non-HLA genotypes	Cases *n* (%)	Controls *n* (%)	OR	95% CI	Test for interaction (*p* value)
	INS[b]					0.67
High risk	I-I	218 (25.2)	20 (2.8)	1.56	0.76–3.21	
	III+	98 (28.6)	14 (2.5)	1		
Intermediate risk	I-I	412 (47.6)	122 (17.1)	2.10	1.52–2.89	
	III+	160 (46.8)	99 (17.4)	1		
Neutral risk	I-I	221 (25.5)	346 (48.5)	2.34	1.74–3.15	
	III+	81 (23.7)	299 (52.6)	1		
Low risk	I-I	14 (1.6)	225 (31.6)	3.26	0.92–11.54	
	III+	3 (0.9)	156 (27.5)	1		
Total		1207	1281			
	PTPN22					0.024
High risk	TT+TC	89 (23.2)	8 (3.1)	1.26	0.55–2.88	
	CC	237 (27.8)	27 (2.6)	1		
Intermediate risk	TT+TC	177 (46.1)	55 (21.2)	1.31	0.91–1.86	
	CC	406 (47.7)	164 (15.6)	1		
Neutral risk	TT+TC	111 (28.9)	121 (46.7)	2.44	1.80–3.31	
	CC	199 (23.4)	536 (51.0)	1		
Low risk	TT+TC	7 (1.8)	75 (29.0)	3.27	1.18–9.06	
	CC	9 (1.1)	323 (30.8)	1		
Total		1235	1309			
	CTLA4					0.53
High risk	TT	122 (24.9)	14 (3.6)	0.90	0.45–1.84	
	TC-CC	210 (27.9)	22 (2.5)	1		
Intermediate risk	TT	241 (49.2)	65 (16.5)	1.59	1.14–2.23	
	TC-CC	346 (46.0)	148 (16.7)	1		
Neutral risk	TT	121 (24.7)	212 (53.8)	1.34	1.01–1.77	
	TC-CC	187 (24.9)	439 (49.7)	1		
Low risk	TT	6 (1.2)	103 (26.1)	1.80	0.63–5.18	
	TC-CC	9 (1.2)	275 (31.1)	1		
Total		1242	1278			

n, Number of cases/controls

[a] HLA risk categories: high risk, *DQA1*03-DQB1*0302/DQA1*05-DQB1*0201* (*DQ8/DQ2*), where *DRB1≠0404*; low risk, at least one *DQB1*0602* allele independent of genotype on the other allele; intermediate risk, *DRB1*0404-DQ8/DR3-DQ2, DR3-DQ2/DR3-DQ2, DR4-DQ8/DR4-DQ8* (excluding homozygous *DRB1 0404*), *DRB1*0401* or *040x-DQ8/X* (*X≠DQB1*0602* or *DR3*). The remaining haplotypes were grouped in the neutral risk category (see Methods)

[b] Represents genotypes III/III and I/III

from a two-locus multiplicative model (3 *df* test; *p*=0.53). This was also the case in the trio data set (ESM Table 4) and was supported by the case-only analysis (ESM Table 5; 3 *df* test; *p*=0.57). Again, there was no association between the two loci among controls (3 *df* test; *p*=0.21).

Joint effects of non-HLA loci There was also no indication of deviation from multiplicative two-locus joint effects of *PTPN22-INS*, *INS-CTLA4* or *PTPN22-CTLA4* in the case-control data (Table 2) or in the case-only analysis (ESM Table 5) (all *p*>0.39). For the trios, the test for interaction between *PTPN22* and *CTLA4* showed *p*=0.046 (ESM Table 6). Taken together with the analysis of the case-control and the case-only data, this weighs against any deviation from a multiplicative two-locus joint effect also of *CTLA4* and *PTPN22*.

Joint effects of more than two susceptibility loci We also tested models with all three-way and four-way interactions involving the four susceptibility loci using logistic regression (categorising all loci in two groups: increased risk genotypes or not), but none of the multi-way interactions were statistically significant (all *p*>0.29). The simultaneous distribution of risk genotypes at all four loci among cases and controls is shown in ESM Table 7. The results show that the more risk loci an individual carries, the higher the relative risk, but the presence or absence of HLA risk loci influences the relative risk much more than the other loci, as expected. For instance, carrying risk genotypes at all three non-HLA loci but not at HLA is associated with a much lower risk than HLA risk genotypes together with low-risk genotypes at all three other loci. The relative risk (OR) conferred by simultaneously carrying high- or moderate-risk HLA and risk genotypes at all the three other loci compared with non-risk-associated genotypes at all four loci was 61. The expected relative risk under a strict multiplicative model involving all four loci was 123 (multiplying all four single-locus effects by each other). The relatively small number of individuals simultaneously carrying all risk genotypes indicates that the observed negative deviation from a four-way multiplicative model was not statistically significant, in accordance with the formal test cited above.
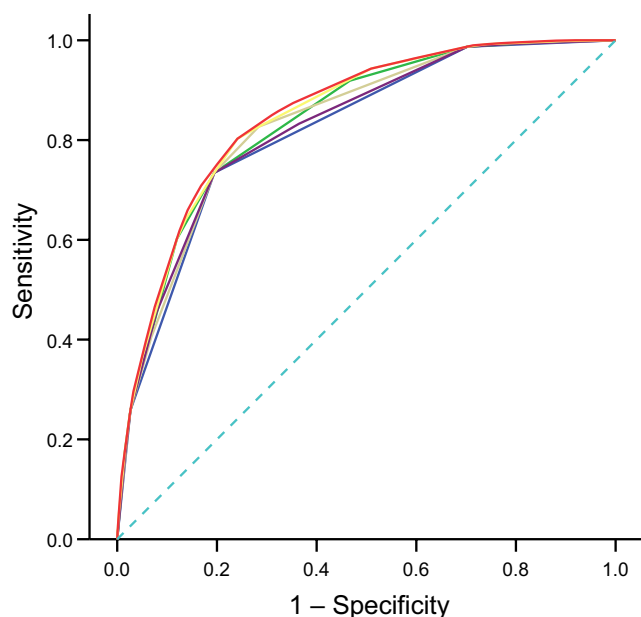
ROC curve Another way to assess the predictive utility of combinations of genetic risk markers is the ROC curve [40]. This utilises the genotypes of all included individuals and assesses the combination of sensitivity and specificity of different combinations of genotypes. ROC curves for HLA alone, pairwise combinations of HLA and non-HLA loci, and multiple genotypes (Fig. 1) showed an area under the curve of 0.82 for HLA alone, which was only marginally increased by adding non-HLA loci.

Table 2 Interaction between *INS-PTPN22*, *INS-CTLA4* and *PTPN22-CTLA4* in the case-control data set using logistic regression

Non-HLA genotypes		Cases	Controls	OR	95% CI	Test for interaction[a] (*p* value)
		n (%)	*n* (%)			
INS	PTPN					0.67
I-I	TT+TC	274 (72.9)	156 (58.4)	1.74	1.38–2.18	
I-I	CC	594 (71.6)	587 (54.8)	1		
III+	TT+TC	102 (27.1)	111 (41.6)	1.89	1.38–2.58	
III+	CC	236 (28.4)	485 (45.2)	1		
Total (*n*)		1,206	1,339			
INS	CTLA4					0.42
I-I	TT	354 (73.1)	222 (53.9)	1.51	1.23–1.86	
I-I	TC-CC	520 (71.0)	492 (54.8)	1		
III+	TT	130 (26.9)	190 (46.1)	1.31	0.99–1.73	
III+	TC-CC	212 (29.0)	405 (45.2)	1		
Total (*n*)		1,216	1,309			
PTPN	CTLA4					0.78
TT+TC	TT	143 (29.7)	75 (17.7)	1.46	1.04–2.05	
TT+TC	TC-CC	239 (31.7)	183 (20.1)	1		
CC	TT	339 (70.3)	348 (82.3)	1.4	1.15–1.67	
CC	TC-CC	514 (68.3)	729 (79.9)	1		
Total (*n*)		1,235	1,335			

The low-risk genotypes were used as reference (CC, TC-CC, TC-CC)

[a] Likelihood ratio tests of whether the OR conferred by one locus is significantly different over strata defined by genotypes in the other locus

n, Number of cases/controls

Fig. 1 ROC curve for HLA genotypes in four categories and for combinations of genotypes defined by HLA and non-HLA susceptibility loci. The area under the curve (95% confidence interval) was 0.820 (0.803–0.836) for HLA (dark blue line), 0.828 (0.811–0.844) for HLA+CTLA4 (purple line), 0.835 (0.819–0.851) for HLA+PTPN22 (grey line), 0.840 (0.824–0.855) for HLA+INS (green line), 0.848 (0.833–0.863) for HLA+INS+PTPN22 (yellow line) and 0.852 (0.837–0.867) for HLA+INS+PTPN22+CTLA4 (red line). Turquoise dashed line, reference line

Age of disease onset and sex We found no significant deviation from a multiplicative model concerning age–locus and sex–locus interaction for any of the genes. This was confirmed in the trio families and case-only analysis (ESM Tables 8 and 9).

Discussion

The present study is a comprehensive evaluation of joint effects of the four most well established type 1 diabetes susceptibility genes in both a large case-control series and family material. The relative risk conferred by PTPN22 was stronger in the lower-risk HLA categories than in the high-risk HLA category, while all other two-locus combinations (*HLA-INS*, *HLA-CTLA4*, *INS-CTLA4*, *INS-PTPN22* and *PTPN22-CTLA4*) were consistent with multiplicative models. Although model-free methods have been developed for gene–gene interaction studies, such as multifactor dimensionality reduction (see [1] and references therein), these methods are designed for the detection of novel susceptibility loci, which was not the goal of our investigation.

Two of the three previous studies of the joint effect of PTPN22 and HLA were in accordance with our results [25, 27] while the other study found no deviation from a multiplicative model [22]. It should be noted that the interaction between PTPN22 and HLA found in the case-control material and case-only analysis was not replicated in our trio data. One of the reasons for this could be lower statistical power in the trio data. Using the Quanto program ([41]; http://hydra.usc.edu/gxe), we found that we had more than 80% power to detect significant two-way gene–gene interaction if the true interaction parameter was 0.5. For the trio design we would need as many trios as we had cases in the case-control study to obtain a similar power. Our number of trios was only about a third of the number of cases in the case-control study, with consequently lower power. The case-only design is known to be the most efficient to detect interaction under certain assumptions. For instance, we had >99% power to detect interaction if the true interaction parameter was 0.5.

The few studies concerning two-locus interaction effects between HLA and CTLA4 and among non-HLA genes [22] have generally indicated multiplicative effects, which is in accordance with our results. Some previous studies have found that the relative risk conferred by INS was similar in subgroups defined by HLA susceptibility genes [11, 12], while three studies have indicated that the effect was stronger in the low risk-HLA categories [16–18], a finding that was only partially supported by our data. On the other hand, one relatively small study has found that the effect of INS was confined to the high-risk *HLA-DR4* group [10].

The reason for diverging results of the joint effect of established type 1 diabetes susceptibility genes in the literature could be that the studies have been performed with varying sample sizes and with different study designs. Linkage studies [14, 15], case-control association studies [10, 12, 13, 16] and family trio designs [16, 18, 22, 25] have all been used in these studies. Smaller studies might be inadequate to reveal significant interactions. Different criteria for categorising the HLA risk groups could potentially influence the result of the joint effect of HLA and other type 1 diabetes susceptibility genes. However, our conclusions were not affected by alternative classification of HLA risk groups, such as into *DR4-DQ8* vs *DR3-DQ2* carriers (data not shown). Studies in different populations and ethnic groups have indicated some heterogeneity in HLA-associated risk of type 1 diabetes and it is also possible that gene–gene interactions may vary across populations. However, despite the observed variations in population risk of type 1 diabetes and in HLA haplotype frequencies across populations, the relative predisposing effects of HLA haplotypes seem to be consistent across populations [42]. In our study all the patients were diagnosed before 15 years of age. The fact that the relative risks associated with both risk genotypes and low-risk genotypes seem to diminish with age above 15 years [43] raises the question whether gene–gene interactions may also differ in different age-groups.

Although no preventive intervention is available for type 1 diabetes today, prediction of disease is an important part of strategies for prevention, both for recruitment of participants for research studies and for identification of target populations for future preventive interventions. Understanding the interacting effect of the established type 1 diabetes susceptibility genes will increase this possibility. In a multiplicative model the relative risk (RR) for a person holding a high-risk genotype at both loci compared with a person with low risk at both loci will be $RR_{locus1} \times RR_{locus2}$. In terms of absolute risk differences, a doubling or tripling of risk due to *INS* or *PTPN22* would be greater for a person with a high-risk HLA genotype than it would for someone with a low-risk HLA genotype. The absolute risk for persons with a given genotype can be estimated by multiplying the average cumulative incidence in the population (0.42% cumulative risk up to age 15 years in Norway [44]) by the ratio of genotype frequency in patients and genotype frequency in controls [18], or using Bayes' formula. For instance, for a person with a low-risk HLA genotype, a high- or low-risk *INS* genotype will define whether the estimated absolute risk is approximately 0.01% or 0.028% (absolute risk difference, 0.027%), whereas for those with the high-risk HLA genotype *INS* will define an estimated risk of approximately 3.0% or 4.7% (absolute risk difference 1.7%).

As discussed in the general setting by Janssens et al. [40], increasing the number of susceptibility loci considered simultaneously generally increases the predictive value for disease. The downside is that the proportion of the population simultaneously carrying multiple risk alleles becomes minute even with a moderate number of susceptibility polymorphisms, and that even with relatively large data sets, as in our study, the absolute risk estimate becomes imprecise. The high-risk HLA genotype is carried by fewer than 3% of population controls, but confers a very high risk of disease. Several practical and scientific aspects of prediction should be considered when evaluating the utility of different prediction regimes. The ROC curve analysis confirms that, despite the higher absolute risk for those few with combinations of several risk markers, adding non-HLA genetic markers only marginally increases the utility of the prediction over that of HLA alone. While up to six susceptibility loci in addition to those studied here have recently been established in type 1 diabetes [45], the magnitude of the effect for each additional locus is very much smaller than that of HLA and even smaller than that of *INS* and *PTPN22*, suggesting that they are likely to add only marginally to the prediction of disease in individuals. Furthermore, an informal assessment of the number needed to be genetically screened in order to obtain a cohort of high-risk individuals, which will give rise to a given number of cases of type 1 diabetes, and the costs connected to the genotyping also suggest limited cost-effectiveness in adding non-HLA genetic markers to the prediction regime (data not shown).

In conclusion, in this comprehensive study of interactions among established type 1 diabetes susceptibility genes, we found that the joint effect of HLA and *PTPN22* was significantly less than multiplicative in the case-control material, while a multiplicative model could not be rejected for *HLA-INS*, *HLA-CTLA*, *PTPN-INS*, *INS-CTLA4* and *PTPN-CTLA4*. Despite near-multiplicative effects for most loci, and the fact that groups with very high relative risk of type 1 diabetes can be identified by testing for multiple susceptibility genes, only a small proportion of the population (and cases with type 1 diabetes) simultaneously carry HLA and multiple non-HLA susceptibility genotypes.

References

1. Moore JH (2003) The ubiquitous nature of epistasis in determining susceptibility to common human diseases. Hum Hered 56:73–82
2. Cordell HJ (2002) Epistasis: what it means, what it doesn't mean, and statistical methods to detect it in humans. Hum Mol Genet 11:2463–2468
3. Li W, Reich J (2000) A complete enumeration and classification of two-locus disease models. Hum Hered 50:334–349
4. Hirschhorn JN, Lohmueller K, Byrne E, Hirschhorn K (2002) A comprehensive review of genetic association studies. Genet Med 4:45–61
5. Undlien DE, Lie BA, Thorsby E (2001) HLA complex genes in type 1 diabetes and other autoimmune diseases. Which genes are involved? Trends Genet 17:93–100
6. Ueda H, Howson JMM, Esposito L et al (2003) Association of the T cell regulatory gene CTLA4 with susceptibility to autoimmune disease. Nature 423:506–511
7. Pugliese A, Miceli D (2002) The insulin gene in diabetes. Diabetes Metab Res Rev 18:13–25
8. Bottini N, Musumeci L, Alonso A et al (2004) A functional variant of lymphoid tyrosine phosphatase is associated with type I diabetes. Nat Genet 36:337–338
9. Thompson WD (1991) Effect modification and the limits of biological inference from epidemiologic data. J Clin Epidemiol 44:221–232
10. Julier C, Hyer RN, Davies J et al (1991) Insulin-IGF2 region on chromosome 11p encodes a gene implicated in HLA-DR4-dependent diabetes susceptibility. Nature 354:155–159

11. Bain SC, Prins JB, Hearne CM et al (1992) Insulin gene region-encoded susceptibility to type 1 diabetes is not restricted to HLA-DR4-positive individuals. Nat Genet 2:212–215

12. van der Auwera BJ, Heimberg H, Schrevens AF, Van WC, Flament J, Schuit FC (1993) 5′ Insulin gene polymorphism confers risk to IDDM independently of HLA class II susceptibility. Diabetes 42:851–854

13. She JX, Bui MM, Tian XH et al (1994) Additive susceptibility to insulin-dependent diabetes conferred by HLA-DQB1 and insulin genes. Autoimmunity 18:195–203

14. Dizier MH, Babron MC, Clerget-Darpoux F (1994) Interactive effect of two candidate genes in a disease: extension of the marker-association-segregation chi square method. Am J Hum Genet 55:1042–1049

15. Cordell HJ, Todd JA, Bennett ST, Kawaguchi Y, Farrall M (1995) Two-locus maximum lod score analysis of a multifactorial trait: joint consideration of IDDM2 and IDDM4 with IDDM1 in type 1 diabetes. Am J Hum Genet 57:920–934

16. Metcalfe KA, Hitman GA, Fennessy MJ, McCarthy MI, Tuomilehto J, Tuomilehto-Wolf E (1995) In Finland insulin gene region encoded susceptibility to IDDM exerts maximum effect when there is low HLA-DR associated risk. Diabetologia 38:1223–1229

17. Laine AP, Hermann R, Knip M, Simell O, Akerblom HK, Ilonen J (2004) The human leukocyte antigen genotype has a modest effect on the insulin gene polymorphism-associated susceptibility to type 1 diabetes in the Finnish population. Tissue Antigens 63:72–74

18. Motzo C, Contu D, Cordell HJ et al (2004) Heterogeneity in the magnitude of the insulin gene effect on HLA risk in type 1 diabetes. Diabetes 53:3286–3291

19. Phillips PC (1998) The language of gene interaction. Genetics 149:1167–1171

20. Risch N (1990) Linkage strategies for genetically complex traits. I. Multilocus models. Am J Hum Genet 46:222–228

21. Wade MJ, Winther RG, Agrawal AF, Goodnight CJ (2001) Alternative definitions of epistasis: dependence and interaction. Trends Ecol Evol 16:498–504

22. Smyth D, Cooper JD, Collins JE et al (2004) Replication of an association between the lymphoid tyrosine phosphatase locus (LYP/PTPN22) with type 1 diabetes, and evidence for its role as a general autoimmunity locus. Diabetes 53:3020–3023

23. van der Auwera B, Vandewalle CL, Schuit FC et al (1997) CTLA-4 gene polymorphism confers susceptibility to insulin-dependent diabetes mellitus (IDDM) independently from age and from other genetic or immune disease markers. Clin Exp Immunol 110:98–103

24. Felner EI, Klitz W, Ham M et al (2005) Genetic interaction among three genomic regions creates distinct contributions to early- and late-onset type 1 diabetes mellitus. Pediatr Diabet 6:213–220

25. Hermann R, Lipponen K, Kiviniemi M et al (2006) Lymphoid tyrosine phosphatase (LYP/PTPN22) Arg620Trp variant regulates insulin autoimmunity and progression to type 1 diabetes. Diabetologia 49:1198–1208

26. Kahles H, Ramos-Lopez E, Lange B, Zwermann O, Reincke M, Badenhoop K (2005) Sex-specific association of PTPN22 1858T with type 1 diabetes but not with Hashimoto's thyroiditis or Addison's disease in the German population. Eur J Endocrinol 153:895–899

27. Steck AK, Liu SY, McFann K et al (2006) Association of the PTPN22/LYP gene with type 1 diabetes. Pediatr Diabetes 7:274–278

28. Stene LC, Joner G, the Norwegian Childhood Diabetes Study Group (2003) Use of cod liver oil during the first year of life is associated with lower risk of childhood-onset type 1 diabetes: a large, population-based, case-control study. Am J Clin Nutr 78:1128–1134

29. Stene LC, Thorsby PM, Berg JP et al (2006) The relation between size at birth and risk of type 1 diabetes is not influenced by adjustment for the insulin gene (-23HphI) polymorphism or HLA-DQ genotype. Diabetologia 49:2068–2073

30. EURODIAB ACE Study G (2000) Variation and trends in incidence of childhood diabetes in Europe. Lancet 355:873–876

31. Witsø E, Stene LC, Paltiel L, Joner G, Rønningen KS (2002) DNA extraction and HLA genotyping using mailed mouth brushes from children. Pediatric Diabetes 3:89–94

32. Helland A, Børresen AL, Kristensen G, Rønningen KS (1994) DQA1 and DQB1 genes in patients with squamous cell carcinoma of the cervix: relationship to human papillomavirus infection and prognosis. Cancer Epidemiol Biomarkers Prev 3:479–486

33. Cinek O, Kolouskova S, Snajderova M et al (2001) HLA class II genetic association of type 1 diabetes mellitus in Czech children. Pediatric Diabetes 2:98–102

34. Olerup O, Aldener A, Fogdell A (1993) HLA-DQB1 and -DQA1 typing by PCR amplification with sequence-specific primers (PCR-SSP) in 2 hours. Tissue Antigens 41:119–134

35. Undlien DE, Friede T, Rammensee HG et al (1997) HLA-encoded genetic predisposition in IDDM: DR4 subtypes may be associated with different degrees of protection. Diabetes 46:143–149

36. Piegorsch WW, Weinberg CR, Taylor JA (1994) Non-hierarchical logistic models and case-only designs for assessing susceptibility in population-based case-control studies. Stat Med 13: 153–162

37. Ewens WJ, Spielman RS (1995) The transmission/disequilibrium test: history, subdivision and admixture. Am J Hum Genet 57:455–464

38. Dudbridge F (2003) Pedigree disequilibrium tests for multilocus haplotypes. Genet Epidemiol 25:115–121

39. Skrondal A (2003) Interaction as departure from additivity in case-control studies: a cautionary note. Am J Epidemiol 158:251–258

40. Janssens AC, Pardo MC, Steyerberg EW, van Duijn CM (2004) Revisiting the clinical validity of multiplex genetic testing in complex diseases. Am J Hum Genet 74:585–588

41. Gauderman WJ (2002) Sample size requirements for association studies of gene-gene interaction. Am J Epidemiol 155:478–484

42. Thomson G, Valdes AM, Noble JA et al (2007) Relative predispositional effects of HLA class II DRB1-DQB1 haplotypes and genotypes on type 1 diabetes: a meta-analysis. Tissue Antigens 70:110–127

43. Sabbah E, Savola K, Ebeling T et al (2000) Genetic, autoimmune, and clinical characteristics of childhood- and adult-onset type 1 diabetes. Diabetes Care 23:1326–1332

44. Aamodt G, Stene LC, Njølstad PR, Søvik O, Joner G (2007) Spatiotemporal trends and age-period-cohort modeling of the incidence of type 1 diabetes among children aged <15 years in Norway 1973–1982 and 1989–2003. Diabetes Care 30:884–889

45. Todd JA, Walker NM, Cooper JD et al (2007) Robust associations of four new chromosome regions from genome-wide analyses of type 1 diabetes. Nat Genet 39:857–864