Heterogeneity in the Magnitude of the Insulin Gene Effect on HLA Risk in Type 1 Diabetes

Costantino Motzo,[1] Daniela Contu,[1] Heather J. Cordell,[2] Rosanna Lampis,[1] Mauro Congia,[1] Maria Giovanna Marrosu,[3] John A. Todd,[2] Marcella Devoto,[4] and Francesco Cucca[1,5]

There is still uncertainty concerning the joint action of the two established type 1 diabetes susceptibility loci, the HLA class II *DQB1* and *DRB1* genes (*IDDM1*) and the insulin gene (*INS*) promoter (*IDDM2*). Some previous studies reported independence, whereas others suggested heterogeneity in the relative effects of the genotypes at these disease loci. In this study, we have assessed the combined effects of the *HLA-DQB1/DRB1* and *INS* genotypes in 944 type 1 diabetic patients and 1,023 control subjects, all from Sardinia. Genotype variation at *INS* significantly influenced disease susceptibility in all HLA genotype risk categories. However, there was a significant heterogeneity ($P = 2.4 \times 10^{-4}$) in the distribution of the *INS* genotypes in patients with different HLA genotypes. The *INS* predisposing genotype was less frequent (74.9%) in high-risk HLA genotype–positive patients than in those with HLA intermediate-risk (86.1%) and low-risk (84.8%) categories. Gene-gene interaction modeling led to rejection of the additive model, whereas a multiplicative model showed a better, albeit still partial, fit to the observed data. These genetic results are consistent with an interaction between the protein products of the HLA and *INS* alleles, in which both the affinity of the various HLA class II molecules for a preproinsulin-derived peptide and the levels of this peptide in the thymus act jointly as key regulators of type 1 diabetes autoimmunity. *Diabetes* 53:3286–3291, 2004

From the [1]Dipartimento di Scienze Biomediche e Biotecnologie, Universita' di Cagliari, Sardinia, Italy; the [2]Juvenile Diabetes Research Foundation/Wellcome Trust Diabetes and Inflammation Laboratory, Cambridge Institute for Medical Research, University of Cambridge, Addenbrooke's Hospital, Cambridge, U.K.; the [3]Centro Sclerosi Multipla, Dipartimento di Neuroscienze, Universita' di Cagliari, Sardinia, Italy; the [4]Department of Research, Nemours Children's Clinic, Wilmington, DE, and the Dipartimento di Oncologia, Biologia, e Genetica, Universita' di Genova, Genova, Italy; and the [5]Centro di Genetica Clinica, Dipartimento di Scienze Biomediche, Universita' di Sassari, Sassari, Italy.

Address correspondence and reprint requests to Francesco Cucca, Universita' di Cagliari, Dipartimento di Scienze Biomediche e Biotecnologie, Via Jenner, Cagliari, 09124 Italy. E-mail: fcucca@mcweb.unica.it.

Received for publication 26 May 2004 and accepted in revised form 3 September 2004.

C.M. and D.C. contributed equally to this work.

Additional information for this article can be found in an online appendix at http://diabetes.diabetesjournals.org.

AFBAC, affected family-based control subject; AR, absolute risk; MHC, major histocompatibility complex; P/C ratio, patient/control ratio; POR, pairwise odds ratio; PPI, preproinsulin.

It is believed that most cases of type 1 diabetes result from a T-cell–dependent selective destruction of the insulin-producing pancreatic β-cells and subsequent irreversible insulin deficiency. The disease is caused by predisposing genetic factors in the presence of a permissive environment. Whole genome linkage scans have shown that the major histocompatibility complex (MHC)/HLA region on chromosome 6p21 contains the major genetic component of the disease (*IDDM1*) (1,2). Within the HLA complex, variation at the *HLA-DQB1* and *-DRB1* loci dominates the association with the disease (3,4). The two loci act as a complex superlocus, with both haplotype- and genotype-specific effects and with additional modifying effects due to variation at other HLA loci (5–9).

Another established disease locus, *IDDM2*, has been mapped to chromosome 11p15.5 and carries a relatively more modest but clearly defined genetic effect (10). Allelic association and functional studies have shown that within *IDDM2*, a minisatellite (VNTR) locus in the insulin gene (*INS*) promoter region is likely to represent the etiologic polymorphism (10–17). Differences in length occur in three discrete classes of VNTR alleles. The shorter class of alleles, named class I, is positively associated with type 1 diabetes, whereas the longer class of alleles, named class III, is negatively associated. Experimental evidence suggests that the VNTR may bind transcriptional regulatory proteins depending on the sequence of the particular VNTR allele (18).

A large body of studies supports the direct involvement of *HLA-DRB1*, *DQB1*, and *INS*-VNTR variation in type 1 diabetes. In the case of *DRB1-DQB1*, a correlation of polymorphic amino acids in the peptide-binding active site of the molecules with susceptibility and resistance to disease was observed (19,20). There are also marked structural cross-species similarities between MHC class II molecules respectively with a predisposing (20–22) and protective effect (20) in human and mouse type 1 diabetes. Data from transgenic animal models (22,23) indicate direct MHC class II–mediated effects on the T-cell repertoire in the thymus on disease susceptibility. Susceptibility may arise through positive selection of autoreactive T-cells by positively associated class II alleles, whereas protection may occur via negative selection of diabetogenic T-cells and/or selection of T regulatory cells, mediated by negatively associated alleles. These events in the thymus could affect both the CD4 and CD8 T-cell repertoires (24,25).

The *INS*-VNTR locus affects expression of the insulin

gene and its precursors in the thymus, consistent with a model in which either positive selection or negative selection of autoreactive T-cell clones against epitopes from preproinsulin (PPI) occurs (12–15). It has been suggested that a two- to threefold decrease in the amount of PPI in the thymus caused by *INS*-VNTR class I alleles can explain their positive association with the disease (12,13,15). This model was greatly strengthened by the discovery that mutations in the transcription factor AIRE cause autoimmune disease in humans, including type 1 diabetes (16,26), and in mice (17) by lowering the expression of peripheral antigens such as PPI in the thymus.

Overall, the biological pathways leading to type 1 diabetes are not fully understood and the nature of the interaction of the products of HLA class II and *INS*-VNTR in the disease process is still unclear.

There is also uncertainty concerning the interaction of the HLA and *INS* loci, both in the results obtained and the nomenclature used. Julier et al. (10), using a French sample set, initially reported evidence of linkage and association of the disease with *INS*-VNTR allelic variation only in *HLA*-DRB1*04–positive patients. This observation supported an etiopathological pathway restricted to a specific HLA class II allotype/PPI peptide complex. However, subsequent reports based on the analysis of U.K. and German samples showed not only that the *INS*-VNTR–encoded susceptibility was independent from *HLA*-DRB1*04, but also that, consistent with the multiplicative model of gene-gene interaction, genotype variation at both loci occurred independently of each other (27,28). Still, other studies, in Belgian (29) and Finnish sample sets (30,31), reported that the predisposing INS class I alleles might have less (30,31) or no (29) effect in individuals with high-risk HLA genotypes, compared with intermediate- and low-risk HLA genotypes. However, it should be noted that the statistical support for the presence of uneven *INS* size effects was limited and that these studies did not take into account the possible confounding effects of population substructure. The uncertainty present in the literature is illustrated by the conflicting way in which the results of previous studies have been reported; the joint action of HLA and *INS* has variously been described as being multiplicative (32,33), additive (34), providing evidence of interaction (32,35), and exhibiting no interaction (31,33,34).

In this study, we analyzed the interaction of HLA and *INS* in a large sample set from Sardinia to help establish the genetic relationship between these two susceptibility loci in the context of contradictory observations of previous reports. The isolated population from the island of Sardinia is particularly well-suited to carry out a two-locus analysis of the type 1 diabetes susceptibility loci, HLA and *INS*. Sardinia is a genetic isolate, and there is no evidence of population substructure within the island (36). In fact, population substructure could cause background differences in the allele frequencies at unlinked loci and thus disrupt the assessment of the gene-gene interactions in both case-control and case-only analyses.

RESEARCH DESIGN AND METHODS

The case subjects consisted of 944 independent type 1 diabetic Sardinian patients including 466 case subjects from a family dataset and 478 additional sporadic patients. The average (arithmetic mean) age at disease onset of the patients (probands) analyzed in this study was 10.6 ± 6.9 years (± SD) with a range of 0.4–40 years and a median of 10 years. The control set consisted of 634 affected family-based control subjects (AFBACs) (see STATISTICAL ANALYSIS below) and 389 blood donors whose average age at the sample collection was 36.2 ± 9.3 years with a range of 19–58 years and a median of 35 years. The blood donors were healthy volunteers who have been selected based on their Sardinian origin avoiding any bias for cultural status, religion, or social condition.

HLA typing. The whole sample set was typed through PCR amplification of the polymorphic second exon of the *HLA-DRB1* and *-DQB1* genes and dot blot analysis of amplified DNA with sequence-specific oligonucleotide probes as described previously (37–41). Based on the almost complete linkage disequilibrium between *INS*-VNTR allele classes and the −23 *Hph*I single nucleotide polymorphism in Caucasians (10,42), we attributed the *INS*-VNTR genotype status by genotyping the −23 *Hph*I single nucleotide polymorphism as a surrogate marker for the VNTR.

Statistical analysis. The association of both the *DRB1-DQB1 (IDDM1)* and *INS*-VNTR (*IDDM2*) genotypes with type 1 diabetes was assessed using a case-control design. The control dataset included 389 blood donors and 634 AFBAC genotypes (43). The AFBAC genotypes were assembled from 352 simplex families with type 1 diabetes and 282 simplex families with multiple sclerosis by selecting in each family the alleles or haplotypes not transmitted from the parents to the affected child at both *HLA-DRB1–DQB1* and *INS*-VNTR (43). In this way, it is possible to assemble in each simplex family a pseudo-control carrying one genotype at *HLA-DRB1-DQB1* and one genotype at *INS*-VNTR. Note that the blood donor and AFBAC sample sets were both in Hardy-Weinberg equilibrium and showed similar allele and genotype frequencies at the loci considered in this study (data not shown). The individual and joint frequencies of the HLA class II and *INS*-VNTR genotypes observed in patients and control subjects were compared by estimating absolute risk (AR), patient/control (P/C) ratio, and pairwise odds ratio (POR). The AR is obtained by multiplying the P/C ratio, defined as the frequency in the patients divided by the frequency in the control subjects, by the observed disease prevalence. In this study, we considered a type 1 diabetes prevalence of 0.459 per 100 that has been previously established in the 0–29 years age range in one of the Sardinian provinces (44). Accordingly, the AR represents the number of individuals per 100 carrying a given genotype or combination of genotypes who will develop the disease in the 0–29 years range.

When a disease is associated with more than one marker (allele, haplotype, or genotype) at a given locus, the association of a test marker is influenced by the other associated markers. In this respect, the standard odds ratios (ORs), computed comparing one marker against all the others grouped, are not appropriate for an accurate estimate of the strength of the disease association. To minimize this problem and to provide a more reliable computation of the relative risk of disease, in this study, we used PORs in which the various genotypes are analyzed relative to one reference genotype. The resulting data points are arranged in a 2×2 contingency table and tested by Pearson's χ^2 test. In this study, we used as the reference genotype the baseline genotype category (represented by the genotype with the lower P/C ratio). In the multilocus analysis, the PORs were also computed by comparing genotype variation at one locus (using the baseline risk category as reference) within each different genotype categories at the other locus.

In the family dataset, we also adapted the transmission/disequilibrium test (45) to test the null hypothesis of equality of transmission of the *INS* allelic variants conditional on the HLA genotype risk category of the patients. Specifically, the transmission and untransmission counts for the *INS* variants were arranged in a 3×2 contingency table and tested for heterogeneity by Pearson's χ^2 test. Only the probands were evaluated in the families with more than one affected sibling in all the association tests performed in this study.

Multilocus statistical models. We tested a multiplicative as well as an additive model of gene-gene interaction. The models are defined as originally described by Risch (46) and have been discussed subsequently (47,48). Briefly, in the multiplicative model of epistasis, the probability of developing disease due to genotypes at one locus increases or decreases by a factor (the multiplicative factor) that is constant (i.e., does not vary according to genotype variation and relative risks provided by the other locus). In the additive model, the probability of developing disease due to genotypes at one locus does not increase or decrease by a constant factor but increases or decreases by a constant amount. We fitted multiplicative models to the observed case/control data using standard software for logistic regression and fitted additive models to the observed case/control data by direct maximization of the prospective likelihood. In this way, we actually fitted multiplicative/additive models for the ORs, which under a rare disease assumption correspond to multiplicativity/additivity on the penetrance scale. The restricted models (multiplicative or additive) were compared with a saturated model in which the ORs are not restricted using a likelihood ratio test.

TABLE 1
Individual associations of the *HLA-DRB1/DQB1* and *INS* genotypes with type 1 diabetes in the Sardinian population

	Patients	Control subjects	P/C ratio	AR	POR	95% CI	*P*
DRB1-DQB1							
High risk	574 (60.8)	77 (7.5)	8.1	3.71	95.3	65.0–139.7	3.5×10^{-185}
Intermediate risk	324 (34.3)	358 (35.0)	1.0	0.45	11.6	8.3–16.2	3.1×10^{-59}
Low risk	46 (4.9)	588 (57.5)	0.08	0.04	1		
Total	944	1,023					
INS							
I/I	748 (79.2)	662 (64.7)	1.2	0.56	2.7	1.6–4.7	1.7×10^{-4}
I/III	177 (18.8)	315 (30.8)	0.6	0.28	1.4	0.8–2.4	>0.05
III/III	19 (2.0)	46 (4.5)	0.4	0.21	1		
Total	944	1,023					

Data are *n* (%) unless otherwise indicated. PORs were calculated using as reference genotypes the baseline categories (the HLA low-risk genotype for *HLA-DRB1/DQB1* and the *VNTR* class III/III genotype for *INS*, respectively).

RESULTS

We initially ranked the two-locus *HLA-DRB1–DQB1* genotypes into three main risk categories based on their P/C ratios (see the online appendix at http://diabetes.diabetesjournals.org): *1*) high-risk or genotypes with a P/C ratio ≥3.5, *2*) intermediate risk or genotypes with a P/C ratio <3.5 and >0.35, and *3*) low risk or genotypes with a P/C ratio ≤0.35 (Table 1). The high-risk category included various combinations of DRB1*03-DQB1*0201 and DRB1*04-DQB1*0302 or DRB1*04-DQB1*0201 (in which DRB1*04 is equal to any DRB1*04 subtype different from DRB1*0403) high-risk haplotypes and overall shows a P/C ratio of 8.1 with an AR of developing the disease in the 0–29 years age-group equal to 3.71%. The intermediate-risk category is constituted mainly by individuals carrying genotypes given by combinations of high-risk and permissive-neutral haplotypes as well as rare genotypes with sparse counts; taken as a whole, it provides a P/C ratio of 1.0 with an AR of 0.45% (which is virtually the same as the general untyped 0–29 years Sardinian population). Finally, the low-risk category encompasses genotypes constituted by two copies of negatively associated haplotypes as well as combinations of negatively associated and neutral haplotypes and overall shows a P/C ratio of 0.08 with an AR of 0.039%. POR computed using the HLA low-risk category as a baseline reference genotype shows values of 95.3 (95% CI 65.0–139.7, $P = 3.5 \times 10^{-185}$) and 11.6 (8.3–16.2, $P = 3.1 \times 10^{-59}$), respectively, for the high- and intermediate-risk genotype categories. Similarly, albeit with a much lower genetic effect, the association of the various *INS* genotypes can be grouped into three classes (Table 1): class I/I VNTR showing a P/C ratio of 1.2 with an AR of 0.56%, class I/III VNTR exhibiting a P/C ratio of 0.6 with an AR of 0.28%, and class III/III VNTR with a P/C ratio of 0.4 and an AR of 0.21%. The two negatively associated genotypes were considered here as an individual category, referred to as class III–positive (III+) genotypes.

Once the associations of the genotypes at the HLA class II and *INS* loci were established individually, we evaluated the net effects of variation at these loci jointly (Table 2). Variation at *INS* is only able to affect the magnitude, but not the direction, of the positive association in the HLA high-risk category with a P/C ratio of 9.7 and an AR of 4.46% in individuals with the I/I genotype against a P/C ratio of 5.4 and an AR of 2.47% in individuals with the class III+ genotypes (in comparison with a population prevalence of 0.46%). Likewise, class I alleles at *INS* only modulate but do not override the negative association observed in individuals carrying HLA low-risk genotypes with ARs ranging from 0.050% in individuals with the I/I genotype (P/C ratio of 0.11) to 0.017% (P/C ratio of 0.04) in individuals with the class III+ genotypes. In contrast, genotype variation at *INS* changes the direction of the disease association in the HLA intermediate-risk category (i.e., whether the risk is increased or decreased relative to the population AR of 0.45%), with the I/I genotype showing a positive association (AR = 0.61%, P/C ratio of 1.3) and the class III–positive genotypes having a negative association with type 1 diabetes (AR = 0.17%, P/C ratio of 0.4) (Table 2).

We also evaluated the relative impact of variation at *INS* within each genotype category at the *HLA-DRB1* and *-DQB1* loci. The *INS* effects were detectable in all of the HLA genotype categories with PORs of *INS* class I/I rela-

TABLE 2
Joint associations of the *HLA-DRB1/DQB1* and *INS* genotypes with type 1 diabetes in the Sardinian population

	INS	Patients	Control subjects	P/C ratio	AR	POR	95% CI	*P*
DRB1-DQB1								
High risk	I-I	430 (45.6)	48 (4.7)	9.7	4.46	1.8	1.1–3.0	1.9×10^{-2}
High risk	III+	144 (15.3)	29 (2.8)	5.4	2.47	1		
Intermediate risk	I-I	279 (29.6)	226 (22.1)	1.3	0.61	3.6	2.5–5.3	8.1×10^{-12}
Intermediate risk	III+	45 (4.8)	132 (12.9)	0.4	0.17	1		
Low risk	I-I	39 (4.1)	388 (37.9)	0.11	0.050	2.9	1.3–6.5	8.8×10^{-3}
Low risk	III+	7 (0.7)	200 (19.6)	0.04	0.017	1		
Total		944	1,023					

Data are *n* (%) unless otherwise indicated. PORs were calculated using as reference the *INS-VNTR* class III+ genotype category within each HLA genotype risk category.

TABLE 3
Transmission of the *INS* class I allele from heterozygous parents to type 1 diabetic children conditional on their HLA genotype risk category

	INS class I			
	T	NT	% T	*P*
DRB1/DQB1				
High risk	78	53	59.5	2.9×10^{-2}
Intermediate risk	57	20	74.0	2.5×10^{-5}
Low risk	9	2	81.8	3.5×10^{-2}

NT, not transmitted; T, transmitted.

tive to *INS* class III+ equal to 1.8 (95% CI 1.1–3.0, $P = 1.9 \times 10^{-2}$) in individuals with high-risk HLA, 3.6 (2.5–5.3, $P = 8.1 \times 10^{-12}$) in individuals with intermediate-risk HLA, and 2.9 (1.3–6.5, $P = 8.8 \times 10^{-3}$) in individuals with low-risk HLA genotypes (Table 2). We did not observe any significant differences in the distribution of *HLA-DRB1/DQB1* and *INS* genotypes according to age of onset of type 1 diabetes nor in the relative associations of *INS* genotypes conditional on HLA genotypes when the patients were divided into two different groups according to their age of disease onset (0–14 vs. 15–30 years; see online appendix). When the goodness-of-fit of the observed data to proposed models of gene-gene interaction was evaluated, an additive model of interaction could be clearly rejected ($P < 1 \times 10^{-6}$). However, we also noted that the *INS* relative risks were uneven or heterogeneous across the three HLA risk categories, and, in fact, a multiplicative model of gene-gene interaction was only marginally accepted ($P = 0.09$ against a multiplicative model). We also analyzed the transmission of the *INS* class I allele in the type 1 diabetes families (Table 3). The transmission of the *INS* class I allele, albeit significantly increased over the random expectation of 50% (Table 3), was significantly heterogeneous in patients belonging to the three different HLA risk categories ($\chi^2 = 5.8$, 2 degrees of freedom [df], $P = 0.05$). These observations suggested unequal *INS* size effects conditional on the HLA genotypes of the patients. We therefore further investigated this issue in a more powerful and robust case-only analysis (49) by comparing the distribution of the *INS* genotypes in the three groups of patients belonging to different HLA risk categories (Table 2). We found significant heterogeneity in the distribution of the *INS* genotypes in these three classes of case subjects, thus showing strong evidence against a multiplicative model ($\chi^2 = 16.7$, 2 df, $P = 2.4 \times 10^{-4}$). Analysis of the basis of this heterogeneity showed that the *INS* predisposing genotype is less common (74.9%) in high-risk HLA genotype–positive patients than in those carrying HLA intermediate- (86.1%) and low-risk (84.8%) genotypes. Importantly, no evidence of heterogeneity was observed in the distribution of the *INS* genotypes in a control-only analysis ($P = 0.61$) in which the distribution of the *INS* genotype was homogeneous in control subjects within different HLA risk categories.

DISCUSSION

In human type 1 diabetes, two unlinked disease loci, *HLA-DRB1/DQB1* and *INS*-VNTR, have been unequivocally established, offering one of the few current opportunities to evaluate gene-gene interactions in a complex multifactorial disorder. We have therefore used a large sample set from the homogeneous population of Sardinia to assess the joint genetic effects of genotype variation at these loci and to attempt to clarify the uncertainty derived from previous studies on different populations. Our results, in agreement with some of the previous studies (27,28,31), prove that the *INS* genotype significantly influences type 1 diabetes risk in all HLA genotype risk categories. However, confirming earlier observations suggesting a heterogeneity in the relative effects of *INS* (29,30), we also provide persuasive evidence that the *INS* predisposing genotype is significantly less frequent in high-risk HLA genotype–positive patients than in those with HLA intermediate- and low-risk categories. Thus, our results highlight a particular feature in the interactions between *INS* and HLA: the effects of *INS* on type 1 diabetes risk are detectable in all of the HLA genotype risk categories, but at the same time, these effects are less pronounced in individuals carrying HLA high-risk genotypes. Similar findings were also obtained, albeit with a lower statistical support, in the Finnish population (31).

When these data were evaluated in the context of the statistical models of gene-gene interaction, the additive model was rejected. However, the multiplicative model is also inadequate in explaining the gene-gene interaction and does not reflect the complexity of the molecular interaction of the protein products of the *INS* and HLA class II genes.

These genetic data could be interpreted in terms of a gradient in the strength of binding the various HLA class II allotypes with a PPI peptide influencing the T-cell avidity for the resulting HLA-PPI peptide complex, thus affecting positive selection, T regulatory cell selection, and negative selection in the thymus. Notably, our results suggest that in the presence of high-risk HLA class II allotypes, the generation and tolerance of autoreactive T-cell clones against PPI would be less affected by variation in the amount of PPI in the thymus than in the presence of intermediate- to low-risk allotypes. However, interpreting the biological mechanisms underlying the HLA-*INS* interactions in type 1 diabetes only in terms of the effects of the products of these genes in the thymus is oversimplistic. Indeed, the functional consequences of *INS*-VNTR variation in type 1 diabetes susceptibility might include extrathymic immunological and metabolic effects. Furthermore, the joint action of *INS* and HLA variation might also be influenced by additional genetic and environmental factors, which could vary in different populations. The complex and erratic nature of these interactions and the use of sample sizes inadequate to detect subtle differences in the size effects of variants at one locus conditional on variants at another locus might help explain the contradictory results of previous studies.

ACKNOWLEDGMENTS

We thank the Juvenile Diabetes Research Foundation, the Wellcome Trust, the Italian Telethon, and the Regione Autonoma Sardegna Assessorato Sanita' for financial support. J.A.T. and F.C. are recipients of a Wellcome Trust Biomedical Research Collaboration Grant.

We wish to thank Antonio Cao, Bryan Barratt, Iain

Eaves, and Cristiana Meloni for help and advice; Efisio Angius, Mario Maioli, Paola Frongia, Margi Chessa, and Rossella Ricciardi for help in collecting the Sardinian type 1 diabetes families and for clinical information; and Maria Melis and Antonella Deidda for drawing blood from the patients and their relatives.

REFERENCES

1. Cox NJ, Wapelhorst B, Morrison VA, Johnson L, Pinchuk L, Spielman RS, Todd JA, Concannon P: Seven regions of the genome show evidence of linkage to type 1 diabetes in a consensus analysis of 767 multiplex families. *Am J Hum Genet* 69:820–830, 2001
2. Nerup J, Pociot F: A genomewide scan for type 1-diabetes susceptibility in Scandinavian families: identification of new loci with evidence of interactions. *Am J Hum Genet* 69:1301–1313, 2001
3. Herr M, Dudbridge F, Zavattari P, Cucca F, Guja C, March R, Campbell RD, Barnett AH, Bain SC, Todd JA, Koeleman BP: Evaluation of fine mapping strategies for a multifactorial disease locus: systematic linkage and association analysis of IDDM1 in the HLA region on chromosome 6p21. *Hum Mol Genet* 9:1291–1301, 2000
4. Zavattari P, Lampis R, Mulargia A, Loddo M, Angius E, Todd JA, Cucca F: Confirmation of the DRB1-DQB1 loci as the major component of IDDM1 in the isolated founder population of Sardinia. *Hum Mol Genet* 9:2967–2972, 2000
5. Erlich HA, Rotter JI, Chang JD, Shaw SJ, Raffel LJ, Klitz W, Bugawan TL, Zeidler A: Association of HLA-DPB1*0301 with IDDM in Mexican-Americans. *Diabetes* 45:610–614, 1996
6. Noble AJ, Valdes AM, Cook M, Klitz W, Thomson G, Erlich HA: The role of HLA classs II genes in insulin-dependent diabetes mellitus: molecular analysis of 180 Caucasian, multiplex families. *Am J Hum Genet* 59:1134–1148, 1996
7. Lie BA, Todd JA, Pociot F, Nerup J, Akselsen HE, Joner G, Dahl-Jorgensen K, Ronningen KS, Thorsby E, Undlien DE: The predisposition to type 1 diabetes linked to the human leukocyte antigen complex includes at least one non-class II gene. *Am J Hum Genet* 64:793–800, 1999
8. Cucca F, Dudbridge F, Loddo M, Mulargia AP, Lampis R, Angius E, De Virgiliis S, Koeleman BP, Bain SC, Barnett AH, Gilchrist F, Cordell H, Welsh K, Todd JA: The HLA-DPB1–associated component of the IDDM1 and its relationship to the major loci HLA-DQB1, -DQA1, and -DRB1. *Diabetes* 50:1200–1205, 2001
9. Zavattari P, Lampis R, Motzo C, Loddo M, Mulargia A, Whalen M, Maioli M, Angius E, Todd JA, Cucca F: Conditional linkage disequilibrium analysis of a complex disease superlocus, IDDM1 in the HLA region, reveals the presence of independent modifying gene effects influencing the type 1 diabetes risk encoded by the major HLA-DQB1, -DRB1 disease loci. *Hum Mol Genet* 10:881–889, 2001
10. Julier C, Hyer RN, Davies J, Merlin F, Soularu P, Briant L, Cathelineau G, Deschamps I, Rotter JI, Froguel P, Boitard C, Bell JI, Lathrop GM: Insulin-IGF2 region on chromosome 11p encodes a gene implicated in HLA-DR4-dependent diabetes susceptibility. *Nature* 354:155–159, 1991
11. Barratt BJ, Payne F, Lowe CE, Hermann R, Healy BC, Harold D, Concannon P, Gharani N, McCarthy MI, Olavasen MG, McCormack R, Guja C, Ionescu-Tirgoviste C, Undlien D, Ronningen K, Gillespie KM, Tuomilehto-Wolf E, Tuomilehto J, Bennett ST, Clayton D, Cordell H, Todd JA: Remapping the insulin gene/IDDM2 locus in type1 diabetes. *Diabetes* 53:1884–1889, 2004
12. Pugliese A, Zeller M, Fernandez A, Zalcberg LJ, Bartlett RJ, Ricordi C, Pietropaolo M, Eisenbarth GS, Bennett ST, Patel DD: The insulin gene is transcribed in the human thymus and transcription levels correlated with allelic variation at the INS VNTR-IDDM2 susceptibility locus for type 1 diabetes. *Nat Genet* 15:293–297, 1997
13. Vafiadis P, Bennett ST, Todd JA, Nadeau J, Grabs R, Goodyer CG, Wickramasinghe S, Colle E, Polychronakos C: Insulin expression in human thymus is modulated by INS VNTR alleles at the IDDM2 locus. *Nat Genet* 15:289–292, 1997
14. Chentoufi AA, Polychronakos C: Insulin expression levels in the thymus modulate insulin-specific autoreactive T-cell tolerance: the mechanism by which the *IDDM2* locus may predispose to diabetes. *Diabetes* 51:1383–1390, 2002
15. Chentoufi AA, Palumbo M, Polychronakos C: Proinsulin expression by Hassall's corpuscles in the mouse thymus. *Diabetes* 53:354–359, 2004
16. Consortium TF-GA: An autoimmune disease, APECED, caused by mutations in a novel gene featuring two PHD-type zinc-finger domains: the Finnish-German APECED Consortium: Autoimmune Polyendocrinopathy-Candidiasis-Ectodermal Dystrophy. *Nat Genet* 17:399–403, 1997
17. Anderson MS, Venanzi ES, Klein L, Chen Z, Berzins SP, Turley SJ, von Boehmer H, Bronson R, Dierich A, Benoist C, Mathis D: Projection of an immunological self shadow within the thymus by the aire protein. *Science* 298:1395–1401, 2002
18. Kennedy C, German M, Rutter W: The minisatellite in the diabetes susceptibility locus IDDM2 regulates insulin transcription. *Nat Genet* 9:293–298, 1995
19. Todd JA, Bell JI, McDevitt HO: HLA-DQ beta gene contributes to susceptibility and resistance to insulin-dependent diabetes mellitus. *Nature* 329:599–604, 1987
20. Cucca F, Lampis R, Congia M, Angius E, Nutland S, Bain SC, Barnett AH, Todd JA: A correlation between the relative predisposition of MHC class II alleles to type 1 diabetes and the structure of their proteins. *Hum Mol Genet* 10:2025–2037, 2001
21. Nepom BS, Nepom GT, Coleman M, Kwok WW: Critical contribution of beta chain residue 57 in peptide binding ability of both HLA-DR and -DQ molecules. *Proc Natl Acad Sci U S A* 93:7202–7206, 1996
22. Wen L, Wong FS, Tang J, Chen NY, Altieri M, David C, Flavell R, Sherwin R: In vivo evidence for the contribution of human histocompatibility leukocyte antigen (HLA)-DQ molecules to the development of diabetes. *J Exp Med* 191:97–104, 2000
23. Schmidt D, Verdaguer J, Averill N, Santamaria P: A mechanism for the major histocompatibility complex-linked resistance to autoimmunity. *J Exp Med* 186:1059–1075, 1997
24. Serreze DV, Holl TM, Marron MP, Graser RT, Johnson EA, Choisy-Rossi C, Slattery RM, Lieberman SM, DiLorenzo TP: MHC class II molecules play a role in the selection of autoreactive class I-restricted CD8 T cells that are essential contributors to type 1 diabetes development in nonobese diabetic mice. *J Immunol* 172:871–879, 2004
25. Morgan DJ, Nugent CT, Raveney BJ, Sherman LA: In a transgenic model of spontaneous autoimmune diabetes, expression of a protective class II MHC molecule results in thymic deletion of diabetogenic CD8+ T cells. *J Immunol* 172:1000–1008, 2004
26. Nagamine K, Peterson P, Scott HS, Kudoh J, Minoshima S, Heino M, Krohn KJ, Lalioti MD, Mullis PE, Antonarakis SE, Kawasaki K, Asakawa S, Ito F, Shimizu N: Positional cloning of the APECED gene. *Nat Genet* 17:393–398, 1997
27. Bain SC, Prins JB, Hearne CM, Rodrigues NR, Rowe BR, Pritchard LE, Ritchie RJ, Hall JR, Undlien DE, Ronningen KS: Insulin gene region-encoded susceptibility to type 1 diabetes is not restricted to HLA-DR4-positive individuals. *Nat Genet* 2:212–215, 1992
28. Walter M, Albert E, Conrad M, Keller E, Hummel M, Ferber K, Barratt BJ, Todd JA, Ziegler AG, Bonifacio E: IDDM2/insulin VNTR modifies risk conferred by IDDM1/HLA for development of type 1 diabetes and associated autoimmunity. *Diabetologia* 46:712–720, 2003
29. van der Auwera B, Schuit F, Lyaruu I, Falorni A, Svanholm S, Vandewalle CL, Gorus FK: Genetic susceptibility for insulin-dependent diabetes mellitus in Caucasians revisited: the importance of diabetes registries in disclosing interactions between HLA-DQ- and insulin gene-linked risk. Belgian Diabetes Registry. *J Clin Endocrinol Metab* 80:2567–2573, 1995
30. Metcalfe KA, Hitman GA, Fennessy MJ, McCarthy MI, Tuomilehto J, Tuomilehto-Wolf E: In Finland insulin gene region encoded susceptibility to IDDM exerts maximum effect when there is low HLA-DR associated risk: DiMe (Childhood Diabetes in Finland) Study Group. *Diabetologia* 38:1223–1229, 1995
31. Laine AP, Hermann R, Knip M, Simell O, Akerblom HK, Ilonen J: 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, 2004
32. Cordell HJ, Todd JA, Bennett ST, Kawaguchi Y, Farral M: Two locus maximum lod score analysis of a multifactorial trait: joint consideration of IDDM2 and IDDM4 with IDDM1 in type I diabetes. *Am J Hum Genet* 57:920–934, 1995
33. Dizier MH, Babron MC, Clerget-Darpoux F: Interactive effect of two candidate genes in a disease: extension of the marker-association-segregation chi(2) method. *Am J Hum Genet* 55:1042–1049, 1994
34. She JX, Bui MM, Tian XH, Muir A, Wakeland EK, Zorovich B, Zhang LP, Liu MC, Thomson G, Maclaren NK: Additive susceptibility to insulin-dependent diabetes conferred by HLA-DQB1 and insulin genes. *Autoimmunity* 18:195–203, 1994
35. Cordell HJ, Wedig GC, Jacobs KB, Elston RC: Multilocus linkage tests based on affected relative pairs. *Am J Hum Genet* 66:1273–1286, 2000
36. Lampis R, Morelli L, Congia M, Macis MD, Mulargia A, Loddo M, De Virgiliis S, Marrosu MG, Todd JA, Cucca F: The inter-regional distribution

of HLA class II haplotypes indicates the suitability of the Sardinian population for case-control association studies in complex diseases. *Hum Mol Genet* 9:2959–2965, 2000
37. Bugawan TL, Erlich HA: Rapid typing of HLA-DQB1 DNA polymorphism using nonradioactive oligonucleotide probes and amplified DNA. *Immunogenetics* 33:163–170, 1991
38. Cucca F, Frau F, Lampis R, Floris M, Argiolas L, Macis D, Cao A, De Virgiliis S, Congia M: HLA-DQB1*0305 and -DQB1*0304 alleles among Sardinians: evolutionary and practical implications for oligotyping. *Hum Immunol* 40:143–149, 1994
39. Horn GT, Bugawan TL, Long CM, Erlich HA: Allelic sequence variation of the HLA-DQ loci: relationship to serology and to insulin-dependent diabetes susceptibility. *Proc Natl Acad Sci U S A* 85:6012–6016, 1988
40. Wordsworth BP, Allsopp CE, Young RP, Bell JI: HLA-DR typing using DNA amplification by the polymerase chain reaction and sequential hybridization to sequence-specific oligonucleotide probes. *Immunogenetics* 32:413–418, 1990
41. Petersdorf EW, Smith AG, Mickelson EM, Martin PJ, Hansen JA: Ten HLA-DR4 alleles defined by sequence polymorphisms within the DRB1 first domain. *Immunogenetics* 33:267–275, 1991
42. Bennett ST, Lucassen AM, Gough SC, Powell EE, Undlien DE, Pritchard LE, Merriman ME, Kawaguchi Y, Dronsfield MJ, Pociot F: Susceptibility to human type 1 diabetes at *IDDM2* is determined by tandem repeat variation at the insulin gene minisatellite locus. *Nat Genet* 9:284–292, 1995
43. Thomson G: Mapping disease genes: family-based association studies. *Am J Hum Genet* 57:487–498, 1995
44. Frongia O, Mastinu F, Sechi GM: Prevalence and 4-year incidence of insulin-dependent diabetes mellitus in the province of Oristano (Sardinia, Italy). *Acta Diabetol* 34:199–205, 1997
45. 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* 52:506–516, 1993
46. Risch N: Linkage strategies for genetically complex traits. II. The power of affected relative pairs. *Am J Hum Genet* 46:229–241, 1990
47. Cordell HJ, Todd JA, Hill NJ, Lord CJ, Lyons PA, Peterson LB, Wicker LS, Clayton DG: Statistical modeling of interlocus interactions in a complex disease: rejection of the multiplicative model of epistasis in type 1 diabetes. *Genetics* 158:357–367, 2001
48. Cordell HJ: Epistasis: what it means, what it doesn't mean, and statistical methods to detect it in humans. *Hum Mol Genet* 11:2463–2468, 2002
49. Clayton D, McKeigue PM: Epidemiological methods for studying genes and environmental factors in complex diseases. *Lancet* 358:1356–1360, 2001