Insulin Gene Region-Encoded Susceptibility to IDDM Maps Upstream of the Insulin Gene

Dag E. Undlien, Simon T. Bennett, John A. Todd, Hanne E. Akselsen, Irma Ikäheimo, Helena Reijonen, Mikael Knip, Erik Thorsby, and Kjersti S. Rønningen

The gene region on chromosome 11p15.5 known to be involved in insulin-dependent diabetes mellitus (IDDM) susceptibility was recently mapped to a 4.1-kilobase region including the insulin gene. The region contains 10 candidate polymorphisms that are in strong linkage disequilibrium. By genotyping 7 of these 10 polymorphisms and the tyrosine hydroxylase microsatellite in Finnish Caucasoid IDDM patients and control subjects, we demonstrate that many of the polymorphisms found to be associated with IDDM in other Caucasoid populations do not show any association in this Finnish population. Of the polymorphisms typed, only those at −23 *Hph* I and the variable number of tandem repeats (VNTR) sites confer significant relative risk. Furthermore, we have demonstrated that the −23 *Hph* I polymorphism cannot explain the association. Comparison of the genotypic patterns observed here and previously suggests that the VNTR is the most likely candidate for *IDDM2*. The VNTR is located adjacent to defined regulatory DNA sequences affecting insulin gene expression, which suggests a possible effect on expression of insulin or one of the neighboring genes, tyrosine hydroxylase or insulin-like growth factor 2. *Diabetes* 44:620–625, 1995

Apart from genes in the human leukocyte antigen (HLA) complex on chromosome 6p21 (1,2), only polymorphisms in the insulin–insulin-like growth factor 2 gene region (INS) on chromosome 11p15 have repeatedly shown association with insulin-dependent diabetes mellitus (IDDM) in Caucasians (3–12). Julier et al. (4) found that the IDDM-associated INS polymorphisms, recently designated *IDDM2* (13,14), were associated only in HLA-DR4$^+$ individuals. However, all other studies have reported that INS polymorphisms confer risk independent of HLA genotype (5–7). The disease-associated polymorphisms were recently mapped to a 4.1-kilobase (kb) region encompassing the whole of the insulin gene (9) (Fig. 1). This region contains 10 common polymorphisms. The polymorphisms are in strong linkage disequilibrium and have in previous studies shown a similar strength of association with IDDM, making it difficult to locate the primary association. Lucassen et al. (9) have sequenced the entire region, except the variable number of tandem repeats (VNTR), from different haplotypes without finding any additional polymorphisms. Owerbach and Gabbay (10) found the region of strongest association to be 5′ of 1,428 *Fok* I (Fig. 1). However, these authors genotyped their data set for only 2 of the 10 candidate polymorphisms (i.e., 1,428 *Fok* I and the VNTR), as defined by Lucassen et al. (9). It can be argued that sequencing data provided by Owerbach and Gabbay (10) and previous reports on strong linkage disequilibria between 1,428 *Fok* I and some of these 10 polymorphisms (4,5,9) provided circumstantial evidence for excluding 805 *Dra* III, 1,127 *Pst* I, and 1,404 *Fnu*4HI. Owerbach and Gabbay (10), however, neither detected the −2,733 A/C, −2,221 *Msp* I, and 1,355 T/C sites nor presented significant evidence to exclude the −23 *Hph* I and 1,140 A/C polymorphisms as primary susceptibility determinants.

Finland has the highest incidence of IDDM in the world (16,17). The present Finnish population is thought to stem mainly from the immigration of a small number of settlers to the southwestern region of the country. The population has remained highly isolated for the past 2,000 years with little subsequent immigration. Epidemiological studies have revealed a marked difference in genetic markers as well as frequencies of genetic diseases in Finland compared with the rest of Europe, as would be expected from a founder effect caused by a population bottleneck (18–21). Linkage disequilibrium or allelic association mapping has been used extensively in the investigation of the role of candidate genes in disease, particularly with respect to HLA and autoimmune disease (22). It has also been used widely to fine map the genes for monogenic traits such as Huntington's disease (23), Friedreich ataxia (24), myotonic dystrophy (25), adult polycystic kidney disease (26), diastrophic dysplasia (20,21), and Wilson's disease (27). More recently, linkage disequilibrium mapping has been applied on a genome-wide basis in founder populations (28) and has been successfully applied to the identification of a new IDDM-associated microsatellite marker locus on chromosome 2q31 (29). This method has a much greater mapping resolution than does linkage analysis. Disease-associated alleles at marker loci indicate that the disease-predisposing polymorphism(s) is probably within 2 cM or 2 Mb of the marker locus and may be as close as 50 kb (30–32), depending on several factors including the history of the population, recombination, and the chromosome location. The Finnish population is a genetically young one in

From the Institute of Transplantation Immunology (D.E.U., H.E.A., E.T., K.S.R.), The National Hospital, Oslo, Norway; Nuffield Department of Surgery (S.T.B., J.A.T.), The Wellcome Trust Centre for Human Genetics, University of Oxford, Oxford, U.K.; the Departments of Medical Microbiology (I.I.) and Pediatrics (M.K.), University of Oulu, Oulu, Finland; and Department of Virology (H.R.), University of Turku, Turku, Finland.

Address correspondence and reprint requests to Dr. Dag E. Undlien, Institute of Transplantation Immunology, The National Hospital of Norway, 0027 Oslo, Norway.

Received for publication 11 October 1994 and accepted in revised form 9 February 1995.

bp, base pair; HLA, human leukocyte antigen; IDDM, insulin-dependent diabetes mellitus; INS, insulin–insulin-like growth factor 2 gene region; kb, kilobase; PCR, polymerase chain reaction; RFLP, restriction fragment-length polymorphism; RR, relative risk; VNTR, variable number of tandem repeats.

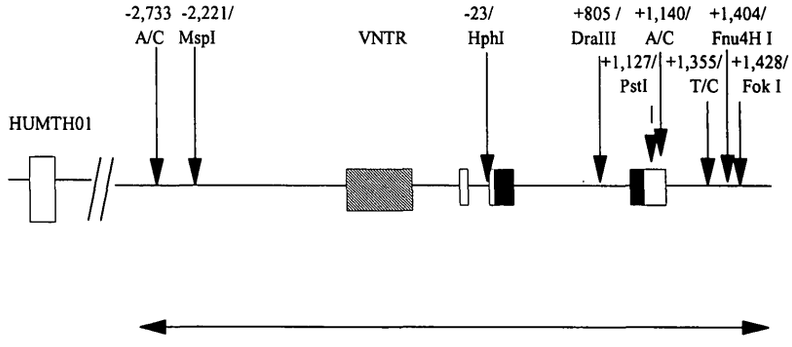

FIG. 1. Location of insulin gene region polymorphisms showing the 4.1-kb region and the 10 polymorphisms associated with IDDM susceptibility (9). Positions are designated denoting the first nucleotide of the protein coding sequence of the insulin gene +1. Polymorphic RFLP sites are indicated by arrows. Untranslated parts of exons are shown as open boxes, and the translated parts of exons are shown as black boxes. The shaded box represents the VNTR. The VNTR was positioned at −596 using the 5′ and 3′ boundaries described by Bell et al. (15) (5′-CTGGGG, ACAG-3′; GenBank accession number J00265).

which linkage disequilibrium mapping has previously proven to be particularly productive (20,21). Comparison of the associations of polymorphisms in different ethnic groups has also been a useful approach in the identification of primary or causative disease polymorphisms (22). By analyzing insulin gene region genotypes in Finnish IDDM patients and control subjects, we have eliminated the −2,221 *Msp* I, −23 *Hph* I, 805 *Dra* III, and 1,127 *Pst* I polymorphisms as primary determinants of disease susceptibility. Furthermore, a comparison of the genotypic pattern of these polymorphisms and others within the susceptibility region (5,9,33) (Fig. 1) suggests that the most likely primary candidate for *IDDM2* is the VNTR.

RESEARCH DESIGN AND METHODS

In total, 105 unrelated Finnish IDDM patients whose diabetes was diagnosed before the age of 16 were included (34). The control population consisted of 204 laboratory and hospital staff, medical students, and randomly selected healthy blood donors.

Polymerase chain reaction (PCR)–restriction fragment-length polymorphism (RFLP) typing of INS polymorphisms. The typing for the −2,221 *Msp* I, −23 *Hph* I, 805 *Dra* III, 1,127 *Pst* I, 1,404 *Fnu*4HI, and 1,428 *Fok* I polymorphisms was performed by PCR-RFLP typing, mainly following previously published methods (4,5,9,12). All samples were typed for −2,221 *Msp* I, −23 *Hph* I, 805 *Dra* III, and 1,127 *Pst* I. In addition, 40 samples, including all that were scored differently on −23 *Hph* I compared with the other RFLP loci (Table 2), were also typed for 1,404 *Fnu*4HI and 1,428 *Fok* I. In accordance with nomenclature from previous studies (4,5,9), those alleles that have previously shown association to IDDM were designated INS+ and those that have shown negative association were designated INS−. The INS+ haplotype has the restriction sites −2,221 *Msp* I, −23 *Hph* I, and 1,404 *Fnu*4HI but lacks the restriction sites 805 *Dra* III, 1,127 *Pst* I, and 1,428 *Fok* I: [(−2,221 *Msp* I+) − (−23 *Hph* I+) − (805 *Dra* III−) − (1,127 *Pst* I−) − (1,404 *Fnu*4HI+) − (1,428 *Fok* I−)]. Conversely, the INS− haplotype is the opposite of the INS+ haplotype: [(−2,221 *Msp* I−) − (−23 *Hph* I−) − (805 *Dra* III+) − (1,127 *Pst* I+) − (1,404 *Fnu*4HI−) − (1,428 *Fok* I+)].

VNTR analysis. The VNTR polymorphism was typed by conventional Southern blots using the probe pHINS 310 on *Pvu* II–digested genomic DNA (3) and by PCR amplification. The PCR was run in a Perkin-Elmer/Cetus DNA thermal cycler using fluorescently labeled primers (either 6-FAM, HEX, or TAMRA; Applied Biosystems). Each PCR product was reduced in length by 450 base pairs (bp) by digesting the PCR product with *Nco* I and then was run on a 4% polyacrylamide gel using an ABI 373A DNA sequencer (Applied Biosystems) with GENESCAN 672 software for fragment size determination. Primer sequences and details of the PCR assay are given elsewhere (35). Under these conditions, only the VNTR class I alleles were amplified. DNA samples having either no or only one VNTR class I allele with this PCR assay were also typed by Southern blot analysis.

Genotyping of the tyrosine hydroxylase microsatellite HUMTH01. This tetranucleotide repeat polymorphism was typed by PCR amplification using fluorescently labeled primers. PCR products were run on a 6% denaturing polyacrylamide gel using an ABI 373A DNA sequencer with GENESCAN 672 software for fragment size determination. The primers used are published elsewhere (35). Five alleles have been recognized: Z-16 (158 bp), Z-12 (162 bp), Z-8 (166 bp), Z-4 (170 bp), and Z (174 bp) (36,37).

HLA class II genotyping. The typing for HLA-DQB1 was performed by PCR–SSO (sequence-specific oligonucleotide) as described previously (34).

Statistical analysis. Relative risks (RRs) were calculated by the method of Wolf (38). Significance was determined using the chi-squared test or, when appropriate, Fisher's exact test. $P < 0.05$ was considered significant. The test for interallelic association between HUMTH01 alleles and genotype A and genotype B, respectively (Table 3), was performed using chi-squared analysis. The test for heterogeneity was calculated using the BMDP statistical software (Program 4F).

RESULTS

INS polymorphisms and haplotypes. Of the eight polymorphisms investigated, only the VNTR (RR = 1.9, $P < 0.05$) and −23 *Hph* I (RR = 1.9, $P < 0.05$) were found to be significantly associated with IDDM (Table 1). Genotypes for the −2,221 *Msp* I and 805 *Dra* III sites were completely concordant (Table 2). The 1,127 *Pst* I polymorphism was also in strong linkage disequilibrium with these markers, with only four samples being discordant. The 40 samples typed for 1,428 *Fok* I and 1,404 *Fnu*4HI were all scored identically as those typed for 1,127 *Pst* I (i.e., INS+/+, INS+/−, or INS−/− at all three sites). Among the IDDM patients, strong linkage disequilibrium between polymorphisms at −23 *Hph* I and those at 805 *Dra* III and −2,221 *Msp* I was observed, with only two discordances (2%). This evidence for linkage disequilibrium was much weaker in control subjects, with as many as 21 discordances being found (10%). The VNTR class I/I, I/III, and III/III genotypes were concordant with INS+/+, INS+/−, and INS−/−, respectively, at −23 *Hph* I, with no discordances found in either the control subjects or the patients (Table 2). The concordance in genotypes implies strong linkage disequilibrium. However, the unequivocal assignment of haplotypes and estimation of haplotype frequencies will require analysis of a large number of families. Our three main genotypes are the same as those reported in previous studies conducted among Caucasians, indicating that the composition of the most common haplotypes in Finns does not differ from other Caucasoid populations (9,10).

TABLE 1
Genotype frequencies and RRs for IDDM of polymorphisms in the insulin gene region

	IDDM patients		Control subjects			
	INS+/+	INS+/− or −/−	INS+/+	INS+/− or −/−	RR	(95% CI)
−2,221 *Msp* I	83 (87)	17 (18)	81 (165)	19 (39)	1.1	(0.6–2.3)
VNTR	82 (86)	18 (19)	71 (144)	29 (60)	1.9	(1.0–3.6)*
−23 *Hph* I	82 (86)	18 (19)	71 (144)	29 (60)	1.9	(1.0–3.6)*
+805 *Dra* III	83 (87)	17 (18)	81 (165)	19 (39)	1.1	(0.6–2.3)
+1.127 *Pst* I	82 (86)	18 (19)	82 (168)	18 (36)	1.0	(0.5–1.9)

Data are % (*n*). INS+/+ equals VNTR class I/I homozygous individuals whereas INS+/− or −/− equals VNTR class I/III or VNTR class III/III individuals. * $P < 0.05$. CI, confidence interval.

Table 2 shows that the only multilocus genotype conferring susceptibility to IDDM is the genotype that was −23 *Hph* I+/+ and VNTR class I/I homozygous (row 1, RR = 1.9, $P < 0.05$). The genotypes that were class I/III and INS+/− at −23 *Hph* I were either close to neutral (row 3; genotype B, RR = 0.9) or significantly protective (row 2; genotype A, RR = 0.1, $P < 0.01$). Critically, the significant protective effect of the genotype VNTR class I/III, −23 *Hph* I+/−, 805 *Dra* III+/+, 1,127 *Pst* I+/+, and −2,221 *Msp* I+/+ genotype indicated that the *Dra* III, *Pst* I, and *Msp* I sites cannot be the sole or primary determinants of IDDM susceptibility. Because the 1,355 T/C polymorphism is known to be concordant with these polymorphisms (9), this polymorphism is unlikely to be involved primarily in IDDM susceptibility. Furthermore, these data imply that the primary determinant of IDDM susceptibility probably corresponds to variation at either the VNTR, the −23 *Hph* I site, the 1,140 A/C site (which is concordant with −23 *Hph* I in Caucasians [9]), the −2,733 A/C site, or a combination of these.

A second important result came from comparison of the RR of genotype A and genotype B (Table 2). The two RRs are significantly different (heterogeneity test for RR for genotype A compared with RR for genotype B, $\chi^2 = 4.84$ [1 df], $P = 0.028$), which implies that even though these two genotypes are both heterozygous at the VNTR and −23 *Hph* I sites, they influence IDDM in significantly different ways. Since the −23 *Hph* I polymorphism and the 1,140 A/C polymorphism only have two variants and are INS+/− on both of these genotypes (assuming complete concordance between −23 *Hph* I and 1,140 A/C [9]), it is unlikely that these polymorphisms can explain the different risks conferred by genotype A compared with genotype B. Thus, this difference appears to be caused by polymorphisms at the −2,733 A/C locus, the VNTR locus, or a combination of both.

Bennett et al. (35) have recently discovered that the Z-8 allele of the microsatellite in intron 1 of the tyrosine hydroxylase gene is in linkage disequilibrium with certain VNTR class III alleles and that these specific HUMTH01 allele Z-8, VNTR class III (Z-8/III) haplotypes may be more protective than other class III haplotypes, for example, those in linkage disequilibrium with the Z-allele at HUMTH01. Therefore, we tested the possibility that the HUMTH01 allele Z-8 was more frequent in genotype A than in genotype B (Table 3) and, conversely, the possibility that the HUMTH01 allele Z was more frequent in the less protective/neutral genotype B than in genotype A. This was indeed the case (Table 3 [$P < 0.001$]; the data for patients are not presented in Table 3 because there was only one patient with genotype A). We observed evidence for strong linkage disequilibrium between VNTR class III alleles and the HUMTH01 alleles Z and Z-8 with only 4 of 79 samples being VNTR class I/III heterozygous or VNTR class III/III homozygous and not carrying one of these HUMTH01 alleles (Table 3). This suggests that there is a strongly protective haplotype (HUMTH01 Z-8/−2,221 *Msp* I+/VNTR class III/−23 *Hph* I−/805 *Dra* III+) and a less protective or neutral haplotype (HUMTH01 Z/−2,221 *Msp* I−/VNTR class III/−23 *Hph* I−/805 *Dra* III−). There was no evidence for a direct effect of the HUMTH01 locus on protection or susceptibility when genotype frequencies or allele frequencies were examined (data not shown). When VNTR class I/I homozygous genotypes were analyzed in a similar manner based on either the presence or the absence of particular HUMTH01 alleles, we found no evidence for differential susceptibility on different VNTR class I haplotypes (data not shown).

Interaction between INS and HLA. The IDDM patients were stratified for the presence or absence of DQB1*0302, which was found on all DR4 haplotypes in the IDDM patients

TABLE 2
Multilocus genotypes in the insulin gene region

	−2,221 *Msp* I	VNTR	−23 *Hph* I	+805 *Dra* III	Patients	Control subjects	RR	(95% CI)
	+/+	Class I/I	+/+	+/+	82 (86)	71 (144)	1.9	(1.0–3.6)*
Genotype A	+/+	Class I/III	+/−	+/+	1 (1)	11 (21)	0.1	(0.0–0.5)†
Genotype B	+/−	Class I/III	+/−	+/−	16 (17)	18 (36)	0.9	(0.5–1.8)
	−/−	Class III/III	−/−	−/−	0 (0)	1 (3)		
	−/−	Class I/III	+/−	−/−	1 (1)	0		
Total *n*					105	204		

Data are % (*n*). Each row represents a particular combination of genotypes at the various loci reported. The RR and *P* values are calculated comparing the frequencies of these genotype combinations between patients and control subjects. Comparing the RR of genotype A versus genotype B gave a significant test for heterogeneity ($\chi^2 = 4.84$ [1 df], $P = 0.028$), suggesting that the RRs are significantly different. The 1,127 *Pst* I, 1,428 *Fok* I, and 1,404 *Fnu*4HI sites were in strong linkage disequilibrium with 805 *Dra* III and −2,221 *Msp* I (see text for details) and are not included in the table. CI, confidence interval. *$P < 0.05$; †$P < 0.01$.

TABLE 3
Subdivision of VNTR class III haplotypes by tyrosine hydroxylase microsatellite analysis

HUMTH01	Genotype A	Genotype B	*P* value
Z+/(Z-8)−	5 (1)	75 (27)	<0.001
(Z-8)+/Z−	67 (14)	0 (0)	<0.001
Z+/(Z-8)+	24 (5)	22 (8)	
Z−/(Z-8)−	5 (1)	3 (1)	

Data are % (*n*). The controls are stratified for the HUMTH01 alleles Z and Z-8. Presence and absence of alleles are denoted by + and −, respectively. Genotype A (n = 21): INS+/+ at −2,221 *Msp* I and 805 *Dra* III combined with INS+/− at −23 *Hph* I and VNTR class I/III heterozygosity. Genotype B (n = 36): INS+/− at −2,221 *Msp* I, −23 *Hph* I and 805 *Dra* III combined with VNTR class I/III. *P* values were obtained when the frequencies of HUMTH01 alleles Z-8 and Z were compared in VNTR class I/III heterozygous controls stratified for being genotype A and genotype B.

(34). The frequencies of INS genotypes in control subjects when DQB1*0302 was stratified for were similar. Therefore, we compared the frequencies in IDDM patients with those in pooled control subjects. The frequency of the −23 *Hph* I+/+-VNTR class I/I genotype was higher in the DQB1*0302$^-$ patients, but the number of patients was small and the difference between the DQB1*0302$^+$ and DQB1*0302$^-$ groups was not statistically significant (P = 0.17, data not shown).

DISCUSSION

In contrast with previous studies in Caucasoids (4,5,9), we find in the Finnish population that polymorphisms at −2,221 *Msp* I, 805 *Dra* III, and 1,127 *Pst* I (and, therefore, 1,404 *Fnu*4HI and 1,428 *Fok* I) are not significantly associated with IDDM. The lack of association at these sites in the Finnish population appears to owe largely to the increased frequencies of the disease-associated alleles in the control data set compared with those in the French (4,9) and U.K. (5) control data sets. We find the same multilocus genotypes that have been reported in other studies in Caucasians, but these genotypes occur with different frequencies. Of the eight polymorphisms investigated, only those at −23 *Hph* I and the VNTR loci were significantly associated with disease. It is unlikely that the 805 *Dra* III, 1,127 *Pst* I, and −2,221 *Msp* I polymorphisms can be involved primarily in IDDM susceptibility, since INS+/+ at these sites occurs on a significantly protective genotype (Table 2). On the basis of reported linkage disequilibria between these sites and 1,355 T/C and 1,404 *Fnu*4HI (9), these polymorphisms can also be excluded. This leaves the VNTR, −23 *Hph* I, the 1,140 A/C, which apparently is in complete linkage disequilibrium with −23 *Hph* I (9), and −2,733 A/C as the most likely candidates for the primary disease determinants in this region (Fig. 1). This finding agrees with a recent report in which it was concluded that one or more of these four polymorphisms were the most likely to be primarily involved in IDDM susceptibility (33). In that study, however, the investigators were unable to determine which of these four polymorphisms is the primary disease determinant(s).

The only significantly protective genotype in this study is −2,221 *Msp* I+/+, VNTR class I/III, −23 *Hph* I+/−, 805 *Dra* III+/+ (genotype A, Table 2). Comparison of RR values for genotypes A and B (test for heterogeneity, χ^2 = 4.84 [1 df], P = 0.028) suggests that these two genotypes influence IDDM in significantly different ways. The data for HUMTH01 (Table 3) suggest that HUMTH01 allele Z-8 is a marker for VNTR class III haplotypes that are strongly protective, whereas VNTR class III haplotypes carrying the HUMTH01 allele Z are less protective. These data agree with results obtained by Bennett et al. (35) while studying transmission in multiplex families. The observation of different class III haplotypes conferring differential protection suggests that −23 *Hph* I and 1,140 A/C cannot be the biologically important polymorphisms conferring protection.

A recent report (33) described common INS genotypes and haplotypes in French Caucasoids. Assuming that similar linkage disequilibria occur in our Finnish population, it seems likely that our genotypes A and B will be mostly INS+/− at the −2,733 A/C locus (probably <10% will be INS−/−). If the same multilocus genotypes and haplotypes as those observed in the French occur in the Finns, then it is unlikely that the −2,733 A/C polymorphism can explain the difference in protection conferred between genotypes A and B. However, analysis of haplotypes or multilocus genotypes, including the −2,733 A/C site, is required to confirm the exclusion of this polymorphism. Our results in the Finns are entirely consistent with those obtained recently by Bennett et al. (35), who have excluded the −2,733 A/C site and have confirmed that the VNTR is *IDDM2*.

The insulin gene region has been sequenced several times (3,4,9,10), and no common polymorphisms in addition to those described here have been found. The least well-characterized polymorphisms in this region are those in the VNTR, and previous studies have found sequence variation within the VNTR classes (10). Taken together, these results suggest that the most plausible explanation for the observed differential protection is that the different VNTR class III haplotypes carry different VNTR class III alleles that influence IDDM susceptibility directly. In this respect, the observation of sequence differences in the 14-bp (15) repeat units on different VNTR alleles (10) is interesting.

Recently, Owerbach and Gabbay (10) have excluded 1,428 *Fok* I as the primary disease determinant (Fig. 1), and they hypothesized that the VNTR might be the primary disease-associated polymorphism. They identified two class III haplotypes (designated class III–3123-A and class III–3123-G, respectively) that appeared to give rise to two different VNTR class I/III heterozygous genotypes (3123-A and 3123-G in their study) corresponding to genotypes A and B, respectively, in this study. In that study, there was also a tendency for genotype A (i.e., the VNTR class I/VNTR class III–3123-A heterozygous genotype) to provide less risk for IDDM (RR = 0.19) than genotype B (i.e., the VNTR class I/VNTR class III–3123-G heterozygous genotype; RR = 0.65). However, these RRs were not statistically significantly different (heterogeneity test: χ^2 = 2.65 [1 df], P = 0.1). Also, Owerbach and Gabbay (10) were not able to show that the differences in transmission of the VNTR class III–3123-A haplotype, compared with the VNTR class III–3123-G haplotype, from VNTR class I/III heterozygous parents to their diabetic offspring were significant. In contrast, we present statistically significant evidence from a case-control study for the two different VNTR class I/III heterozygous genotypes conferring different risks for IDDM development, suggesting that polymorphisms within the VNTR are the primary disease determinant.

The VNTR is located adjacent to defined regulatory DNA sequences affecting insulin gene expression (39,40). In addi-

tion, the VNTR itself has been implicated in gene regulation (39,40). It has been shown that in vitro the VNTR adopts an unusual quadriplex structure, and it has been hypothesized that this may affect gene expression in vivo (41). This hypothesis has some support in reports finding correlation between INS genotype and plasma insulin levels in vivo (42–44). The data from the different studies are, however, conflicting. Although the most likely explanation of our findings is that the biologically important polymorphisms are within the VNTR, we cannot rule out the possibility that a combination of allelic variants at several loci may also be involved.

Regarding INS-HLA interactions, we found no support for INS+/+ at −23 *Hph* I and VNTR class I/I homozygosity conferring risk only in HLA-DR4$^+$ individuals (4), nor was there a tendency for INS conferring stronger susceptibility in DR4$^+$ individuals (9).

In conclusion, this study maps the insulin gene region-associated IDDM susceptibility, *IDDM2*, to the upstream region of the insulin gene and implicates the VNTR as having a direct role in disease development.

Note added in proof: Most recently it has been shown in three studies that the VNTR can influence insulin gene expression in vitro (Kennedy GC, German MS, Rutter WJ: The minisatellite in the diabetes susceptibility locus *IDDM2* regulates insulin transcription. *Nat Genet* 9:293–298, 1995; Lucassen A, Screaton G, Julier C, Elliot T, Lathrop M, Bell J: Regulation of insulin gene expression by the IDDM associated insulin locus haplotypes. *Hum Mol Genet* 4:501–506, 1995; D. Owerbach and K.H. Gabbay, unpublished observations) as well as in vivo (35). Even though one has to await future studies to identify the mechanism by which *IDDM2* influences IDDM development, these studies support the hypothesis that the VNTR is involved in regulation of insulin gene transcription and that this somehow may influence IDDM susceptibility.

ACKNOWLEDGMENTS

This work was supported by The Research Council of Norway, Division NAVF; The Norwegian Diabetes Association; The Novo Nordisk Foundation; and The Wellcome Trust.

We thank Dr. Tobias Gedde-Dahl, Jr., for comments on the manuscript.

REFERENCES

1. Todd JA, Bell JI, McDevitt HO: HLA-DQβ gene contributes to susceptibility and resistance to insulin-dependent diabetes mellitus. *Nature* 329:599–604, 1987
2. Thorsby E, Rønningen KS: Particular HLA-DQ molecules play a dominant role in determining susceptibility or resistance to type 1 (insulin-dependent) diabetes mellitus. *Diabetologia* 36:371–373, 1993
3. Bell GI, Horita S, Karam JH: A polymorphic locus near the human insulin gene is associated with insulin-dependent diabetes mellitus. *Diabetes* 33:176–183, 1984
4. Julier C, Hyer RN, Davies J, Merlin F, Soularue 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-DR-4 dependent diabetes susceptibility. *Nature* 354:155–159, 1991
5. Bain SC, Prins J-B, Hearne CM, Rodrigues NR, Rowe BR, Pritchard LE, Ritchie RJ, Hall JRS, Undlien DE, Rønningen KS, Dunger DB, Barnett AH, Todd JA: Insulin gene region-encoded susceptibility to type 1 diabetes is not restricted to HLA-DR4-positive individuals. *Nat Genet* 2:212–215, 1992
6. Raffel LJ, Vadheim CE, Klein R, Moss SE, Riley WJ, Maclaren NK, Rotter JI: HLA-DR and the 5′ insulin gene polymorphism in insulin-dependent diabetes mellitus. *Metabolism* 40:1244–1248, 1991
7. Van der Auwera B, Heimberg H, Schrevens AF, Van Wayenberge C, Flament J, Schuit FC: 5′ insulin gene polymorphism confers risk to IDDM independently of HLA class II susceptibility. *Diabetes* 42:851–854, 1993
8. Hitman GA, Tarn AT, Winter RM, Drummond V, Williams LG, Jowett NI, Bottazzo GF, Galton DJ: Type 1 (insulin-dependent) diabetes and a highly variable locus close to the insulin gene on chromosome 11. *Diabetologia* 28:218–222, 1985
9. Lucassen AM, Julier C, Beressi JP, Boitard C, Froguel P, Lathrop M, Bell JI: Susceptibility to insulin-dependent diabetes mellitus maps to a 4.1 kb segment of DNA spanning the insulin gene and associated VNTR. *Nat Genet* 4:305–310, 1993
10. Owerbach D, Gabbay KH: Localization of a type I diabetes susceptibility locus to the variable tandem repeat region flanking the insulin gene. *Diabetes* 42:1708–1714, 1993
11. Owerbach D, Gabbay KH: Linkage of the VNTR/insulin-gene and type 1 diabetes mellitus: increased gene sharing in affected sibling pairs. *Am J Hum Genet* 54:909–912, 1994
12. Undlien DE, Hamaguchi K, Kimura A, Tuomilehto-Wolf E, Swai ABM, McLarty DG, Tuomilehto J, Thorsby E, Rønningen KS: IDDM susceptibility associated with polymorphisms in the insulin gene region: a study of blacks, Caucasians and Orientals. *Diabetologia* 37:745–749, 1994
13. Davies JL, Kawaguchi Y, Bennett ST, Copeman JB, Cordell HJ, Pritchard LE, Reed PW, Gough SCL, Jenkins SC, Palmer SM, Balfour KM, Rowe BR, Farrall M, Barnett AH, Bain SC, Todd JA: A genome-wide search for human type 1 diabetes susceptibility genes. *Nature* 371:130–136, 1994
14. Pritchard LE, Kawaguchi Y, Reed PW, Copeman JB, Davies JL, Barnett AH, Bain SC, Todd JA: Analysis of the CD3 gene region and type 1 diabetes: application of fluorescence-based technology to linkage disequilibrium mapping. *Hum Mol Genet* 4:197–202, 1995
15. Bell GI, Selby MJ, Rutter WJ: The highly polymorphic region near the human insulin gene is composed of simple tandemly repeating sequences. *Nature* 295:31–35, 1982
16. Green A, Gale EAM, Patterson CC: Incidence of childhood-onset insulin-dependent diabetes mellitus: the EURODIAB ACE study. *Lancet* 339:905–909, 1992
17. Karvonen M, Tuomilehto J, Libman I, LaPorte R: A review of the recent epidemiological data on the worldwide incidence of type 1 (insulin-dependent) diabetes mellitus. *Diabetologia* 36:883–892, 1993
18. Nevanlinna HR: The Finnish population structure: a genetic and geneological study. *Hereditas* 71:195–236, 1972
19. Norio R, Nevanlinna HR, Perheentupa J: Hereditary diseases in Finland: rare flora in rare soil. *Ann Clin Res* 5:109–141, 1973
20. Hästbacka J, de la Chapelle A, Kaitila I, Sistonen P, Weaver A, Lander E: Linkage disequilibrium mapping in isolated founder populations: diastrophic dysplasia in Finland. *Nat Genet* 2:204–211, 1992
21. Hästbacka J, de la Chapelle A, Mahtani MM, Clines G, Reeve-Daly MP, Daly M, Hamilton BA, Kusumi K, Trivedi B, Weaver A, Coloma A, Lovett M, Buckler A, Kaitila I, Lander ES: The diastrophic dystrophy gene encodes a novel sulfate transporter: positional cloning by fine-structure linkage disequilibrium mapping. *Cell* 78:1073–1087, 1994
22. Todd JA, Mijovic C, Fletcher J, Jenkins D, Bradwell AR, Barnett AH: Identification of susceptibility loci for insulin-dependent diabetes mellitus by trans-racial gene mapping. *Nature* 338:587–589, 1989
23. MacDonald ME, Lin C, Srinidhi L, Bates G, Altherr M, Whaley WL, Lehrach H, Wasmuth J, Gusella JF: Complex pattern of linkage disequilibrium in the Huntington disease region. *Am J Hum Genet* 49:723–734, 1991
24. Fujita R, Hanauer A, Sirugo G, Heilig R, Mandel J-L: Additional polymorphisms at marker loci D9S5 and D9S15 generate extended haplotypes in linkage disequilibrium with Friedreich ataxia. *Proc Natl Acad Sci USA* 87:1796–1800, 1990
25. Harley HG, Brook JD, Floyd J, Rundle SA, Crow S, Walsh KV, Thibault M-C, Harper PS, Shaw DJ: Detection of linkage disequilibrium between the myotonic dystrophy locus and a new polymorphic DNA marker. *Am J Hum Genet* 49:68–75, 1991
26. Snarey A, Thomas S, Schneider MC, Pound SE, Barton N, Wright AF, Somlo S, Germino GG, Harris PC, Reeders ST, Frischauf A-M: Linkage disequilibrium in the region of the autosomal dominant polycystic kidney disease gene (PKD1). *Am J Hum Genet* 55:365–371, 1994
27. Bowcock AM, Tomfohrde J, Weissenbach J, Bonne-Tamir B, St. George-Hyslop P, Giagheddu M, Cavalli-Sforza LL, Farrer LA: Refining the position of Wilson disease by linkage disequilibrium with polymorphic microsatellites. *Am J Hum Genet* 54:79–87, 1994
28. Houwen RHJ, Baharloo S, Blankenship K, Raeymaekers P, Juyn J, Sandkuijl LA, Freimer NB: Genome screening by searching for shared segments: mapping a gene for benign recurrent intrahepatic cholestasis. *Nat Genet* 8:380–386, 1994
29. Copeman J, Cucca F, Hearne CM, Cornall RJ, Reed PW, Rønningen KS, Undlien DE, Nistico L, Buzzetti R, Tosi R, Pociot F, Nerup J, Cornelis F, Barnett AH, Bain SC, Todd JA: Linkage disequilibrium mapping of a type 1 diabetes susceptibility gene (*IDDM7*) to human chromosome 2q31-q33. *Nat Genet* 9:80–85, 1995
30. Watkins WS, Zenger R, O'Brien E, Nyman D, Eriksson AW, Renlund M, Jorde LB: Linkage disequilibrium patterns vary with chromosomal location: a case study from the von Willebrand factor region. *Am J Hum Genet* 55:348–355, 1994
31. Jorde LB, Watkins WS, Viskochil D, O'Connel P, Ward K: Linkage disequilibrium in the neurofibromatosis 1 (NF1) region: implications for

gene mapping. *Am J Hum Genet* 53:1038–1050, 1993
32. Jorde LB, Watkins WS, Carlson M, Groden H, Albertsen A, Thliveris A, Leppert M: Linkage disequilibrium predicts physical distance in the adenomatous polyposis coli region. *Am J Hum Genet* 54:884–898, 1994
33. Julier C, Lucassen A, Villedieu P, Delepine M, Levy-Marchal C, Danzé PM, Bianchi F, Boitard C, Froguel P, Bell J, Lathrop GM: Multiple DNA variant association analysis: application to the insulin gene region in type 1 diabetes. *Am J Hum Genet* 55:1247–1254, 1994
34. Reijonen H, Ilonen J, Knip M, Åkerblom HK: HLA-DQB1 alleles and absence of Asp 57 as susceptibility factors of IDDM in Finland. *Diabetes* 40:1640–1644, 1991
35. Bennett ST, Lucassen AM, Gough SCL, Powell EE, Undlien DE, Pritchard LE, Merriman ME, Kawaguchi Y, Dronsfield M, Pociot F, Nerup J, Bouzekri N, Cambon-Thomsen A, Rønningen KS, Barnett AH, Bain SC, Todd JA: 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
36. Kimpton CP, Gill P, Walton A, Urquhart A, Millican ES, Adams M: Automated DNA profiling employing multiplex amplification of short tandem repeat loci. *PCR Methods Appl* 3:13–22, 1993
37. Hearne CM, Ghosh S, Todd JA: Microsatellites for linkage analysis of genetic traits. *Trends Genet* 8:288–294, 1992
38. Tiwari JL, Terasaki P: The data and statistical analysis. In *HLA and Disease Associations*. New York, Springer-Verlag, 1985, p. 18–27
39. Clark AR, Docherty K: The insulin gene. In *Insulin: Molecular Biology to Pathology*. Ashcroft FM, Ashcroft SJH, Eds. Oxford, U.K., Oxford Univ. Press, 1992, p. 37–63
40. Docherty K: 1992 R.D. Lawrence Lecture: the regulation of insulin gene expression. *Diabetic Med* 9:792–798, 1992
41. Hammond-Kosack M, Dobrinski B, Lurz R, Docherty K, Kilpatrick MW: The human insulin gene linked polymorphic region exhibits an altered DNA structure. *Nucleic Acids Res* 20:231–236, 1992
42. Weaver JU, Kopelman PG, Hitman GA: Central obesity and hyperinsulinaemia in women are associated with polymorphism in the 5′ flanking region of the human insulin gene. *Eur J Clin Invest* 22:265–270, 1992
43. Cocozza S, Riccardi G, Monticelli A, Capaldo B, Genovese S, Krogh V, Celentano E, Farinaro E, Varrone S, Avvedimento VE: Polymorphism at the 5′ end flanking region of the insulin gene is associated with reduced insulin secretion in healthy individuals. *Eur J Clin Invest* 18:582–586, 1988
44. Knowler WC, Pettit DS, Vasquez B, Rotwein PS, Andreone TL, Permutt MA: Polymorphism in the 5′ flanking region of the human insulin gene: relationship with noninsulin-dependent diabetes mellitus, glucose and insulin concentrations, and diabetes treatment in the Pima Indians. *J Clin Invest* 74:2129–2135, 1984