Genetics of Type 1 Diabetes

Andrea K. Steck1* and Marian J. Rewers1

BACKGROUND: Type 1 diabetes, a multifactorial disease with a strong genetic component, is caused by the autoimmune destruction of pancreatic β cells. The major susceptibility locus maps to the HLA class II genes at 6p21, although more than 40 non-HLA susceptibility gene markers have been confirmed.

CONTENT: Although HLA class II alleles account for up to 30%–50% of genetic type 1 diabetes risk, multiple non-MHC loci contribute to disease risk with smaller effects. These include the insulin, PTPN22, CTLA4, IL2RA, IFIH1, and other recently discovered loci. Genomewide association studies performed with high-density single-nucleotide–polymorphism genotyping platforms have provided evidence for a number of novel loci, although fine mapping and characterization of these new regions remain to be performed.

Children born with the high-risk genotype HLADR3/4-DQ8 comprise almost 50% of children who develop antiislet autoimmunity by the age of 5 years. Genetic risk for type 1 diabetes can be further stratified by selection of children with susceptible genotypes at other diabetes genes, by selection of children with a multiple family history of diabetes, and/or by selection of relatives that are HLA identical to the proband.

SUMMARY: Children with the HLA-risk genotypes DR3/4-DQ8 or DR4/DR4 who have a family history of type 1 diabetes have more than a 5% risk for developing islet autoantibodies during childhood, and children with the same HLA-risk genotype but no family history have approximately a 1 in 20 risk. Determining extreme genetic risk is a prerequisite for the implementation of primary prevention trials, which are now underway for relatives of individuals with type 1 diabetes.
in the general population, and thus $\lambda_s = 6/0.4 = 15$ (9, 10).

The risk for T1D in siblings of patients is 15-fold higher than the risk for T1D in the general population, which suggests that genetic factors play an important role in disease susceptibility. The pattern of inheritance is complex, and the development of disease is thought to be determined by an interaction between genetic predisposition and environmental triggers. Concordance rates for T1D in monozygous twins with long-term follow-up is $50\%$, compared to $6\%$–$10\%$ in dizygous twins, which is similar to what is found in nontwin siblings. With long-term (>30 years) follow-up, at least two-thirds of initially discordant monozygous twins show evidence of persistent $\beta$-cell autoantibodies and/or diabetes (Fig. 2) (11).

Among first-degree relatives, siblings are at a higher risk ($5\%$–$10\%$ risk by age 20) than offspring; offspring of diabetic fathers are at a higher risk (approximately $12\%$) than offspring of diabetic mothers (approximately $6\%$) (12). In addition, DR3/4 siblings who share both HLA haplotypes identical by descent with their diabetic DR3/4 sibling are at a higher risk than those sharing 1 or no haplotypes (13).

When the first twin of a twin pair develops T1D after age 25 years, the risk of the second monozygotic twin developing T1D is $<5\%$ with long-term follow-up (14), whereas when the first twin develops diabetes before the age of 6 years the risk of the second twin developing diabetes is at least $60\%$. For monozygotic twins of patients with T1D, expression of antiislet autoantibodies directly correlates with progression to overt diabetes. Essentially all such twins who express “biochemical” antiislet autoantibodies progress to diabetes, even though this may not occur until after decades of follow-up (14, 15). Similar results were found in a population-based twin cohort of 22 650 twin pairs from Finland (9). With longer follow-up, the cumulative incidence of diabetes among monozygotic twins who were initially discordant for diabetes was $65\%$, and persistent autoantibody positivity or T1D developed in $78\%$ by age 60 years (Fig. 2) (11).

### Linkage and Association Studies

The 2 primary approaches used to identify risk loci for T1D have been linkage studies and association studies. Linkage studies, typically using affected sibling pairs, can identify regions of the genome that are shared more frequently among affected relatives. Linkage studies use markers spanning the genome at a modest density and are the method of choice when the risk factors have large effect sizes but are relatively rare. They provide broad information on chromosomal regions that may contribute to T1D risk. The most important and consistent evidence of linkage for T1D has been in the HLA region on chromosome 6p21 (10, 16). A metaanalysis done by the Type 1 Diabetes Genetics Consortium provided a total sample of 1435 families with 1636 affected sibling pairs (17). In addition to the HLA re-
region (nominal $P = 2.0 \times 10^{-52}$), 9 non–HLA-linked regions showed some evidence of linkage to T1D (nominal $P < 0.01$), including 3 at (or near) genomewide significance ($P < 0.05$): 2q31-q33, 10p14-q11, and 16q22-q24.

In contrast to linkage studies, association studies can detect alleles with much more modest effects on risk as long as those alleles are relatively common. Association studies use specifically selected markers in genes of interest that are genotyped in case patients and unaffected control individuals or, in some studies, in case patients and their parents. Association studies in T1D have initially focused on candidate genes. All of the 4 well-established risk loci, including HLA, INS, CTLA4, and PTPN22, were identified in candidate gene association studies. Recently, GWA studies have revolutionized genetic studies by the development of new, high-throughput SNP genotyping platforms and dense maps of SNPs.

HLA Class II Alleles

The major T1D susceptibility locus maps to the class II loci HLA-DRB1 and HLA-DQB1 on chromosome 6p21. The highest risk DR/DQ haplotypes for T1D are DR3-DQA1*0501-DQB1*0201 (DR3) and DR4-DQA1*0301-DQB1*0302 (DR4), and these alleles account for 30%–50% of genetic T1D risk (1).

Although this genotype confers extremely high risk, there is a spectrum of risk associated with HLA-DR/DQ genotypes—from increased, to neutral, to protective. For instance, the HLA-DQA1*0102, DQB1*0602 haplotype confers dominant protection...
from T1D, even in the presence of islet autoantibodies (21). Too few HLA-DQB1*0602 individuals expressing multiple islet autoantibodies have been studied to know the exact magnitude of protection, but 1% of new-onset patients expressing islet autoantibodies have DQB1*0602 vs 20% of the general US population.

Risk of diabetes is also influenced by DRB1*04 variants and DQ alleles on DR4 haplotypes, with higher risk from DRB1*0405 [odds ratio (OR) = 11.4], DRB1*0401 (OR = 8.4), DRB1*0402 (OR = 3.6), and DRB1*0404 (OR = 1.6), whereas DRB1*0403 is protective (OR = 0.27) (6). Similarly, for DRB1*0401, variation of DQB1 influences risk, because haplotypes with DQB1*0302 (OR = 8.4) are highly susceptible, whereas those with DQB1*0301 (OR = 0.35) are modestly protective.

The incidence of T1D has been increasing worldwide by approximately 3% per year (7), with the highest increase occurring in young children (8). Interestingly, however, as T1D incidence increases, the percentage of those cases with the high-risk HLA-DR3/4 genotype seems to be decreasing (22, 23). These temporal changes in HLA genotypes suggest increased environmental pressure with higher disease progression rate in individuals with lower-risk HLA genotypes and/or contribution of other non-HLA class II alleles or non-MHC–related alleles to T1D risk.

Several investigators have reported some association of HLA-DPB1 alleles with T1D (24, 25). Results of these studies have shown that the frequency of the DPB1*0402 allele is lower in patients than in control groups and that transmission of DPB1*0402 to T1D patients is decreased. On average, the relative risk for DPB1*0402 overall from these studies is 0.56. DPB1*0301 and DPB1*0202 have been reported to be associated with a predisposition to T1D. Among HLA DR3/DR4-DQB1*0302 children in the general population, we can now identify a risk of 20% of activating islet autoimmunity by excluding both the protective alleles DP*0402 and DRB1*0403 (Fig. 3) (24).

**Fig. 3.** Protective effects of DPB1*0402 and DRB1*0403 in DR3/4-DQB1*0302 newborn/general population children (NECs) and progression to autoantibody positivity (A) and TID (B).

Cumulative incidence of diabetes and autoantibody positivity according to life-table analysis in NECs. Survival analyses of progression to diabetes and antislet autoimmunity according to age are shown. The numbers of children still followed at each time point are shown. HR (high risk), no protective alleles; LR (low risk), DRB1*0403 and/or DPB1*0402. Reprinted with permission from the American Diabetes Association, © 2007; Baschal et al. (24).

**HLA Class I and Class III Alleles**

Recently, evidence has been found for susceptibility loci within or linked to the MHC independent of HLA-DR/DQ, such as HLA class I alleles (26–28). In the Diabetes Autoimmunity Study in the Young (DAISY) cohort, siblings of children with T1D who have HLA DR3/DR4-DQB1*0302 and are identical by descent for both HLA haplotypes with their diabetic proband siblings were observed to have a 65% risk for developing islet autoantibodies by age 7 years and a 50% risk of developing diabetes by age 10 years (13). These findings suggest that additional MHC-linked genes determine T1D risk.

HLA-A is associated with T1D independent of DR/DQ (29). Allele A24 has been associated with younger age of onset; allele A30 is higher risk whereas allele A1 is lower risk than other DR3 haplotypes. HLA-B and HLA-C alleles have also been associated...
with T1D independent of DR/DQ (26). After adjustment for linkage disequilibrium with DR/DQ, the alleles HLA-B18, B39, B44, C3, C8, and C16 remained associated with T1D (26). In a study by Nejentsv et al., both HLA-A and HLA-B survived correction for HLA-DR and DQ, and in particular, HLA-B39 was observed to confer high risk for T1D and was associated with lower age of onset (27).

MIC-A is an MHC class I–like molecule that activates certain T cells, natural killer cells. MIC-A allele 5.1 introduces a stop codon into the molecule, interrupting production of the membrane-spanning portion of the molecule, and thus a functional change. MIC-A has been reported to be associated with T1D when the DR/DQ genotype is fixed (within DR3/4 first-degree relatives) (30).

Non-MHC Genetic Factors

More than 40 non–HLA-susceptibility gene markers have been confirmed (5, 31). At present, polymorphisms of the INS gene and PTPN22 genes contribute most to diabetes risk after HLA alleles. Adding high-risk alleles of these genetic markers to HLA class II genotyping can improve risk prediction, but the effect is small even for the strongest loci, with ORs between 1.7 and 2.0.

INS GENE

The INS gene located on chromosome 11p15.5 confers about 10% of the genetic susceptibility to T1D. Both a variable number of tandem repeats located approximately 0.5 kb upstream of INS (2) and other polymorphisms in tight linkage disequilibrium such as −23Hph1 and +1140A/C (32) have been implicated as etiological factors. All of the polymorphisms lie outside coding sequences, suggesting that diabetes susceptibility derives from modulation of INS transcription (33). The short class I variable number of tandem repeats alleles (26–63 repeats) were associated with predisposition to T1D, whereas class III alleles (140–210 repeats) were dominantly protective. The allelic variant associated with protection from T1D has been associated with greater expression of insulin messenger RNA within the thymus (33, 34). Results of studies in the NOD (nonobese diabetic) mouse model of T1D have implicated insulin as the primary autoantigen, where decreased insulin expression in the thymus also correlates with risk (35).

PTPN22 Gene

The lymphoid-specific phosphatase LYP, encoded by the PTPN22 gene on chromosome 1p13, is involved in preventing spontaneous T cell activation. An association of a nonsynonymous SNP in PTPN22 at position 1858 with T1D has been reported from many populations, with an OR of 2–3 for the homozygous TT genotype (4, 36–38) (Fig. 4). The C1858T SNP results in a missense mutation (R620W) that changes an arginine at position 620 to a tryptophan and thereby abrogates the ability of the molecule to bind to the signaling molecule Csk (4,39–41). The polymorphism has been associated with a gain-of-function mutation, apparently decreasing T-cell–receptor signaling (40). The Trp620 variant is associated with other autoimmune disorders including Graves disease (42), rheumatoid arthritis (39), and systemic lupus erythematosus (43).

CTLA4 Gene

The CTLA4 gene encodes a molecule that functions as a negative regulator of T-cell activation. The G allele of the first exon (Ala17Thr) has been most consistently associated with T1D (44) and reduced control of T-cell proliferation (45). The G/G variant in the first exon has been associated with decreased expression of a soluble variant of CTLA4 that may have an influence on immune function, especially in light of CTLA4 polymorphism associated with diabetes of the NOD mouse (46), Graves disease (47), and Addison disease (48). The (AT)n microsatellite marker in the 3′ untranslated region, in strong linkage disequilibrium with Ala17Thr, is also associated with T1D (49) and Graves disease (50). CTLA4, located at 2q33, has an OR of 1.1–1.2, but the association has been replicated in multiple studies (50, 51). The locus is associated primarily with patients with both thyroid autoimmunity and T1D, concordant with its stronger association with thyroid autoimmunity.

Other Defined Loci, GWA, and Microarrays

Besides these well-established non–HLA loci, a number of other associations with T1D have been reported for IL2RA/CD25 (52, 53), SUMO4 (small ubiquitinlike modifier 4) (54) and IFIH1 (55, 56). Of these 3 genes, the clearest confirmation is for IL2RA (10p13), with the finding of more than 1 SNP associated with T1D. The IL2RA/CD25 locus is implicated in a number of autoimmune disorders including multiple sclerosis, but in that case a different SNP is associated with increased risk. The IL2RA/CD25 SNPs associated with risk for T1D have been shown to have different circulating concentrations of CD25; despite extensive overlap between concentrations, this variation in concentration is highly significant when a large number of individuals are studied (53). The mechanism by which the SNPs contribute to diabetes risk is currently unknown, and may be complex given 2 associated SNPs in the same locus, neither of which alters the coding sequence.
The minor allele of rs1990760 in the IFIH1 gene was reported to be associated with T1D, with a risk ratio of 0.86 in a large study of 4253 cases and 5842 controls, and with an additional analysis of 2134 parent-child trios (55). The Wellcome Trust analysis used a different SNP (rs3788964) and revealed a very modest association with T1D (31). The gene (on chromosome 2q24) is of particular interest in that it may relate the innate immune system to the development of disease, presumably mediated by the adaptive immune system, and animal models are available in which activation of innate immunity, and interferon γ, are associated with induction of autoimmune diabetes (57).

Several GWA studies of T1D have now been performed and reported (58–60). A GWA study for T1D was completed in 2007 by the Wellcome Trust Case Control Consortium, which reported signals at known loci [6p21 (MHC region), 11q15 (INS), 2q33 (CTLA4), 1p13 (PTPN22), 10p15 (IL2RA) and 2q24 (IFIH1)] and also described signals at several novel loci [12q13, 12q24, 16p13, 4q27, 12p13, and 18p11] (31). The associations at 12q24, 12q13, 16p13, and 18p11 were confirmed in other recent independent studies (5, 59, 61–64). Additional follow-up studies (65) and metaanalyses of GWA (60, 66) have increased the list of new T1D loci to more than 40.

Microarray technology and bioinformatics allow the comparison of gene expression profiles (or transcriptomes). This new technology seems promising because it enables investigators to study the behavior of many genes simultaneously. Studies in humans with T1D have been rather small so far and limited to peripheral blood mononuclear cells (67). Some authors (68) have reported an overexpression of IL-1–regulated genes in both type 1 and type 2 diabetic patients, whereas others (69) have reported on an expression signature that includes IL-1 cytokine family members and chemokines in both new-onset T1D patients and at-risk siblings who later developed the disease. Gene expression profiles in pancreatic lymph nodes have been described in the NOD mouse. Future research in humans might be possible with new collaborative efforts such as the international Network for Pancreatic Organ Donors with Diabetes (www.jdrfnpod.org).
Genetic Markers in Prediction of T1D

A number of large population-based studies have been carried out stratifying individuals at birth by HLA genotype and insulin gene polymorphisms. Children born in Denver with the highest risk genotype DR3/4-DQ8 comprise 2.4% of newborns and almost 50% of children developing antiislet autoimmunity by age 5 (DAISY study) (18, 70). The BabyDiab study of offspring of patients with T1D in Germany and the Diabetes Prediction and Prevention study from Finland provide similar information concerning the risk associated with specific HLA genotypes and insulin gene polymorphisms (71, 72). Risk can be further stratified by selection of children with susceptible genotypes at other diabetes genes (71), by selection of children with a multiple family history of diabetes (73), and by selection of relatives that are HLA identical to the proband (13). In the DAISY cohort, siblings of children with T1D who have the highest risk HLA DR3/DR4-DQ8 and are identical by descent with their diabetic proband sibling were found to have a 65% risk for developing islet autoantibodies by age 7 years and a 50% risk of developing diabetes by age 10 years (13) (Fig. 5). These findings suggest that additional MHC-linked genes determine T1D risk.

More than 90% of T1D patients have no affected relatives, so to have a major impact preventive efforts will require inclusion of the general population. Among HLA DR3/DR4-DQB1*0302 general-population children, we can now identify a risk of 20% of activating islet autoantibodies by excluding both DP*0402 and DRB1*0403 alleles (Fig. 4) (24). This risk of 20% is, however, much lower than the risk of 80% conferred for siblings who have DR3/DR4-DQB1*0302 identical by descent with their sibling proband and is similar to a 20% risk for those DR3/DR4-DQB1*0302 siblings sharing only 1 (or no) haplotype identical by descent (13).

Conclusions

The major T1D susceptibility locus maps to the class II loci HLA-DRB1 and HLA-DQB1 on chromosome 6p21. With the advent of whole-genome SNP genotyping studies in the past several years, many additional non-MHC loci have been identified that contribute to T1D risk. A majority of these loci appear to have effects in the immune system, particularly on T cells. In addition, non-DR/DQ loci within the MHC have been identified that contribute to prediction of T1D in the general population (i.e., HLA-DP).

Preventive therapies can be applied before diabetes onset or even before the appearance of islet autoimmunity in individuals who are genetically at risk for T1D. Children who have a family history of T1D and the HLA risk genotypes DR3/4-DQ8 or DR4/DR4 have a more than 1 in 5 risk for developing islet autoanti-
bodies during childhood (71, 74) and approximately a 1 in 20 risk without a relative with T1D (74). Primary prevention trials such as Pre-POINT (Primary Oral Insulin Trial) are currently underway in siblings with high genetic risk for T1D. The Pre-POINT trial will determine the feasibility of performing a primary autoantigen vaccination trial in children with high genetic risk for T1D in a dose-escalation primary-intervention pilot study (75). In this study oral insulin will be administered to high-risk siblings who have not yet developed islet autoantibodies to provide protection before islet autoimmunity starts. Eligible children are those who have multiple first-degree relatives with T1D or those who have the HLA-DR3/DR4-DQ8 genotype inherited identical by descent with a sibling proband; such siblings have T1D risk as high as 80% (13). This multicenter, placebo-controlled, double-blinded trial will enroll up to 25 children randomized to receive increasing doses of oral insulin (2.5, 7.5, 22.5, or 67.5 mg per day) to determine a dose that is safe and bioavailable to the immune system. Children will be monitored for the development of islet autoantibodies, diabetes, and protective immune responses to insulin. Pre-POINT will proceed to the POINT study if and when the dose-finding committee identifies a dose that can be tested for efficacy.

The oral insulin study conducted by the Type 1 Diabetes TrialNet consortium is currently enrolling first-degree relatives ages 3–45 years and second-degree relatives ages 3–20 years. The primary objective is to determine whether administration of oral insulin will prevent or delay the development of T1D in individuals at risk. Eligible individuals have to be insulin-autoantibody positive on 2 samples, with at least 1 other islet autoantibody present, have oral glucose tolerance test results within reference intervals, and meet certain criteria on first-phase insulin response by intravenous glucose tolerance test.

However, only 10% of new-onset T1D patients have a relative with T1D. About 90% of new cases of T1D occur in the general population. Determining extreme genetic risk in the general population is a prerequisite for the implementation of primary prevention trials in the general population, where most new cases of T1D arise.

Author Contributions: All authors confirmed they have contributed to the intellectual content of this paper and have met the following 3 requirements: (a) significant contributions to the conception and design, acquisition of data, or analysis and interpretation of data; (b) drafting or revising the article for intellectual content; and (c) final approval of the published article.

Authors’ Disclosures or Potential Conflicts of Interest: Upon manuscript submission, all authors completed the Disclosures of Potential Conflict of Interest form. Potential conflicts of interest:

Employment or Leadership: None declared.
Consultant or Advisory Role: M.J. Rewers, GlaxoSmithKline.
Stock Ownership: None declared.
Honoraria: None declared.
Research Funding: M.J. Rewers, NIH R37 DK32493 (DAISY), U01 DK63821 (TEDDY); A.K. Steck, Juvenile Diabetes Research Foundation Early Career Patient Oriented Research Award 11-2010-206.
Expert Testimony: M.J. Rewers, Astra Zeneca.

Role of Sponsor: The funding organizations played no role in the design of study, choice of enrolled patients, review and interpretation of data, or preparation or approval of manuscript.

References


