Genetic and Epigenetic Factors in Etiology of Diabetes Mellitus Type 1

abstract

Diabetes mellitus type 1 (T1D) is a complex disease resulting from the interplay of genetic, epigenetic, and environmental factors. Recent progress in understanding the genetic basis of T1D has resulted in an increased recognition of childhood diabetes heterogeneity. After the initial success of family-based linkage analyses, which uncovered the strong linkage and association between HLA gene variants and T1D, genome-wide association studies performed with high-density single-nucleotide polymorphism genotyping platforms provided evidence for a number of novel loci, although fine mapping and characterization of these new regions remains to be performed. T1D is one of the most heritable common diseases, and among autoimmune diseases it has the largest range of concordance rates in monozygotic twins. This fact, coupled with evidence of various epigenetic modifications of gene expression, provides convincing proof of the complex interplay between genetic and environmental factors. In T1D, epigenetic phenomena, such as DNA methylation, histone modifications, and microRNA dysregulation, have been associated with altered gene expression. Increasing epidemiologic and experimental evidence supports the role of genetic and epigenetic alterations in the etiopathology of diabetes. We discuss recent results related to the role of genetic and epigenetic factors involved in development of T1D. *Pediatrics* 2013;132:1–11
Diabetes mellitus type 1 (T1D) is a complex disease resulting from the interplay of genetic, epigenetic, and environmental factors.1 Worldwide, the T1D epidemic represents an increasing global public health burden, and the incidence of T1D among children has been rising.2 Over the last few decades, there has been an overall increase in the incidence of T1D of ∼3% to 5% per year, and it is estimated that there are ∼65 000 new cases per year in children <15 years old.3–6 This significant worldwide increase in the incidence of T1D suggests the importance of interactions between genetic predisposition and environmental factors in the multifactorial etiology of T1D.7–10 Environmental factors implicated so far include infection with enteroviruses or endogenous retroviruses and consumption of milk proteins, as well as the influence of environmental pollutants, variations in gut flora, and vitamin D exposure.2,10,11 Only 10% to 15% of newly diagnosed patients have a positive family history of T1D. However, increased susceptibility to T1D can be inherited, because the average prevalence risk in children of a person with T1D is ∼6%, and ∼7% in siblings, whereas 12% to 67.7% of identical twins of patients with T1D are at risk, compared with 0.4% of the general population.12,13 In fact, T1D is one of the most heritable common diseases, with the sibling relative risk estimated at 15 and the largest range in concordance rates in monzygotic twins among autoimmune diseases (AIDs).8,13 This, coupled with evidence of various epigenetic modifications of gene expression, provides important and convincing proof of the complex interplay between genetic and environmental factors (Fig 1). In T1D, epigenetic phenomena, such as DNA methylation, histone modifications, and microRNA (miRNA) dysregulation, have been associated with altered gene expression.14

Extensive familial and population genetic studies conducted over the past 30 years have provided an explanation for nearly 80% of the heritability of T1D, suggesting the primary role of several risk alleles in diabetes susceptibility.15 The main susceptibility genes currently accepted for T1D are the HLA class II alleles, which account for up to 50% of genetic T1D risk. Among HLA class I genes, one of the strongest associations with T1D is reported for B*39 allele, which significantly increases the risk of T1D.16 Multiple non-HLA loci contribute to disease risk with smaller effects. These include the insulin gene, CTLA4, PTPN22, interleukin 2 receptor α (IL2RA), the interferon induced with helicase C domain 1 (IFIH1), the CAPSL-IL7R block, the lectin CLEC16A, the Th1 transcription factor STAT4, the tyrosine phosphatase PTPN2, and other recently discovered loci.17,18

Common variants account for only part of the risk associated with complex disease phenotypes.9 The remaining genetic susceptibility (ie, missing heritability) can be accounted for by the influence of additional variants that remain underrepresented in current genotyping platforms applied in underpowered studies, which are insufficient for detecting gene–gene (epistatic) and gene–environment interactions.19 These biological interactions are critical for gene regulation, signal transduction, biochemical networks, and numerous other physiologic and developmental pathways. The concept of epistasis implies that analyses of haplotype combinations and their interactions with environmental factors are needed to establish the risk profile for T1D.19 Therefore, the application of statistical and computational methods that detect patterns of epistasis across the genome, implemented in a system biology framework, may help solve the missing heritability problem.20

An important example of a 2-locus epistatic interaction is the susceptibility conferred by HLA and PTPN22 loci. The effect of PTPN22 coding for an R620W variant of LYP protein, measured by relative risk, is higher in low-risk HLA
genotypes. In vivo experiments in mice suggest that this common PTPN22 variant promotes degradation of IYP protein, indicating a loss-of-function allele. However, when main effects are included and the results converted to absolute risks, it can be seen that the additional risk due to PTPN22 is largest in the high-risk HLA group.21 Compared with PTPN22, there is no evidence for a statistical interaction with HLA class II genotypes for INS locus variants, further supporting the hypothesis that rs2476601/R620W is the most likely causal variant for T1D.22

The latest advances in personalized medical genomics that are translated to the clinic are aimed at comprehensive sequencing of the disease genome, exome, epigenome, and transcriptomes. Next-generation sequencing technology has contributed to the precise mapping and quantification of chromatin features, DNA modifications, and several specific steps in the cascade of information from transcription to translation.23,24

Increasing epidemiologic and experimental evidence supports the role of genetic and epigenetic alterations in the etiopathogenesis of diabetes. This review aims to briefly discuss recent results related to the role of genetic and epigenetic factors in the development of T1D.

MULTIFACTORIAL ETIOLOGY OF T1D

T1D results from deficient insulin production, as a consequence of autoimmune T-cell–mediated destruction of pancreatic β cells.25 The initiating events in T1D pathogenesis that trigger the provision of β-cell antigens (the first described target antigen is insulin) to dendritic cells and presentation to T cells for their activation are yet to be elucidated.26 As depicted in the Fig 2, proinflammatory cytokines, produced by infiltrating leukocytes and islet cells, including interleukin-1β (IL-1β), tumor necrosis factor-α (TNF-α), and interferon γ (IFN-γ), play a central role in β-cell failure and the development of diabetes.27 The destruction is caused by infiltration of the islets of Langerhans by dendritic cells, macrophages, and T lymphocytes, is specific for insulin-producing β cells, and does not affect glucagon-producing (α) or somatostatin-producing (δ) cells. Prolonged exposure to cytokines leads to a decreased capacity of β cells to produce and release insulin in response to secretagogues and, in the long term, destruction of the cells by apoptosis or necrosis.25,28 T1D typically affects young children, with onset as early as 1 year of age; most cases are diagnosed before the patient is 18 years old. The autoimmune process starts even earlier, as evidenced by the presence of autoantibodies in serum against T1D-specific antigens. The major autoantigens involved are insulin itself, glutamic acid decarboxylase, 65-kDa isoform (GAD65), insulin autoantigen 2 (an intracellular phosphatase), and zinc transporter 8.29,30 The autoantibodies, such as insulin autoantibody, glutamic acid decarboxylase antibody, insulinoma-associated protein 2 autoantibody, and zinc transporter 8 antibody, are markers of T-lymphocyte–mediated β-cell destruction.29 Both CD4+ (helper) and CD8+ (cytotoxic) T-lymphocyte subsets are important in T1D. The former recognize extracellular antigens and promote inflammation through cytokine release, whereas the latter respond to endogenously synthesized antigens and directly kill target cells. By the time symptoms leading to clinical diagnosis appear, most of the β-cell insulin secretory capacity has been lost. Therefore, prediction and prevention of T1D are of crucial importance, and it is hoped that better knowledge of the genetic underpinnings of T1D will greatly facilitate both.31

In addition to T-cell–mediated effectors of β-cell death, other apoptotic stimuli also contribute to β-cell apoptosis. Pancreatic β cells surrounded by inflammatory infiltrate are exposed to multiple cell stress mediators, hyperglycemia, and reactive oxygen species, and in such a cellular environment endoplasmic reticulum (ER) stress may also develop.25,32 The link between ER stress and T1D mellitus has been suggested by several investigators.33,34 A balance between ER protein load and folding capacity is needed for the proper folding of insulin within the ER. Large biosynthetic load, defects in ER folding machinery, and disturbances in the regulation of Ca2+ can all disrupt ER homeostasis, leading to accumulation of misfolded proteins within the ER lumen and ER stress.34,35 ER stress activates an elaborate adaptive process called the unfolded protein response.36 Unfolded protein response protects against developing T1D by ensuring that efficient function is maintained in β cells, despite large fluctuations in ER protein flux.37

OVERVIEW OF GENETIC FACTORS INVOLVED IN T1D ETIOPATHOGENESIS

Recent progress in understanding the genetic basis of T1D has led to an increased recognition of childhood diabetes heterogeneity.8 It has become apparent that not all diabetes presenting in childhood is type 1.58,39 None–type 1 diabetes mellitus can have marked differences in treatment and complications from T1D. Although the monogenic forms of diabetes remain mainly unrecognized, their molecular diagnosis has provided insight into the genes controlling human β-cell function, and their study has revealed the basic mechanisms of diabetes pathogenesis.40,41 These rare forms of monogenic diabetes have significant
Susceptibility to T1D is determined by complex interactions between several genetic loci and environmental factors. Common allelic variants at the HLA class II loci account for the major T1D genetic risk in children and young adults (primarily HLA-DRB1, HLA-DQA1, and HLA-DQB1 genes). Alleles at the HLA locus on chromosome 6p21, which encode the highly polymorphic antigen-presenting proteins, explain up to 50% of the familial clustering of T1D through a large variety of protective and predisposing haplotypes, and the remainder is contributed to by multiple loci, of which only 4 were known until recently. Two HLA class II haplotypes, DR4–DQ8 and DR3–DQ2, are present in ~90% of children T1D. Within the major histocompatibility complex class II haplotypes, the DRB1*1501–DQA1*0102–DQB1*0602 allele is strongly associated with T1D protection. The protection associated with DQA1*0102–DQB1*0602 is dominant over the susceptibility of all other DQ haplotypes, and this protection in children is thought to occur even in the presence of islet cell antibodies. Early-onset diabetes (EOD) (diagnosed before 5 years of age) is characterized by significant...
differences in the frequencies of HLA alleles and haplotypes from other T1D age groups. None of the 40 patients with EOD analyzed in a study by Hathout et al.52 (average age 2.6 years) had either of the protective (DRB1*1501 or DQB1*0602) alleles. In this EOD cohort, there was a negative correlation between GAD65 and the predisposing haplotype DR3/DQ2. Other established loci associated with T1D risk have smaller effects than HLA, conferring a relative risk ranging from odds ratio (OR) = 2.38 (11p15.5, INS) to OR = 1.05 (17q21.2, SMARCE1).53–56 A comprehensive list of genetic loci variants predisposing to T1D is provided in Supplemental Table 1.

Since completion of the HapMap project and the development of massively parallel array-based genotyping technologies, the results of genome-wide association studies (GWASs) have uncovered dozens of additional solidly replicated loci.53–58 Several consortia analyzed a remarkable number of individuals within cohorts of cases and controls (with trios in 1 study), using 61 A comprehensive list of genetic loci variants predisposing to T1D is provided in Supplemental Table 1. Since completion of the HapMap project and the development of massively parallel array-based genotyping technologies, the results of genome-wide association studies (GWASs) have uncovered dozens of additional solidly replicated loci.53–58 Several consortia analyzed a remarkable number of individuals within cohorts of cases and controls (with trios in 1 study), using 61 A comprehensive list of genetic loci variants predisposing to T1D is provided in Supplemental Table 1. Since completion of the HapMap project and the development of massively parallel array-based genotyping technologies, the results of genome-wide association studies (GWASs) have uncovered dozens of additional solidly replicated loci.53–58 Several consortia analyzed a remarkable number of individuals within cohorts of cases and controls (with trios in 1 study), using >200,000 markers for GWASs. The results identified only 5 loci associated with T1D, and in second-stage typing in additional cohorts, 18 loci reached genome-wide significance (P < 10\(^{-7}\)).55,56 These GWAS results are enriched by recent examination of associations from the largest meta-analysis to date between T1D and 2.54 million single-nucleotide polymorphisms (SNPs), involving a combined cohort of 9934 cases and 16,956 controls originating from 6 consortia and resulting in replication of 3 loci. Targeted follow-up of 63 SNPs, with fixed-effects P < .00001, in a replication set of 1120 affected case–parent trios from the T1DGC consortium, uncovered 3 new loci associated with T1D that reached genome-wide significance.59 GWASs have identified 140 regions with polymorphic variants that have statistically significant association with AID susceptibility, including T1D.60,61 Numerous examples of susceptibility variants shared by multiple AIDs are summarized in Fig 3. It is estimated that at least 44% of SNPs associated with 1 AID are associated with another, providing an example of biological pleiotropy (a genetic variant or gene that has a direct biological influence on >1 phenotypic trait).64 The common loss-of-function allele of PTPN22 decreases the risk of Crohn disease but increases the risk of rheumatoid arthritis (RA) and T1D.65 Another association analysis was performed with multiple SNPs in an allele-specific fashion, to compare genetic variation profiles of multiple sclerosis (MS), ankylosing spondylitis, autoimmune thyroid disease, RA, Crohn disease, and T1D with 5 non-AIDs. This approach resulted in identification of specific SNPs and genes with opposite risk profiles in 2 classes of diseases: RA and ankylosing spondylitis form one such class, and MS and autoimmune thyroid disease another.66 The HLA-related genes associated with T1D that reside in HLA class II DR and DQ loci are shared by other AIDs, such as RA, MS, systemic lupus erythematosus, and celiac disease. Analysis of celiac disease and T1D risk loci in a large sample of patients with RA and controls yielded significant evidence for association of the TAGAP locus, which is associated with risk of celiac disease and protection against T1D.66,67 According to recent cross-phenotype meta-analysis, which statistically supports the hypothesis that each independent SNP has multiple pleiotropic associations, 2 gene clusters were associated with T1D and other AIDs.68 This analysis was extended to protein–protein interactions, revealing that variant-encoded proteins within groups were more likely to interact with each other than proteins encoded by variants in other groups, providing more evidence for biological pleiotropy in AIDs.69 The immunochip is a cost-effective approach in AIDs that share several common loci, including T1D. To genotype common and rare variants, to map remaining heritability, and to fine-map established loci, an array of 200,000 known SNPs has been developed. However, it will probably be necessary to supplement it with novel SNPs identified from recent resequencing data.8 To understand T1D pathogenesis using GWAS data, it is important to apply an integrative network- or pathway-based approach. Data from pilot 1000 Genome Project studies and other sequencing data showed that next-generation deep sequencing in candidate regions has been successful in discovering the 4 protective rare variants in the IFIH1 gene.58,60 This combined approach of next-generation sequencing technologies and GWASs strengthened the hypothesis of disease pathogenesis involving virus–genetic interplay and raised type 1 interferon levels as a cofactor in β-cell destruction.53

Another interesting insight into T1D pathogenesis was provided by the mapping of T1D-associated linkage disequilibrium (LD) block with the IL-2–IL-21 ligand gene and variant in the α subunit of the IL-2 receptor gene IL2RA. The protein encoded by IL2RA is upregulated in effector T cells, and the protein expressed from the protective allele is higher in CD4+ memory cells, as an example of loss-of-function of activation of the adaptive immune response that is attenuated in T1D. Decreased expression of IL-2 cytokine from the risk allele in nonobese diabetic (NOD) mice results in suppression of T regulatory (T\(_{reg}\)) cells, indicating complex and cell-specific regulation of self-tolerance.19

Initiation of an immune response by recognition of a specific antigen by transmembrane T-cell receptors (TcRs) may result in autoimmune attack,
further potentiated by gene variants that impair antigen presentation or T-cell signaling during thymic self-tolerance. Such variants, which downregulate the TcR activation response, are mapped in *PTPN22*, *UBASH3A*, and *CTLA4* loci. Like the *PTPN22 R620W* variant, which may interfere with proper tolerance induction in the thymus through lower secretion of IL-2 and IL-10 by T cells on TcR activation, a T1D-associated *A17T CTLA4* variant similarly affects T-cell activation.69 *UBASH3A* is specifically expressed in lymphocytes, with a risk allele associated with gain of function through overexpression of protein in lymphoblastoid lines of patients with T1D.256 These insights into alterations of genetic pathways associated with risk variants in T1D represent targets for translational medicine efforts, exemplified by prevention and reversal of diabetes in mice by IL-2 supplementation and by discovery of small-molecule inhibitors of the *PTPN22* protein product Lyp.70,71

Functional insights into the role of T1D susceptibility loci revealed that candidate genes are involved in functions related to T-cell–mediated adaptive immune response and tolerance mechanisms and also to innate immunity in recognition of β-cell
Recent progress in understanding genetic susceptibility in T1D is in a complex dynamic with other hypotheses proposed to explain the increasing incidence of T1D, including infections; socioeconomic, lifestyle, and dietary factors; pollutants; prenatal environment; gut microbiome variability; and epigenetic factors such as DNA methylation and histone modification.74–76 β-cell development, maintenance, metabolism, and regeneration can all be influenced by epigenetic mechanisms. Also, immune responses, including the activation of T cells and induction of Treg cells, rely on appropriate epigenetic regulation. Furthermore, insulin and glucose metabolism influence the epigenome of tissues such as the liver, and potentially the pancreas, which could contribute to T1D associated pathologies. The hope is that advances in T1D epigenomics will provide new understanding of the pathogenesis, potential treatment, or even prevention of T1D.77

**EPIGENETIC CONTROL OF GENE TRANSCRIPTION IN DIABETES**

DNA methylation is a tissue-specific epigenetic mechanism of cellular response to stress essential for regulating the expression of genes. DNA methyltransferases catalyze the covalent addition of a methyl group to the C5 position in cytosine guanine dinucleotides (CpG), usually found in the promoter regions of genes, thus inducing transcriptional gene silencing. In AIDs, including T1D, defects in CpG methylation, resulting from decreased DNA methyltransferase activity, are associated with dysregulated gene expression, chromosomal instability, and increased susceptibility to disease development. Several nutritional factors, including diets deficient in methionine, folic acid, and choline, are associated with global hypomethylation.18 DNA tissue-specific hypomethylation in hepatocytes is observed in animal models as a result of diabetes-induced perturbations of methyl group and homocysteine metabolism, leading to functional methyl deficiency.19 In addition, a recent study of DNA methylation patterns of 7 CpGs proximal to the transcription start site in the INS gene promoter revealed that patients with T1D have a reduced methylation of CpG-19, CpG-135, and CpG-234 ($P = 2 \times 10^{-16}$) and increased methylation of CpG-180 compared with controls.79 A novel method, the highly sensitive methylation-specific quantitative polymerase chain reaction assay, was successfully applied in the detection of circulating β-cell DNA in peripheral blood from diabetic mice and holds great promise for future monitoring of β-cell death and the prediction of diabetes onset.80 Another study, an epigenome-wide association study, involving 27,458 different CpG sites within 14,475 promoters in the monocytes of 15 T1D-discordant monozygotic twin pairs, identified 132 different CpG sites where differences significantly correlated with diabetic state.81 This study included identification of T1D-specific methylation variable positions in patients before the appearance of T1D autoantibodies, thus demonstrating that T1D-specific methylation variable positions antedate clinical disease, including epigenetic changes in HLA class II gene, HLA-DQB1, or in GAD2, which encodes GAD65, a major T1D autoantigen involved in T1D etiology.82

Several classes of small RNA can regulate genes by targeting transcripts in the cytoplasm, such as miRNAs, small interfering RNAs, repeat-associated RNAs, and germline-specific RNAs.82 The human genome encodes numerous noncoding molecules, including >1600 miRNA precursors, generating up to 2237 mature miRNAs. Binding of miRNAs to 3’ untranslated regions (UTRs) of target mRNAs is consistent with a role for miRNAs as key negative regulators of gene expression, involved in the
control of cellular processes and disease pathogenesis. The ability of noncoding RNA molecules to associate with 3’ UTRs of target mRNAs and to repress their translation has been suggested to be necessary for proper β-cell development and function. During the initial stages of T1D, β cells are submitted to diverse types of stress stimuli, causing dysfunction and death that may be associated with modifications in miRNA localization, potentially resulting in β-cell dysfunction. Mature miRNAs can be released by cells into the extracellular compartment. Based on an experimental model, increased levels of circulating miR-375 can be used as a marker of β-cell death and a potential predictor of diabetes, before the onset of hyperglycemia. In serum samples from children with new-onset T1D and age-matched healthy controls, residual β-cell function and glycemic control after 3 months of clinical disease were associated with miR-25 overexpression, present soon after diagnosis, providing evidence for the role of miRNAs as clinically applicable biomarkers in T1D. The expression levels of miR-326 are significantly higher in peripheral blood lymphocytes from T1D antibody–positive patients than in patients with T1D without autoantibodies. Analysis of miRNA expression in Treg cells in patients with T1D revealed that miR-342 and miR-19 were downregulated, whereas miR-510 showed significantly higher expression. In light of these results, measurement of specific miRNA levels may be useful for identifying people at risk for developing T1D, possibly preventing the development of the disease. Lentiviral transgenesis has been used to generate NOD mice in which the Slc11a1 T1D candidate gene was silenced by RNA interference. Silencing reduced the frequency of T1D, thus demonstrating a role for Slc11a1 in modifying susceptibility to T1D.

**PROFILES OF EPIGENETIC HISTONE POSTTRANSLATIONAL MODIFICATIONS IN DIABETES**

The chromatin status of gene promoters determines gene expression modulation as a response to various stimuli. Within the identified network in lymphocytes, posttranslational modifications (PTMs) at the N-terminal amino acids of histones, including acetylation, methylation, phosphorylation, or ubiquitination, might affect the etiology of T1D and its complications. Recent studies have shown that diabetic stimuli, including hyperglycemia, can alter the levels of key histone PTMs, including H3K9Ac, H3K9me2, H3K9me3, and H3K9me1, at the promoters of several genes related to diabetes pathogenesis. CLTA4, a known T1D susceptibility gene, is also overexpressed in the lymphocytes of patients with T1D compared with controls, in addition to displaying differential H3K9me2 methylation. Recent data from key histone PTM profiling, obtained by chromatin immunoprecipitation linked to microarrays, showed marked variations in H3K9Ac levels at upstream regions of HLA-DRB1 and HLA-DQB1 within the T1D locus in monocytes of patients with T1D relative to controls.

In summary, T1D is a polygenic AID with a complex origin, caused by the interplay of genetic, epigenetic, and environmental factors. Our current knowledge of T1D etiopathogenesis has been significantly improved by GWAS and recent epigenome-wide association study analyses aimed at identifying epigenetic mechanisms with the potential to reveal specific networks that may be involved in gene–environment interactions. Advances in epigenetic research may enable the identification and potential use of circulating miRNAs as epigenetic blood biomarkers for the detection of β-cell death and diabetes prediction.

**ACKNOWLEDGMENTS**

Supported by the Ministry of Education and Science, Republic of Serbia, project no. 41012. The authors are grateful to Dr Edward Petri for editing assistance.

**REFERENCES**


22. Smyth DJ, Cooper JD, Howson JM, et al. PTNP22 Trp620 explains the association of chromosome 1p13 with type 1 diabetes and shows a statistical interaction with HLA class II genotypes. *Diabetes*. 2008;57(6):1730–1737


50. Sanjeevi CB, Landin-Olsson M, Kockum I, Dahlquist G, Lernmark A. Effects of the second HLA-DQ haplotype on the


54. Wellcome Trust Case Control Consortium. Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. Nature. 2007;447(7145):661–678


88. Kissler S, Stern P, Takahashi K, Hunter K, Peterson LB, Wicker LS. In vivo RNA interference demonstrates a role for Nramp1 in...


(Continued from first page)
Genetic and Epigenetic Factors in Etiology of Diabetes Mellitus Type 1
Karmen Stankov, Damir Benc and Dragan Draskovic
Pediatrics; originally published online November 4, 2013;
DOI: 10.1542/peds.2013-1652

The online version of this article, along with updated information and services, is located on the World Wide Web at:
/content/early/2013/10/30/peds.2013-1652