

CHAPTER 12

GENETICS OF TYPE 1 DIABETES

Stephen S. Rich, PhD, Henry Erlich, PhD, and Patrick Concannon, PhD

Dr. Stephen S. Rich is Director of the Center for Public Health Genomics, University of Virginia, Charlottesville, VA. Dr. Henry Erlich is Senior Scientist at the Children's Hospital Oakland Research Institute, Oakland, CA, and past Director of Human Genetics and Vice-President of Discovery Research, Roche Molecular Systems, Pleasanton, CA. Dr. Patrick Concannon is Director of the Genetics Institute, University of Florida, Gainesville, FL.

SUMMARY

Type 1 diabetes is a complex disease that has both genetic and environmental determinants. Based on twin and family studies, the estimated contribution of genetic factors to type 1 diabetes risk is ~50%. Genes and their variants within the human major histocompatibility complex (MHC), including the human leukocyte antigen (HLA) class I (HLA-A, -B, and -C) and class II (HLA-DR, -DQ, and -DP) loci, account for about one-half of the genetic risk of type 1 diabetes. Three amino acid positions in HLA-DQ and HLA-DR define ~90% of the variation in the MHC, with evidence of interactions between pairs of HLA haplotypes that affect antigen binding. Other major contributors to type 1 diabetes genetic risk have been identified through candidate gene and linkage studies and include variants in or near the *INS*, *CTLA4*, *IL2RA*, and *PTPN22* genes. Genome-wide association approaches have revealed additional loci containing common variants with relatively small individual effects on type 1 diabetes risk. International efforts led by the Type 1 Diabetes Genetics Consortium and others have identified over 40 non-MHC loci and narrowed the likely candidate genes and variants substantially. The majority of non-MHC variants affect gene regulation rather than directly altering protein structure. Both analytic and molecular work are required to assess the functional significance of the variants in type 1 diabetes susceptibility genes in order to identify critical biologic pathways that could lead to novel interventions and therapeutics.

INTRODUCTION

The natural history of type 1 diabetes is based on precipitating events in an individual with genetic susceptibility. The evidence for genetic factors contributing to type 1 diabetes risk comes from twin and family studies that estimate the familial aggregation of the disease based on risk in relatives of an affected individual. In monozygotic twins, who share 100% of their genes, when a member of the pair has type 1 diabetes, the risk to the co-twin is ~50%, suggesting that both genetic and nongenetic factors contribute to risk (1). This concordance increases to 65% by age 60 years and 89% for auto-antibody-positive pairs (2). Presence of high-risk genotypes of the human leukocyte antigen (HLA) loci (HLA-DR3, -DR4) in the major histocompatibility complex (MHC) (3) and the *INS* gene (4) are found in higher frequency in concordant pairs, suggesting a major impact of genetic factors and heterogeneity in concordance rates. The population risk is decreased to ~8% in dizygotic twin pairs, similar to the risk observed for siblings, who also share 50% of their genes, but have a lesser extent of common

environmental exposures than twins. In population studies, the estimated prevalence of type 1 diabetes in non-Hispanic whites is approximately 4 per 1,000; thus, the increased risk in siblings (8%) relative to the prevalence in the general population (4/1,000) is consistent with a major genetic contribution to type 1 diabetes risk (5,6).

Type 1 diabetes is most common in non-Hispanic whites (i.e., populations of European ancestry); it also occurs in those of African, Hispanic, and Asian ancestry but at decreasing prevalence and incidence. Despite extensive epidemiologic data on prevalence of type 1 diabetes in varied ethnic groups, there is a critical absence of data on risk to siblings in diverse populations, preventing estimation of genetic impact on type 1 diabetes risk in nonwhite populations. Although nearly one-half of type 1 diabetes in non-Hispanic white populations is diagnosed after age 20 years, there is a similar absence of information on the risk to siblings of adult-onset type 1 diabetes. This lack of data can be attributed to the concept that onset

of diseases (e.g., type 1 diabetes) early in life is more likely to be "genetic" than when the disease has onset later in life. As a result, estimates of the impact of genetic factors on risk of type 1 diabetes in adults (of any population ancestry) are lacking.

The genetic risk ratio in siblings (λ_s), defined by the ratio of the sibling risk to the population prevalence, is ~16 in type 1 diabetes of European ancestry (7), much higher than that in type 2 diabetes, for example, yet this figure provides no insight into either the number of genes contributing to type 1 diabetes risk or the sizes of their effects. Although ~50% of the risk of type 1 diabetes can be attributed to genetic (familial) factors, an additional question can be asked about the extent of this genetic risk accounted by specific genes and their variants. For example, in families with two children with type 1 diabetes, the expectation under the hypothesis that a variant in a gene has no effect on type 1 diabetes risk is that 25% of sibling pairs would share both copies of the variant (same genotype),

50% would share one copy of the variant in the genotype, and 25% would share no copies of the variant (completely different genotypes). In fact, the risk to siblings who share two HLA haplotypes is 55%, much greater than 25%, indicating that a significant proportion of the total genetic risk can be attributed to the HLA region (7).

The recognition that ~50% of risk for type 1 diabetes was genetic and approximately one-half of the genetic risk was due to factors in the MHC provided the stimulus for two subsequent research paths. The first was to delineate the specific genes and mechanisms in the MHC (the HLA genes and also other genes and variants)

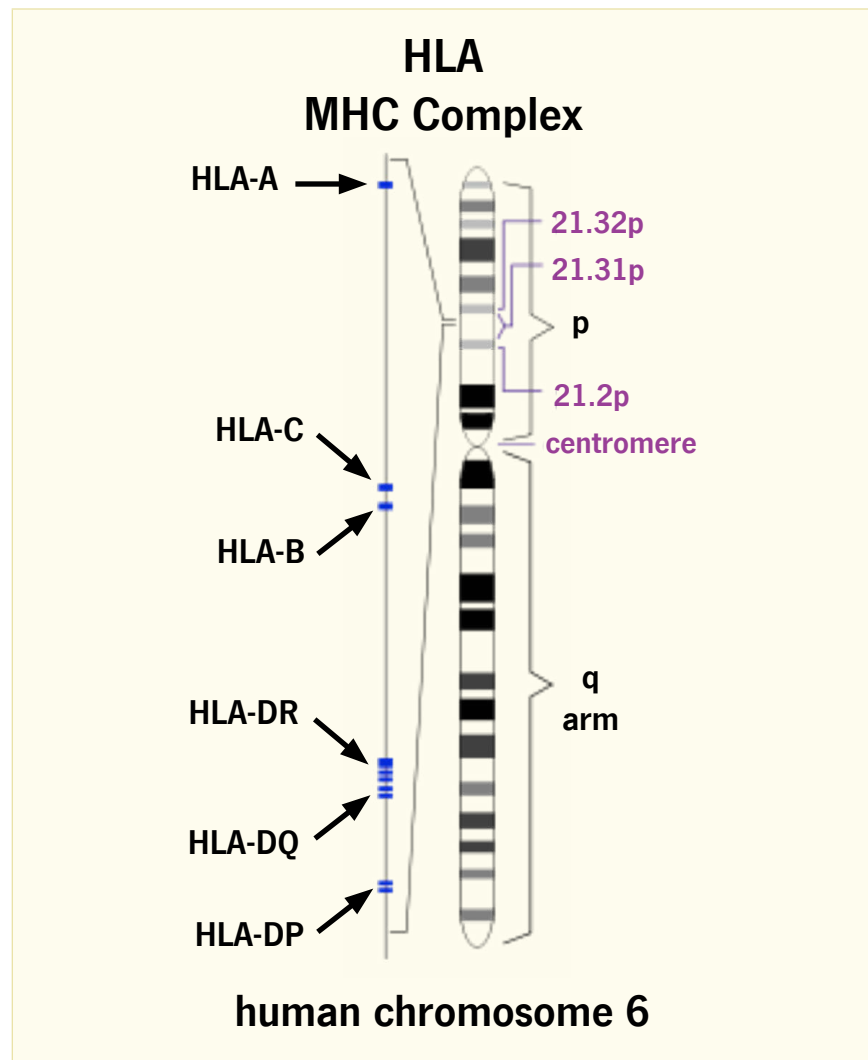
that account for the majority of type 1 diabetes genetic risk. The second was to identify the non-MHC genes that account for the remaining 50% of genetic risk for type 1 diabetes. Two primary study designs were used for these efforts: family-based (typically families with two parents and two type 1 diabetes-affected children) and case-control (a series of type 1 diabetes cases and a series of unaffected controls). The family-based designs were used for linkage analyses (co-segregation of the alleles transmitted from the parents to the affected children, with evidence of increased sharing of genotype in the affected children) or for association in the presence of linkage

analyses (alleles transmitted from parents to affected offspring are “case” alleles, while those not transmitted are “control” alleles). The case-control approach was used to determine the frequency of the alleles in cases and controls, with a significant difference in frequency suggesting an association with type 1 diabetes for that genetic variant. As collection of large numbers of affected sibpair (ASP) families (for robust statistical power) was difficult, much of the genetic evaluation of genes contributing to type 1 diabetes employed a case-control design. The following sections expand on the evolution of the genetic technology and findings related to the genetic basis of type 1 diabetes.

MAJOR HISTOCOMPATIBILITY COMPLEX AND TYPE 1 DIABETES RISK

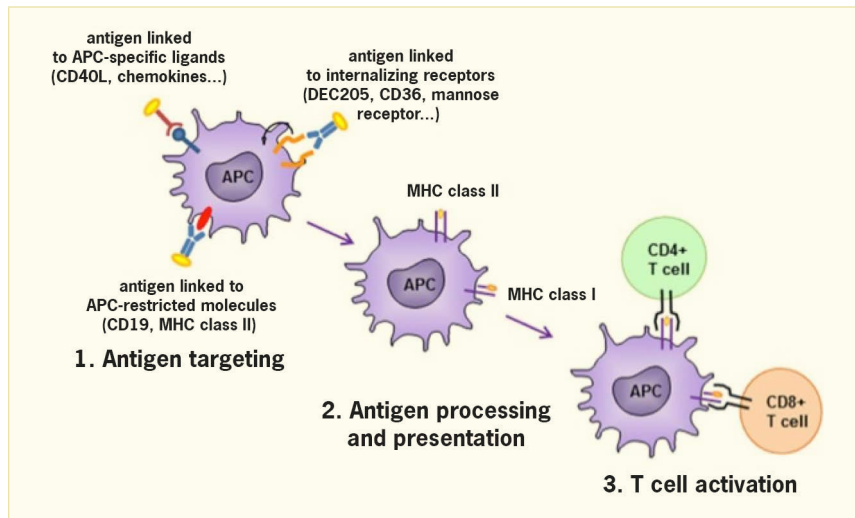
Early efforts (i.e., 1970s to 2000) to localize and identify genes that contribute to the occurrence of type 1 diabetes, as well as other autoimmune diseases, focused on genes involved in the immune response. Obvious candidates, in this regard, were genes encoding the highly polymorphic HLA molecules that play critical roles in the immunologic distinction between self and non-self, as well as in the presentation of antigens to the cellular immune system. In humans, the MHC is a gene-rich region on chromosome 6p21.3 that includes genes encoding the HLA class I (HLA-A, -B, and -C) and class II (HLA-DR, -DQ, and -DP) molecules (Figure 12.1). The importance of the MHC in type 1 diabetes risk is likely through its role in the presentation of peptide antigens to T cells; genetic variation in this system could act centrally by interfering with the tolerance of lymphocytes during their maturation in the thymus, or peripherally by altering the repertoire of antigens presented. The immune response is centered on antigen presentation, in which foreign antigens are recognized by antigen-presenting cells (APCs), processed into peptides, complexed to the MHC, and presented on the cell surface, where they can potentially be recognized by T cells (Figure 12.2). As the T cell does not recognize “free” foreign peptides, the processing requires APCs and the action

FIGURE 12.1. Human Leukocyte Antigen Major Histocompatibility Complex



Genomic localization of the human major histocompatibility complex (MHC) on chromosome 6p21.3 and positions of the human leukocyte antigen (HLA) class I (HLA-A, -B, -C) and class II (HLA-DR, -DQ, -DP) loci are shown.

SOURCE: Reference 44

FIGURE 12.2. Antigen Presentation and T Cell Activation

Antigen is linked to APC via multiple mechanisms (step 1) followed by MHC class II recognition, antigen processing (step 2), and presentation to T cells for removal (step 3). APC, antigen-presenting cell; MHC, major histocompatibility complex.

SOURCE: Reference 45, copyright © 2012 Frontiers, reprinted with permission

of a number of other genes, some also located in the HLA region, to complex the peptide with MHC. Thus, a critical aspect of protection is the ability of the MHC to provide a broad array of options to receive the foreign peptide, with these options defined by genetic variation.

The genes in the MHC contain many alternative forms (many alleles) and thereby provide extensive variation in humans, perhaps the most variable in the human genome. This variation is critical, as the MHC controls a major part of the human immune response through the interactions of its cell surface molecules with other molecules or peptides (both from itself and from external sources). The presence of many alleles in each of the HLA genes (class I and class II) makes it highly likely that any two individuals in a population (except for monozygotic twins) have a different combination of MHC alleles. Most cells in the body can use MHC class I molecules to present foreign peptide to (CD8+) T cells and “act” as APCs. Typically, the APCs internalize the foreign peptide and display a fragment bound to an MHC class II molecule on the cell surface (Figure 12.2). The T cell will then recognize this complex and initiate T cell activation. This becomes critical as the focus of the MHC is on immune surveillance. Thus, while there can be relatively broad ability for the MHC to recognize a broad array

of peptides, there can also be specificity as to which peptides are bound. The variation permits both a broad response to peptides with a few highly variable genes and also provides protection in a population of individuals from a new foreign peptide due to extreme diversity.

The initial report of a strong association of HLA with type 1 diabetes occurred in 1973 with alleles of the class I HLA-A (or, then, HL-A) locus (8). Many reports confirmed and extended the association of type 1 diabetes with antigens/alleles of the class I (HLA-A and -B) and class II (HLA-DR, -DQ, and -DP) loci. Importantly, the structure of the MHC on 6p21.3 contains a cluster of HLA loci that are physically close (within 4 Mb, Figure 12.1) and, therefore, genetically correlated (in linkage disequilibrium, LD). The extent of the LD in the MHC spans the ~4 Mb interval and results in the transmission of HLA “haplotypes” from parent to child. As LD is the occurrence of some combinations of alleles at adjacent loci more (or less) often than expected from the frequencies of the individual alleles, groups of alleles on a chromosomal segment can be inherited as a unit, or as a haplotype. The human MHC exhibits strong LD and specific HLA haplotypes are associated with type 1 diabetes risk and provide a “fingerprint” of the transmission of disease-associated variation across populations.

Although statistical methods can be used to construct HLA haplotypes from unrelated individuals, the most precise method for observation and estimation of haplotypic association comes from family studies. As noted above, family studies have been conducted in non-Hispanic white populations, with a high prevalence and typically restricted to those probands with onset of type 1 diabetes before age 16 years (but onset up to age 35 years in siblings). A common measure of association is the odds ratio (OR), an epidemiologic statistic that measures the relationship between an exposure (in this case, genotype or haplotype) and an outcome (type 1 diabetes). As the odds ratio is used to compare the relative odds of the outcome given the exposure, it can be interpreted as “risk” (OR >1), “protection” (OR <1), or neutral (OR 1). A family study of type 1 diabetes (9) demonstrated that the strongest class I associations with type 1 diabetes occurred with the HLA-B8 (OR 3.20) and HLA-B15 (OR 3.69) alleles and the class II associations with HLA-DR2 (OR 0.21), HLA-DR3 (OR 3.54), HLA-DR4 (OR 6.81), HLA-DR5 (OR 0.30), and HLA-DR7 (OR 0.24) alleles. Presence of the HLA-B8, -B15, -DR3, and -DR4 alleles increased type 1 diabetes risk, while the presence of HLA-DR2, -DR5, and -DR7 decreased type 1 diabetes risk. However, the haplotypes formed by HLA-B7-DR2 (OR 0.10) and HLA-B15-DR4 (OR 7.55) exhibited greater extremes of association than the individual alleles, suggesting genetic complexity within the MHC association with type 1 diabetes.

It should be noted that these type 1 diabetes-associated alleles and haplotypes were obtained from studies that primarily recruited Caucasians. The association between HLA alleles and haplotypes with type 1 diabetes remains strong across populations of non-Caucasian ancestry (e.g., in Asians, a high risk is seen for DRB1*09:01, an allele found at low frequency in Caucasians and, therefore, not associated with Caucasian type 1 diabetes risk). There are also ethnic differences in HLA allele frequencies that differ from Caucasian populations, e.g., African-specific HLA-DR3 and HLA-DR7

haplotypes have opposite effects on type 1 diabetes susceptibility than their Caucasian counterparts. In addition, there are difficulties in assembling large cohorts of non-Caucasian type 1 diabetes cases or families, given differences in disease risk among races/ethnicities. The study of type 1 diabetes in non-Caucasian populations, both from an epidemiologic and a genetic perspective, is an important area of research that could provide insights on type 1 diabetes pathogenesis.

Subsequent studies have investigated the association of class I and class II alleles with type 1 diabetes, adjusting for LD within the MHC in numerous populations. The most comprehensive evaluation of these associations has occurred at the HLA and Immunogenetics Workshop (10). The class I alleles most significantly associated with type 1 diabetes are HLA-B*57:01 (protective) and HLA-B*39:06 (susceptible). The HLA-B*57:01 allele has been shown to be less frequent in cases with type 1 diabetes than controls in several populations and is associated with extreme sensitivity to abacavir (11,12), a treatment for HIV-1 that slows the spread of the virus. The HLA-B*39:06 allele increases the risk on HLA-DR4 susceptibility haplotypes in Scandinavian populations, as well as other risk haplotypes (DR3, DR1, and DR8) in larger cohorts. Other significantly associated class I alleles include HLA-A*24:02, HLA-A*02:01, HLA-B*18:01, and

HLA-C*05:01 (susceptibility); HLA-A*11:01, HLA-A*32:01, HLA-A*66:01, HLA-B*07:02, HLA-B*44:03, HLA-B*35:02, HLA-C*16:01, and HLA-C*04:01 (protective) (10,13).

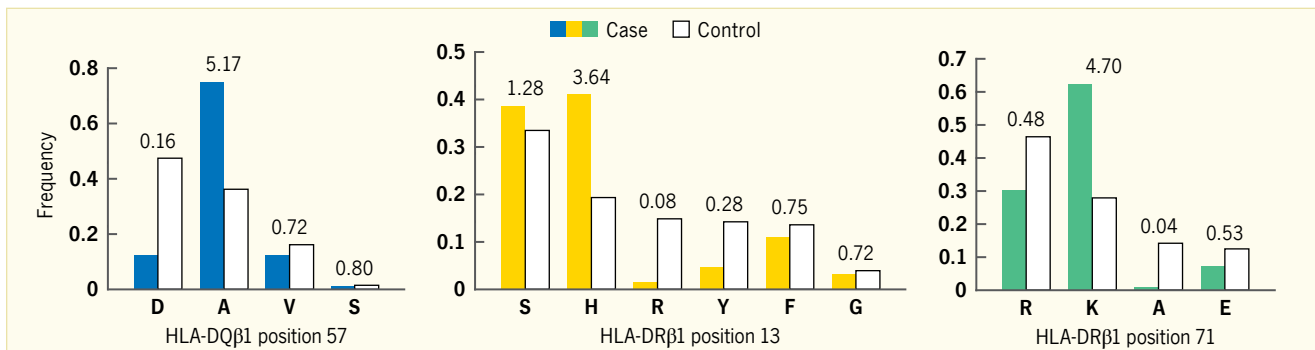
HLA allele and haplotype frequencies differ among populations, perhaps contributing to the differences in type 1 diabetes prevalence across geographic areas. A North-South gradient in type 1 diabetes prevalence has been observed, with Scandinavians having the highest type 1 diabetes risk. Sardinia, however, has a very high type 1 diabetes risk, due in part to the founder effect that established particularly high frequencies of the type 1 diabetes-susceptible HLA-DR3 and HLA-DR4 alleles and haplotypes and the low frequency of the type 1 diabetes-protective HLA-DR2 allele and haplotype (14).

Further complexity in the MHC and risk of type 1 diabetes has been uncovered. Variation in the genes encoding the HLA-DR and HLA-DQ molecules is associated with risk of type 1 diabetes with odds ratios in excess of 10 for susceptible genotypes (e.g., DR3/DR4 or DRB1*03:01-DQB1*02:01/DRB1*04-DQB1*03:02) or less than 0.1 for protective genotypes (e.g., DR2/DRX, where "X" is neither DR3 nor DR4 [or DR2]; the dominantly protective haplotype is DRB1*15:01-DQB1*06:02). The class II molecules encoded by these genes combine to form heterodimeric ($\alpha\beta$) protein receptors that are typically

expressed on the surface of APC. HLA-DQ is determined by a polymorphic α -chain (encoded by HLA-DQA1 locus) and a polymorphic β -chain (encoded by HLA-DQB1 locus). In contrast, HLA-DR is determined by a monomorphic α -chain (encoded by the HLA-DRA locus) paired with a highly polymorphic β chain. Most DR haplotypes in the population contain two loci: DRB1 and a second DRB locus (either DRB3, DRB4, or DRB5). A few DR haplotypes (e.g., DRB1*01, DRB1*08, and DRB1*10) do not have a second DRB locus. A heterozygous individual could encode up to four different DRB molecules (15). Thus, type 1 diabetes risk is determined by specific genotype combinations; for example, the genotype HLA-DRB1*03(HLA-DQB1*02)/HLA-DRB1*04(HLA-DQB1*03:02) confers the highest risk and has the highest frequency in type 1 diabetes cases with youngest onset.

The allelic variations in the critical HLA class II genes (HLA-DRB1, HLA-DQA1, and HLA-DQB1) are thought to alter specific amino acid residues that affect binding of foreign peptides leading to initiation of the autoimmune disease process. Only recently have independent amino acid positions and large-scale genetic studies been available to examine the contribution of HLA class II alleles and their interactions in type 1 diabetes (16). In a large series of European cases, controls and families, over 7,000 single nucleotide polymorphisms (SNPs) and

FIGURE 12.3. Effect Sizes for Amino Acid Residues



Case (filled bars) and control (unfilled bars) frequencies, as well as unadjusted univariate odds ratio estimates, are shown for each residue at HLA-DQB1 position 57, HLA-DRβ1 position 13, and HLA-DRβ1 position 71. The number above the paired (case/control) bars is the unadjusted univariate odds ratio. The letter underneath each set of paired (case/control) bars is the single letter code for the amino acid residue: A, alanine; D, aspartic acid; E, glutamic acid; F, phenylalanine; G, glycine; H, histidine; K, lysine; R, arginine; S, serine; V, valine; Y, tyrosine.

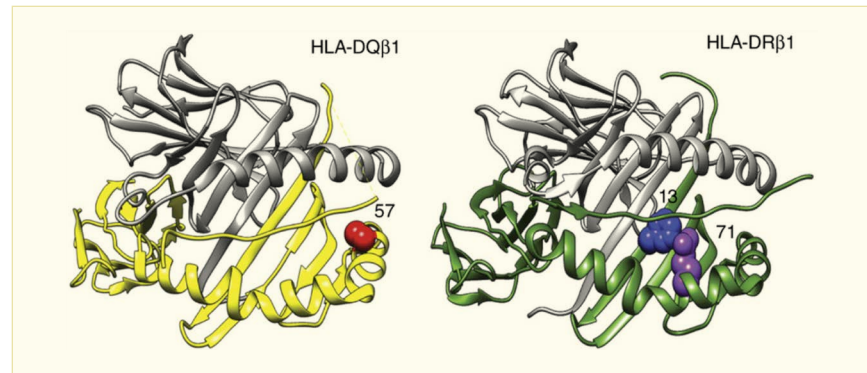
SOURCE: Reference 16, copyright © 2015 Nature Publishing Group, reprinted with permission

amino acid residues at 399 positions for eight HLA genes were examined. As expected, the single most associated risk variant was HLA-DQB1*03:02 that encodes an alanine at HLA-DQ β 1 position 57, while the aspartic acid at this position is most protective. The second independent association is at HLA-DR β 1 position 13, where a histidine residue (tagging HLA-DR4) confers risk (arginine is protective), as does a serine residue that tags HLA-DR3 (tyrosine is protective). The HLA-DR β 1 residue at position 71 is a third independent association, with lysine conferring strong risk (tagging HLA-DR3 and HLA-DR4), while an alanine residue at this site is protective (Figure 12.3) (16).

The HLA-DQ β 1 position 57, HLA-DR β 1 position 13, and HLA-DR β 1 position 71 are each located in the peptide-binding groove of the HLA molecule (Figure 12.4) (16). HLA-DR β 1 positions 13 and 71 line the P4 pocket of HLA-DR. The HLA-DQ β 1 position 57 alone accounts for ~15% of the total risk of type 1 diabetes, and HLA-DR β 1 position 13 and HLA-DR β 1 position 71 account for an additional 12% of risk (16). Thus, together these three amino acid positions capture ~27% of type 1 diabetes risk, or ~80% of the MHC-associated risk. The total type 1 diabetes risk explained by independent (additive) effects in the classical HLA genes is ~34% (Figure 12.5). Structurally, the known association of HLA-DQ β 1 position 57 would alter the HLA-DQ P9 pocket; however, the novel associations with the HLA-DR β 1 position 13 and HLA-DR β 1 position 71 sites would alter the HLA-DR P4 pocket. This structural site may be important in binding specific autoantigens that are associated with type 1 diabetes risk.

Although the influence of HLA class II genes on type 1 diabetes risk remains unquestioned, significant residual evidence of association with type 1 diabetes can be detected in the MHC region after controlling for the effects of HLA class II loci (16). Analysis of HLA haplotypes suggest that non-additive effects are common within the MHC. HLA-DQ β 1 position 57 and HLA-DR β 1

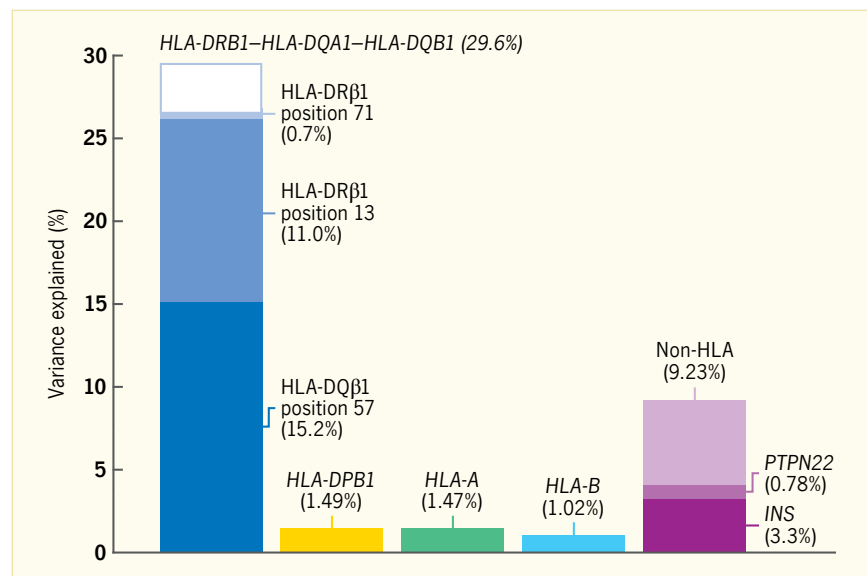
FIGURE 12.4. Structural Model of HLA-DQ and HLA-DR Molecules



HLA-DQ β 1 position 57, HLA-DR β 1 position 13, and HLA-DR β 1 position 71 are each located in the respective molecule's peptide-binding groove. HLA-DR β 1 positions 13 and 71 line the P4 pocket of the HLA-DR molecule.

SOURCE: Reference 16, copyright © 2015 Nature Publishing Group, reprinted with permission

FIGURE 12.5. Phenotypic Variance in the *HLA-DRB1*–*HLA-DQA1*–*HLA-DQB1* Locus



Assuming the liability threshold model and a global type 1 diabetes prevalence of 0.4%, all haplotypes in *HLA-DRB1*–*HLA-DQA1*–*HLA-DQB1* together explain 29.6% of total phenotypic variance. HLA-DQ β 1 position 57 alone explains 15.2% of the variance; the addition of HLA-DR β 1 position 13 and HLA-DR β 1 position 71 increases the explained proportion to 26.9%. Therefore, these three amino acid positions together capture over 90% of the signal within *HLA-DRB1*–*HLA-DQA1*–*HLA-DQB1*. In contrast, variation in *HLA-A*, *HLA-B*, and *HLA-DPB1* together explain approximately 4% of total variance. Genome-wide independently associated SNPs outside the HLA together explain about 9% of variance; rs678 (in *INS*) and rs2476601 (in *PTPN22*) explain 3.3% and 0.78%, respectively. HLA, human leukocyte antigen; SNP, single nucleotide polymorphism.

SOURCE: Reference 16, copyright © 2015 Nature Publishing Group, reprinted with permission

position 13 are the strongest contributors to both additive and interactive effects on type 1 diabetes risk in the MHC (17). These two strongest interacting positions are in separate HLA molecules (HLA-DQ for position 57 and HLA-DR for position 13). In a detailed interaction analysis of the MHC, the effect of interaction within a site (dominance) was shown to have a small (~1%), but significant impact on the contribution to risk of type 1 diabetes (17). Including interactions at the individual haplotypes, rather than alleles at a given

site, resulted in an additional 3% of risk of type 1 diabetes. Together, the additive and interaction effects of the classical HLA sites and amino acid residues account for over 90% of the MHC-type 1 diabetes association.

NON-MAJOR HISTOCOMPATIBILITY COMPLEX RISK LOCI

The fundamental process by which genetic variation is associated with type 1 diabetes risk is through assessing DNA sequence variation that differs in type 1 diabetes cases and controls or through co-segregation with type 1 diabetes in families. One such DNA variant occurs when a single nucleotide in the genome differs in a population with greater than 5% frequency. This variant is then termed a single nucleotide polymorphism (SNP). There are millions of SNPs throughout the human genome, roughly one SNP every 2,000 bases.

LINKAGE MAPPING

The primary focus of genome-wide linkage scans in type 1 diabetes was to identify regions of the genome that exhibited a significant deviation of “sharing” (identity-by-descent, IBD) by sibling pairs from that expected for Mendelian segregation of chromosomes (e.g., 25% sharing two alleles IBD, 50% sharing one allele IBD, 25% sharing no alleles IBD). In type 1 diabetes, this process was typically performed on multiple affected siblings in a family, with a large number of families required to achieve robust statistical power given the expected size of risk in non-MHC genes (18,19). This strategy effectively addressed an underlying hypothesis of detecting a gene with variant alleles that have a large effect on risk of type 1 diabetes.

Unlike the MHC, for which a sample of 100 ASP families could detect significant evidence of linkage, the estimated risks for other genes suggested a sample size of as many as 4,000 ASP families might be required. The Type 1 Diabetes Genetics Consortium (T1DGC) was created to assemble a collection of ASP families for the purpose of conducting genome-wide linkage scans for type 1 diabetes (20). Although the resolution of the linkage approach was limited to ~1 Mb regions, the first major analysis of 2,658 ASPs confirmed the contribution of the HLA region, insulin (*INS*) gene, and *CTLA4* (18). With additional ASP family collection, analysis of 4,422 ASP families continued to identify the contribution of the MHC

and the *INS* locus, but no compelling evidence for other susceptibility genes with similar large effects was observed (19). These results suggested that additional loci contributing as much as 50% of the genetic risk have smaller effects, may be common (allele frequency greater than 5%) in the population or private (unique) to individual families, and require a different experimental approach for detection.

GENOME-WIDE ASSOCIATION SCANS

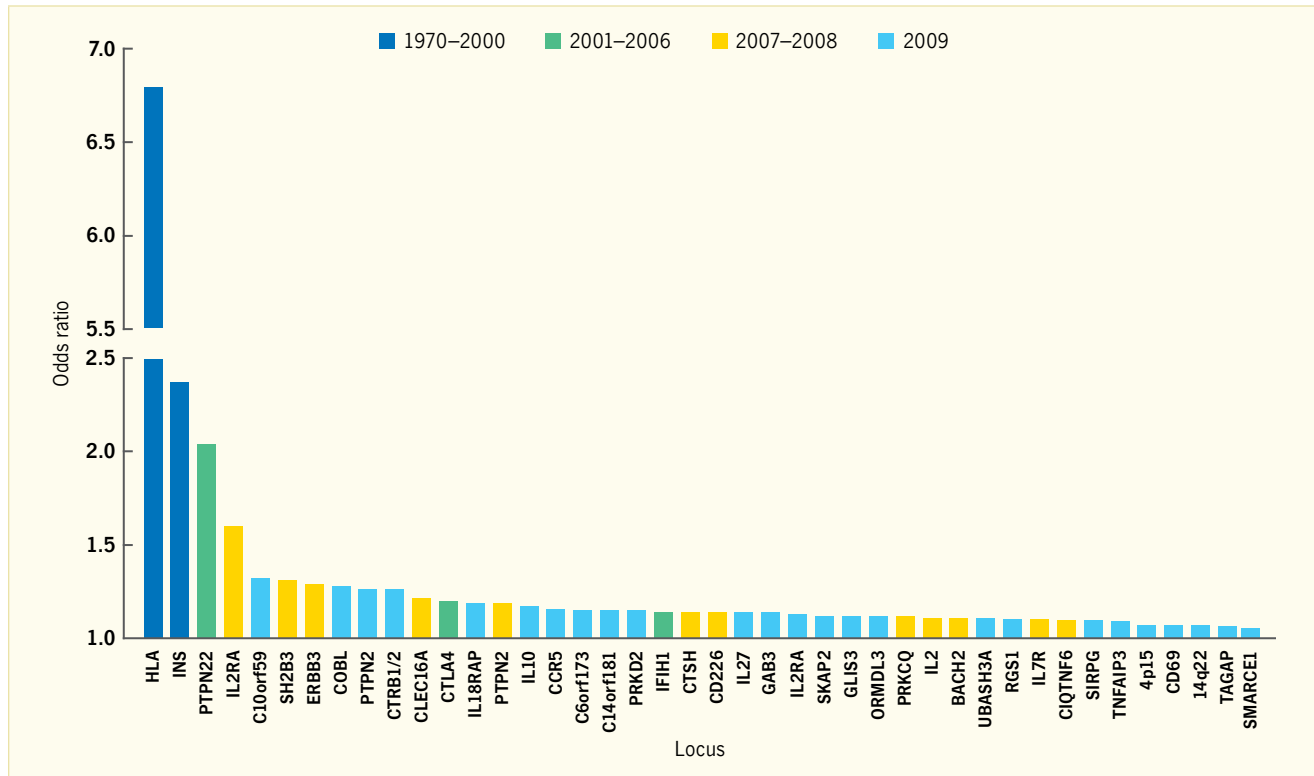
With the improved genomic coverage through SNP discovery from the International HapMap Project (21) and decreasing costs of genotyping, the use of genome-wide association scans (GWAS) with robust statistical power is feasible. The primary source of participants for GWAS is unrelated type 1 diabetes cases and controls, as the statistical method is association and not linkage. Through association, the SNP that exhibits a difference in frequency between cases and controls (type 1 diabetes-associated variant) is thought to be in LD with the “true” type 1 diabetes causal variant. However, family-based association analyses provide important complementary results because of their resistance to stratification bias, caused (for example) when genetic differences in cases and controls occur due to sampling from populations of different ancestry, and are unrelated to disease.

The first GWAS publications in type 1 diabetes focused on nonsynonymous SNPs (those that change the amino acid sequence of protein), with the assumption that variants resulting in an amino acid change would increase the likelihood of detecting a functional change associated with type 1 diabetes risk (22). From this design, a novel gene, *IFIH1*, was identified as having multiple rare protein coding variants associated with type 1 diabetes. The function of *IFIH1* was not in the obvious immune pathway of cytokines or beta cell proteins—*IFIH1* is a cytoplasmic sensor of viral RNA that is part of the innate immune system and the response to viral infection. Notably, *IFIH1* is involved in the response to picornaviruses, a family that includes viruses

previously implicated in type 1 diabetes risk (23). Rare, inactivating alleles of *IFIH1* are underrepresented among those with type 1 diabetes, suggesting that loss-of-function is protective from type 1 diabetes. In this manner, GWAS results can lead to discovery of genes that identify novel biologic pathways and avenues to novel therapeutic (or interventional) drug targets to prevent or reduce the burden of type 1 diabetes.

The Wellcome Trust Case-Control Consortium (WTCCC) conducted the first extensive GWAS using the Affymetrix GeneChip 500K Mapping Array Set in participants from the United Kingdom (24). Approximately 2,000 subjects for each of seven diseases and a shared set of approximately 3,000 controls were genotyped. Case-control analyses identified seven significant, independent associations with type 1 diabetes. Six genes/regions had strong preexisting statistical support for a role in type 1 diabetes susceptibility: MHC, *INS*, *CTLA4*, *PTPN22*, *IL2RA/CD25*, and *IFIH1*. Regions on chromosomes 12q13, 12q24, and 16p13 exhibited strong, yet not statistically significant, evidence of association that warrant additional investigation. The WTCCC study demonstrated the power of the GWAS approach to identify novel loci associated with type 1 diabetes in an unbiased manner and suggested that larger samples with greater density of SNP genotyping could uncover additional novel genes and pathways related to type 1 diabetes risk. The 12q13 and 12q24 loci map to regions containing several functional candidates, including *ERBB3* (12q13) and *SH2B3* (12q24). The 16p13 region contained only two genes of unknown function, one of which (*KIAA0350*) was identified in a second GWAS (25) and as a novel locus of susceptibility for other autoimmune diseases, consistent with pervasive sharing of genetic risk across immune-mediated disorders (26).

The T1DGC expanded the findings from the WTCCC by initiating a new GWAS of 3,983 type 1 diabetes cases and

FIGURE 12.6. Progress Mapping Type 1 Diabetes Susceptibility Genes Based on Era of Genomic Search

SOURCE: Reference 26

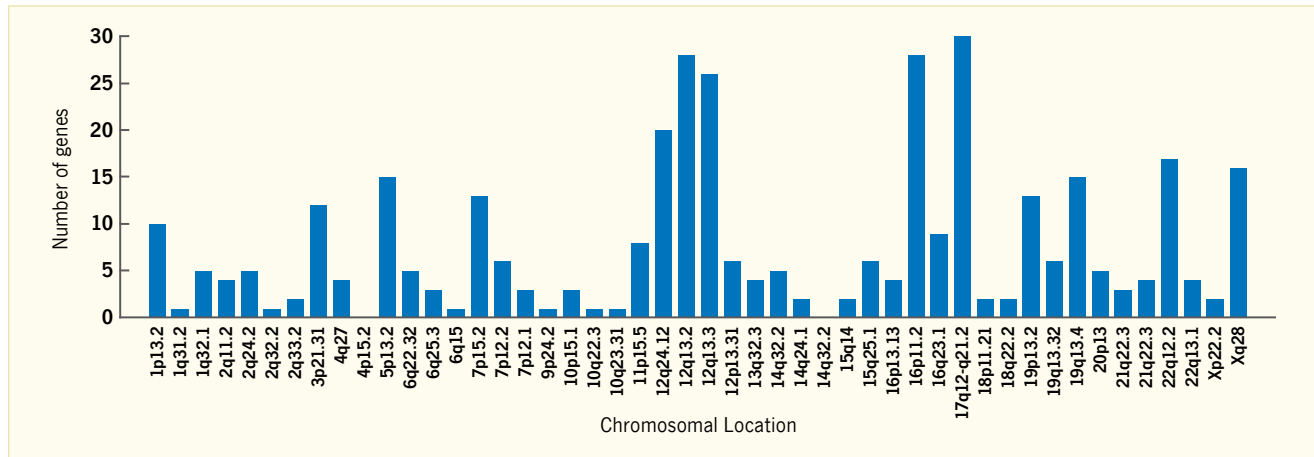
3,999 controls, followed by a combined meta-analysis with previously published studies that added a further 3,715 cases and 5,069 controls, with information from more than 841,622 SNPs. From the meta-analysis of 7,514 cases and 9,045 controls with confirmation in 4,267 cases, 4,670 controls, and 2,319 ASP families, 28 novel regions provided initial significant evidence for association to type 1 diabetes risk (26). Eighteen novel regions were strongly replicated from the GWAS meta-analysis.

In total, the number of loci contributing significantly to risk of type 1 diabetes risk had now reached 42, although each locus varied in size of the interval, the number of candidate genes in the locus, the size of the effect on type 1 diabetes risk for the most associated SNP, and the strength of the evidence supporting the most likely candidate (Figure 12.6). This figure demonstrates several points. The first is that the genes (loci) of largest impact on type 1 diabetes risk were identified in the 30-year period from early 1970s (HLA) to 2000 (*INS*),

reflecting the larger effects identified only with smaller sample sizes. These two regions account for over 50% of the genetic variation in type 1 diabetes risk and were identified as the result of small case-control and family studies. In 2001–2006, with the advent of large-scale genotyping arrays and assembly of larger case-control series, several additional loci were identified (including *PTPN22*, *CTLA4*, and *IFIH1*) that slightly increased the amount of genetic risk identified. In 2007–2008, the first well-powered GWAS by the WTCCC greatly increased the number of loci but was of modest (e.g., *IL2RA*) to small contribution to risk. In 2009, the T1DGC GWAS meta-analysis (26) nearly doubled the number of loci, (with 8,000 cases and 9,000 controls plus replication) with each having small effect. Of interest, the meta-analysis confirmation in the T1DGC ASP families was consistent in direction of the risk allele effect, but the size of the effect was consistently less in the families than in the case-control comparison, perhaps indicating the greater role of the MHC in type 1 diabetes risk when multiple family

members express disease; statistical evidence for replication in families was not reached for 11 of the 18 loci.

The T1DGC further examined the 18 novel loci by expansion of the family sample to account for potential population stratification and replication genotyping (27). The analyses of these 2,322 additional families, combined with the original 2,319 families, provided improved protection from population stratification bias and increased power to provide further replication support for the associations of the 18 novel susceptibility loci. With this larger sample of families, only one (*C14orf181/14q24.1*) of the 18 novel loci failed to reach statistical significance. Further, all of the novel type 1 diabetes risk loci had consistent direction of effects (e.g., risk or protective for type 1 diabetes risk) with the original GWAS meta-analysis (26,27) with no evidence of heterogeneity in the disease associations across family collections, despite there being significant SNP genotype frequency differences. After unequivocal replication of type 1

FIGURE 12.7. Gene Content of Type 1 Diabetes Susceptibility Loci Identified by Genome-Wide Association Study Meta-Analysis

SOURCE: References 26 and 28

diabetes risk loci, the next steps involve dense SNP mapping to better define causal genes and variants. Identification of causal type 1 diabetes genes with both positional and functional evidence for a role in type 1 diabetes pathogenesis could reveal the pathways and early precursors or biomarkers underlying type 1 diabetes.

FINE MAPPING WITHIN TYPE 1 DIABETES SUSCEPTIBILITY LOCI

Most common variants identified through a GWAS are located between genes (intergenic), in regions that contain regulatory DNA, or (if located in a gene) not in a protein-coding region (intronic). Relatively few robustly type 1 diabetes-associated variants with likely deleterious effects have been located in coding regions (exons). Thus, the variants that are associated with type 1 diabetes from the GWAS typically do not have obvious functional consequences. This result suggests the tentative conclusion that many of the disease-associated variants affect the regulation of gene expression, although the possibility that these variants are correlated with one or more causative variants within a coding region cannot typically be excluded based upon the initial GWAS results. GWAS actually provide limited information on the likely causal gene given the sparse coverage of a given region of the genome even with high-density SNP genotyping arrays. Thus, prior to detailed examination of genes and regulation,

dense SNP mapping of the type 1 diabetes-associated locus is required to firmly establish the most likely associated SNPs and causal genes.

The T1DGC GWAS meta-analysis (26), when combined with previous studies, identified more than 40 loci that were significantly associated with type 1 diabetes risk; however, within each type 1 diabetes locus, there was an average of seven genes with a range of 0–27 (Figure 12.7) (26,28). Some regions harbored potential candidate genes (such as those involved in the immune response), but in other cases, only a single gene without obvious functional relevance to type 1 diabetes or autoimmunity was identified. Indeed, for a few regions, no coding genes were present (Figure 12.7). The novel loci identified in the GWAS meta-analysis typically had modest-small effects on type 1 diabetes risk, with *C10orf159* (*RNLS*, renalase) having the largest effect (OR ~1.3) and others with decreasing effect size, consistent with the dependence of their detection on large sample sizes. A large number of previously detected loci had smaller effects, including those with strong or confirmed candidate causal genes (e.g., *CTLA4*, *IFIH1*).

In most GWAS regions, the locus associated with disease is typically defined by few (perhaps 100) SNPs. The size of the average type 1 diabetes susceptibility locus from the meta-analysis

(26) was defined by an *ad hoc* algorithm based upon the most associated SNP at the center with boundaries for that region defined by site of high recombination. With multiple candidate genes in each type 1 diabetes region, functional studies are necessary to establish a role for a particular gene and/or allelic variant in type 1 diabetes pathogenesis. Such studies require significant time and resources and, in the absence of unambiguous identification of at least the disease-relevant gene within a region, could well target the incorrect gene or variant, elucidating regulatory mechanisms not relevant to type 1 diabetes. In order to refine GWAS regions of susceptibility for type 1 diabetes and address this concern, dense genotyping (fine-mapping) was required that could comprehensively assess the impact of thousands of SNPs in a region.

A GWAS seeks to capture genomic architecture with high information content using a standard set of SNPs. In contrast, fine-mapping requires saturation of specific regions of the genome, with the predetermined boundaries of each region defining the set of SNPs to be genotyped. In order to meet the needs of type 1 diabetes and other autoimmune diseases with genome-wide significant regions to be interrogated, the ImmunoChip consortium was established. The consortium consisted of investigators who provided data (including unpublished) to design a genotyping array (the ImmunoChip) for

fine-mapping autoimmune disease susceptibility genes. The ImmunoChip is a custom Illumina Infinium high-density genotyping array consisting of 196,524 SNPs over 186 regions of the human genome.

For type 1 diabetes, the initial analyses of ImmunoChip data included 6,670 type 1 diabetes cases, 9,416 controls, 2,601 T1DGC ASP families, and 69 T1DGC trio (affected child and both parents) families (29). A large proportion of these samples (98% of cases, 76% of controls, and 57% of family samples) were used in the previous GWAS meta-analysis that initially identified many of the type 1 diabetes regions (26). Including those loci that

were significantly associated with another autoimmune disease (but only nominally with type 1 diabetes), 55 non-MHC loci were evaluated. Of the non-MHC type 1 diabetes-associated regions, 36 were significantly associated with type 1 diabetes risk from the ImmunoChip data (including 33 of 45 that had been reported by the T1DGC or others previously); in addition, there were three loci that were nominally associated with type 1 diabetes that were also robustly associated with other autoimmune diseases (30) and were, therefore, included as type 1 diabetes loci (Table 12.1). Some evidence of association remained for 13 of the 19 regions that failed to reach

statistical significance, and some of these 19 loci may contain valid type 1 diabetes susceptibility genes. Three new type 1 diabetes loci were identified: 1q32.1, also associated with ulcerative colitis (31,32,33) and Crohn's disease (34); 4q32.3, also implicated in celiac disease (35); and 5p13.2, also associated with multiple sclerosis (36). In general, the number of coding variants and noncoding variants that potentially contribute to type 1 diabetes risk varies by gene region (Figure 12.7). These data suggest that much work remains to be accomplished to identify the causal genes, the variants within the genes that modify risk, and the functional impact of those variants.

TABLE 12.1. Type 1 Diabetes-Associated Regions From ImmunoChip Meta-Analysis

CHROMOSOME	POSITION	SNP	ALLELES	MAF	OR	P	CANDIDATE GENE(S)	OTHER DISEASES†
1p13.2	114,377,568	rs2476601	G>A	0.09	1.89	<10 ⁻¹⁰⁰	<i>PTPN22</i>	ATD, CRO, RA, SLE, VIT
1q32.1*	200,814,959	rs6691977	T>C	0.19	1.13	4.3 × 10 ⁻⁸		
1q32.1	206,939,904	rs3024505	G>A	0.16	0.86	6.4 × 10 ⁻⁸	<i>IL10</i>	CRO, IBD, SLE, UC
2q11.2	100,764,087	rs13415583	T>G	0.35	0.90	1.1 × 10 ⁻⁷	<i>AFF3</i>	RA
2q13*	111,615,079	rs4849135	G>T	0.29	0.89	4.4 × 10 ⁻⁸		
2q24.2	163,110,536	rs2111485	G>A	0.39	0.85	3.8 × 10 ⁻¹⁸	<i>IFIH1</i>	
2q24.2	163,124,637	rs35667974	T>C	0.02	0.59	9.3 × 10 ⁻⁹	<i>IFIH1</i>	PSO
2q24.2	163,136,942	rs72871627	A>G	0.01	0.61	2.4 × 10 ⁻⁶	<i>IFIH1</i>	
2q33.2	204,738,919	rs3087243	G>A	0.45	0.84	7.4 × 10 ⁻²¹	<i>CTLA4</i>	ATD, CEL, RA
3p21.31	46,457,412	rs113010081	T>C	0.11	0.85	4.6 × 10 ⁻⁸	<i>CCR5</i>	CEL
4q27	123,243,596	rs75793288	C>G	0.36	1.15	5.6 × 10 ⁻¹³	<i>IL2, IL21</i>	CEL
4q32.3*	166,574,267	rs2611215	G>A	0.15	1.18	1.8 × 10 ⁻¹¹		
5p13.2*	35,883,251	rs11954020	C>G	0.39	1.11	4.4 × 10 ⁻⁸	<i>IL7R</i>	
6q15	90,976,768	rs72928038	G>A	0.17	1.20	6.4 × 10 ⁻¹⁴	<i>BACH2</i>	
6q22.32	126,752,884	rs1538171	C>G	0.45	1.12	7.4 × 10 ⁻¹⁰		
7p12.2	50,465,830	rs62447205	A>G	0.28	0.89	2.5 × 10 ⁻⁸	<i>IKZF1</i>	
7p12.1	51,028,987	rs10277986	A>T	0.04	0.76	1.4 × 10 ⁻⁷		
9p24.2	4,290,823	rs6476839	A>T	0.40	1.12	1.0 × 10 ⁻⁹	<i>GLIS3</i>	
10p15.1	6,094,697	rs61839660	C>T	0.10	0.62	2.8 × 10 ⁻³⁹	<i>IL2RA</i>	VIT
10p15.1	6,108,340	rs10795791	A>G	0.41	1.16	5.6 × 10 ⁻¹¹	<i>IL2RA</i>	
10p15.1	6,129,643	rs41295121	C>T	0.01	0.49	4.9 × 10 ⁻⁸	<i>IL2RA</i>	
10q23.31	90,035,654	rs12416116	C>A	0.28	0.85	3.9 × 10 ⁻¹⁵		
11p15.5	2,182,224	rs689	T>A	0.30	0.42	<10 ⁻¹⁰⁰	<i>INS</i>	
11p15.5	2,198,665	rs72853903	C>T	0.38	0.85	6.2 × 10 ⁻¹⁰	<i>INS</i>	
12p13.31	9,905,851	rs917911	A>C	0.36	1.10	1.9 × 10 ⁻⁷	<i>CD69</i>	MS
12q13.2	56,435,504	rs705705	G>C	0.34	1.25	4.4 × 10 ⁻³²	<i>IKZF4</i>	
12q24.12	112,007,756	rs653178	T>C	0.48	1.30	1.6 × 10 ⁻⁴⁴	<i>SH2B3</i>	CEL, JIA, PBC, PSC, RA, VIT
13q32.3	100,081,766	rs9585056	T>C	0.24	1.12	3.3 × 10 ⁻⁸	<i>GPR183</i>	
14q32.2	98,488,007	rs1456988	T>G	0.27	1.12	2.9 × 10 ⁻⁸		

Table 12.1 continues on the next page.

TABLE 12.1. (continued)

CHROMOSOME	POSITION	SNP	ALLELES	MAF	OR	P	CANDIDATE GENE(S)	OTHER DISEASES†
14q32.2	101,306,447	rs56994090	T>C	0.41	0.88	1.1×10^{-11}		
15q14	38,847,022	rs72727394	C>T	0.19	1.15	3.6×10^{-10}	RASGRP1	
15q25.1	79,234,957	rs34593439	G>A	0.10	0.78	9.0×10^{-14}	CTSH	
16p11.2	28,505,660	rs151234	G>C	0.12	1.19	4.8×10^{-11}	IL27	
16p13.13	11,194,771	rs12927355	C>T	0.32	0.82	3.0×10^{-22}	DEXI	MS, PBC
16p13.13	11,351,211	rs193778	A>G	0.25	1.14	4.4×10^{-10}	DEXI	CRO
16q23.1	75,252,327	rs8056814	G>A	0.07	1.32	3.0×10^{-19}	BCAR1	
17q12	38,053,207	rs12453507	G>C	0.49	0.90	1.0×10^{-8}	IKZF3, ORMDL3, GSDMB	RA
17q21.2	38,775,150	rs757411	T>C	0.36	0.90	1.1×10^{-7}	CCR7	
17q21.31*	44,073,889	rs1052553	A>G	0.24	0.89	8.2×10^{-8}		
18p11.21	12,809,340	rs1893217	A>G	0.16	1.21	1.2×10^{-15}	PTPN2	CEL, CRO, IBD, UC
18p11.21	12,830,538	rs12971201	G>A	0.39	0.89	2.1×10^{-6}	PTPN2	
18q22.2	67,526,644	rs1615504	C>T	0.47	1.13	1.8×10^{-11}	CD226	CEL, MS
19p13.2	10,463,118	rs34536443	G>C	0.04	0.67	4.4×10^{-15}	TYK2	MS, PBC, PSO, RA
19p13.2	10,469,975	rs12720356	A>C	0.09	0.82	3.7×10^{-7}	TYK2	
19q13.32	47,219,122	rs402072	T>C	0.16	0.87	4.7×10^{-8}		
19q13.33	49,206,172	rs516246	T>C	0.49	0.87	5.2×10^{-14}	FUT2	
20p13	1,616,206	rs6043409	G>A	0.35	0.88	3.0×10^{-10}		
21q22.3	43,825,357	rs11203202	C>G	0.33	1.16	1.2×10^{-15}	UBASH3A	RA, VIT
21q22.3*	45,621,817	rs6518350	A>G	0.18	0.88	9.6×10^{-8}	ICOSLG	
22q12.2	30,531,091	rs4820830	T>C	0.38	1.14	1.2×10^{-12}		
22q12.3	37,587,111	rs229533	A>C	0.43	1.11	1.8×10^{-8}	C1QTNF6, RAC2	

The most strongly associated SNP in a region is shown, together with the effect of the minor allele relative to the major allele. Alleles are shown as major allele>minor allele. rs689 (11p15.5, *INS*) data were obtained from previous TaqMan genotyping. Named candidate genes are genes for which there was additional evidence that they might be causal or which encode proteins with known immune functions that are part of the immune pathways already identified as being involved in type 1 diabetes pathogenesis. Because SNPs may alter enhancer sequences distant from the target gene, a gene (or a noncoding RNA) is not named if the only evidence for a causal role was that the peak of SNP association lies in or very near a gene (unless the SNPs alter coding sequence or splicing signals in a potentially functional way). For example, *RNLS* at 10q23.31 has no established role in the immune system, and there is currently no specific functional data linking this gene to type 1 diabetes etiology. *ATD*, autoimmune thyroid disease; *CEL*, celiac disease; *CRO*, Crohn's disease; *IBD*, inflammatory bowel disease; *JIA*, juvenile idiopathic arthritis; *MAF*, minor allele frequency; *MS*, multiple sclerosis; *OR*, odds ratio; *PBC*, primary biliary cirrhosis; *PSC*, primary sclerosing cholangitis; *PSO*, psoriasis; *RA*, rheumatoid arthritis; *SLE*, systemic lupus erythematosus; *SNP*, single nucleotide polymorphism; *UC*, ulcerative colitis; *VIT*, vitiligo. * Loci newly identified by the ImmunoChip analysis (at $P < 3.23 \times 10^{-7}$)

† Evidence that SNPs in the same locus, but not necessarily the same SNP, are associated with other autoimmune disease.

SOURCE: Adapted from Reference 29, copyright © 2015 Nature Publishing Group, reprinted with permission

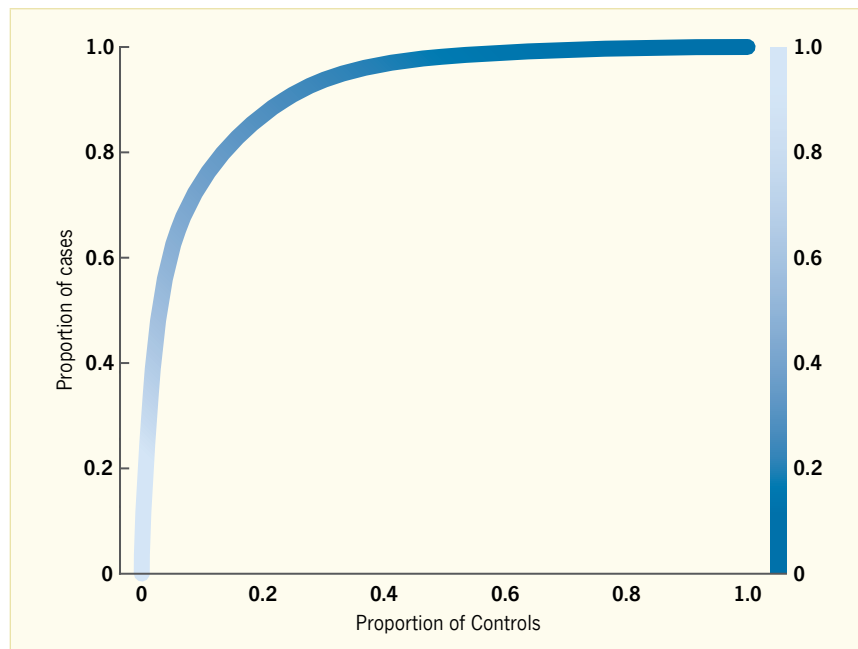
ACCOUNTING FOR THE GENETIC RISK OF TYPE 1 DIABETES

The genetic prediction of type 1 diabetes risk, unlike other autoimmune diseases, is relatively effective due, in part, to modest risk to relatives (8% sibling risk) and low disease prevalence (4/1,000). Following fine-mapping with the ImmunoChip, 27 MHC SNPs could account for the majority of independent additive effects in the region (tagging the classical HLA alleles) and 46 non-MHC SNPs could account for the remaining independent associations. These SNPs were used to determine the area under the Receiver Operator Characteristic (ROC) curve for prediction of type 1 diabetes (W.-M. Chen, personal communication). The area under the

ROC curve for MHC SNPs is 0.741, which under a polygenic model, corresponds to a sibling recurrence risk ratio (λ_s) of 1.52, consistent with the estimate from the previous GWAS meta-analysis. When the 46 non-MHC SNPs are combined with 27 SNPs in the MHC, the area under the ROC curve is increased to 0.918 (Figure 12.8). Using this genetic risk score in an independent set of data, the sensitivity of these 27 MHC SNPs and 46 non-MHC SNPs was 86% with a specificity of 81%. These results are consistent with those initially determined from the GWAS meta-analysis, but with greater sensitivity and specificity from novel loci detected and identification

of SNPs more associated with type 1 diabetes in each locus (37).

These estimates of sensitivity and specificity of the type 1 diabetes-associated SNPs are high, but results should be considered with caution, as they are based on Caucasian samples with well-defined case-control status and not in a clinical setting. Further, the predictive values (positive and negative) will be low, using SNPs only, as the prevalence of type 1 diabetes is low (~4/1,000). Thus, the use of genetic data to predict risk of type 1 diabetes will be useful only in the context of other (e.g., autoantibody) predictive markers.

FIGURE 12.8. Receiver Operator Characteristics in Type 1 Diabetes Case-Control Validation Study

The receiver operator characteristics (ROC) curve represents the proportion of cases and controls in a validation data set with the risk scores greater than a threshold. When the average of the median risk score from the training set (2,330 cases, 2,330 controls) is used as the threshold in the validation set (1,190 cases, 3,470 controls), the sensitivity is 86% and specificity is 81%. Risk score is based on 27 SNPs in the MHC and 46 non-MHC SNPs. MHC, major histocompatibility complex; SNP, single nucleotide polymorphism.

SOURCE: Original figure created by W.-M. Chen.

FROM GENE TO FUNCTION

The proximity of a disease-associated SNP to a particular gene is often thought of as providing evidence that the gene plays a role in disease pathogenesis. While this is one possible hypothesis, it is equally likely that a variant could affect the expression of one or more genes at a distance, even on other chromosomes. For many of the non-MHC type 1 diabetes-associated loci identified and fine-mapped in GWAS studies, the most significant type 1 diabetes-associated SNP is located within the transcription unit of a gene (29); however, in only four cases is the most associated SNP a coding variant assumed to alter the amino acid sequence of a particular gene product (*rs2476601*, *PTPN22*; *rs35667974*, *IFIH1*; *rs34536443*, *TYK2*; *rs6043409*, *SIRPG*). Findings such as these encourage functional studies of

the allelic forms of the specific proteins in the four genes with coding variants to determine whether they impact pathways that could plausibly contribute to type 1 diabetes pathogenesis.

As shown by the T1DGC ImmunoChip analyses (29), the majority of type 1 diabetes-associated SNPs are in DNA regulatory regions, specifically those regions that have clusters of transcription factors (enhancers) that have tissue-specific function. In type 1 diabetes, these cell types are those most likely relevant to autoimmunity—CD4+ T cells, CD8+ T cells, B cells, and CD34+ stem cells (29). Variants that affect transcription factors and enhancers require consideration of other possible mechanisms of action. These mechanisms include effects on

the coding region of a previously unrecognized gene, the transcription of one or several nearby genes, or effects on genes that are distant from the most associated SNP, most recently shown as a fundamental mechanism in obesity (38,39). Evidence of gene-gene interaction would be suggestive of long-range regulatory effects. One approach (among many) to examining the putative functional effects of disease-associated genetic variants is to consider whether they alter the primary sequence of known regulatory motifs. As more information is obtained from functional research at the genomic level, the interpretation of the genetic variants discovered as important in type 1 diabetes will, perhaps, identify novel mechanisms of risk that can be used for prevention and therapeutic targets.

GENES AND INITIATION OF TYPE 1 DIABETES

While most studies of the genetic basis of type 1 diabetes focus on the presence or absence of type 1 diabetes

in independent subjects or family members, the process by which type 1 diabetes develops in a genetically

susceptible individual is also amenable to genetic study. Diagnosis of type 1 diabetes follows a preclinical phase

of undetermined length during which autoimmunity reduces pancreatic islet beta cell mass. Biomarkers of this process, such as metabolic measures of blood glucose or insulin release, and immunologic measures, such as the timing of appearance, diversity, or titer of autoantibodies directed at islet antigens in a given patient, offer opportunities to explore the genetic contribution to disease progression. The analysis of initiation of type 1 diabetes and the discovery of genes and variants contributing to variation in initiation have been limited due to the limited number of at-risk participants being followed for evidence of change in biomarkers and transition to pre-type 1 diabetes states (defined by autoantibodies) and clinical type 1 diabetes. Past studies have involved type 1 diabetes cases with longer disease duration in whom autoantibody status would not reflect true prevalence prior to type 1 diabetes diagnosis or at disease initiation. The Environmental Determinants of Diabetes in the Young (TEDDY) study (40) suggests that close follow-up of participants at risk of type 1 diabetes with biomarkers and genetic data should provide insights in initiation and progression of type 1 diabetes.

The first GWAS of autoantibody positivity in type 1 diabetes was conducted in T1DGC cases with measurement of two

anti-islet autoantibodies (glutamate decarboxylase [GAD] and insulinoma-associated antigen 2 [IA-2A]), antibodies to the autoimmune thyroid (Graves' disease autoantigen thyroid peroxidase, and antibodies against gastric parietal cells (PCA) that are associated with autoimmune gastritis (41). Genome-wide, variants in *FCRL3*/1q23 were associated with IA-2A positivity, and variants in *ABO*/9q34 were associated with presence of PCA (41). These studies were limited, however, as the T1DGC samples were obtained from prevalent cases, and a more appropriate design would be to assess the development of islet autoimmunity over time in the context of genetic factors (as in prospective cohort studies, such as TEDDY, the Diabetes and Autoimmunity Study in the Young [DAISY], BABYDIAB, and the Diabetes Prediction and Prevention study [DIPP]). Critically, genome-wide (rather than candidate gene) approaches will be necessary to interrogate and discover those variants contributing to initiation, as well as progression, of the autoimmune process.

Age at onset of type 1 diabetes has often been used as a simple measure of disease initiation in type 1 diabetes. There have been conflicting reports of possible effects of type 1 diabetes risk variants on age at onset, largely from

studies of modest sample size. However, many reports have recognized the role of HLA alleles, particularly HLA-DR and HLA-DQ, on age at onset (42), with those having onset at age <6 years more likely to have the HLA-DR3 allele and the HLA-DR3/DR4 genotype. The association of type 1 diabetes-associated SNPs with age at onset was reexamined utilizing information on risk variants identified from GWAS and fine-mapping studies, as well as a sufficiently large sample size to provide adequate statistical power to test for interaction effects (43). Two loci, *RNLS* at 10q23.31 and *IL2* at 4q27, displayed significant associations with age at onset. For *IL2*/4q27, the G allele for the rs2069763 variant is "protective," so the age at diagnosis increases with the number of protective alleles (mean age at diagnosis for those with the TT genotype is 8.7 years, the GT genotype 9.0 years, and the GG genotype 9.3 years). This effect is seen similarly for the *RNLS*/10q23.3 rs10500540 variant, in that the mean age at diagnosis increases with the number of protective C alleles (mean age at diagnosis for those with the TT genotype is 8.9 years, the TC genotype 9.3 years, and the CC genotype 9.4 years).

CONCLUSIONS

Type 1 diabetes is a complex disorder that results from the actions, and possibly the interactions, of multiple genetic and environmental factors. The advent of high-throughput genotyping methodologies enabling GWAS combined with large collections of biospecimens from type 1 diabetes cases and unaffected control individuals has facilitated the identification of substantial numbers of genetic loci that contribute to type 1 diabetes risk. Fine-mapping studies building upon GWAS results have narrowed the number of possible causative genes to one or a few in most of these regions. Some genes involved in type 1 diabetes risk have been identified and have consistent

and strong effects (e.g., HLA class I and class II loci or *INS*). Other genes have been shown to contribute to risk, but their individual contributions to risk are small. Nevertheless, they may represent components in common biologic pathways where there may be rate-limiting reactions or products amenable to manipulation for therapy or, at least, facilitate better predictive models for type 1 diabetes.

Future studies of the genetics of type 1 diabetes need to focus on the relationship of genotype and phenotype, elucidating the immediate effects of individual risk variants on the relevant genes, the cellular effects of modulating the function or

expression of these genes, and, most importantly, the organismic effects of such changes at the level of the patients. These studies, building upon unbiased searches for risk genes and variants hold the promise of providing both novel insights into the pathogenesis of type 1 diabetes and targets for potential preventive therapies.

LIST OF ABBREVIATIONS

APC	antigen-presenting cell
ASP	affected sibpair
DNA	deoxyribonucleic acid
GAD	glutamate decarboxylase
GWAS	genome-wide association study
HLA	human leukocyte antigen
IA-2A	insulinoma-associated antigen 2
IBD	identity-by-descent
LD	linkage disequilibrium
MHC	major histocompatibility complex
OR	odds ratio
PCA	parietal cell antibodies
ROC	receiver operator characteristics
SNP	single nucleotide polymorphism
T1DGC	Type 1 Diabetes Genetics Consortium
TEDDY	The Environmental Determinants of Diabetes in the Young study
WTCCC	Wellcome Trust Case-Control Consortium

DUALITY OF INTEREST

Drs. Rich, Erlich, and Concannon reported no conflicts of interest.

ACKNOWLEDGMENTS/ FUNDING

This work utilized resources provided by the Type 1 Diabetes Genetics Consortium, a collaborative clinical study sponsored by the National Institute of Diabetes and Digestive and Kidney Diseases, the National Institute of Allergy and Infectious Diseases, the National Human Genome Research Institute, the Eunice Kennedy Shriver National Institute of Child Health and Human Development, and the JDRF. Further support was provided to Dr. Rich by grants from the National Institute of Diabetes and Digestive and Kidney Diseases (DK062418 and DK096926) and the JDRF (9-2011-530 and 17-2012-542), and to Dr. Concannon by grants from the National Institute of Diabetes and Digestive and Kidney Diseases (DK046635 and DK085678).

REFERENCES

- Barnett AH, Eff C, Leslie RD, Pyke DA: Diabetes in identical twins. A study of 200 pairs. *Diabetologia* 20:87–93, 1981
- Redondo MJ, Jeffrey J, Fain PR, Eisenbarth GS, Orban T: Concordance for islet autoimmunity among monozygotic twins. *New Engl J Med* 359:2849–2850, 2008
- Johnston C, Pyke DA, Cudworth AG, Wolf E: HLA-DR typing in identical twins with insulin-dependent diabetes: differences between concordant and discordant pairs. *Br Med J* 286:253–255, 1983
- Metcalfe KA, Hitman GA, Rowe RE, Hawa M, Huang X, Stewart T, Leslie RD: Concordance for type 1 diabetes in identical twins is affected by insulin genotype. *Diabetes Care* 24:838–842, 2001
- Rewers M: The changing face of the epidemiology of insulin-dependent diabetes mellitus (IDDM): research designs and models of disease causation. *Ann Med* 23:419–426, 1991
- Rewers M, LaPorte RE, King H, Tuomilehto J: Trends in the prevalence and incidence of diabetes: insulin-dependent diabetes mellitus in childhood. *World Health Stat Q* 41:179–189, 1988
- Rich SS: Mapping genes in diabetes. A genetic epidemiological perspective. *Diabetes* 39:1315–1319, 1990
- Singal DP, Blajchman MA: Histocompatibility (HLA) antigens, lymphocytotoxic antibodies and tissue antibodies in patients with diabetes mellitus. *Diabetes* 22:429–432, 1973
- Rich SS, Weitkamp LR, Barbosa J: Genetic heterogeneity of insulin-dependent (type 1) diabetes mellitus: evidence from a study of extended haplotypes. *Am J Hum Genet* 36:1015–1023, 1984
- Steenkiste A, Valdes AM, Feolo M, Hoffman D, Concannon P, Noble J, Schoch G, Hansen J, Helmsberg W, Dorman JS, Thomson G, Pugliese A; 13th IHWS 1 Diabetes Component participating investigators: 14th International HLA and Immunogenetics Workshop: report on the HLA component of type 1 diabetes. *Tissue Antigens* 69(Suppl 1):214–225, 2007
- Mallal S, Nolan D, Witt C, Masel G, Martin AM, Moore C, Sayer D, Castley A, Mamotte C, Maxwell D, James I, Christiansen FT: Association between presence of HLA-B*5701, HLA-DR7, and HLA-DQ3 and hypersensitivity to HIV-1 reverse-transcriptase inhibitor abacavir. *Lancet* 359:727–732, 2002
- Hetherington S, Hughes AR, Mosteller M, Shortino D, Baker KL, Spreen W, Lai E, Davies K, Handley A, Dow DJ, Fling ME, Stocum M, Bowman C, Thurmond LM, Roses AD: Genetic variations in HLA-B region and hypersensitivity reactions to abacavir. *Lancet* 359:1121–1122, 2002
- Noble JA, Valdes AM: Genetics of the HLA region in the prediction of type 1 diabetes. *Curr Diab Rep* 11:533–542, 2011
- Cucca F, Muntoni F, Lampis R, Frau F, Argiolas L, Silvetti M, Angius E, Cao A, De Virgiliis S, Congia M: Combinations of specific DRB1, DQA1, DQB1 haplotypes are associated with insulin-dependent diabetes mellitus in Sardinia. *Hum Immunol* 37:85–94, 1993
- Erlich HA, Valdes AM, McDevitt SL, Simen BB, Blake LA, McGowan KR, Todd JA, Rich SS, Noble JA; Type 1 Diabetes Genetics Consortium (T1DGC): Next

- generation sequencing reveals the association of DRB3*02:02 with type 1 diabetes. *Diabetes* 62:2618–2622, 2013
16. Hu X, Deutsch AJ, Lenz TL, Onengut-Gumuscu S, Han B, Chen WM, Howson JM, Todd JA, de Bakker PI, Rich SS, Raychaudhuri S: Additive and interaction effects at three amino acid positions in HLA-DQ and HLA-DR molecules drive type 1 diabetes risk. *Nat Genet* 47:898–905, 2015
 17. Lenz TL, Deutsch AJ, Han B, Hu X, Okada Y, Eyre S, Knapp M, Zernakova A, Huizinga TW, Abecasis G, Becker J, Boeckstaens GE, Chen WM, Franke A, Gladman DD, Gockel I, Gutierrez-Achury J, Martin J, Nair RP, Nothen MM, Onengut-Gumuscu S, Rahman P, Rantapaa-Dahlqvist S, Stuart PE, Tsoi LC, van Heel DA, Worthington J, Wouters MM, Klareskog L, Elder JT, Gregersen PK, Schumacher J, Rich SS, Wijmenga C, Sunyaev SR, de Bakker PI, Raychaudhuri S: Widespread non-additive and interaction effects within HLA loci modulate the risk of autoimmune diseases. *Nat Genet* 47:1085–1090, 2015
 18. Concannon P, Chen WM, Julier C, Morahan G, Akolkar B, Erlich HA, Hilner JE, Nerup J, Nierras C, Pociot F, Todd JA, Rich SS; Type 1 Diabetes Genetics Consortium: Genome-wide scan for linkage to type 1 diabetes in 2,496 multiplex families from the Type 1 Diabetes Genetics Consortium. *Diabetes* 58:1018–1022, 2009
 19. Morahan G, Mehta M, James I, Chen WM, Akolkar B, Erlich HA, Hilner JE, Julier C, Nerup J, Nierras C, Pociot F, Todd JA, Rich SS; Type 1 Diabetes Genetics Consortium: Tests for genetic interactions in type 1 diabetes: linkage and stratification analyses of 4,422 affected sib-pairs. *Diabetes* 60:1030–1040, 2011
 20. Rich SS, Akolkar B, Concannon P, Erlich H, Hilner JE, Julier C, Morahan G, Nerup J, Nierras C, Pociot F, Todd JA: Overview of the Type 1 Diabetes Genetics Consortium. *Genes Immun* 10(Suppl 1):S1–S4, 2009
 21. International HapMap Consortium: The International HapMap Project. *Nature* 426:789–796, 2003
 22. Smyth DJ, Cooper JD, Bailey R, Field S, Burren O, Smink LJ, Guja C, Ionescu-Tirgoviste C, Widmer B, Dunger DB, Savage DA, Walker NM, Clayton DG, Todd JA: A genome-wide association study of nonsynonymous SNPs identifies a type 1 diabetes locus in the interferon-induced helicase (IFIH1) region. *Nat Genet* 38:617–619, 2006
 23. Nejentsev S, Walker N, Riches D, Egholm M, Todd JA: Rare variants of IFIH1, a gene implicated in antiviral responses, protect against type 1 diabetes. *Science* 324:387–389, 2009
 24. The Wellcome Trust Case Control Consortium: Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. *Nature* 447:661–678, 2007
 25. Cotsapas C, Voight BF, Rossin E, Lage K, Neale BM, Wallace C, Abecasis GR, Barrett JC, Behrens T, Cho J, De Jager PL, Elder JT, Graham RR, Gregersen P, Klareskog L, Siminovitch KA, van Heel DA, Wijmenga C, Worthington J, Todd JA, Hafler DA, Rich SS, Daly MJ; FOCIS Network of Consortia: Pervasive sharing of genetic effects in autoimmune disease. *PLoS Genet* 7:e1002254, 2011
 26. Barrett JC, Clayton DG, Concannon P, Akolkar B, Cooper JD, Erlich HA, Julier C, Morahan G, Nerup J, Nierras C, Plagnol V, Pociot F, Schuilenburg H, Smyth DJ, Stevens H, Todd JA, Walker NM, Rich SS; Type 1 Diabetes Genetics Consortium: Genome-wide association study and meta-analysis find that over 40 loci affect risk of type 1 diabetes. *Nat Genet* 41:703–707, 2009
 27. Cooper JD, Howson JM, Smyth D, Walker NM, Stevens H, Yang JH, She JX, Eisenbarth GS, Rewers M, Todd JA, Akolkar B, Concannon P, Erlich HA, Julier C, Morahan G, Nerup J, Nierras C, Pociot F, Rich SS; Type 1 Diabetes Genetics Consortium: Confirmation of novel type 1 diabetes risk loci in families. *Diabetologia* 55:996–1000, 2012
 28. Concannon P, Rich SS, Nepom GT: Genetics of type 1A diabetes. *New Engl J Med* 360:1646–1654, 2009
 29. Onengut-Gumuscu S, Chen WM, Burren O, Cooper NJ, Quinlan AR, Mychaleckyj JC, Farber E, Bonnie JK, Szpak M, Schofield E, Achuthan P, Guo H, Fortune MD, Stevens H, Walker NM, Ward LD, Kundaje A, Kellis M, Daly MJ, Barrett JC, Cooper JD, Deloukas P; Type 1 Diabetes Genetics Consortium, Todd JA, Wallace C, Concannon P, Rich SS: Fine mapping of type 1 diabetes susceptibility loci and evidence for colocalization of causal variants with lymphoid gene enhancers. *Nat Genet* 47:381–386, 2015
 30. Trynka G, Hunt KA, Bockett NA, Romanos J, Mistry V, Szperl A, Bakker SF, Bardella MT, Bhaw-Rosun L, Castillejo G, de la Concha EG, de Almeida RC, Dias KR, van Diemen CC, Dubois PC, Duerr RH, Edkins S, Franke L, Fransen K, Gutierrez J, Heap GA, Hrdlickova B, Hunt S, Plaza Izurieta L, Izzo V, Joosten LA, Langford C, Mazzilli MC, Mein CA, Midah V, Mitrovic M, Mora B, Morelli M, Nutland S, Nunez C, Onengut-Gumuscu S, Pearce K, Platteel M, Polanco I, Potter S, Ribes-Koninckx C, Ricano-Ponce I, Rich SS, Rybak A, Santiago JL, Senapati S, Sood A, Szajewska H, Troncone R, Varade J, Wallace C, Wolters VM, Zernakova A; Spanish Consortium on the Genetics of Coeliac Disease (CEGEC); PreventCD Study Group; Wellcome Trust Case Control Consortium (WTCCC), Thelma BK, Cukrowska B, Urcelay E, Bilbao JR, Mearin ML, Barisani D, Barrett JC, Plagnol V, Deloukas P, Wijmenga C, van Heel DA: Dense genotyping identifies and localizes multiple common and rare variant association signals in celiac disease. *Nat Genet* 43:1193–1201, 2011
 31. Raychaudhuri S, Thomson BP, Remmers EF, Eyre S, Hinks A, Guiducci C, Catanese JJ, Xie G, Stahl EA, Chen R, Alfredsson L, Amos CI, Ardlie KG; BIRAC Consortium, Barton A, Bowes J, Burtt NP, Chang M, Coblyn J, Costenbader KH, Criswell LA, Crusius JB, Cui J, De Jager PL, Ding B, Emery P, Flynn E, Harrison P, Hocking LJ, Huizinga TW, Kastner DL, Ke X, Kurreeman FA, Lee AT, Liu X, Li Y, Martin P, Morgan AW, Padyukov L, Reid DM, Seielstad M, Seldin MF, Shadick NA, Steer S, Tak PP, Thomson W, van der Helm-van Mil AH, van der Horst-Bruinsma IE, Weinblatt ME, Wilson AG, Wolbink GJ, Wordsworth P; YEAR Consortium, Altschuler D, Karlson EW, Toes RE, de Vries N, Begovich AB, Siminovitch KA, Worthington J, Klareskog L, Gregersen PK, Daly MJ, Plenge RM: Genetic variants at CD28, PRDM1 and CD2/CD58 are associated with rheumatoid arthritis risk. *Nat Genet* 41:1313–1318, 2009
 32. Anderson CA, Boucher G, Lees CW, Franke A, D'Amato M, Taylor KD, Lee JC, Goyette P, Imielinski M, Latiano A, Lagace C, Scott R, Amininejad L, Bumpstead S, Baidoo L, Baldassano RN, Barclay M, Bayless TM, Brand S, Buning C, Colombel JF, Denson LA, De Vos M, Dubinsky M, Edwards C, Elinghaus D, Fehrmann RS, Floyd JA, Florin T, Franchimont D, Franke L, Georges M, Glas J, Glazer NL, Guthery SL, Haritunians T, Hayward NK, Hugot JP, Jobin G, Laukens D, Lawrance I, Lemann M, Levine A, Libioulle C, Louis E, McGovern DP, Milla M, Montgomery GW, Morley KI, Mowat C, Ng A, Newman W, Ophoff RA, Papi L, Palmieri O, Peyrin-Biroulet L, Panes J, Phillips A, Prescott NJ, Proctor DD, Roberts R, Russell R, Rutgeerts P, Sanderson J, Sans M, Schumm P, Seibold F, Sharma Y, Simms LA, Seielstad M, Steinhart AH, Targan

- SR, van den Berg LH, Vatn M, Verspaget H, Walters T, Wijmenga C, Wilson DC, Westra HJ, Xavier RJ, Zhao ZZ, Ponsioen CY, Andersen V, Torkvist L, Gazouli M, Anagnou NP, Karlsen TH, Kupcinskas L, Svventoraityte J, Mansfield JC, Kugathasan S, Silverberg MS, Halfvarson J, Rotter JI, Mathew CG, Griffiths AM, Geary R, Ahmad T, Brant SR, Chamaillard M, Satsangi J, Cho JH, Schreiber S, Daly MJ, Barrett JC, Parkes M, Annesse V, Hakonarson H, Radford-Smith G, Duerr RH, Vermeire S, Weersma RK, Rioux JD: Meta-analysis identifies 29 additional ulcerative colitis risk loci, increasing the number of confirmed associations to 47. *Nat Genet* 43:246–252, 2011
33. McGovern DP, Gardet A, Torkvist L, Goyette P, Essers J, Taylor KD, Neale BM, Ong RT, Lagace C, Li C, Green T, Stevens CR, Beauchamp C, Flesher PR, Carlson M, D'Amato M, Halfvarson J, Hibberd ML, Lordal M, Padyukov L, Andriulli A, Colombo E, Latiano A, Palmieri O, Bernard EJ, Deslandres C, Hommes DW, de Jong DJ, Stokkers PC, Weersma RK; NIDDK IBD Genetics Consortium, Sharma Y, Silverberg MS, Cho JH, Wu J, Roeder K, Brant SR, Schumm LP, Duerr RH, Dubinsky MC, Glazer NL, Haritunians T, Ippoliti A, Melmed GY, Siscovick DS, Vasiliauskas EA, Targan SR, Annesse V, Wijmenga C, Pettersson S, Rotter JI, Xavier RJ, Daly MJ, Rioux JD, Seielstad M: Genome-wide association identifies multiple ulcerative colitis susceptibility loci. *Nat Genet* 42:332–337, 2010
34. Franke A, McGovern DP, Barrett JC, Wang K, Radford-Smith GL, Ahmad T, Lees CW, Balschun T, Lee J, Roberts R, Anderson CA, Bis JC, Bumpstead S, Ellinghaus D, Festen EM, Georges M, Green T, Haritunians T, Jostins L, Latiano A, Mathew CG, Montgomery GW, Prescott NJ, Raychaudhuri S, Rotter JI, Schumm P, Sharma Y, Simms LA, Taylor KD, Whiteman D, Wijmenga C, Baldassano RN, Barclay M, Bayless TM, Brand S, Buning C, Cohen A, Colombel JF, Cottone M, Stronati L, Denson T, De Vos M, D'Inca R, Dubinsky M, Edwards C, Florin T, Franchimont D, Geary R, Glas J, Van Gossom A, Guthery SL, Halfvarson J, Verspaget HW, Hugot JP, Karban A, Laukens D, Lawrence I, Lemann M, Levine A, Libioulle C, Louis E, Mowat C, Newman W, Panes J, Phillips A, Proctor DD, Regueiro M, Russell R, Rutgeerts P, Sanderson J, Sans M, Seibold F, Steinhart AH, Stokkers PC, Torkvist L, Kullak-Ublick G, Wilson D, Walters T, Targan SR, Brant SR, Rioux JD, D'Amato M, Weersma RK, Kugathasan S, Griffiths AM, Mansfield JC, Vermeire S, Duerr RH, Silverberg MS, Satsangi J, Schreiber S, Cho JH, Annesse V, Hakonarson H, Daly MJ, Parkes M: Genome-wide meta-analysis increases to 71 the number of confirmed Crohn's disease susceptibility loci. *Nat Genet* 42:1118–1125, 2010
35. International Multiple Sclerosis Genetics Consortium; Wellcome Trust Case Control Consortium 2, Sawcer S, Hellenthal G, Pirinen M, Spencer CC, Patsopoulos NA, Moutsianas L, Dilthey A, Su Z, Freeman C, Hunt SE, Edkins S, Gray E, Booth DR, Potter SC, Goris A, Band G, Oturai AB, Strange A, Saarela J, Bellenguez C, Fontaine B, Gillman M, Hemmer B, Gwilliam R, Zipp F, Jayakumar A, Martin R, Leslie S, Hawkins S, Giannoulatos E, D'Alfonso S, Blackburn B, Martinelli Boneschi F, Liddle J, Harbo HF, Perez ML, Spurkland A, Waller MJ, Mycko MP, Ricketts M, Comabella M, Hammond N, Kockum I, McCann OT, Ban M, Whittaker P, Kempainen A, Weston P, Hawkins C, Widaa S, Zajicek J, Dronov S, Robertson N, Bumpstead SJ, Barcellos LF, Ravindrarajah R, Abraham R, Alfredsson L, Ardlie K, Aubin C, Baker A, Baker K, Baranzini SE, Bergamaschi L, Bergamaschi R, Bernstein A, Berthele A, Boggild M, Bradfield JP, Brassat D, Broadley SA, Buck D, Butzkueven H, Capra R, Carroll WM, Cavalla P, Celius EG, Cepok S, Chiavacci R, Clerget-Darpoux F, Clysters K, Comi G, Cossburn M, Cournu-Rebeix I, Cox MB, Cozen W, Cree BA, Cross AH, Cusi D, Daly MJ, Davis E, de Bakker PI, Debouverie M, D'hooghe MB, Dixon K, Dobosi R, Dubois B, Ellinghaus D, Elovaara I, Esposito F, Fontenille C, Foote S, Franke A, Galimberti D, Ghezzi A, Glessner J, Gomez R, Gout O, Graham C, Grant SF, Guerini FR, Hakonarson H, Hall P, Hamsten A, Hartung HP, Heard RN, Heath S, Hobart J, Hoshi M, Infante-Duarte C, Ingram G, Ingram W, Islam T, Jagodic M, Kabesch M, Kermod AG, Kilpatrick TJ, Kim C, Klopp N, Koivisto K, Larsson M, Lathrop M, Lechner-Scott JS, Leone MA, Leppa V, Liljedahl U, Bomfim IL, Lincoln RR, Link J, Liu J, Lorentzen AR, Lupoli S, Macciardi F, Mack T, Marriott M, Martinelli V, Mason D, McCauley JL, Mentch F, Mero IL, Mihalova T, Montalban X, Mottershead J, Myhr KM, Naldi P, Ollier W, Page A, Palotie A, Pelletier J, Piccio L, Pickersgill T, Piehl F, Pobywajlo S, Quach HL, Ramsay PP, Reunanen M, Reynolds R, Rioux JD, Rodegher M, Roesner S, Rubio JP, Ruckert IM, Salvetti M, Salvi E, Santaniello A, Schaefer CA, Schreiber S, Schulze C, Scott RJ, Sellebjerg F, Selmaj KW, Sexton D, Shen L, Simms-Acuna B, Skidmore S, Sleiman PM, Smestad C, Sorensen PS, Sondergaard HB, Stankovich J, Strange RC, Sulonen AM, Sundqvist E, Syvanen AC, Taddeo F, Taylor B, Blackwell JM, Tienari P, Bramon E, Tourbah A, Brown MA, Tronczynska E, Casas JP, Tubridy N, Corvin A, Vickery J, Jankowski J, Villoslada P, Markus HS, Wang K, Mathew CG, Wason J, Palmer CN, Wichmann HE, Plomin R, Willoughby E, Rautanen A, Winkelmann J, Wittig M, Trembath RC, Yaouanq J, Viswanathan AC, Zhang H, Wood NW, Zuvich R, Deloukas P, Langford C, Duncanson A, Oksenberg JR, Pericak-Vance MA, Haines JL, Olsson T, Hillert J, Iverson AJ, De Jager PL, Peltonen L, Stewart GJ, Hafler DA, Hauser SL, McVean G, Donnelly P, Compston A: Genetic risk and a primary role for cell-mediated immune mechanisms in multiple sclerosis. *Nature* 476:214–219, 2011
36. Zhang Z, Duvefelt K, Svensson F, Masterman T, Jonasdottir G, Salter H, Emahazion T, Hellgren D, Falk G, Olsson T, Hillert J, Anvret M: Two genes encoding immune-regulatory molecules (LAG3 and IL7R) confer susceptibility to multiple sclerosis. *Genes Immun* 6:145–152, 2005
37. Clayton DG: Prediction and interaction in complex disease genetics: experience in type 1 diabetes. *PLoS Genet* 5:e1000540, 2009
38. Smemo S, Tena JJ, Kim KH, Gamazon ER, Sakabe NJ, Gomez-Marin C, Aneas I, Gredidio FL, Sobreira DR, Wasserman NF, Lee JH, Puvindran V, Tam D, Shen M, Son JF, Vakilt NA, Sung HK, Naranjo S, Acemel RD, Manzanares M, Nagy A, Cox NJ, Hui CC, Gomez-Skarmeta JL, Nobrega MA: Obesity-associated variants within FTO form long-range functional connections with IRX3. *Nature* 507:371–375, 2014
39. Claussnitzer M, Dankel SN, Kim KH, Quon G, Meuleman W, Haugen C, Glunk V, Sousa IS, Beaudry J, Puvindran V, Abdennur NA, Liu J, Svensson PA, Hsu YH, Drucker DJ, Mellgren G, Hui CC, Hauner H, Kellis M: FTO obesity variant circuitry and adipocyte browning in humans. *N Engl J Med* 373:895–907, 2015
40. Elding Larsson H, Vehik K, Gesualdo P, Akolkar B, Hagopian W, Krischer J, Lernmark A, Rewers M, Simell O, She JX, Ziegler A, Haller MJ; TEDDY Study Group: Children followed in the TEDDY study are diagnosed with type 1 diabetes at an early stage of disease. *Pediatr Diabetes* 15:118–126, 2014
41. Plagnol V, Howson JM, Smyth DJ, Walker N, Hafler JP, Wallace C, Stevens H, Jackson L, Simmonds MJ; Type 1

- Diabetes Genetics Consortium, Bingley PJ, Gough SC, Todd JA: Genome-wide association analysis of autoantibody positivity in type 1 diabetes cases. *PLoS Genet* 7:e1002216, 2011
42. Tait BD, Harrison LC, Drummond BP, Steward V, Varney MD, Honeyman MC: HLA antigens and age at diagnosis of insulin-dependent diabetes mellitus. *Hum Immunol* 42:116–122, 1995
43. Howson JM, Cooper JD, Smyth DJ, Walker NM, Stevens H, She JX, Eisenbarth GS, Rewers M, Todd JA, Akolkar B, Concannon P, Erlich HA, Julier C, Morahan G, Nerup J, Nierras C, Pociot F, Rich SS; Type 1 Diabetes Genetics Consortium: Evidence of gene-gene interaction and age at diagnosis effects in type 1 diabetes. *Diabetes* 61:3012–3017, 2012
44. HLA MHC Complex [article online], 2009. Available from <http://commons.wikimedia.org/wiki/File:HLA.svg#mediaviewer/File:HLA.svg>. Accessed 27 December 2015
45. Lopez-Requena A, Burrone OR, Cesco-Gaspere M: Idiotypes as immunogens: facing the challenge of inducing strong therapeutic immune responses against the variable region of immunoglobulins. *Front Oncol* 2:159, 2012