Searching for the Autoimmune Thyroid Disease Susceptibility Genes: From Gene Mapping to Gene Function

Yaron Tomer and Terry F. Davies


To subscribe to Endocrine Reviews or any of the other journals published by The Endocrine Society please go to: http://edrv.endojournals.org//subscriptions/
Searching for the Autoimmune Thyroid Disease Susceptibility Genes: From Gene Mapping to Gene Function

YARON TOMER AND TERRY F. DAVIES

Division of Endocrinology, Diabetes, and Bone Diseases, Department of Medicine, Mount Sinai School of Medicine, New York, New York 10029

The autoimmune thyroid diseases (AITD) are complex diseases that are caused by an interaction between susceptibility genes and environmental triggers. Genetic susceptibility, in combination with external factors (e.g., dietary iodine), is believed to initiate the autoimmune response to thyroid antigens. Abundant epidemiological data, including family and twin studies, point to a strong genetic influence on the development of AITD. Various techniques have been used to identify the genes contributing to the etiology of AITD, including candidate gene analysis and whole genome screening. These studies have enabled the identification of several loci (genetic regions) that are linked with AITD, and in some of these loci putative AITD susceptibility genes have been identified. Some of these genes/loci are unique to Graves' disease (GD) and Hashimoto's thyroiditis (HT), and some are common to both diseases, indicating that there is a shared genetic susceptibility to GD and HT. The putative GD and HT susceptibility genes include both immune modifying genes (e.g., human leukocyte antigen, cytotoxic T lymphocyte antigen-4) and thyroid-specific genes (e.g., TSH receptor, thyroglobulin). Most likely these loci interact, and their interactions may influence disease phenotype and severity. It is hoped that in the near future additional AITD susceptibility genes will be identified and the mechanisms by which they induce AITD will be unraveled. (Endocrine Reviews 24: 694–717, 2003)

I. Introduction

II. The Autoimmune Thyroid Diseases Are Familial
A. Population data (Table 1)
B. Family studies
C. Sibling risk ratio (as)
D. Segregation analyses
E. Twin studies (Table 2)

III. Tools Used to Map and Identify Complex Disease Genes
A. Linkage and association
B. Candidate gene analysis
C. Whole genome screens
D. Markers used in linkage and association studies
E. A suggested algorithm for searching for complex disease genes

IV. Immune Regulatory Genes Studied in AITD
A. The role of HLA in the genetic susceptibility to AITD
B. The role of CTLA-4 in the genetic susceptibility to AITD
C. Other immune-related genes

V. Thyroid-Specific Genes Studied in AITD

VI. Linkage Studies of Candidate Chromosomal Regions (Table 7)
A. Chromosome X
B. The IDDM loci
C. The 14q region
D. The 20q region

VII. Whole Genome Screening in AITD

VIII. Mechanisms by Which Genes Can Induce Thyroid Autoimmunity
A. General principles
B. HLA
C. CTLA-4
D. Hypothetical mechanisms by which Tg could induce susceptibility to AITD

IX. Conclusions and Future Directions

Abbreviations: AITD, Autoimmune thyroid disease(s); APC, antigen-presenting cell(s); CTLA-4, cytotoxic T lymphocyte antigen-4; DM, diabetes mellitus; DZ, dizygotic; EAT, experimental autoimmune thyroiditis; GD, Graves' disease; GO, Graves' ophthalmopathy; HLA, human leukocyte antigen; HLOD, heterogeneity LOD score; HT, Hashimoto's thyroiditis; IDDM, insulin-dependent IDDM; IgH, IgG heavy chain; LD, linkage disequilibrium; LOD, logarithm of odds; MHC, major histocompatibility complex; MLS, maximum LOD score; MZ, monozygotic; NHANES III, National Health and Nutrition Examination Survey; RR, relative risk; SNP, single nucleotide polymorphism; TAb, thyroid antibody or antibodies; TDT, transmission disequilibrium test; Tg, thyroglobulin; Tg-Ab, Tg antibodies; Tgms2, Tg microsatellites in intron 27; TPO, thyroid peroxidase; TPO-Ab, TPO antibodies; TSHR, TSH receptor; UTR, untranslated region.
Amiodarone (for review, see Ref. 6), and interferon-α (7, 8), thyroiditis accompanying the polyglandular autoimmune syndromes (reviewed in Refs. 9 and 10), and the presence of thyroid antibodies (TAb) [please note that we use the abbreviation TAb for thyroid peroxidase (TPO) and thyroglobulin (Tg) antibodies only] with no apparent clinical disease (11) (in this review we use the term AITD to denote GD and HT only). Although the exact etiology of the autoimmune response to the thyroid remains unknown, there is solid evidence for a major genetic influence on the development of AITD (reviewed in Refs. 12 and 13). Therefore, the current paradigm is that AITD are complex diseases in which susceptibility genes and environmental triggers act in concert to initiate the autoimmune response to the thyroid. In this review, we will summarize the recent advances in our understanding of the genetic contribution to the development of AITD. We will limit the discussion to the genetics of GD, HT, and thyroid autoantibodies.

II. The Autoimmune Thyroid Diseases Are Familial

A. Population data (Table 1)

A recently published, large, epidemiological study in the United States population, which is iodine sufficient [National Health and Nutrition Examination Survey (NHANES III)], found the prevalence of thyrotoxicosis from any cause to be 1.3% (0.5% clinical and 0.7% subclinical) (14). Another recent large study from the United States (The Colorado Thyroid Disease Prevalence Study) found the prevalence of subclinical and clinical hyperthyroidism to be 2.1 and 0.1%, respectively (15). Additionally, a comprehensive survey in the town of Whickham in the United Kingdom found the prevalence of thyrotoxicosis to be 2.7% in women, and 10-fold less in men (0.16–0.23%) (16). Because the commonest cause of thyrotoxicosis in iodine-sufficient Western countries is GD, accounting for 70–80% of cases (17–19), these frequencies are probably good estimates of the prevalence of GD in the United Kingdom and United States. However, the level of dietary iodine has been shown to markedly affect the incidence of thyrotoxicosis and GD (20). In iodine-deficient regions, thyrotoxicosis is more prevalent, and GD accounts for a lower percentage of all thyrotoxicosis cases. Indeed, a recent well-designed large study from Denmark has shown the incidence of thyrotoxicosis to be 65.4/100,000/yr in a mild iodine deficiency region, and 92.9/100,000/yr in an area with moderate iodine deficiency (21). In another study, the incidence of GD was higher (19.7/100,000/yr) in Iceland (a high iodine intake region) when compared with a region in Denmark with low average iodine intake (14.8/100,000/yr) (20).

Epidemiological surveys from various, mostly iodine-sufficient regions have shown a relatively similar incidence of GD in Caucasian populations (18, 19, 22–24). The annual incidence of GD in these studies was approximately 20–25 per 100,000 (Table 1).

In mostly iodine-sufficient regions, similar prevalence and incidence trends have also been observed for HT. In the NHANES III study, the prevalence of hypothyroidism was found to be 4.6% (0.3% clinical and 4.3% subclinical) (14). In the Whickham survey, the prevalence of spontaneous hypothyroidism was 1.5% in females and less than 0.1% in males (16). These prevalence rates are similar to those reported in Finland (25), Japan (26), and in another U.S. survey (27). A higher prevalence of hypothyroidism was reported in the Colorado study (0.4% clinical and 9.0% subclinical), but this study included an older population with a higher percentage of women than in the general population of Colorado (15), and this might have increased the prevalence of subclinical hypothyroidism in this study (28). It should be emphasized that the phenotype definitions of GD and HT in these epidemiological studies are not always specified and

Table 1. Prevalence and incidence of clinical thyrotoxicosis/GD and hypothyroidism/HT in different geographic regions

<table>
<thead>
<tr>
<th>Country</th>
<th>Phenotype</th>
<th>Years of survey</th>
<th>Prevalence (F/M)</th>
<th>Incidence per 100,000/yr (F/M)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sweden</td>
<td>GD</td>
<td>1970–1974</td>
<td>NA</td>
<td>17.2 (27/27.4)</td>
</tr>
<tr>
<td>Sweden</td>
<td>GD</td>
<td>1988–1990</td>
<td>NA</td>
<td>22.2 (34.4/6.8)</td>
</tr>
<tr>
<td>United States</td>
<td>GD</td>
<td>1935–1967</td>
<td>NA</td>
<td>34.8/3.3</td>
</tr>
<tr>
<td>United States</td>
<td>HT</td>
<td>1935–1967</td>
<td>NA</td>
<td>40.7/1.0</td>
</tr>
<tr>
<td>Iceland</td>
<td>Thyrotoxicosis</td>
<td>1980–1982</td>
<td>NA</td>
<td>23.6 (35.8/8.9)</td>
</tr>
<tr>
<td>New Zealand</td>
<td>Thyrotoxicosis</td>
<td>1983–1985</td>
<td>NA</td>
<td>25.8 (40.7/10.5)</td>
</tr>
<tr>
<td>United Kingdom</td>
<td>Thyrotoxicosis</td>
<td>1972–1974</td>
<td>(2.7%/0.16%)</td>
<td>NA</td>
</tr>
<tr>
<td>United Kingdom</td>
<td>Thyrotoxicosis</td>
<td>1972–1992</td>
<td>(3.9%/0.2%)</td>
<td>(80/0)*</td>
</tr>
<tr>
<td>United States</td>
<td>Thyrotoxicosis</td>
<td>1988–1994</td>
<td>0.5%</td>
<td>NA</td>
</tr>
<tr>
<td>United States</td>
<td>Thyrotoxicosis</td>
<td>1995</td>
<td>0.1%</td>
<td>NA</td>
</tr>
<tr>
<td>United States</td>
<td>Thyrotoxicosis</td>
<td>1977</td>
<td>0.31%</td>
<td>NA</td>
</tr>
<tr>
<td>United States</td>
<td>Thyrotoxicosis</td>
<td>1965–1995</td>
<td>1.2%</td>
<td>NA</td>
</tr>
<tr>
<td>United States</td>
<td>Hypothyroidism</td>
<td>1972–1974</td>
<td>(1.9%/0.1%)</td>
<td>NA</td>
</tr>
<tr>
<td>United States</td>
<td>Hypothyroidism</td>
<td>1972–1992</td>
<td>(7.7%/1.3%)</td>
<td>(350/60)*</td>
</tr>
<tr>
<td>United States</td>
<td>Hypothyroidism</td>
<td>1988–1994</td>
<td>0.3%</td>
<td>NA</td>
</tr>
<tr>
<td>United States</td>
<td>Hypothyroidism</td>
<td>1995</td>
<td>0.4%</td>
<td>NA</td>
</tr>
<tr>
<td>United States</td>
<td>Hypothyroidism</td>
<td>1977</td>
<td>0.5%</td>
<td>NA</td>
</tr>
<tr>
<td>United States</td>
<td>Hypothyroidism</td>
<td>1965–1995</td>
<td>0.8%</td>
<td>NA</td>
</tr>
</tbody>
</table>

The incidence/prevalence of subclinical disease is not included in the table. See text (Section II.A) for details. F, female; M, male; NA, not available.

a The incidence was calculated only in the survivors of the original study and may not reflect the incidence in the general population.

b Clinical hypo/hyperthyroidism.

c This is a metaanalysis of several studies.
that variations in phenotype definitions may significantly affect the population data.

The comparable prevalence and incidence of GD and HT in geographically different populations that are iodine sufficient suggests a significant genetic effect on the development of these diseases because these populations are exposed to different environmental factors (other than dietary iodine). Moreover, it may imply that different Caucasian populations share some of their susceptibility genes for AITD.

Longitudinal surveys can also point to the relative importance of genetic factors in the etiology of a disease. An early longitudinal study from the Mayo Clinic (1935–1967) showed no significant change in the incidence of GD over the 33 yr of the study (29). The stable incidence of GD with time points to a genetic susceptibility to GD because the genetic makeup of a population does not change over several decades, but environmental factors would be expected to vary over time. Therefore, the incidence rates of diseases with strong environmental influences (e.g., lung cancer) are expected to change over time, whereas the incidence rates of diseases with strong genetic influences should be more stable with time, although complex diseases in which environmental and genetic factors interact can also vary with time. The Mayo Clinic observations were supported by a more recent study from Sweden (30) that showed no significant change in the incidence of GD in Malmö over a 20-yr period (the incidence was 17.7/100,000/yr in 1970–1974 and 22.2/100,000/yr in 1988–1990). However, the Swedish study also found an increased incidence of GD in females younger than 50 yr (30), demonstrating that environmental effects also play a role in the etiology of GD. In contrast to GD, the early Mayo survey (1935–1967) found a significant increase in the incidence of HT over the 33 yr of the survey (29). This may have reflected a stronger environmental influence on the development of HT or, more likely, a change in the diagnostic criteria over time. Indeed, the diagnosis of HT is much more subject to variation (e.g., goitrous vs. atrophic) than the diagnosis of GD (31).

B. Family studies

The familial occurrence of AITD has been reported by investigators for many years. Earlier studies showing familial aggregation of AITD were mostly observational, based on careful family histories from patients. These early studies reported a family history of thyroid disease in up to 60% of patients with GD (32, 33). Later, in the 1960s, Hall and Stanbury (34) showed that 33% of siblings of patients with GD or HT developed AITD themselves. Additionally, they found that 56% of siblings of AITD patients had TAb, and in almost all cases at least one of the parents of an affected individual had TAb, suggesting dominant inheritance of the TAb trait (34). More recently, several groups have reported a high frequency of thyroid abnormalities in relatives of patients with GD (35–38), most commonly the presence of thyroid autoantibodies that were reported in up to 50% of the siblings of patients with GD (reviewed in Ref. 39; also see Refs. 36 and 40). A recent survey by our own group revealed that 41 of 114 (36%) GD patients with ophthalmopathy reported a family history of AITD, and 26 of 114 (23%) had a first-degree relative with AITD (41).

C. Sibling risk ratio (Rs)

Familial clustering of a disease does not necessarily mean that the disease is genetic. Familial clustering can be the result of random chance, shared extrinsic (environmental) factors, shared intrinsic (genetic) factors, or a combination of these. Coincidental familial clustering may occur in cases in which a disease is common and, therefore, may occur in multiple members of the same family by chance alone. For example, essential hypertension is common in Western populations and, therefore, may occur in multiple members in the same family. However, this does not necessarily mean that it is caused by genetic factors. Shared external factors include diet (e.g., iodine deficiency can cause familial clustering of cretinism and goiter), infections [e.g., exposure to certain viruses can cause familial clustering of subacute thyroiditis (reviewed in Ref. 42)], and other environmental exposures [e.g., exposure to radiation fallout can cause familial clustering of childhood thyroid cancer (43)]. These situations have to be distinguished from genetic causes of familial clustering.

Several methods are available to determine whether familial clustering of a disease is the result of genetic susceptibility or nongenetic factors. One method is to calculate the Rs, which is the ratio of the prevalence of the disease in siblings of affected individuals compared with the prevalence of the disease in the general population (44). The Rs expresses the increased risk of developing the disease in an individual who has a sibling with the disease and is a quantitative measure of the genetic contribution to the disease. A Rs greater than 5 usually indicates a significant genetic contribution to the pathogenesis of a disease (44). In the case of the AITD, the Rs was estimated to be between 5.9 and more than 10 (12, 41, 45, 46), supporting a strong genetic influence on the development of AITD.

D. Segregation analyses

Another approach for determining whether familial clustering of a disease reflects genetic predisposition is to perform segregation analyses. Here, the segregation of the disease in families is analyzed to see whether it occurs at random or demonstrates Mendelian or a complex pattern of inheritance. This analysis has the advantage of determining both the existence of genetic predisposition to a disease and the mode of inheritance. Two segregation analyses have been performed in families with TAB and have suggested a Mendelian dominant pattern of inheritance for the tendency to develop TAB (47–49). In keeping with these observations, it was recently found that recognition of particular TPO epitopes within the autoantibody immunodominant region may be transmitted within families (50).

E. Twin studies (Table 2)

The most powerful method for evaluating genetic predisposition to complex diseases involves twins. Twin studies are based on comparison of the concordance (simultaneous oc-
curcurrence) of a given disease among monozygotic (MZ; i.e., near identical) twins with the concordance among dizygotic (DZ; i.e., fraternal) twins. MZ twins have almost identical genetic makeup, whereas DZ twins share 50% of their genes (like siblings). Therefore, if the concordance is higher in the MZ twins when compared with the DZ twins, it suggests that the disease has an inherited component, although the definition of identical twins may not be as simple as first thought (for a review, see Ref. 51). Any discordance among the MZ twin pairs is usually interpreted to mean that the gene or genes concerned show reduced penetrance, i.e., certain nongenetic events must occur or certain environmental factors must be present before the disease becomes manifest. It must be emphasized that MZ twins are not identical in their immune repertoire due to somatic recombinations that T and B cells undergo throughout life, as well as individual immune experiences that influence the immune repertoire (52). Therefore, part of the observed discordance between MZ twins may also be due to the discordance in their immune repertoire. Several large twin studies have been reported from Denmark showing a higher concordance of AITD in MZ twins when compared with DZ twins (Table 2). For GD, the concordance was 35% in MZ twins and 3% in DZ twins (53, 54). A recent GD twin study from California confirmed the Danish twin study results (55). Brix et al. (54) have suggested that genetic factors are responsible for 79% of the susceptibility to develop GD, whereas environmental factors contribute the remaining 21%, but this requires further evaluation.

Twin studies in HT have shown concordance rates of 55 and 0% in MZ and DZ twins, respectively (56). The concordance rates for thyroid autoantibodies (TAb) were also reported to be higher in MZ twins compared with DZ twins (Table 2). In a recent study from the United Kingdom, the concordance rates for Tg antibodies (Tg-Ab) were 59 and 23% for MZ and DZ twins, respectively (57). The concordance rates for TPO antibodies (TPO-Ab) were 47 and 29% for MZ and DZ twins, respectively (57). The Danish twin studies also found higher concordance rates for TAb in MZ twins (80%) when compared with DZ twins (40%) (56). These twin data confirm with remarkable clarity the presence of a substantial inherited susceptibility to AITD.

III. Tools Used to Map and Identify Complex Disease Genes

Based on the abundant epidemiological evidence for a strong genetic effect on the development of AITD, several groups have been trying to map and identify the AITD susceptibility genes. The two basic strategies used for mapping complex disease genes include linkage and association studies of candidate genes and whole genome screening. These tools have been successful in mapping classical Mendelian disorders [e.g., Pendred’s syndrome (58)], and more recently have enabled the identification of the first complex disease genes for type 1 diabetes (59) and Crohn’s disease (60). These strategies have also been used in the study of the genetic susceptibility to AITD.

A. Linkage and association

1. Linkage. Genetic linkage techniques are powerful tools for analyzing complex disease-related genes because they detect genes that have a major influence on the development of a disease (61). However, linkage studies are less sensitive than association studies because they do not detect less influential genes (61). A linkage study, therefore, may be negative in the absence of major genes contributing to disease susceptibility. The principle of linkage analysis is based on the fact that if two genes or markers are close together on a chromosome, they will cosegregate because the likelihood that a recombination will occur between them during meiosis is low. Therefore, if a tested marker is close to a disease susceptibility gene, its alleles will cosegregate with the disease in families (Fig. 1). The logarithm of odds (LOD) score is the measure of the likelihood of linkage between a disease and a genetic marker (62). The LOD score is the base-10 logarithm of the odds ratio in favor of linkage. In Mendelian disorders, a LOD score of greater than 3 (i.e., odds ratio greater than 1000) is considered strong evidence for linkage (62). The classical linkage tests are model based (parametric), i.e., different modes of inheritance and penetrance have to be tested when calculating the likelihood of linkage. The parametric tests are the most powerful statistical tests for linkage (63, 64), and they can be used to test for heterogeneity within a dataset (heterogeneity exists in a dataset when more than one gene causes the same disease phenotype; see Section III.A.3) (65). In complex diseases, the mode of inheritance is often unknown, and therefore, simpler model-independent methods (nonparametric) have also been widely used (44). One such method is sib-pair analysis (44). In this method, siblings that are both affected by the disease being studied are tested for sharing of alleles at a marker locus. By random chance alone, the sibs would be expected to share one allele of the marker 50% of the time and two alleles 25% of the time. If affected sib-pairs share a significantly higher than expected proportion of alleles at the marker locus, this suggests that the region containing the marker locus also contains the disease gene. The observed to expected allele sharing can then be converted to a LOD score equivalent.

2. Guideline for measuring linkage in complex diseases. In simple Mendelian disorders, a maximum LOD score (MLS) of more than 3 is considered strong evidence of linkage (62). However, the inheritance of complex diseases (e.g., AITD) does not follow a simple Mendelian pattern. These diseases are likely to be caused by several genes and have reduced penetrance (i.e., not all the individuals inheriting the gene will
develop the disease). In addition, it is likely that different genes may cause almost identical phenotypes (genetic heterogeneity). This results in non-Mendelian transmission of the disease in pedigrees and makes mapping the susceptibility genes for complex diseases difficult. Therefore, Lander and Kruglyak (66) have suggested guidelines for genetic linkage studies in complex traits. According to their guidelines, in complex diseases a LOD score of greater than 1.9 is suggestive of linkage, and a LOD score of greater than 3.3 indicates significant linkage in studies using the parametric approach. For nonparametric sib-pair studies, the cutoff LOD scores are higher (66). Linkage is confirmed if evidence for linkage is replicated in two separate datasets (66). Conversely, a LOD score lower than −2.0 has been used to exclude linkage.

3. Phenotype definitions and genetic heterogeneity. Phenotype definitions are important in genetic studies because different phenotypes are likely to be caused by different genes and analyzing them together would make identification of these genes more difficult. Although the GD phenotype is relatively homogenous, GD can occur with or without a large goiter and/or ophthalmopathy and/or pretibial myxedema, and these should be analyzed separately. However, our segregation analysis has shown that the ophthalmopathy phenotype is most likely not determined by genetic factors (41), and therefore, it is possible that the subsetting by ophthalmopathy may not be necessary. In the case of HT, the phenotype is even more variable (e.g., the presence of goiter, TPO-Ab, and the need for thyroid hormone replacement therapy) (31), and therefore, strict definitions of HT should be used in genetic studies to ensure a uniform HT phenotype. However, even when the phenotype is uniform, genetic heterogeneity can still exist. Genetic heterogeneity exists when different genotypes give rise to indistinguishable phenotypes (67). If heterogeneity exists in a dataset of families with a given phenotype (e.g., AITD), the dataset may include only a subset of families that are linked with a tested marker. The existence of genetic heterogeneity in a dataset can be tested for by the Admixture Test, which calculates the likelihood that a proportion of the families in a dataset are linked to the marker (the fraction of linked families is estimated by the e-statistic) (65, 68, 69). For example, the heterogeneity LOD score (HLOD) function of the Genehunter computer program performs the Admixture Test and can be used to detect heterogeneity in a given dataset (70).

4. Association. Linkage studies are excellent for screening the whole genome. However, they have limited resolution (~2–3 cM) because as the linked interval is narrowed all markers in the region will show linkage (61). Association studies are more sensitive than linkage studies and therefore are better for fine-mapping linked genetic regions, because the association signal [i.e., the relative risk (RR)] increases as the markers get closer to the susceptibility gene. Association analysis is highly sensitive and may detect genes contributing less than 5% of the total genetic contribution to a disease (71).

Association analyses are performed by comparing the frequency of the allele studied [e.g., human leukocyte antigen (HLA)-DR3] in unrelated patients and in unrelated, ethnically matched controls. This is usually performed by typing each individual patient and each control for the tested marker, but recently, methods for DNA pooling have been developed, which could simplify large-scale association studies (72). If the allele tested is associated with the disease, it will appear significantly more frequently in patients than in controls. The probability of having the disease in an individual positive for the allele compared with an individual negative for the allele is estimated by the RR (73). There are at least two possible explanations for the existence of an association between an allele and a disease: 1) the associated allele itself is the genetic variant causing an increased risk for the disease; and 2) the associated allele itself is not causing the disease but rather a gene in linkage disequilibrium (LD) with it (74). Linkage disequilibrium exists when chromosomes with the mutant allele at the disease locus carry certain marker alleles more often than expected.

This population-based association method may produce spurious associations if the patients and controls are not
accurately matched (population stratification) (75). Therefore, new association tests have been developed that are family-based and use an internal control group from within each family, thus avoiding the necessity to match patients and controls altogether. The most widely used family-based association test is the transmission disequilibrium test (TDT) (75–77). The TDT is based on comparison of parental marker alleles that are transmitted and those that are not transmitted to affected children (Fig. 2). Assuming two heterozygous parents for a certain tested marker, the four parental alleles in each family are categorized into two groups: those transmitted to a child with the disease (T alleles), and those not transmitted to an affected child (N alleles). The same allele may belong to the T group or the N group in different families. The frequency of the T alleles vs. the N alleles is then compared by a χ² test. An association between a certain allele and the disease exists if there is an excess occurrence of this allele in the T group compared with the N group. The TDT can also serve as a linkage test if there is a known LD between the tested marker and the disease.

5. Power calculations. Before performing a linkage/association study, power calculations should determine whether the proposed dataset size is sufficient to detect/reject linkage/association with the marker. The weaker the association or linkage, the larger the dataset required to detect it. Power calculations for case-control association studies are performed by simulating a dataset similar to the one being used in the study. Data for a hypothetical marker locus associated with the disease are generated using a variety of odds ratios to generate proportions of patients and controls, with and without the marker. The association analyses for that marker determine the odds ratio at which an association would not be detected with the given dataset. For linkage studies, power calculations are performed by simulating a large number (>1000) of datasets consisting of families similar (or identical) to the ones being used and simulating a marker that is linked to the disease locus in these families. It is also possible to simulate another marker that is not linked to the disease in the simulated families to determine whether one can reject linkage with the given marker in the chosen families. The linkage analysis is then run on each of the datasets to find the percentage of times at which the simulated linked marker will show linkage, and similarly the percentage of times at which the simulated nonlinked marker will show evidence against linkage (78, 79). Genetic studies should always provide this analysis to show that the dataset is appropriate.

B. Candidate gene analysis

Candidate genes are of known sequence and location that by virtue of their physiological functions may be involved in disease pathogenesis. For example, one can hypothesize that the TSH receptor (TSHR) may be a candidate gene for GD because the hallmark of the disease is the presence of TSHR antibodies. If a candidate locus is indeed the cause of a disease, then markers in that locus should show association and linkage with the disease (80). Because the basic abnormality in AITD is an immune response against thyroid antigens, possible candidate genes for AITD include genes that control immune responses [e.g., the major histocompatibility complex (MHC) genes, the T cell receptor genes, and antibody genes], and genes encoding the target autoantigens in AITD (Tg, TPO, iodide transporter, TSHR). Many of these genes have now been studied for a possible role in the genetic susceptibility to AITD (see Sections IV and V).

C. Whole genome screens

Another approach is to screen the whole human genome for linkage with a disease without any assumptions on disease pathogenesis (whole genome screening) (59). Whole genome screening is performed by testing a panel of markers that span the entire human genome for linkage with a disease in a given dataset. If one or more of the markers shows evidence for linkage with the disease according to the guidelines of Lander and Kruglyak (66), these regions may harbor susceptibility genes for the disease studied. These linked regions can then be fine-mapped, and the genes can be identified (see Section III.E). The two requirements for performing parametric linkage from a whole genome screen in a complex disease are: 1) the availability of a sizeable and well-validated dataset of multiplex families (large families with more than one individual affected); and 2) the availability of a map of highly polymorphic, closely-spaced markers covering the whole genome, which can be tested for linkage with the disease.

D. Markers used in linkage and association studies

1. Microsatellites. The most widely used polymorphic markers for whole genome screening are microsatellite markers. Mi-
crosatellites are regions in the genome that are composed of short sequence repeats, most commonly two-base CA repeats (81) and usually have no known physiological function. Microsatellite loci are highly polymorphic (i.e., have many alleles) because the number of repeats in each individual is variable. Moreover, they are extremely abundant and uniformly distributed throughout the genome at distances of less than 1 million bp (81). Therefore, microsatellites serve as excellent markers in whole genome linkage studies. Two independent whole genome screens have now been completed for the AITDs (see Section VII). Microsatellites can also be useful in candidate gene studies if a microsatellite is identified within the candidate gene or very close to it.

2. Single nucleotide polymorphisms (SNPs). SNPs are single base pair positions in genomic DNA at which different sequence alternatives (alleles) exist in normal individuals. Although four alleles are theoretically possible (A, C, T, G), in humans, most SNPs are di- allelic (~70% are A/G; reviewed in Ref. 82). SNPs are very abundant, and current data suggest that their frequency is about one SNP per 1000 bp (83). Because SNPs are less informative than microsatellites (SNPs have only two alleles, and microsatellites usually have more than five alleles), they are not useful in linkage studies. However, because SNPs are much more abundant and closely spaced than microsatellites, they are ideal for fine-mapping genes in linked regions using association studies. The importance of SNPs stems from the fact that many have the potential to change the amino acid sequence of a gene product and be directly involved in the susceptibility to complex diseases because they cause changes in gene function (82). Thus, if a SNP allele inside a gene is found to be significantly associated with a disease, it may be the actual causative allele, increasing susceptibility to the disease (84).

E. A suggested algorithm for searching for complex disease genes

Unlike the search for genes causing simple Mendelian disorders, it is still not known what is the best approach to identify susceptibility genes for complex diseases. However, several groups have successfully identified putative susceptibility genes in complex disorders, notably the identification of NOD2 as a major gene for Crohn’s disease (60, 84), and the approach used in these studies is gaining popularity (85, 105). It is based on five steps that combine genetic linkage and association analyses with functional-biological studies:

1. Identifying linked loci. This is achieved by whole genome screening using microsatellite markers at an average distance of less than 10 cM.

2. Confirming linked loci. A linked locus should be confirmed by finding evidence for linkage in two independent datasets (66). Confirmed loci most likely harbor susceptibility genes, e.g., HLA in type 1 diabetes mellitus (DM) (59).

3. Fine-mapping confirmed loci. Several methods are available to fine-map a linked region. The most popular approach for fine-mapping loci linked with complex disease is LD mapping. LD mapping is based on association studies with markers that saturate the region of interest. The marker that shows the strongest association with the disease (i.e., the lowest P value and highest RR) is probably closest to the disease gene. This method can narrow down the region of interest to a few hundred kilobases. LD mapping may be limited in populations without abundant founder haplotypes. However, in carefully selected populations with homogenous disease etiology, it may be very effective (64, 83). In addition, this method is effective in regions showing strong evidence for linkage with the disease (83).

4. Testing genes in the linked region. After the linked region has been narrowed to less than 1 cM, the known and unknown genes in this region can be analyzed. The Human Genome Project identified most of the genes in a linked region, although it is still possible that an unidentified gene may not be annotated in the Human Genome Project maps. Known genes can be analyzed by identifying SNPs in them and testing these SNPs for association with the disease. If a certain SNP shows a consistently significant association with the disease, it may be the susceptibility allele in the region, although LD with another disease-causing polymorphism cannot be ruled out.

5. Functional studies. To demonstrate that an associated allele is a true susceptibility allele, it is necessary to show that it affects the function of the gene in a way that increases the risk of developing disease. This provides indirect evidence that it may be the actual susceptibility allele for the disease. For example, the SNP in the NOD2 gene that is associated with Crohn’s disease was shown to influence the activation of nuclear factor-κB, thereby influencing the activation of intestinal T cells (84). Although there is no absolute way to prove that a certain SNP is the true susceptibility allele for a disease, as the function of the SNP and its influence on disease development is clarified the role of the identified SNP can be substantiated.

IV. Immune Regulatory Genes Studied in AITD

A. The role of HLA in the genetic susceptibility to AITD

1. The MHC. The MHC region, encoding the HLA glycoproteins, consists of a complex of genes located on chromosome 6p21. The MHC region also encodes various additional proteins, most of which are associated with immune responsiveness (86). The MHC locus itself encodes genes that are grouped into three classes: 1) class I genes, including the HLA antigens A, B, and C; 2) class II genes, including the HLA-DR, DP, and DQ genes; and 3) class 3 genes, including several complement components (e.g., C4), TNFα, heat shock protein 70, and several other genes (for a review, see Ref. 87). Because the HLA region is highly polymorphic and contains many immune response genes, it was the first candidate genetic region to be studied for association and linkage with AITD.

2. Association of HLA with GD. GD was initially found to be associated with HLA-B8 in Caucasians (88) (Table 3). In these early studies, HLA-B8 was associated with RRs for GD ranging from 1.5–3.5 (89). Subsequently, it was found that GD was more strongly associated with HLA-DR3, which is now
known to be in LD with HLA-B8 (Table 3; reviewed in Ref. 90). The frequency of DR3 was generally 40–55% in GD patients and approximately 15–30% in the general population, giving a RR for people with HLA-DR3 of up to 4.0 (Table 3) (89, 91, 92; reviewed in Ref. 93). A recent family-based study from the United Kingdom using TDT analysis con-

**TABLE 3. Some of the important HLA association studies in GD performed in Caucasians**

<table>
<thead>
<tr>
<th>Country</th>
<th>No. of patients</th>
<th>HLA allele</th>
<th>RR</th>
<th>Ref.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Belgium</td>
<td>194</td>
<td>DRB1*0301</td>
<td>2.53</td>
<td>110</td>
</tr>
<tr>
<td>Canada</td>
<td>175</td>
<td>B8</td>
<td>3.1</td>
<td>89</td>
</tr>
<tr>
<td></td>
<td></td>
<td>DR3</td>
<td>5.7</td>
<td></td>
</tr>
<tr>
<td>Denmark</td>
<td>86</td>
<td>B8</td>
<td>2.80</td>
<td>88</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Dw3</td>
<td>3.94</td>
<td></td>
</tr>
<tr>
<td>France</td>
<td>94</td>
<td>B8</td>
<td>3.4</td>
<td>117</td>
</tr>
<tr>
<td></td>
<td></td>
<td>DR3</td>
<td>4.2</td>
<td></td>
</tr>
<tr>
<td>Germany</td>
<td>253</td>
<td>DR3</td>
<td>2.52</td>
<td>114</td>
</tr>
<tr>
<td>Hungary</td>
<td>256</td>
<td>DR3</td>
<td>3.48</td>
<td>290</td>
</tr>
<tr>
<td>Ireland</td>
<td>86</td>
<td>B8</td>
<td>2.5</td>
<td>116</td>
</tr>
<tr>
<td></td>
<td></td>
<td>DR3</td>
<td>2.6</td>
<td></td>
</tr>
<tr>
<td>Canada</td>
<td>133</td>
<td>DR3</td>
<td>2.63</td>
<td>291</td>
</tr>
<tr>
<td>Sweden</td>
<td>78</td>
<td>B8</td>
<td>4.4</td>
<td>115</td>
</tr>
<tr>
<td></td>
<td></td>
<td>DR3</td>
<td>3.9</td>
<td></td>
</tr>
<tr>
<td>United Kingdom</td>
<td>127</td>
<td>B8</td>
<td>2.77</td>
<td>118</td>
</tr>
<tr>
<td></td>
<td></td>
<td>DR3</td>
<td>2.13</td>
<td></td>
</tr>
<tr>
<td>United Kingdom</td>
<td>65</td>
<td>DR3</td>
<td>3.20</td>
<td>292</td>
</tr>
<tr>
<td>United Kingdom</td>
<td>101</td>
<td>DR3</td>
<td>1.10</td>
<td>293</td>
</tr>
<tr>
<td>United States</td>
<td>65</td>
<td>DR3</td>
<td>3.28</td>
<td>92</td>
</tr>
<tr>
<td>United States</td>
<td>92</td>
<td>DRB1*03</td>
<td>2.6</td>
<td>294</td>
</tr>
<tr>
<td>United States</td>
<td>92</td>
<td>DRB1*08</td>
<td>3.2</td>
<td></td>
</tr>
<tr>
<td>United States</td>
<td>60</td>
<td>DR3</td>
<td>3.4</td>
<td>124</td>
</tr>
<tr>
<td>United Kingdom</td>
<td>120</td>
<td>DQA1*0501</td>
<td>3.8</td>
<td>107</td>
</tr>
<tr>
<td>United Kingdom</td>
<td>228</td>
<td>DRB1*0301</td>
<td>2.7</td>
<td>94</td>
</tr>
<tr>
<td>United States</td>
<td>94</td>
<td>DQA1*0501</td>
<td>3.71</td>
<td>108</td>
</tr>
</tbody>
</table>

**TABLE 4. Major HLA association studies in GD performed in non-Caucasian populations**

<table>
<thead>
<tr>
<th>Country</th>
<th>Ethnic group</th>
<th>No. of patients</th>
<th>HLA allele(s)</th>
<th>RR/P value</th>
<th>Ref.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hong Kong</td>
<td>Chinese</td>
<td>132</td>
<td>Bw46</td>
<td>4.8</td>
<td>295</td>
</tr>
<tr>
<td>Hong Kong</td>
<td>Chinese</td>
<td>67 (children)</td>
<td>DQB1*0303</td>
<td>4.2</td>
<td>296</td>
</tr>
<tr>
<td>Hong Kong</td>
<td>Chinese</td>
<td>97</td>
<td>B46</td>
<td>2.3</td>
<td>101</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>DR9</td>
<td>2.2</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>DQB1*0303</td>
<td>3.2</td>
<td></td>
</tr>
<tr>
<td>Singapore</td>
<td>Chinese</td>
<td>35</td>
<td>B46</td>
<td>8.2</td>
<td>102</td>
</tr>
<tr>
<td>Singapore</td>
<td>Chinese</td>
<td>159</td>
<td>Bw46</td>
<td>4.2</td>
<td>104</td>
</tr>
<tr>
<td>Japan</td>
<td>Japanese</td>
<td>33</td>
<td>Bw35</td>
<td>P &lt; 0.02</td>
<td>95</td>
</tr>
<tr>
<td>Japan</td>
<td>Japanese</td>
<td>106</td>
<td>B46</td>
<td>P &lt; 0.0004</td>
<td>97</td>
</tr>
<tr>
<td>Japan</td>
<td>Japanese</td>
<td>30</td>
<td>DR5</td>
<td>8.14</td>
<td>297</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>DRw8</td>
<td>3.07</td>
<td></td>
</tr>
<tr>
<td>Japan</td>
<td>Japanese</td>
<td>62</td>
<td>DRB1*0803</td>
<td>P &lt; 0.02</td>
<td>98</td>
</tr>
<tr>
<td>Japan</td>
<td>Japanese</td>
<td>76</td>
<td>A2</td>
<td>2.86</td>
<td>298</td>
</tr>
<tr>
<td>Korea</td>
<td>Korean</td>
<td>128</td>
<td>DPB1*0501</td>
<td>5.32</td>
<td>299</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>B13</td>
<td>3.8</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>DR5</td>
<td>4.4</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>DRw8</td>
<td>2.3</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>B8</td>
<td>4.1</td>
<td>300</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Dqw2</td>
<td>5.4</td>
<td></td>
</tr>
<tr>
<td>United States</td>
<td>African-American</td>
<td>73</td>
<td>No association</td>
<td></td>
<td>301</td>
</tr>
<tr>
<td>United States</td>
<td>African-American</td>
<td>49</td>
<td>DRB3*0202</td>
<td>3.6</td>
<td>302</td>
</tr>
<tr>
<td>South Africa</td>
<td>African</td>
<td>103</td>
<td>DR1</td>
<td>3.5</td>
<td>303</td>
</tr>
<tr>
<td>Brazil</td>
<td>Mixed</td>
<td>75</td>
<td>DRB1*0301</td>
<td>2.8</td>
<td>106</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>DQA1*0501</td>
<td>3.74</td>
<td></td>
</tr>
</tbody>
</table>

firmed the results of the case control studies (94). The HLA genes were shown to be associated with GD in non-Caucasians, as well, although the associated alleles were different (Table 4). Studies in the Japanese population have shown associations of GD with HLA-B35 (95, 96). However, other HLA alleles have also been reported to be increased in Japanese GD patients (97–99). In the Chinese population, an increased frequency of HLA-Bw46 has been reported (100–104), and in African-Americans an increased frequency of HLA DRB3*0202 was reported (Table 4) (302). Interestingly, one study of a mixed population in Brazil showed association with HLA-DR3, implying that this allele may confer susceptibility in other ethnic groups and not just Caucasians (106). Alternatively, this Brazilian population was mostly of European ancestry. Among Caucasians, HLA-DQA1*0501 was also shown to be associated with GD (RR, 3.8; Table 3) (Refs. 107–109), but recent studies have suggested that the primary susceptibility allele in GD is indeed HLA-DR3 (HLA-DRB1*03) (110). HLA-DR3 has over 20 subtypes (111). We recently subtyped, by direct sequencing, HLA-DR3 in a population of DR3-positive GD patients and controls (112). The allelic frequency of DRB1*0311 was significantly lower in patients than in controls (112). These results suggested that GD is associated with specific sequences of the DR3 allele. The pattern of transmission of HLA alleles from parents to offspring was also studied. A recent study suggested a preferential transmission of HLA susceptibility alleles from fathers to affected offspring, whereas maternal susceptibility alleles were not transmitted more frequently than expected (113). This may suggest parental imprinting in the transmission of HLA susceptibility alleles to affected offspring.

The role of HLA polymorphisms in the clinical expression of GD has also been explored. Some groups reported an association between the likelihood of relapse of GD and
HLA-DR3, but most other investigators were unable to confirm this observation (114–116). Studies of HLA associations in Graves’ ophthalmopathy (GO) have produced conflicting results, with some workers reporting increased frequency of HLA-DR3 in patients with GO, and others reporting no difference in the distribution of HLA-DR alleles between GD patients with and without ophthalmopathy (88, 89, 117, 118). These results were not surprising in view of our recent segregation analysis that showed no genetic influences on the development of GO (41). Likewise, no difference in the DR3 frequency was found in GD patients with and without pre-tibial myxedema (89). Some workers have suggested that local factors such as orbital pressure play an important role in the development of GO and pre-tibial myxedema (119).

3. Association of HLA with HT. Data on HLA haplotypes in HT have been less definitive than in GD. A general methodological problem has been disease definition. Although the diagnosis of GD may be relatively straightforward, the definition of HT has been more controversial. HT encompasses a spectrum of manifestations, ranging from the simple presence of thyroid autoantibodies with focal lymphocytic infiltration, which may be of no functional consequence (asymptomatic autoimmune thyroiditis), to the presence of goitrous or atrophic thyroiditis, characterized by gross thyroid failure (31). Initial studies failed to demonstrate an association between goitrous HT and HLA A-, B-, or C-antigens (120). Later studies showed an association of goitrous HT with HLA DR5 (RR, 3.1) (121) and of atrophic HT with DR3 (RR, 5.1) (122) (Table 5). Association of HT with HLA DR3 in Caucasians has been confirmed in subsequent studies (123, 124) and further supported by studies of transgenic mice (125). An association between HT and HLA-DQw7 (DQB1*0301) has also been reported in Caucasians (126, 127). In non-Caucasian ethnic groups, different HLA haplotypes were reported to be associated with HT, e.g., HLA-DRw53 in Japanese (128) and HLA-DR9 in Chinese (129) (Table 5).

4. HLA linkage studies in AITD. Linkage studies of HLA in AITD have been largely negative (130–134). Only one recent study from the United Kingdom showed weak evidence for linkage between GO and the HLA region (135), and an additional study reported linkage only when conditioning on DR3 (136). It is difficult to explain why the HLA genes show consistent association with GD but no evidence for linkage. The lack of linkage means that HLA-DR3, as measured, did not cause the familial segregation of GD, whereas the relatively strong association showed that HLA-DR3 conferred a generalized increase in risk for GD in the general population. Indeed, we were able to show that HLA was associated with GD in both sporadic GD patients and probands from GD families, giving similar RRs (our unpublished data).

B. The role of CTLA-4 in the genetic susceptibility to AITD

1. The cytotoxic T lymphocyte antigen-4 (CTLA-4) costimulatory molecule. CTLA-4 is an important costimulatory molecule that participates in the interaction between T cells and antigen-presenting cells (APC). APC activate T cells by presenting to the T cell receptor an antigenic peptide bound to an HLA class II protein on the cell surface. However, a second signal is also required for T cell activation, and these costimulatory signals may be provided by the APC themselves or other local cells (137). The costimulatory signals are provided by a variety of proteins that are expressed on APC (e.g., B7-1, B7-2, B7h, CD40) and interact with receptors (CD28, CTLA-4, and CD40L) on the surface of CD4+ T lymphocytes during antigen presentation (137) (Fig. 3). Whereas the binding of B7 to CD28 on T cells costimulates T cell activation, the presence of CTLA-4, which has a higher affinity for B7, down-regulates T cell activation by competing for the binding of B7 to CD28. The suppressive effects of CTLA-4 on T cell activation have raised the possibility that mutations altering CTLA-4 expression and/or function could result in an exaggerated T cell activation and lead to the development of autoimmunity.

2. CTLA-4 association studies in GD. Recently, there have been several reports demonstrating an association between the CTLA-4 gene and theAITDs (Table 6) (138–142). The initial studies found an association between a microsatellite marker located at the 3′ untranslated region (UTR) of the CTLA-4 gene and GD, giving a RR of 2.1–2.8 (138, 141). Later, two SNPs were also identified in the CTLA-4 gene: 1) at position 49 in the CTLA-4 leader peptide (A/G49), resulting in an alanine/threonine polymorphism; and 2) in the promoter of
CTLA-4 at position −318 (C/T−318). Case-control studies from several groups, including our own group, have shown an association between the alanine (G) polymorphism and GD with a RR of approximately 2.0 (Table 6) (41, 140, 143–145). The association between GD and the CTLA-4 3’ UTR microsatellite and A/G49 SNP has been consistent across populations of different ethnic backgrounds such as Caucasians (138), Japanese (145, 146), and Koreans (147). The association of CTLA-4 and GD has also been confirmed in a family-based study using TDT analysis (148). In contrast, association studies using the C/T−318 SNP of CTLA-4 have been less consistent, with some showing association (144) and others not (149).

3. CTLA-4 association studies in HT. Because CTLA-4 is a nonspecific costimulatory molecule, it is expected to confer susceptibility to AITD and autoimmunity in general and not specifically to GD (150). Indeed, CTLA-4 was reported to be associated and linked with all forms of AITD (GD, HT, and TAb; see Section IV.B.4), and with many autoimmune diseases such as type 1 diabetes (139, 140, 151), Addison’s disease (152), and myasthenia gravis (153). CTLA-4 has been reported to be associated with HT in various populations including Caucasians (141, 143, 154) and Japanese (146, 155). There have also been two reports of no association of HT with CTLA-4, most likely due to lack of power (144, 156).

4. CTLA-4 and TAb. Two studies have now shown that CTLA-4 confers susceptibility to the production of thyroid antibodies. Our group has shown strong evidence for linkage between the CTLA-4 gene region and the production of thyroid antibodies with a MLS of 4.2 (157). Recently, another report has described an association between the G allele of the CTLA-4 A/G49 SNP and the thyroid autoantibody diagnosis (158). Because the development of thyroid antibodies often represents the preclinical stage of AITD (159), it is possible that CTLA-4 predisposes, nonspecifically, to the development of thyroid autoimmunity. Additional genetic and/or environmental factors must be necessary for the development of the specific GD/HT phenotypes (150).

5. CTLA-4 and disease severity. Several studies have examined whether CTLA-4 polymorphisms influence disease severity. Heward et al. (148) reported that the CTLA-4 A/G49 SNP G allele was associated with more severe thyrotoxicosis at diagnosis (as reflected by higher free T4 levels). Similar findings were reported by Park et al. (147) but not by Zaletel et al. (158). It has also been reported that the frequency of the G allele and the GG genotype of the CTLA-4 A/G49 SNP was significantly higher in GD patients who did not go into remission after 5 yr on antithyroid medications (160). In addition, CTLA-4 has been shown to be associated with GD in children (161).

![Fig. 3. T cell activation by APC. The APC presents a peptide antigen bound to HLA class II molecules, and the peptide is recognized by the T cell receptor. This interaction will not lead to T cell activation without additional costimulation provided by costimulatory molecules. Engagement of B7 molecules with CD28 provides costimulation, whereas CTLA-4 binding to the B7 molecules blocks CD28 activation by B7. In addition, CTLA-4 engagement with B7 directly suppresses T cell activation. Additional costimulation is provided by activation of inducible costimulator by its ligand (B7 h), and by CD40-CD40-ligand (CD40L or CD154) interaction. CD40 is also expressed on B cells, and its activation results in B cell differentiation, Ig production, and isotype switching.](image)

Table 6. Major CTLA-4 association studies in AITD

<table>
<thead>
<tr>
<th>CTLA-4 polymorphism</th>
<th>Country</th>
<th>Ethnic group</th>
<th>Disease</th>
<th>No. of patients</th>
<th>RR/P value</th>
<th>Ref.</th>
</tr>
</thead>
<tbody>
<tr>
<td>CTLA-4(−318)</td>
<td>United Kingdom</td>
<td>Caucasians</td>
<td>GD</td>
<td>135</td>
<td>2.82</td>
<td>138</td>
</tr>
<tr>
<td>CTLA-4(−318)</td>
<td>United Kingdom</td>
<td>Caucasians</td>
<td>GD</td>
<td>132</td>
<td>2.1</td>
<td>141</td>
</tr>
<tr>
<td>CTLA-4(AT)</td>
<td>Hong Kong</td>
<td>Chinese</td>
<td>GD</td>
<td>94</td>
<td>2.0</td>
<td>140</td>
</tr>
<tr>
<td>CTLA-4(AT)</td>
<td>Japan</td>
<td>Japanese</td>
<td>GD</td>
<td>92</td>
<td>4.2</td>
<td>139</td>
</tr>
<tr>
<td>CTLA-4(AT)</td>
<td>Japan</td>
<td>Japanese</td>
<td>GD</td>
<td>92</td>
<td>4.2</td>
<td>139</td>
</tr>
<tr>
<td>CTLA-4(AT)</td>
<td>United Kingdom</td>
<td>Caucasians</td>
<td>GD</td>
<td>92</td>
<td>4.2</td>
<td>139</td>
</tr>
<tr>
<td>Thr/Ala (AT)</td>
<td>United Kingdom</td>
<td>Caucasians</td>
<td>GD</td>
<td>305</td>
<td>2.0</td>
<td>140</td>
</tr>
<tr>
<td>Thr/Ala (AT)</td>
<td>United Kingdom</td>
<td>Caucasians</td>
<td>GD</td>
<td>379</td>
<td>2.0</td>
<td>140</td>
</tr>
<tr>
<td>Thr/Ala (AT)</td>
<td>United Kingdom</td>
<td>Caucasians</td>
<td>GD</td>
<td>484</td>
<td>2.0</td>
<td>140</td>
</tr>
<tr>
<td>Thr/Ala (AT)</td>
<td>United States</td>
<td>Caucasians</td>
<td>GD</td>
<td>85</td>
<td>1.6</td>
<td>140</td>
</tr>
<tr>
<td>Thr/Ala (AT)</td>
<td>United States</td>
<td>Caucasians</td>
<td>GD</td>
<td>73</td>
<td>1.6</td>
<td>140</td>
</tr>
<tr>
<td>Thr/Ala (AT)</td>
<td>Italy</td>
<td>Caucasians</td>
<td>GD</td>
<td>126</td>
<td>2.0</td>
<td>140</td>
</tr>
<tr>
<td>Thr/Ala (AT)</td>
<td>United Kingdom</td>
<td>Caucasians</td>
<td>HT</td>
<td>158</td>
<td>2.0</td>
<td>140</td>
</tr>
<tr>
<td>Thr/Ala (AT)</td>
<td>Slovenia</td>
<td>Caucasians</td>
<td>TAb’s</td>
<td>67</td>
<td>2.0</td>
<td>140</td>
</tr>
<tr>
<td>Thr/Ala (AT)</td>
<td>Japan</td>
<td>Japanese</td>
<td>GD</td>
<td>153</td>
<td>2.0</td>
<td>140</td>
</tr>
<tr>
<td>Thr/Ala (AT)</td>
<td>Korea</td>
<td>Korean</td>
<td>GD</td>
<td>97</td>
<td>1.6</td>
<td>140</td>
</tr>
<tr>
<td>C/T (−318)</td>
<td>United Kingdom</td>
<td>Caucasians</td>
<td>GD</td>
<td>188</td>
<td>2.0</td>
<td>140</td>
</tr>
<tr>
<td>C/T (−318)</td>
<td>United Kingdom</td>
<td>Caucasians</td>
<td>GD</td>
<td>90</td>
<td>2.0</td>
<td>140</td>
</tr>
<tr>
<td>C/T (−318)</td>
<td>Hong Kong</td>
<td>Chinese</td>
<td>GD</td>
<td>98</td>
<td>2.0</td>
<td>140</td>
</tr>
<tr>
<td>C/T (−318)</td>
<td>Germany</td>
<td>Caucasians</td>
<td>GD</td>
<td>125</td>
<td>2.0</td>
<td>140</td>
</tr>
</tbody>
</table>

NS, Not significant.
Together, these studies may suggest that CTLA-4 may influence both the initiation of AITD and the severity of the phenotype.

CTLA-4 polymorphisms have also been tested for association with GO. Most studies have been negative and did not show that CTLA-4 conferred a specific risk for GO beyond that conferred for GD (41, 158, 162). However, two groups have reported an association between GO and CTLA4 (147, 163). It is most likely that the reported CTLA-4 associations with GO reflected an association between CTLA-4 and more severe GD and not a specific association with the GO phenotype. Indeed, our recent segregation analysis showed no evidence for any genetic susceptibility to GO (41).

6. Family linkage studies. Vaidya et al. (135) reported linkage to the CTLA-4 gene region on chromosome 2q33 in families with GD using nonparametric linkage analysis. The linkage became stronger when families with AITD, rather than just GD, were included in the study, again demonstrating that CTLA-4 most likely confers general susceptibility to thyroid autoimmunity and not to a specific AITD phenotype. As discussed earlier, and in keeping with the view that the CTLA-4 gene predisposes to thyroid autoimmunity rather than to one specific disease, we found strong linkage between the CTLA-4 gene region and TAb (157). However, the region on chromosome 2q33 containing the CTLA-4 gene is replete with candidate immune regulatory genes for thyroid autoimmunity (e.g., CD28 and inducible costimulator), and it was unclear whether the CTLA-4 gene itself or another immune regulatory gene in the region was involved in the genetic susceptibility to AITD. Recently, we and others tested additional genes and markers in the 2q33 region, and the strongest association was with the CTLA-4 markers (151, 157). These results were in keeping with results obtained in type 1 DM (164, 165).

C. Other immune-related genes

1. T cell receptor genes. The initiation of GD involves infiltration of the thyroid by T cells reactive with thyroid antigens. These T cells are oligoclonal and are restricted by their T cell receptor V gene use (166). Thus, the T cell receptor seemed a likely candidate gene for thyroid autoimmunity and has, therefore, been studied extensively. Earlier studies reported an association between a restriction fragment length polymorphism of the T cell receptor β-chain and GD (167). However, these results were not confirmed by other groups (92, 168). We have also studied the T cell receptor α-genes (on chromosome 14) and β-genes (on chromosome 7) by linkage studies using microsatellite markers located at or close to these genes (133). These markers gave highly negative LOD scores with GD, HT, and AITD, thus ruling out linkage of the T cell receptor α- and β-genes to AITD (133). Therefore, it can be concluded that the T cell receptor genes are not major susceptibility genes for AITD.

2. The IgG heavy chain (IgH) gene. Another good candidate immune regulatory gene for GD was the IgH gene, because the hallmark of GD is the production of stimulating TSHR IgG autoantibodies. Indeed, the IgH gene was one of the first genes studied in AITD. Early studies reported an association between IgH Gm allotypes and AITD in the Japanese population (169–171). However, these results have not been reproduced in other populations (172, 173). Similarly, we were unable to confirm linkage between the IgH gene and GD in a Caucasian dataset of families (174).

3. Immune regulatory genes. Other genes tested for association with GD include the IL-1 receptor antagonist gene (175–178), TNFα gene (133, 179), interferon-γ gene (180), and the transporters associated with antigen presentation genes (181). However, none of these produced replicable associations with GD. Recently, an association between a promoter polymorphism of the IL-4 gene and GD has been reported, but a subsequent study could not replicate these results (182). The vitamin D receptor, which may have some immune modulatory functions, has also been reported to be associated with GD (183), and another study reported an association of GD with a vitamin D binding protein (184). These results need to be confirmed, and it cannot be excluded that other genes in LD with these genes are the susceptibility genes at these loci.

V. Thyroid-Specific Genes Studied in AITD

A. TSH receptor

The hallmark of GD is the production of the TSHR antibodies. Therefore, the TSHR gene was long thought to be a likely candidate gene for GD. Three common germline SNPs of the TSHR have been described (185). Two of them are located in the extracellular domain of the TSHR: an aspartic acid to histidine substitution at position 36 (D36H), and a proline to threonine substitution at position 52 (P52T). The third SNP is a substitution of glutamic acid for aspartic acid (D727E) within the intracellular domain of the receptor. Most studies on the contribution of the TSHR gene to the genetic susceptibility to GD have focused on the SNPs in the extracellular domain of the TSHR (186–191) because this domain is responsible for TSH and TSHR antibody binding. Amino acid changes in the extracellular domain of the TSHR could theoretically change the amino acid sequence of TSHR T cell epitopes (192). Initial studies suggested that the P52T SNP was associated with GD in females (186). However, other authors were unable to confirm the association between the P52T SNP and GD in Caucasians (187–191). The D36H SNP has also been reported not to be associated with GD (189). Linkage studies in GD families using three microsatellite markers within introns 2 and 7 of the TSHR gene have also been negative in Caucasians (193, 194). Association studies using these markers also have been negative in Caucasians (41). However, an association between AITD and TSHR microsatellite markers has been reported in the Japanese (146, 155), and recently, the D727E SNP was reported to be associated with GD in a Caucasian Russian population (195), but these results were not replicated in a subsequent study by another group (196). We recently tested whether the D727E SNP was associated with GD in general and with disease severity. The results of our study did not show an association between the D727E SNP and GD and did not show an effect of the D727E SNP on the GD phenotype or disease severity (197). The frequency of the G allele was not increased in
patients with more severe forms of GD (i.e., ophthalmopathy and goiter) and in patients with early disease onset. However, our study and other negative TSHR studies could not exclude a weak association between GD and the TSHR gene because very large datasets are needed to detect associations with low RRs. We, therefore, performed a meta-analysis combining our data with the data reported in the previous two negative TSHR studies. The results of the meta-analysis showed a weak association between the D7Z7E SNP E allele and GD ($P = 0.03$; $RR = 1.6$) (197). Therefore, it remains possible that the TSHR is a minor susceptibility gene for GD.

B. Thyroid peroxidase

Antibodies to TPO are the most specific marker for HT, and therefore, TPO is another possible candidate gene for AITD. It was recently found that autoantibodies recognizing immunodominant epitopes of TPO are genetically transmitted within families (50). One possible explanation of the genetic transmission of TPO-Ab epitopes is that this is caused by changes in TPO gene sequence and/or expression in the thyroid. The TPO gene was tested for linkage and association with AITD in two studies using a microsatellite inside the TPO gene. However, these studies showed no evidence of linkage and/or association of the TPO gene with AITD (174, 198). Therefore, the TPO gene does not seem to be a major susceptibility gene for AITD, although a minor role cannot be excluded.

C. Thyroglobulin

1. Tg plays a central role in the development of thyroid autoimmunity. Tg is the major protein product synthesized in the thyroid gland, and Tg autoantibodies are common in AITD (reviewed in Ref. 199). There is abundant, indirect, evidence that Tg plays an important role in the etiology of AITD including: 1) anti-Tg-Ab are detected in almost all patients with AITD, both GD and HT (200, 201), and there is evidence that Tg-Ab of AITD patients are restricted in their epitope specificity in contrast to the polyclonal nature of Tg-Ab found in healthy individuals (202); 2) immunization with Tg induces autoimmune thyroiditis in experimental animals (203), and the induction of experimental autoimmune thyroiditis (EAT) in mice by Tg is well known to be MHC-dependent (204, 205); this last observation implies an interaction between the Tg molecule and the MHC glycoproteins in the induction of thyroiditis; 3) Tg peptides containing the hormonogenic sites at positions 5, 2553, and 2567 (206–208), by changes in TPO gene sequence and/or expression in the thyroid. The TPO gene was tested for linkage and association with AITD in two studies using a microsatellite inside the TPO gene. However, these studies showed no evidence of linkage and/or association of the TPO gene with AITD (174, 198). Therefore, the TPO gene does not seem to be a major susceptibility gene for AITD, although a minor role cannot be excluded.

VI. Linkage Studies of Candidate Chromosomal Regions (Table 7)

A. Chromosome X

The AITDs are 5–10 times more common in women (159). The increased female preponderance of AITD may potentially be explained by the effects of estrogenic sex steroids in promoting autoimmunity (215), by genetic factors, or as a consequence of pregnancy and the resulting maternal microchimerism (216). Although there are studies supporting a role for estrogen in the induction of autoimmunity (215, 217–220) these results have not been consistent, and some studies have even shown suppression of thyroiditis (221) and other autoimmune diseases (222) by estrogens. However, there are also indirect data suggesting that chromosome X abnormalities might be responsible for the increased incidence of AITD in women. These data come mostly from studies of patients with Turner’s syndrome. There is a strong association of Turner’s syndrome with the production of thyroid autoantibodies and autoimmune thyroiditis (223–225). Up to 50% of patients with Turner’s syndrome develop thyroid antibodies in early childhood (224, 225), and up to 20% develop clinical disease (223, 225). Studies on the correlation between the karyotype and AITD have shown that 83% of patients with Turner’s syndrome who had X-isochromosomes developed TAb, and 57% developed clinical AITD (223, 226). These results suggested that a gene on chromosome Xq may play a role in the development of AITD (226). Two studies have subsequently examined the X chromosome for linkage with AITD. We showed evidence for
linkage at chromosome Xq21 (227), but this locus was not confirmed in our second dataset (228). Another study showed evidence for linkage at Xp11 with a maximum non-parametric LOD score of 2.2 (229) (Table 7). More studies of the X chromosome are needed to confirm whether a gene on the X chromosome confers susceptibility to AITD.

There is a number of possible mechanisms whereby the X chromosome could influence the development of AITD. Females have two X chromosomes (one paternal and one maternal), whereas males have only one X chromosome (maternal). Therefore, females are twice as likely to inherit an X chromosome AITD susceptibility gene as males. Several immune regulatory genes are located on the X chromosome (e.g., the CD40 ligand gene), but their involvement in autoimmunity has not been studied. Another possible mechanism is through X-inactivation. X-inactivation in females results in the production of two classes of cells that differ in the transcription of X chromosome-encoded genes, including genes coding for self-antigens. If these two cell classes extend to the thymic cells responsible for tolerizing T cells in embryonic life, the immune repertoire will not be entirely tolerized to one version of the two self-antigens encoded by the X chromosome. Such lymphocytes would be autoreactive to that antigen and could induce an autoimmune response (230). Indeed, some workers have suggested that skewed X-chromosome inactivation in the thymus may lead to inadequate thymic deletion and autoimmunity (231). Although this is an attractive theory that could help explain the female preponderance of autoimmune conditions (because this escape mechanism from tolerance can occur only in females), there are no data to support it.

B. The IDDM loci

Type 1 diabetes [insulin-dependent DM (IDDM)] is known to be associated with AITD (232–235; reviewed in Ref. 236). Up to 20% of patients with IDDM have thyroid antibodies (237), and 50% of IDDM patients with TAb will develop clinical AITD (238). In addition, postpartum thyroiditis is twice as common in women with IDDM (239–241). Therefore, it is possible that AITD and IDDM share genetic susceptibility. Moreover, a recent study has shown that 35.9% of IDDM probands from the Familial Autoimmune Diabetes Study had AITD (26.6% had HT and 9.3% had GD) (242). These, results implied that the association between IDDM and AITD is even stronger in cases of familial IDDM and may point to common susceptibility genes. To date, more than 15 type 1 diabetes loci have been suggested (59, 85, 139, 243–245), and some of these sites have also been evaluated in patients with AITD. One locus that has been found to be shared by type 1 diabetes and AITD is the CTLA-4 gene region discussed earlier. This genetic region is linked with both type 1 diabetes (139) and AITD (135). Moreover, in both type 1 diabetes and AITD it was shown that the CTLA-4 gene is most likely the susceptibility gene at this locus (157, 164, 165), although the specific polymorphism of CTLA-4 responsible for this association is not known. Other IDDM loci have also been tested in AITD. IDDM-2 on chromosome 11p, IDDM-4 on chromosome 11q, IDDM-5 on chromosome 6q25, and IDDM-8 on chromosome 6q27 did not show evidence for linkage and/or association with AITD (174, 246, 247). On the other hand, a study from the United Kingdom showed evidence for linkage of IDDM-6 on chromosome 18q with AITD (247) (Table 7), although we did not find this site linked to AITD in our own whole genome screening.

C. The 14q region

The chromosome 14q region caught the attention of several investigators because it harbors the TSHR gene, and therefore, it has been tested for linkage with AITD. In our preliminary screen of the 14q region using 56 multiplex families, we obtained evidence for linkage with GD at this locus (designated GD-1) with an MLS of 2.5 (174, 248). This locus was distinct from the TSHR gene, which did not show strong evidence for linkage with GD (248). We have recently expanded our dataset to include 102 families (540 individuals) and were able to replicate the evidence for linkage at the 14q region (228). Imrie et al. (229) also tested the 14q region and reported no evidence for linkage at GD-1. However, they tested only a few markers in a very narrow region that did not include the area of MLS obtained in our expanded data.
set. Thus, additional linkage studies are needed to confirm the evidence for linkage at this locus. The GD-I locus contains several potential positional candidate genes. One of them is a newly isolated growth factor, suppressor of lin-12-like protein (249), which was tested, but no association was found with GD (250). Another interesting positional candidate gene in this locus is the TGFβ3 gene.

D. The 20q region

Chromosome 20 was one of the first chromosomes tested in our preliminary genome scan. Because it showed strong evidence for linkage at 20q, this region was studied in more detail. Using a dataset of 56 multiplex families, we identified a locus on chromosome 20q11 showing strong evidence for linkage with GD with a MLS of 3.5 (79, 193). This GD locus was not linked to HT, because analysis of the data for the HT families gave strongly negative LOD scores. Moreover, in families with GD- and HT-affected individuals, the locus was linked only with GD, demonstrating its high specificity for GD. This region was recently tested by another group from the United Kingdom. Although this locus did not show evidence for linkage in the whole United Kingdom dataset, subsetting the families by transmission showed evidence for linkage with GD in the families with vertical transmission (consistent with dominant complex inheritance) of disease (251). This further supported the 20q11 locus as an important susceptibility locus for GD.

Recently, we have expanded our dataset to 102 multiplex families and used this expanded dataset to fine-map the 20q11 locus using 10 densely spaced microsatellite markers. Linkage analysis in our expanded dataset gave a MLS of 1.2 with heterogeneity. However, a subset of the families (Caucasian families of non-Italian origin) gave a MLS (without heterogeneity) of 3.3 (252). Thus, the 20q region was linked with a subset of our dataset (252) and a subset of the United Kingdom dataset (251). It is still unclear what is unique to these subsets, but they may represent a different phenotype of GD. The CD40 gene, an important immune modulator, is located within the linked region on chromosome 20q11, and therefore, it was a likely positional candidate gene for GD.

CD40 is a transmembrane glycoprotein that is expressed predominantly on B cells, but also on monocytes, dendritic cells, epithelial cells, and other cells (reviewed in Ref. 253). It is a member of the TNF receptor superfamily, and it binds to a ligand (CD40L or CD 154) that is expressed mainly on activated T cells. Binding of CD40L to CD40 induces B cells to proliferate and undergo Ig isotype switching (254). CD40 activated T cells. Binding of CD40L to CD40 induces B cells to proliferate and undergo Ig isotype switching (254). CD40 activated T cells. CD40 is a transmembrane glycoprotein that is expressed predominantly on B cells, but also on monocytes, dendritic cells, epithelial cells, and other cells (reviewed in Ref. 253). It is a member of the TNF receptor superfamily, and it binds to a ligand (CD40L or CD 154) that is expressed mainly on activated T cells. Binding of CD40L to CD40 induces B cells to proliferate and undergo Ig isotype switching (254). CD40 has shown to play an essential role in the regulation of humoral immunity, central and peripheral T cell tolerance, and APC function (reviewed in Ref. 255). Moreover, in vivo blockade of CD40 has shown to suppress the induction of EAT (256). Therefore, we tested whether CD40 was the GD susceptibility gene on chromosome 20q11. Sequencing of the CD40 gene revealed a C/T SNP in the Kozak sequence of the gene. Analysis of the CD40 Kozak SNP in 154 Caucasian GD patients and 118 Caucasian controls showed an association between the CC genotype and GD but with a low RR of 1.6 (252). TDT analysis also showed preferential transmission of the C allele of the CD40 Kozak SNP to affected individuals (252). However, it is possible that other polymorphisms in the CD40 gene or another gene in LD with CD40 is the GD susceptibility gene in this region.

VII. Whole Genome Screening in AITD

Recently, two whole genome scans were reported in AITD, one in 56 multigenerational Caucasian families (193) and another in 123 Japanese sib-pair families (134). In both studies, the linkage was analyzed for three different phenotypes: GD, HT, and AITD (i.e., GD+HT), and in both studies many loci gave low positive LOD scores (LOD scores >1.0 and <2.0), and only a few loci gave significant LOD scores (>2.0).

The Japanese whole genome screen identified two loci giving strong evidence for linkage (i.e., MLS > 2.0). One locus on chromosome 5q31 showed evidence for linkage with AITD with an MLS of 3.14, and a second locus on chromosome 8q24 showed evidence for linkage with both AITD (MLS = 2.31) and HT (MLS = 3.77) (134). The identification of the Tg gene as the susceptibility gene in this locus has been discussed earlier (see Section V.C). The 5q31 locus contains a cytokine gene cluster, and therefore, several positional candidate genes exist in this locus and need to be examined.

Our genome screen in the Caucasian families was recently expanded to include 102 multigenerational families (540 individuals) (228). In the expanded dataset seven loci were identified, four of them loci that were identified in the original dataset of 56 families and replicated in the expanded dataset (228). Three loci on chromosomes 6p, 8q, and 10q showed evidence for linkage with both GD and HT (MLS = 2.0, 3.5, and 4.1, respectively); three loci showed evidence for linkage with GD, i.e., on 7q (MLS = 2.3), 14q (MLS = 2.1), and 20q (MLS = 3.3); and one locus on 12q showed evidence for linkage with HT (MLS = 3.4) (228). Four of the loci (the AITD locus on 6p, the GD loci on 14q and 20q, and the HT locus on 12q) were identified in the original dataset and replicated in the expanded dataset (228). These four loci represent strong candidates for harboring susceptibility genes for AITD.

Two additional whole genome screens, performed in two single large pedigrees, were also recently reported. One study was performed in a large family with multiple members affected with vitiligo and HT (307). The study identified an autoimmunity locus on 1p31–32 and an additional HT locus that was mapped to the same position as the 6p locus identified in our whole genome screen (designated AITD-1) (Table 7). The second study was performed in a Chinese-American family and identified two loci on chromosomes 9 and 11 that were different from loci reported in Caucasian families (257). Another very recent whole genome scan showed evidence for linkage with 5q31 (304) replicating the results of the Japanese study (134).

VIII. Mechanisms by Which Genes Can Induce Thyroid Autoimmunity

A. General principles

In classical monogenic diseases, the genetic defect changes the action of a gene by decreasing its effectiveness [e.g., the
Pendrin gene in Pendred syndrome (58)] or causing overactivity in the gene [e.g., the RET protooncogene in multiple endocrine neoplasia 2 (258, 259)]. However, in multifactorial diseases such as the AITD, the situation is more complex. The genetic defect may cause subtle changes in the function of one or more genes that when combined increase the likelihood of an individual to develop the disease. Therefore, even when a gene causing a common disease is mapped, proving that the polymorphism is also biologically meaningful can be difficult. But only after it is demonstrated that alterations in the functions of this gene are involved in the pathogenesis of the disease can the gene be declared conclusively a susceptibility gene for the disease.

Conclusive evidence for genes that confer susceptibility to complex diseases has been found. The two best examples for genes that have been shown to predispose to complex diseases are the HLA-DR and DQ genes in type 1 DM (for review, see Ref. 260) and the NOD2 gene in Crohn’s disease (60, 84). Recently, some functional studies have been published in AITD, although the results are not yet conclusive (261, 262).

### B. HLA

The mechanisms by which HLA molecules confer susceptibility to autoimmune diseases are now beginning to be understood. T cells recognize and respond to an antigen by interacting with a complex between an antigenic peptide and an HLA molecule (reviewed in Ref. 263). It is thought that different HLA alleles have different affinities for peptides from autoantigens (e.g., thyroid antigens) that are recognized by T cell receptors on cells that have escaped tolerance (87). Thus, certain alleles may permit the autoantigenic peptide to fit into the antigen binding groove inside the HLA molecule and to be recognized by the T cell receptor, whereas others may not (264). This would determine whether an autoimmune response to that antigen will develop.

Studies on the structure of HLA polymorphisms associated with type 1 DM provided strong evidence in support of this hypothesis. Sequencing of the HLA DQ genes showed that an aspartic residue at position 57 of the DQβ chain played a key role in the genetic susceptibility to type 1 diabetes (260). Individuals who did not have Asp (57) on both of their DR alleles were at high risk for type 1 diabetes (RR > 50) (265). Moreover, it has been shown that an aspartic acid at position 57 on the DQβ chain influences the antigen binding properties of the HLA-DQαβ heterodimer (260, 266). Lack of aspartic acid at position 57 on the DQβ chain permitted immunogenic insulin peptides to fit into the antigen binding groove inside the HLA molecule and to be recognized by the T cell receptor (267, 268). In contrast, the presence of aspartic acid at position 57 of the DQβ chain prevented insulin peptides from fitting and hence prevented autoantigen presentation to the T cell receptor (264).

GD is associated with HLA-DR3 (Table 3). We sequenced the DRβ3 chain to identify critical amino acids for the susceptibility to GD (112). Analysis of the frequencies of DR3-specific amino acids occupying the peptide binding pockets in GD patients and controls showed that lack of arginine at position 74 of the DRβ chain (DRβ 74Arg) was significantly more frequent in the DR3-positive controls than in the DR3-positive patients (112). Although some HLA-DR binding studies have shown higher affinity of HLA-DR3 to TSHR immunodominant peptides than to TSHR nonimmunodominant peptides, this approach has not been applied to the different DR3 subtypes associated with GD. Nevertheless, such data have suggested that certain DR sequences influence the binding and presentation of TSHR peptides (269), and this may provide a mechanism by which DRβ 74Arg influences the susceptibility to GD (Fig. 4). These results, if confirmed, may indicate a general principle in HLA-induced susceptibility to autoimmune diseases as seen in type 1 DM.

For thyroid autoantigens to be presented by HLA molecules to T cells, a mechanism of autoantigen presentation must exist within the thyroid gland or the draining lymph nodes of the gland. One potential intrathyroidal mechanism not using professional APC may be through aberrant expression of HLA class II molecules on thyrocytes (270–272). Unlike in normal thyroids, the thyroid epithelial cells from patients with GD and HT have been shown to express HLA class II antigen molecules similar to those normally expressed on APC such as macrophages and dendritic cells (270, 273) (reviewed in Ref. 274). This aberrant expression of HLA class II molecules on thyroid cells may initiate thyroid autoimmunity via direct thyroid autoantigen presentation (270, 275) or a secondary event after cytokine secretion by invading T cells. Consistent with the former possibility was the fact that thyroid cell MHC class II antigen expression could be induced by certain viral infections in vitro (276, 277) and that mice constitutively expressing thyroid cell MHC class II antigens developed thyroiditis after immunization with human Tg (125). Furthermore, a murine model of GD has been shown to depend on TSHR antigen presentation on cells expressing MHC class II molecules (278, 279). Coculture of peripheral blood mononuclear cell from GD patients with homologous thyrocytes induced T cell activation (280) as well as interferon-γ production and thyroid cell HLA class II antigen expression (281). Such cytokine secretion may be the common cause of HLA class II antigen expression by thyroid cells in AITD (271, 282; reviewed in Ref. 283).

### C. CTLA-4

The CTLA-4 gene polymorphisms have also been studied for their effects on CTLA-4 function. CTLA-4 is an important costimulatory molecule that participates in the presentation of peptides to T cells. APC activate T cells by presenting to the T cell receptor an antigenic peptide bound to an HLA class II antigen molecule on the cell surface. However, a second signal is also required for T cell activation, and these costimulatory signals may be provided by the APC themselves or other local cells (137). The costimulatory signals are provided by a variety of proteins (e.g., B7-1, B7-2, CD40) that are expressed on APC and interact with receptors (CD80, CTLA-4, and CD40L) on the surface of CD4+ T lymphocytes during antigen presentation (137). Whereas the binding of B7 to CD28 on T cells costimulates T cell activation, the higher affinity binding of B7 to CTLA-4 down-regulates T cell activation and induces tolerance (Fig. 3). The suppressive effects of CTLA-4 on T cell activation have raised the possibility that the
Diseases.

microsatellite and AITD, as well as other autoimmune

lation for the association between the short alleles of the AT

mRNA (284, 285). This could provide an attractive explana-

of the CTLA-4 gene influenced the half-life of the CTLA-4

thenia gravis showed that the AT microsatellite at the 3

A/G49 polymorphism. Indeed, preliminary data in myas-

another CTLA-4 polymorphism that was in LD with the

alleles. However, the differences they found could be due to

tion of T cells taken from individuals with the A and the G

Arginine at position 74 of the DR

TSHR-specific T cells. In contrast, HLA-DR molecules lacking the

sequences. HLA-DR molecules containing Arginine at position 74 of

/ H9252

/ H11032

74Arg) form a peptide binding pocket that

/ H11032

/ H9252

environmental factor [e.g., infection, dietary factors (iodine),

vironmental triggers (Fig. 5). There are sufficient epide-

miological data to support an important genetic contribution

to the development of AITD, and in the past few years several

loci and genes have shown evidence for linkage and/or

association with AITD. Thus, the genetic susceptibility to

AITD seems to involve several genes with varying effects.

With the completion of the Human Genome Project and the

establishment of large SNP databases, the identification of

additional AITD susceptibility genes will become more fea-

sible. One approach that was proposed is to perform whole

genome association studies using several thousand SNPs by

pooling DNA samples (72).

The AITD loci identified so far show that some putative

ence.

FIG. 4. Possible mechanism of induction of GD by specific HLA-DR

sequences. HLA-DR molecules containing Arginine at position 74 of

the DRβ 1 chain (DRβ 74Arg) form a peptide binding pocket that

enables presentation of TSHR immunogenic peptides which stimulate

TSHR-specific T cells. In contrast, HLA-DR molecules lacking the

Arginine at position 74 of the DRβ 1 chain cannot fit the TSHR

immunogenic peptides, and therefore TSHR-specific T cells are not

stimulated.

FIG. 5. A proposed mechanism for the development of AITD. Indi-

viduals who are born with susceptibility genes are predisposed to

develop AITD. The Tg gene may predispose to AITD in a number of

ways, for example: 1) sequence changes in Tg may change its

antigenicity, making certain Tg peptides more immunogenic;

2) sequence changes in Tg may change its interaction with

HLA molecules; and 3) sequence changes in or near the Tg

gene may alter its expression, and this may lead to reduced

immune tolerance to the Tg molecule. In addition, alterations

in Tg could possibly explain interactions between genetic and

environmental factors in the etiology of AITD, because Tg is

iodinated to form thyroid hormones, and dietary iodine

may influence the development of AITD (286–288). Indeed,

as noted above, the Tg hormonogenic sites were shown to

contain the autoepitopes in EAT, although the role of iodine

is still controversial in experimental thyroiditis (206, 207).

CTLA-4 polymorphisms associated with AITD decreased its

expression and/or function, thereby promoting the develop-

ment of autoimmunity (151).

As discussed earlier, two CTLA-4 polymorphisms have

been shown to be associated with AITD, a 3’ UTR micro-

satellite and an A/G polymorphism in the leader sequence

of the gene. One recent study (261) examined the effects of

the A and G alleles of the CTLA-4 A/G49 SNP on the inhibi-

tory function of CTLA-4. The authors showed that blocking

of CTLA-4 on T cells isolated from individuals with the G

allele had less effect on reducing the inhibitory function of

CTLA-4 than blocking CTLA-4 on T cells isolated from in-

dividuals with the A allele (261). This could imply that the

A and G alleles of the CTLA-4 leader sequence influenced its

function and/or expression. We have examined the effects of

the CTLA-4 A/G49 SNP using an in vitro assay by transfecting

T cell lines lacking CTLA-4 with CTLA-4 cDNA having the

A or the G allele. When T cells were transfected with CTLA-4

cDNA carrying the G or A allele, there was no difference in

the expression and inhibitory function of CTLA-4 (262). We

concluded that the A and G alleles of the CTLA-4 A/G49 SNP

did not directly influence its function. Other polymorphisms

must be responsible for the association of CTLA-4 with

AITD. However, our results did not necessarily contradict

the published data because they used a dissimilar study

design. Kouki et al. (261) studied the differences in the func-

tion of T cells taken from individuals with the A and the G

alleles. However, the differences they found could be due to

another CTLA-4 polymorphism that was in LD with the

A/G49 polymorphism. Indeed, preliminary data in myas-

thenia gravis showed that the AT microsatellite at the 3’ UTR

of the CTLA-4 gene influenced the half-life of the CTLA-4

mRNA (284, 285). This could provide an attractive explana-

tion for the association between the short alleles of the AT

microsatellite and AITD, as well as other autoimmune
diseases.
AITD susceptibility genes may be immune modifying genes that increase the susceptibility to autoimmunity in general (e.g., HLA, CTLA-4), whereas others may be specific to AITD (e.g., TSHR, Tg). The next step in investigating the role of these genes in the development of AITD is by functional studies and genotype-phenotype correlations. Preliminary functional studies have been performed for HLA (112, 269) and CTLA-4 (151, 261, 262). More functional studies are needed for these and other genes that have shown association with AITD.

It is most likely that the susceptibility loci for AITD interact and that their interactions may influence disease phenotype and severity (193). The molecular basis for the interactions between susceptibility genes in complex diseases is unknown. These interactions could represent the cumulative effect of increased statistical risk, or alternatively, there may be molecular interactions between the susceptibility genes or their products, which ultimately determine disease phenotype. Another unresolved question is how do environmental factors interact with susceptibility genes to modify the risk for disease, as well as the disease phenotype. We are slowly progressing toward identification of the AITD susceptibility genes, and once they are identified we will begin to understand the underlying molecular mechanisms by which they induce thyroid autoimmunity.

Acknowledgments

We thank Dr. David Greenberg for his teaching, support, and ever ready help in our joint studies.

Address all correspondence and requests for reprints to: Yaron Tomer, M.D., Division of Endocrinology, Diabetes and Bone Diseases, Box 1055, Mount Sinai Medical Center, One Gustave L. Levy Place, New York, New York 10029. E-mail: Yaron.Tomer@mssm.edu

This work was supported in part by National Institute of Diabetes and Digestive and Kidney Diseases Grants DK53764, DK45011, and DK52464 (to T.F.D.), and DK6159 and DK58072 (to Y.T.).

References

17. Phillips DI, Barker DJ, Rees Smith B, Didcote S, Morgan D 1985 The geographical distribution of thyrotoxicosis in England according to the presence or absence of TSH-receptor antibodies. Clin Endocrinol (Oxf) 23:283–287
29. Furszyfer J, Kurland LT, McConahey WM, Woolner LB, Elveback LR 1972 Epidemiologic aspects of Hashimoto’s thyroiditis and...


43. Williams D 1996 Thyroid cancer and the Chernobyl accident. J Clin Endocrinol Metab 81:6–8


106. Maciel LM, Rodrigues SS, Dibbern RS, Navarro PA, Donadi EA 2001 Association of the HLA-DRB1*0301 and HLA-DQA1*0501 alleles with Graves’ disease in a population representing the gene contribution from several ethnic backgrounds. Thyroid 11:31–35


113. Segni M, Pani MA, Pasquino AM, Badenhoop K 2002 Familial clustering of juvenile thyroid autoimmunity: higher risk is conferred by human leukocyte antigen DR3-DQ2 and thyroid peroxidase antibody status in fathers. J Clin Endocrinol Metab 87:3779–3782


712 Endocrine Reviews, October 2003, 24(5):694–717 Tomer and Davies • AITD Susceptibility Genes
relapse of hyperthyroid Graves’ disease after antithyroid drugs and HLA haplotypes. J Clin Endocrinol Metab 57:719–722


150. Tomer Y. 2001 Unraveling the genetic susceptibility to autoimmune thyroid diseases: CTLA-4 takes the stage. Thyroid 11:167–169


ulotar-1 (AIRE-1) genes in sporadic autoimmune Addison's disease. J Clin Endocrinol Metab 85:688–691


156. Petrone A, Giorgi G, Mesturino CA, Capizzi M, Cescino N, Nistico L, Osborn J, Di Mario U, Buzzetti R 2001 Association of DRB1*04-DQB1*0301 haplotype and lack of association of two polymorphic sites at CTLA-4 gene with Hashimoto's thyroiditis in an Italian population. Thyroid 11:171–175


173. Cuddihy RM, Dutton CM, Bahn RS 1995 A polymorphism in the extracellular domain of the thyrotropin receptor is highly associated with autoimmune thyroid disease in females. Thyroid 5:89–95


Tomer and Davies • AITD Susceptibility Genes
Endocrine Reviews, October 2003, 24(5):694–715


192. Rapoport B, Chazenbalk GD, Jaume JC, McLachlan SM 1998 The thyrotropin (TSH) receptor: interaction with TSH and autoanti-
bodies. Thyroid Rev 19:673–716


195. Chistiakov DA, Savost’anov KV, Turakulov RI, Petunina N, Babolink ML, Nosikov VV 2000 Further studies of genetic suscepti-
ibility to Graves’ disease in a Russian population. Med Sci 89:578–584

196. Muhlberg T, Herrmann K, Joba W, Kirchberger M, Heberling HJ, Heufelder AE 2000 Lack of association of nonautoimmune hyper-
fuctioning thyroid disorders and a germline polymorphism of codon 727 of the human thyrotropin receptor in a European Cau-
sasian population. J Clin Endocrinol Metab 85:2640–2643

197. Tomer Y, Greenberg DA, Concepcion ES, Davies TF 2001 Thyrotropin is a thyroid specific gene for the familial autoimmune
thyroid disease. J Clin Endocrinol Metab 87:404–407

198. Grossman CJ, Roselle GA, Mendenhall CL 1991 Sex steroid reg-

199. Ando T, Imaizumi M, Graves PN, Unger P, Davies TF 2002 In-
tra-thyroidal fetal microchimerism in Graves’ disease. J Clin Endo-
crinol Metab 87:3313–3320


204. Ansar AS, Young PR, Penhale WJ 1983 The effects of female sex steroids on the development of autoimmune thyroiditis in thymec-


206. de Kerndet M, Lucas J, Lemee F, Lecornu M 1994 Turner’s syn-
drome with X-isochromosome and Hashimoto’s thyroiditis. Clin Endocrinol (Oxf) 41:673–676


208. Neleiros CC, Marini SH, Baptista MT, Guerra Jr G, Maciel-
Guerra AT 2000 Turner’s syndrome and thyroid disease: a trans-


215. Westman AP, Jenkins RC 2002 Disease associations with autoimmune thyroid disease. Thyroid 12:977–988


