

## GENOMIC MEDICINE

W. Gregory Feero, M.D., Ph.D., and Alan E. Guttmacher, M.D., *Editors*

# Genomics and the Multifactorial Nature of Human Autoimmune Disease

Judy H. Cho, M.D., and Peter K. Gregersen, M.D.

From the Departments of Medicine and Genetics, Section of Digestive Diseases, Yale University, New Haven, CT (J.H.C.); and the Robert S. Boas Center for Genomics and Human Genetics, Feinstein Institute for Medical Research, Manhasset, NY (P.K.G.). Address reprint requests to Dr. Cho at the Section of Digestive Diseases, Department of Internal Medicine, Yale University School of Medicine, 333 Cedar St., 1080 LMP, New Haven, CT 06520-8019, or at judy.cho@yale.edu.

N Engl J Med 2011;365:1612-23.  
Copyright © 2011 Massachusetts Medical Society.

**T**HE MAJOR AUTOIMMUNE DISEASES, INCLUDING RHEUMATOID ARTHRITIS, systemic lupus erythematosus, multiple sclerosis, type 1 diabetes mellitus, psoriasis, and inflammatory bowel disease, share epidemiologic, clinical, and therapeutic features. In each of these diseases, chronic and often intermittent inflammation contributes over time to the destruction of target organs that house inciting antigens or are the sites of immune-complex deposition. For some of these disorders, such as inflammatory bowel disease, the contribution of autoimmune mechanisms is questioned, but the overlap of genetic associations that have been identified during the past 5 years suggests a shared immune pathogenesis. At the same time, genetic data also support some distinct pathways of pathogenesis for the various disorders.

Although the adaptive immune system has long been a focus of attention, innate immune mechanisms are now viewed as being central to the pathogenesis of these disorders. In addition, the concept of quantitative thresholds for immune-cell signaling has emerged in the past decade as a potential way of understanding how multiple genetic factors of relatively small effect may combine to create a state of susceptibility to autoimmune activation. The new genetic findings have also emphasized that the identification of the environmental components that interact with host genetic factors will be critical in developing a deeper understanding of autoimmunity, as well as new approaches to prevention and treatment.

## GENETICS OF HUMAN AUTOIMMUNE DISEASE — THE STATUS QUO

Genomewide association studies (see Glossary) have opened up an exciting new window on the genetics of complex diseases.<sup>1</sup> Perhaps nowhere have such studies been more fruitful than in the area of autoimmune diseases: more than 200 genetic loci have been shown to be associated with one or more autoimmune disorders.<sup>2-26</sup> In many cases, the precise causal alleles or genes that drive these associations have not been identified. Nevertheless, some loci that show associations with multiple diseases belong to categories of associated genes that, in turn, implicate the involvement of particular functional pathways. These include the intracellular signaling that drives the activation of T and B cells, signaling by cytokines and cytokine receptors, and pathways that mediate innate immunity and microbial responses (Table 1). Although the majority of genomewide association studies have focused on case series of European ancestry, studies of other populations show that some associations are observed across populations, and such associations point to pathways that may be particularly important in disease pathogenesis. Each new genetic finding can suggest multiple hypotheses that need to be fit into an overall scheme of pathogenesis.

In this review, we will focus on recent findings from genomewide association studies that are most robust, that most clearly implicate a disease mechanism, or

that suggest a new diagnostic or therapeutic approach. Nevertheless, it should be emphasized that for most of these disorders, genes within the major histocompatibility complex (MHC) have by far the strongest single genetic effect, and many of these associations have been known for decades.<sup>27</sup> It is likely that most of these genetic associations reflect the immunoregulatory effects of the HLA molecules themselves, although the exact mechanisms that underlie these effects are still a matter of some debate. Therefore, although the MHC is not a major focus of our discussion, the new genetic findings in autoimmunity must always be considered in the context of the important contributions of the HLA complex to disease susceptibility and pathogenesis.

AUTOIMMUNITY — A COMPLEX QUANTITATIVE TRAIT

The original idea of autoimmunity derives from Paul Ehrlich's realization that a functional immune system must have "horror autotoxicus," which he conceived as having "certain contrivances" that would prevent immune attack against the self.<sup>28</sup> Recent genetic findings emphasize that these "contrivances" are multiple and complex. Several decades of increasingly sophisticated basic immunologic studies in mice have provided an elegant platform for interpretation of the genetic data and hypothesis generation. It is convenient to divide the immune system into innate and adaptive systems, although in reality they are highly integrated and interdependent. The innate immune system is phylogenetically older and is designed for immediate engagement of pathogens by a highly conserved set of pattern-recognition receptors, such as toll-like receptors, coupled with a prompt defensive response by the cell. In contrast, the adaptive immune system consists primarily of T and B cells, which use highly diverse receptor systems selected somatically for antigen recognition (T-cell receptor and surface immunoglobulin, respectively) that can recognize millions of distinct foreign antigens, and by the formation of immunologic memory. This immediately raises the problem of selecting functional receptors that do not lead to uncontrolled self-reactivity.

Self-reactive B and T cells are a normal component of the immune system, but they are kept in check by a variety of mechanisms, many of which appear to be altered by genetic loci implicated in autoimmunity (Fig. 1). Some are central mecha-

Glossary
<b>Allele:</b> One of two or more versions of a genetic sequence at a particular location in the genome.
<b>Genomewide association study:</b> An approach used in genetics research to look for associations between many (typically hundreds of thousands) specific genetic variations (most commonly single-nucleotide polymorphisms) and particular diseases.
<b>Kinase:</b> An enzyme that transfers a phosphate group to a substrate.
<b>Locus:</b> The specific chromosomal location of a gene or other DNA sequence of interest.
<b>Loss-of-function mutation:</b> A mutation that decreases the production or function of a protein (or does both).
<b>Missense mutation:</b> The alteration of a single DNA nucleotide so that the resulting codon specifies a different amino acid.
<b>Nonsense mutation:</b> The alteration of a single DNA nucleotide so that the resulting codon signals a termination of translation, thus leading to truncation of the encoded protein.

nisms in the thymus and bone marrow that delete or disable self-reactive clones; others are peripheral and include specialized regulatory cells, such as regulatory T cells.<sup>29</sup>

Several mendelian disorders directly illustrate the importance of these mechanisms. For example, mutations affecting the transcription factor autoimmune regulator lead to a relaxing of selection against self-reactivity by T cells in the thymus, giving rise to a rare, aggressive autoimmune disease, autoimmune polyendocrine syndrome 1.<sup>30</sup> The autoimmune regulator controls the ectopic expression of self-antigens within the thymus<sup>31</sup> and thus is critical to the negative selection of T cells reactive with these antigens. Genetic studies indicate the presence of more limited defects in selection against reactivity with self-antigens, such as insulin, in the predisposition to type 1 diabetes.<sup>3</sup> In contrast to these defects in central tolerance, a loss of the FOXP3 transcription factor in the mendelian disorder IPEX (immune dysregulation, polyendocrinopathy, enteropathy, X-linked)<sup>32,33</sup> causes aggressive autoimmunity as a result of defects in the function of regulatory T cells. A milder defect in the regulatory activity of such cells may be caused by quantitative changes in the expression of CD25 (IL2RA),<sup>34</sup> encoding the receptor for interleukin-2, a cytokine that is critical to the survival of regulatory T cells and that such cells cannot produce themselves.<sup>29</sup> Analogous control mechanisms are active at numerous checkpoints in the B-cell portion of the immune system<sup>35</sup> (Fig. 1). For example, pre-B cells in the bone marrow are highly autoreactive but become less so on differentiation into naive B cells

**Table 1. Selected Major Association Signals in Autoimmune Diseases.\***

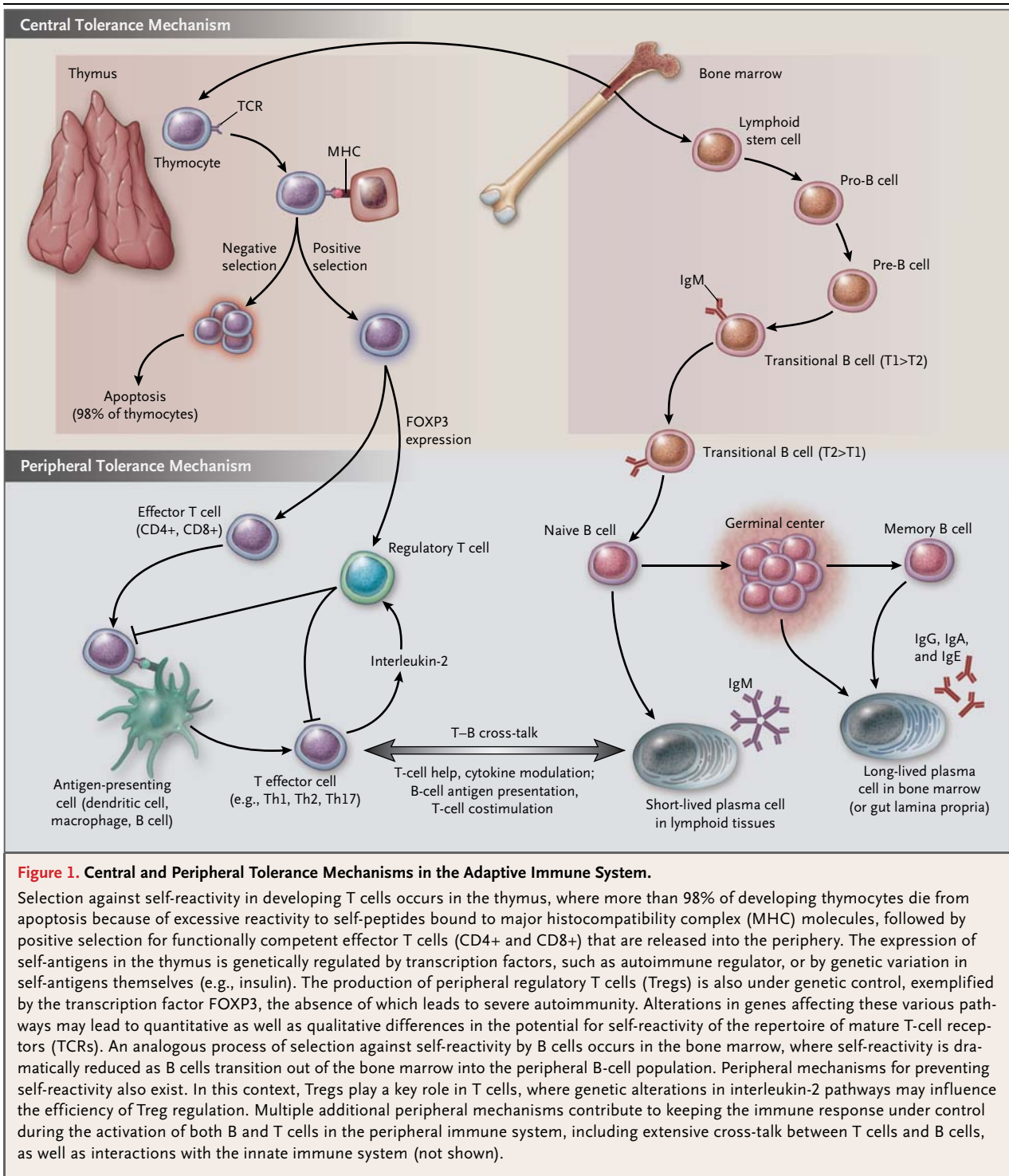
Candidate Gene	Chromosome Location	Possible Functions and Mechanisms of Action	Major Associated Diseases
<b>Lymphocyte activation and intracellular signaling</b>			
Major histocompatibility complex (HLA)	6p21	Antigen presentation; complex, often disease-specific association signals that finely modulate antigen presentation	Most autoimmune disorders
Protein tyrosine phosphatase nonreceptor type 22 ( <i>PTPN22</i> )	1p13	Modulation of lymphocyte receptor activation; a polymorphism resulting in an Arg620Trp substitution drives the association	Type 1 diabetes mellitus, rheumatoid arthritis, systemic lupus erythematosus, Graves' disease, † Crohn's disease
Cytotoxic lymphocyte-associated protein 4 ( <i>CTLA4</i> )	2q33	Transmission of inhibitory signals in T cells	Type 1 diabetes mellitus, rheumatoid arthritis, celiac disease, alopecia areata
T-cell activation, Rho-GTPase-activating protein ( <i>TAGAP</i> )	6q25	Expression in activated T cells	Rheumatoid arthritis, Crohn's disease, celiac disease, type 1 diabetes mellitus
Protein tyrosine phosphatase nonreceptor type 2 ( <i>PTPN2</i> )	18p11	Expression in T cells; role in cell growth and differentiation	Type 1 diabetes mellitus, Crohn's disease, celiac disease
Tyrosine protein kinase 2 ( <i>TYK2</i> )	19p13	Janus kinase downstream of cytokine receptors	Psoriasis, type 1 diabetes mellitus, systemic lupus erythematosus, Crohn's disease, multiple sclerosis
Tumor necrosis factor $\alpha$ -induced protein 3 ( <i>TNFAIP3</i> )	6q23	Regulation of ubiquitination; down-regulation of nuclear factor $\kappa$ B activation	Rheumatoid arthritis, systemic lupus erythematosus
TNFAIP3-interacting protein ( <i>TNIP1</i> )	5q33	Down-regulation of nuclear factor $\kappa$ B activation; function of TNIP1 is dependent on ubiquitin-binding domain	Systemic lupus erythematosus, psoriasis
Tumor necrosis factor receptor superfamily member 5 ( <i>CD40</i> )	20q13	Costimulatory molecule for B-cell activation; interaction with T cells through CD40 ligand ( <i>CD154</i> ); broadly expressed	Rheumatoid arthritis
Protein kinase C theta ( <i>PRKCC</i> )	10p15	T-cell activation and signaling through c-Rel	Type 1 diabetes mellitus, rheumatoid arthritis
<b>Cytokines and cytokine receptors</b>			
Interleukin-23 receptor gene ( <i>IL23R</i> ) region	1p31	Enhancement of select cell subsets, including Th17 cells; multiple association signals (e.g., Arg381Gln polymorphism)	Crohn's disease, ulcerative colitis, psoriasis, ankylosing spondylitis, primary biliary cirrhosis†
Interleukin-2 receptor, subunit alpha ( <i>IL2RA</i> )	10p15	One component of interleukin-2 receptor signaling; linkage of disease-associated genotypes with decreased <i>IL2RA</i> expression	Type 1 diabetes mellitus, multiple sclerosis, rheumatoid arthritis, Crohn's disease, vitiligo, alopecia areata
Interleukin-2/21 gene region	4q26	T-cell trophic growth factors; multiple associations flanking both cytokines	Celiac disease, Crohn's disease, ulcerative colitis, rheumatoid arthritis, type 1 diabetes mellitus
Interleukin-7 receptor ( <i>IL7R</i> )	5p13	Differentiation and activation of T cells affected by interleukin-7 signaling	Multiple sclerosis, primary biliary cirrhosis, alopecia areata
Interleukin-12B, p40 ( <i>IL12B</i> )	5q33	Cytokine subunit common to interleukin-12 and interleukin-23	Crohn's disease, ulcerative colitis, psoriasis, psoriatic arthritis, systemic lupus erythematosus
Interleukin-10 gene ( <i>IL10</i> ) region	1q32	Down-regulation of cytokines, MHC class II and costimulatory molecules	Systemic lupus erythematosus, type 1 diabetes mellitus, Crohn's disease, ulcerative colitis

<b>Innate immunity and microbial recognition</b>	
Nucleotide oligomerization domain 2 ( <i>NOD2</i> )	16q12 Sensing of bacterial peptidoglycan in nuclear factor $\kappa$ B activation; loss-of-function, uncommon missense polymorphisms Crohn's disease
Interferon regulatory factor 5 ( <i>IRF5</i> )	7q32 Inducement of interferons, regulation of activation of pattern-recognition receptor; multiple associated polymorphisms affecting splicing and messenger RNA levels Systemic lupus erythematosus, rheumatoid arthritis, primary biliary cirrhosis, systemic sclerosis
Interferon-induced helicase C domain-containing protein 1 ( <i>IFIH1</i> )	2q24 Recognition of single-stranded RNA from picornaviruses; protection against disease conferred by extremely rare missense mutations Type 1 diabetes mellitus, psoriasis, selective IgA deficiency
Autophagy-like 16L1 ( <i>ATG16L1</i> )	2q37 Targeting of intracellular components to lysosomes Crohn's disease
PR domain zinc finger protein 1 ( <i>PRDM1</i> ); autophagy protein 5 ( <i>ATG5</i> )	6q21 Expression of genes encoding beta-interferons repressed by <i>PRDM1</i> (also known as <i>BLIMP1</i> ), which is a key regulator of B-cell differentiation; <i>ATG5</i> part of autophagy complex, with major association between <i>PRDM1</i> and <i>ATG5</i> Systemic lupus erythematosus, rheumatoid arthritis, Crohn's disease
<b>Transcription factors</b>	
Signal transducer and activator of transcription 4 ( <i>STAT4</i> )	2q32 Mediation of multiple cytokine signals, including interleukin-12 Systemic lupus erythematosus, rheumatoid arthritis, primary biliary cirrhosis, systemic sclerosis
Signal transducer and activator of transcription 3 ( <i>STAT3</i> )	17q21 Mediation of multiple cytokine signals (e.g., interleukins 6, 10, 22, and 23); gene-rich region of association Crohn's disease, multiple sclerosis
c-Rel ( <i>REL</i> )	2p16 Transcription factor, a component of nuclear factor $\kappa$ B Rheumatoid arthritis, Crohn's disease, ulcerative colitis, celiac disease, psoriasis
<b>Other pathways or mechanisms</b>	
Endoplasmic reticulum aminopeptidase 1 ( <i>ERAP1</i> )	5q15 Trimming of peptides for HLA class I presentation; interactive associations with class I alleles observed in MHC class I–pre-dominant diseases Psoriasis, ankylosing spondylitis
Fc fragment of IgG, low affinity IIa, receptor ( <i>FCGR2A</i> )	1p23 Cell-surface receptor on phagocytic cells; associations including <i>His131Arg</i> polymorphism Systemic lupus erythematosus, rheumatoid arthritis, ulcerative colitis
Chemokine (C-C motif) receptor 6 ( <i>CCR6</i> )	6q27 Expression on immature dendritic cells and memory T cells; involvement in lymphocyte trafficking Crohn's disease, rheumatoid arthritis, Graves' disease, vitiligo
Integrin alpha M precursor ( <i>ITGAM</i> )	16p11 Immune complex clearance and leukocyte adhesion; amino acid change implicated as one causal allele Systemic lupus erythematosus
Ubiquitin-associated and SH3 domain-containing protein A ( <i>UBASH3A</i> )	21q22 Association with ubiquitin and SH3 domain-containing protein Type 1 diabetes mellitus, rheumatoid arthritis, celiac disease
Ubiquitin-conjugating enzyme E2L3 ( <i>UBE2L3</i> )	22q11 Ubiquitin-conjugating enzyme Rheumatoid arthritis, celiac disease, systemic lupus erythematosus
Insulin locus ( <i>INS</i> )	11p15 Targeting autoantigen; expression polymorphism; possible role in thymic selection Type 1 diabetes mellitus

\* MHC denotes major histocompatibility complex.

† A potentially distinct association within this implicated genetic region has been shown for this disease.

‡ Genetic associations with this disease have been observed in patients of Asian or European ancestry.



in the periphery, a process that is influenced by the gene encoding protein tyrosine phosphatase non-receptor type 22 (*PTPN22*) and other genes associated with autoimmunity.<sup>36</sup>

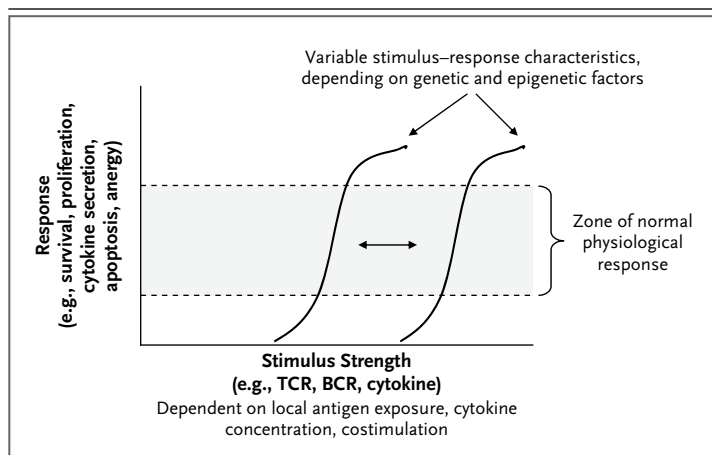
Overall, these processes of selection and regulation of T and B cells are controlled by cell-signaling events that are normally active within a range of potency that may vary among persons and



among cell types, owing in large part to genetic or epigenetic diversity in the population. This leads to a general concept of immune responsiveness and regulation as a trait that exists on a continuum (a quantitative trait), setting thresholds for cell activation and response.<sup>37</sup> Indeed, the original discovery of the MHC, which encodes HLA, as a locus controlling immune responses was described as a quantitative trait — HLA-regulated immune responses are generally high or low, as opposed to just absent or present (Fig. 2). For example, given an antigenic stimulus of a given strength (say, presentation of self-antigen by self-HLA in the thymus), responsiveness can vary among persons. Such variability can result in more or less efficient signaling in the T-cell receptor, leading in turn to apoptosis and negative selection over a range of self-reactivity. These events directly influence the diversity of the mature T-cell repertoire after thymic selection is complete. Similar quantitative effects are likely to be operating in cytokine pathways<sup>34</sup> and in many signaling pathways in both the adaptive and innate systems of the immune system.

INTRACELLULAR SIGNALING PATHWAYS

Many of the genes that have recently been implicated in autoimmunity contribute to immune signaling pathways involving T-cell and B-cell receptors, costimulatory molecules and cytokines, and pattern-recognition receptors, such as toll-like receptors or nucleotide-binding oligomerization domain (NOD) receptors that are involved in innate immune responses. Although the exact causative alleles are not known in most cases, there are some exceptions. For example, the causative change in the intracellular phosphatase PTPN22 is almost certainly due to a specific amino acid substitution (of arginine by tryptophan) at position 620 of the protein. This amino acid change disrupts binding between PTPN22 and an intracellular kinase called Csk (Fig. 3), and this in turn alters the responsiveness of both T and B cells to receptor stimulation. Mice lacking Ptpn22 have dramatically increased T-cell activation.<sup>38</sup> Recent data indicate that a similar phenotype of enhanced lymphocyte responsiveness is associated with the PTPN22 risk allele,<sup>39,40</sup> although data from humans continue to conflict.<sup>41,42</sup> In any case, there is little doubt that alterations in PTPN22 change the thresholds for receptor signaling by T and B cells; the mechanism by

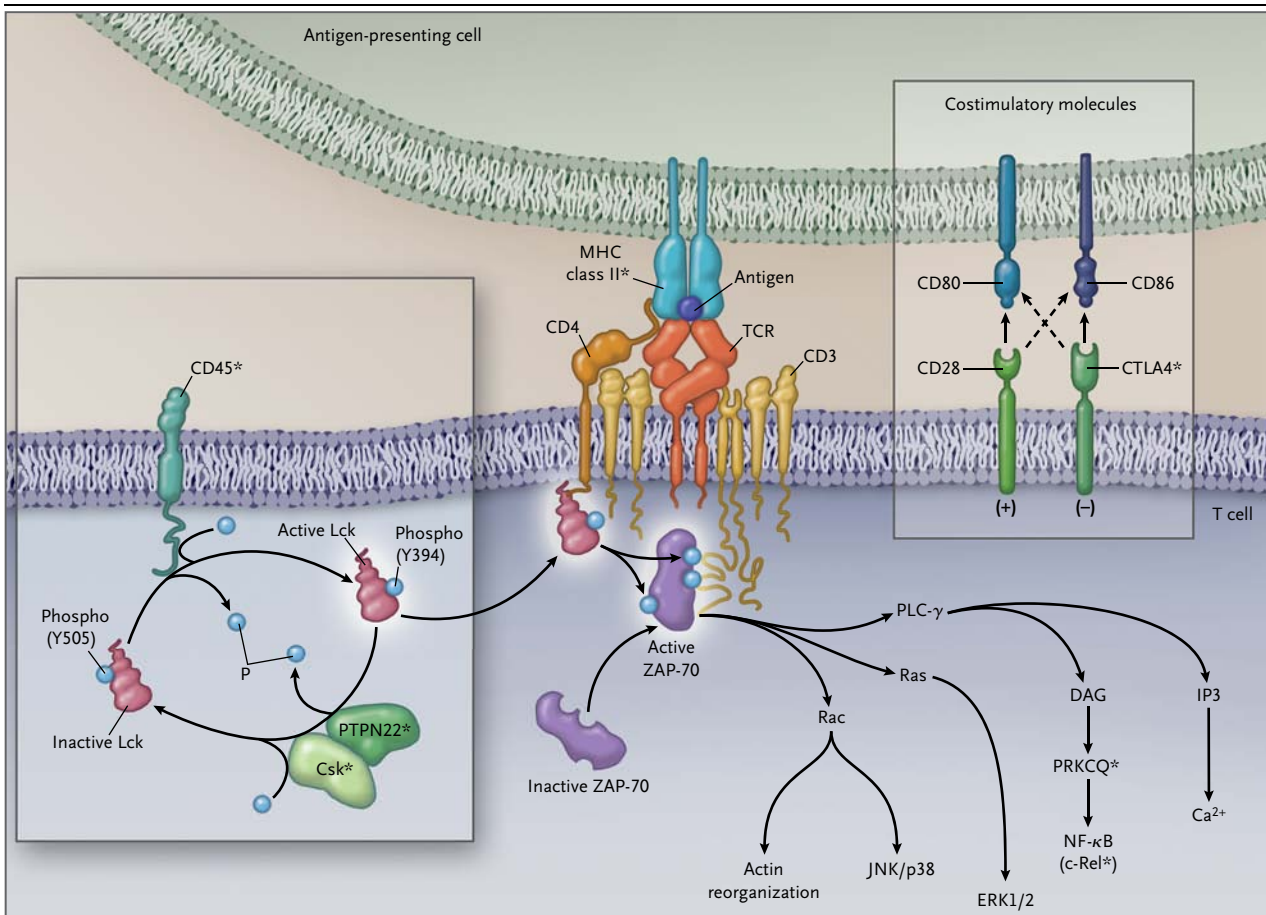


**Figure 2. Stimulus-Response Thresholds and Immune Recognition as a Quantitative Trait.**

Cells of the adaptive immune system are selected for stimulus-response characteristics over a range of values, and these may differ for different individuals or for different cells, depending in part on genetic or other host factors that influence receptor signaling pathways. This can lead to individual variation in the overall thresholds of activation for a diverse array of immune stimuli and effector responses, with associated susceptibility or resistance to autoimmune disease. These threshold effects have been most clearly shown for thymic and early B-cell selection events, but it is likely that individual variation in stimulus-response characteristics influences many signaling systems relevant to autoimmunity. BCR denotes B-cell receptor, and TCR T-cell receptor.

which they lead to autoimmunity remains unclear. Furthermore, the consequences of a specific variant may be different for various diseases. For example, the PTPN22 620 tryptophan risk allele confers an increased risk of a variety of humoral autoimmune disorders, such as rheumatoid arthritis, thyroid disease, type 1 diabetes, and many others,<sup>43</sup> but protects against Crohn's disease.<sup>6</sup> At the same time, it has no effect on the risk of multiple sclerosis.<sup>44</sup> This same allele also appears to influence the outcome for certain infectious diseases.<sup>45,46</sup> Thus, the genetics of PTPN22 shows the profound effect of subtle changes in intracellular-signaling thresholds on individual disease susceptibility.<sup>43</sup>

The precise causative variations that have the greatest effect on other intracellular-signaling pathways involved in autoimmunity are much less clear than for PTPN22, and in some cases the new genetic findings have focused attention on newly emerging signaling mechanisms. A good example of this concerns the role of protein ubiquitination. Analogous to protein phosphorylation, the addition and removal of ubiquitin from proteins is a



**Figure 3. Highly Simplified Scheme for Signaling by T-Cell Receptors, Emphasizing the Proximal Events in Signal Transduction.**

After a T-cell receptor (TCR) is triggered by a peptide and major histocompatibility complex (MHC), one of the earliest events is activation of lymphocyte-specific protein tyrosine kinase (Lck) by removal of an inhibitory phosphate (phospho Y505) by membrane phosphatase CD45 and autophosphorylation of Lck at phospho Y394, as shown at the left of the figure. These events convert Lck to an active state, in which it can phosphorylate substrates, such as 70-kD zeta-associated protein (ZAP-70) and immunoreceptor tyrosine-based activation motif in the TCR complex itself. The risk allele of *PTPN22* encodes an amino acid substitution (Arg620Trp) that disrupts binding between PTPN22 protein and c-src tyrosine kinase (Csk). This disruption may either lower or increase the activation state of Lck, depending on the experimental situation. A number of distinct downstream signaling pathways are regulated by Lck–ZAP-70 signaling, including the activation of kinases, such as mitogen-activated protein kinase 8 (MAPK8, or JNK) and MAPK p38 and extracellular signal-related kinase 1 and 2 (ERK1/2), activation of nuclear factor  $\kappa$ B (NF- $\kappa$ B) through protein kinase C theta (PRKCQ), as well as calcium release and cytoskeletal reorganization. Genetic variants affecting CD45, PRKCQ, and c-Rel have also been associated with autoimmune diseases, which emphasizes that multiple variants may contribute to subtle variation in signaling thresholds. A role for costimulatory molecules in T-cell activation is also indicated at the right of the figure. In this case, genetic variation affecting a negative regulator of T-cell activation (e.g., CTLA4) can alter lymphocyte responses and predispose to autoimmunity. Many other such costimulatory molecules (not shown) have been revealed as potential regulators of autoimmunity in genetic studies (see Table 1). DAG denotes diacylglycerol, IP3 inositol triphosphate, and PLC- $\gamma$  phospholipase C- $\gamma$ . Asterisks indicate gene products for which genetic associations with autoimmunity have been documented.

common means of regulating immune signaling pathways.<sup>47</sup> Several of the genes associated with autoimmunity have a role in regulating ubiquitination (*TNFAIP3*), binding to ubiquitinated proteins (*TNIP1* and *UBASH3A*), or regulating enzymatic events in ubiquitination (*UBE2L3*) (Table 1).

A large number of the signaling pathways that are implicated in autoimmunity have a component of ubiquitin-mediated regulation.<sup>47</sup> This area of investigation is likely to dramatically expand our understanding of autoimmune mechanisms in the coming years.

---

 GENETIC VARIATION AND CYTOKINE PATHWAYS
 

---

Given the central role of modulation of cytokine function in the development of new treatments for autoimmunity, it is not surprising that many of the strongest genetic associations have implicated cytokine pathways. Activities of key cytokine signaling pathways are modulated through complex mechanisms. For example, interleukin-2 is a central mediator of T-cell growth, and its signaling is finely regulated in part by the relative affinity of different combinations of its receptor components. Specifically, the high-affinity interleukin-2 receptor is composed of alpha (IL2RA, or CD25), beta (IL2RB), and gamma (IL2RG) components, whereas a low-affinity interleukin-2 receptor is composed of IL2RA homodimers. The central role of IL2RA expression in autoimmunity is underscored by the associations between the *IL2RA* locus and diseases such as type 1 diabetes,<sup>48</sup> multiple sclerosis,<sup>49</sup> rheumatoid arthritis,<sup>5</sup> and Crohn's disease.<sup>6</sup> Disease-associated DNA polymorphisms in *IL2RA* are associated with altered expression in messenger RNA.<sup>34</sup> Modulation of interleukin-2 signaling can affect the relative growth and survival of subgroups of regulatory and proinflammatory T cells. Furthermore, genetic markers in and near the gene encoding interleukin-2 have shown association with celiac disease,<sup>50</sup> ulcerative colitis,<sup>51</sup> rheumatoid arthritis,<sup>5</sup> and type 1 diabetes<sup>52</sup> (Table 1).

The cellular infiltrates characterizing inflammatory lesions are influenced by a diverse array of proinflammatory and antiinflammatory cytokines. In particular, a central proinflammatory, pathogenic role for subgroups of CD4+ type 1 helper T cells is under the control of interleukin-12 and has long been implicated in a variety of human and murine models of autoimmunity. An essential role for a related cytokine, interleukin-23,<sup>53</sup> has been shown in a variety of autoimmunity models, including mouse models of multiple sclerosis, inflammatory bowel disease, collagen-induced arthritis, and dermatitis.<sup>54</sup> Interleukin-23 is required for the expansion and survival of type 17 helper T cells, which have a key role in mediating mucosal immunity and defense against extracellular pathogens. These immunologic advances coincided with genetic discoveries of major association signals encompassing the interleukin-23 receptor (IL23R) to inflammatory bowel disease,<sup>55</sup> psoriasis,<sup>56</sup> and ankylosing spondylitis.<sup>57</sup> The interleukin-23

and interleukin-12 signaling pathways share components at their cytokines, receptors, and downstream signaling intermediates.<sup>58</sup> A striking number of these components are encoded by genes, variants of which are associated with autoimmunity (Fig. 4). However, the intermediary signaling molecules are not specific to these cytokine pathways.

---

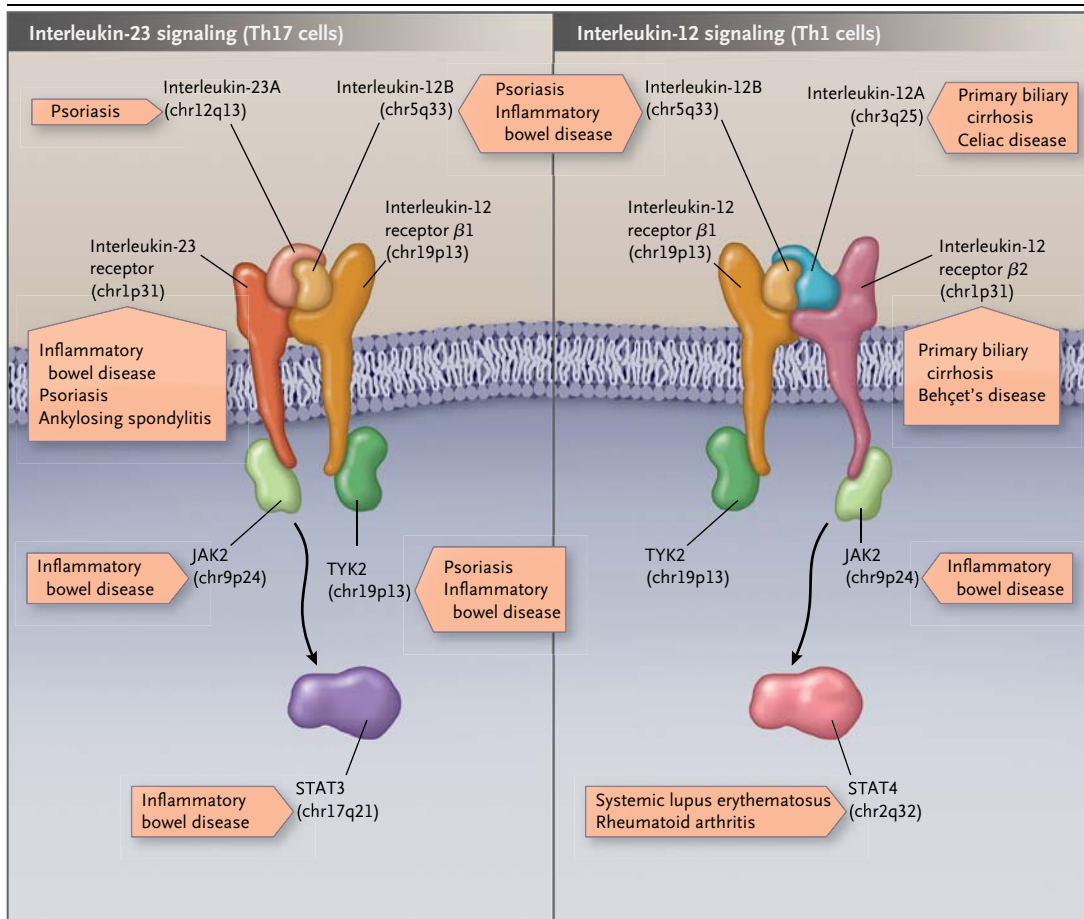
 INNATE IMMUNITY AND MICROBIAL RESPONSES
 

---

The human immune response to microbial pathogens has been shaped throughout evolutionary history by natural selection. It is likely that the increase in the prevalence of various autoimmune disorders that have been observed during the past century has occurred largely through changes in environmental and microbial exposures. Recent genetic advances have highlighted specific differences among individuals in innate immune-system responses that are characteristic of various subtypes of chronic inflammatory diseases. For example, one of the first genetic associations that was reported in complex genetic disorders was that between variants of *NOD2* and Crohn's disease<sup>59</sup>; this association represents the most substantial contributor to the overall heritability of Crohn's disease. *NOD2* functions as an intracellular sensor for bacterial peptidoglycan, and variants associated with Crohn's disease are loss-of-function polymorphisms. Further implicating a primary role for innate immunity, variants of autophagy genes (e.g., *ATG16L1*) that target intracellular components, including microbes, to lysosomes have also been associated with Crohn's disease.<sup>59</sup> The increase in the prevalence of Crohn's disease during the past century may well reflect corresponding changes in the composition of the intestinal microbiota resulting from changes in hygiene, such as eradication of intestinal parasites.<sup>60</sup>

A central role for the intestinal microbiota in inflammatory bowel disease provides the rationale for the therapeutic use of microbiome-altering probiotics.<sup>61</sup> However, the role of the intestinal microbiome is probably not limited to the intestinal immune system. Studies of mouse models of autoimmune diabetes suggest that host differences in the capacity to sense intestinal microbes<sup>62</sup> and the specific composition of the intestinal microbiota modulate susceptibility to diabetes. The dynamic cross-talk between the host immune





**Figure 4. Genetic Associations of Autoimmune Diseases along the Interleukin-12 and Interleukin-23 Signaling Pathways.**

Shown are the major signaling components implicated for each cytokine in selected autoimmune diseases. It should be noted that multiple JAK–STAT components have been implicated in interleukin-12 and interleukin-23 signaling. Conversely, each JAK–STAT component is capable of signaling downstream of multiple cytokines. In inflammatory bowel disease, psoriasis, and ankylosing spondylitis, association signals on chromosome 1p31 in the region encoding the C terminus of interleukin-23 receptor (IL23R) extend into the intergenic region between *IL23R* and *IL12RB2*, encoding a subunit of the interleukin-12 receptor. In contrast, the associations with primary biliary cirrhosis and Behçet's disease predominate more in the *IL12RB2* intergenic region. Taken together, particularly strong genetic associations implicate variant interleukin-12 and interleukin-23 signaling in psoriasis and inflammatory bowel disease.

response and the intestinal microbiota will be an important focus of future research.

GENES, ENVIRONMENT,  
AND AUTOIMMUNITY

In addition to variability in bacterial responses, differences among individuals in antiviral responses contribute to autoimmunity. As noted above, interferon pathways mediate both autoimmunity and viral defense,<sup>63</sup> and genetic associations exist between the increased expression of interferon regulatory factor 5 (IRF5) and lupus,<sup>64</sup> rheumatoid

arthritis,<sup>5</sup> and primary biliary cirrhosis.<sup>65,66</sup> Although IRF5 broadly regulates innate immune responses through toll-like receptors<sup>67</sup> and type 1 interferons, other key genetic polymorphisms have been linked to more specific viral groups. For example, IFIH1 (interferon-induced helicase C domain-containing protein 1) recognizes the single-stranded RNA of picornaviruses, and genetic associations between *IFIH1* and type 1 diabetes mellitus,<sup>68</sup> psoriasis,<sup>24</sup> and selective IgA deficiency<sup>17</sup> have been reported (Table 1). Specifically, uncommon missense mutations, including a non-sense mutation, confer protection against type 1

diabetes mellitus.<sup>68</sup> Previous infection with enteroviruses (a genera of picornaviruses) is more frequent in patients with newly diagnosed type 1 diabetes mellitus.<sup>69</sup> More broadly, it may be speculated that changing microbial exposures contributes to alterations in the prevalence of autoimmune diseases. With new genetic associations, we are now in position to address the mechanisms by which genetic variation controls the response to infection and thereby predisposes to the development of autoimmunity.

Beyond infectious agents, the specific environmental factors that contribute to autoimmunity have been extremely challenging to identify and include such factors as silica dust, cooking oil, sun exposure, and smoking.<sup>70,71</sup> In some cases, the new genetics has made such studies more compelling and potentially more tractable by revealing interactions between environmental and genetic factors. Perhaps nowhere is this more evident than in the role of smoking as a risk factor for rheumatoid arthritis. Thus, although smoking is associated with a modest risk of rheumatoid arthritis (approximate risk ratio, 1.5), when variants of HLA and *PTPN22* are taken into consideration, the combined risk in smokers is increased by more than a factor of 20.<sup>72</sup> Even more strikingly, this interaction between smoking and genetic factors is seen only in the major serologic subgroups of rheumatoid arthritis that have antibodies against particular citrullinated peptides (in 50 to 70% of patients with rheumatoid arthritis), whereas virtually no risk is observed for such patients who do not carry these antibodies.<sup>73</sup> Furthermore, there is evidence that smoking induces the appearance of citrullinated autoantigens in the lung, thus providing the beginnings of a causal pathway linking an environmental exposure (smoking, and possibly air pollution) to the production of pathogenic antibodies (anticitrullinated autoantibodies) in genetically predisposed persons.<sup>74</sup> This emphasizes that the power of environmental studies can be dramatically improved by focusing on specific genetic and phenotypic subgroups of a disease.

---

#### IMPLICATIONS OF THE NEW GENETICS FOR DIAGNOSIS AND TREATMENT

---

The most obvious near-term potential use of genetic data is to improve diagnostic accuracy, as well as to permit disease stratification for risk assessment and treatment selection. However, most

disease-associated alleles individually carry very modest degrees of risk. Indeed, for almost all autoimmune diseases, associations with the presence of HLA alleles are the strongest, and even these have not proved to be useful in clinical situations. Attempts to use data from genomewide association studies to determine drug response also have so far been disappointing.<sup>75</sup> Thus, it is likely that genetic data will need to be combined with other biomarkers to identify clinically meaningful subgroups of patients to guide the treatment of patients. Such an approach may be particularly useful for early detection of persons at risk for autoimmune disease, because serologic autoimmunity may be present for many years,<sup>76</sup> even though overt clinical disease develops in only a subgroup of such persons.

The new genetic data implicate new pathways that suggest new therapeutic targets. Indeed, many of the new biologic therapies appear to be appropriately matched to these pathways. These include inhibitors of tumor necrosis factor, anti-B-cell agents such as rituximab, and the fusion protein abatacept, which targets B7–CTLA4 interactions involved in T-cell activation (Fig. 3). One of the best correlates between new therapies and genetics has been the use of anti-p40 monoclonal antibodies in patients with psoriasis,<sup>77</sup> which is characterized by an extremely strong interleukin-23 genetic signature (Fig. 4). New therapeutic agents are emerging that target lymphocyte signaling through interferon inhibition, lymphocyte trafficking, or kinase inhibition.<sup>78,79</sup> Given the multiplicity of gene associations implicating multiple disease pathways, future therapeutic advances may require the application of combination therapies. However, the blockade of multiple immune pathways is also likely to increase susceptibility to infectious complications. Although each genetic association represents a potentially new therapeutic target, future therapeutic advances will require increasingly refined and integrated models of disease pathogenesis. High-resolution comparison of association patterns across autoimmunity will soon be possible with data obtained from common genotyping platforms and may provide insight into which therapeutic targets should be prioritized for various disease subtypes.

There is much more to be learned about the genetics of autoimmunity; more than half the genetic contribution to autoimmunity still remains to be identified. The nature of this missing heritability remains a matter of intense debate.<sup>80</sup> It may

reside in many hundreds or even thousands of additional common variants with very modest effect, or there may be groups of rare, more highly penetrant variants that explain disease in subgroups of patients,<sup>68,81,82</sup> analogous to the situation of some lipoprotein phenotypes.<sup>83</sup> The rapid technical progress toward sequencing whole human genomes is likely to lead to another wave of genetic discovery in autoimmunity in the near future.

Disclosure forms provided by the authors are available with the full text of this article at NEJM.org.

## REFERENCES

- Manolio TA, Collins FS, Cox NJ, et al. Finding the missing heritability of complex diseases. *Nature* 2009;461:747-53.
- De Jager PL, Jia X, Wang J, et al. Meta-analysis of genome scans and replication identify CD6, IRF8 and TNFRSF1A as new multiple sclerosis susceptibility loci. *Nat Genet* 2009;41:776-82.
- Concannon P, Rich SS, Nepom GT. Genetics of type 1A diabetes. *N Engl J Med* 2009;360:1646-54.
- Flesher DL, Sun X, Behrens TW, Graham RR, Criswell LA. Recent advances in the genetics of systemic lupus erythematosus. *Expert Rev Clin Immunol* 2010;6:461-79.
- Stahl EA, Raychaudhuri S, Remmers EF, et al. Genome-wide association study meta-analysis identifies seven new rheumatoid arthritis risk loci. *Nat Genet* 2010;42:508-14.
- Franke A, McGovern DP, Barrett JC, et al. Genome-wide meta-analysis increases to 71 the number of confirmed Crohn's disease susceptibility loci. *Nat Genet* 2010;42:1118-25.
- Zhernakova A, Stahl EA, Trynka G, et al. Meta-analysis of genome-wide association studies in celiac disease and rheumatoid arthritis identifies fourteen non-HLA shared loci. *PLoS Genet* 2011;7(2):e1002004.
- Evans DM, Spencer CC, Pointon JJ, et al. Interaction between ERAP1 and HLA-B27 in ankylosing spondylitis implicates peptide handling in the mechanism for HLA-B27 in disease susceptibility. *Nat Genet* 2011;43:761-7.
- Hüffmeier U, Uebe S, Ekici AB, et al. Common variants at TRAF3IP2 are associated with susceptibility to psoriatic arthritis and psoriasis. *Nat Genet* 2010;42:996-9.
- Jin Y, Birlea SA, Fain PR, et al. Variant of TYR and autoimmunity susceptibility loci in generalized vitiligo. *N Engl J Med* 2010;362:1686-97.
- Jin Y, Birlea SA, Fain PR, et al. Common variants in FOXP1 are associated with generalized vitiligo. *Nat Genet* 2010;42:576-8.
- Mells GF, Floyd JA, Morley KI, et al. Genome-wide association study identifies 12 new susceptibility loci for primary biliary cirrhosis. *Nat Genet* 2011;43:329-32.
- Quan C, Ren YQ, Xiang LH, et al. Genome-wide association study for vitiligo identifies susceptibility loci at 6q27 and the MHC. *Nat Genet* 2010;42:614-8.
- Radstake TR, Gorlova O, Rueda B, et al. Genome-wide association study of systemic sclerosis identifies CD247 as a new susceptibility locus. *Nat Genet* 2010;42:426-9.
- Anderson CA, Boucher G, Lees CW, et al. Meta-analysis identifies 29 additional ulcerative colitis risk loci, increasing the number of confirmed associations to 47. *Nat Genet* 2011;43:246-52.
- Barrett JC, Clayton DG, Concannon P, et al. Genome-wide association study and meta-analysis find that over 40 loci affect risk of type 1 diabetes. *Nat Genet* 2009;41:703-7.
- Ferreira RC, Pan-Hammarström Q, Graham RR, et al. Association of IFIH1 and other autoimmunity risk alleles with selective IgA deficiency. *Nat Genet* 2010;42:777-80.
- Gateva V, Sandling JK, Hom G, et al. A large-scale replication study identifies TNIP1, PRDM1, JAZF1, UHRF1BP1 and IL10 as risk loci for systemic lupus erythematosus. *Nat Genet* 2009;41:1228-33.
- Han JW, Zheng HF, Cui Y, et al. Genome-wide association study in a Chinese Han population identifies nine new susceptibility loci for systemic lupus erythematosus. *Nat Genet* 2009;41:1234-7.
- Hunt KA, Zhernakova A, Turner G, et al. Newly identified genetic risk variants for celiac disease related to the immune response. *Nat Genet* 2008;40:395-402.
- Kochi Y, Okada Y, Suzuki A, et al. A regulatory variant in CCR6 is associated with rheumatoid arthritis susceptibility. *Nat Genet* 2010;42:515-9.
- Petukhova L, Duvic M, Hordinsky M, et al. Genome-wide association study in alopecia areata implicates both innate and adaptive immunity. *Nature* 2010;466:113-7.
- Reveille JD, Sims AM, Danoy P, et al. Genome-wide association study of ankylosing spondylitis identifies non-MHC susceptibility loci. *Nat Genet* 2010;42:123-7.
- Strange A, Capon F, Spencer CC, et al. A genome-wide association study identifies new psoriasis susceptibility loci and an interaction between HLA-C and ERAP1. *Nat Genet* 2010;42:985-90.
- Sun LD, Cheng H, Wang ZX, et al. Association analyses identify six new psoriasis susceptibility loci in the Chinese population. *Nat Genet* 2010;42:1005-9.
- Australia and New Zealand Multiple Sclerosis Genetics Consortium (ANZgene). Genome-wide association study identifies new multiple sclerosis susceptibility loci on chromosomes 12 and 20. *Nat Genet* 2009;41:824-8.
- Gregersen PK. The major histocompatibility complex. In: Goldman L, Schafer AI, eds. *Cecil medicine*. 24th ed. Philadelphia: Saunders Elsevier, 2011:45.
- Silverstein AM. Paul Ehrlich, archives and the history of immunology. *Nat Immunol* 2005;6:639.
- Wing K, Sakaguchi S. Regulatory T cells exert checks and balances on self tolerance and autoimmunity. *Nat Immunol* 2010;11:7-13.
- Shikama N, Nusspaumer G, Hollander GA. Clearing the AIRE: on the pathophysiological basis of the autoimmune polyendocrinopathy syndrome type-1. *Endocrinol Metab Clin North Am* 2009;38:273-88.
- Mathis D, Benoist C. Aire. *Annu Rev Immunol* 2009;27:287-312.
- Bennett CL, Christie J, Ramsdell F, et al. The immune dysregulation, polyendocrinopathy, enteropathy, X-linked syndrome (IPEX) is caused by mutations of FOXP3. *Nat Genet* 2001;27:20-1.
- Wildin RS, Ramsdell F, Peake J, et al. X-linked neonatal diabetes mellitus, enteropathy and endocrinopathy syndrome is the human equivalent of mouse scurfy. *Nat Genet* 2001;27:18-20.
- Dendrou CA, Plagnol V, Fung E, et al. Cell-specific protein phenotypes for the autoimmune locus IL2RA using a genotype-selectable human bioresource. *Nat Genet* 2009;41:1011-5.
- von Boehmer H, Melchers F. Checkpoints in lymphocyte development and autoimmune disease. *Nat Immunol* 2010;11:14-20.
- Menard L, Saadoun D, Isnardi I, et al. The PTPN22 allele encoding an R620W variant interferes with the removal of developing autoreactive B cells in humans. *J Clin Invest* 2011 August 1 (Epub ahead of print).
- Liston A, Lesage S, Gray DH, Boyd RL, Goodnow CC. Genetic lesions in T-cell tolerance and thresholds for autoimmunity. *Immunol Rev* 2005;204:87-101.
- Hasegawa K, Martin F, Huang G, Thomas D, Diehl L, Chan AC. PEST domain-enriched tyrosine phosphatase (PEP) regulation of effector/memory T cells. *Science* 2004;303:685-9.
- Zhang J, Zahir N, Jiang Q, et al. The autoimmune disease-associated PTPN22 variant promotes calpain-mediated Lyp/

- Pep degradation associated with lymphocyte and dendritic cell hyperresponsiveness. *Nat Genet* 2011 August 14 (Epub ahead of print).
40. Zikherman J, Hermiston M, Steiner D, Hasegawa K, Chan A, Weiss A. PTPN22 deficiency cooperates with the CD45 E613R allele to break tolerance on a non-autoimmune background. *J Immunol* 2009;182:4093-106.
  41. Arechiga AF, Habib T, He Y, et al. Cutting edge: the PTPN22 allelic variant associated with autoimmunity impairs B cell signaling. *J Immunol* 2009;182:3343-7.
  42. Vang T, Congia M, Macis MD, et al. Autoimmune-associated lymphoid tyrosine phosphatase is a gain-of-function variant. *Nat Genet* 2005;37:1317-9.
  43. Gregersen PK, Lee HS, Batliwalla F, Begovich AB. PTPN22: setting thresholds for autoimmunity. *Semin Immunol* 2006;18:214-23.
  44. Begovich AB, Caillier SJ, Alexander HC, et al. The R620W polymorphism of the protein tyrosine phosphatase PTPN22 is not associated with multiple sclerosis. *Am J Hum Genet* 2005;76:184-7.
  45. Chapman SJ, Khor CC, Vannberg FO, et al. PTPN22 and invasive bacterial disease. *Nat Genet* 2006;38:499-500.
  46. Lamsyah H, Rueda B, Baassi L, et al. Association of PTPN22 gene functional variants with development of pulmonary tuberculosis in Moroccan population. *Tissue Antigens* 2009;74:228-32.
  47. Bhoj VG, Chen ZJ. Ubiquitylation in innate and adaptive immunity. *Nature* 2009;458:430-7.
  48. Lowe CE, Cooper JD, Brusko T, et al. Large-scale genetic fine mapping and genotype-phenotype associations implicate polymorphism in the IL2RA region in type 1 diabetes. *Nat Genet* 2007;39:1074-82.
  49. The International Multiple Sclerosis Genetics Consortium. Risk alleles for multiple sclerosis identified by a genomewide study. *N Engl J Med* 2007;357:851-62.
  50. van Heel DA, Franke L, Hunt KA, et al. A genome-wide association study for celiac disease identifies risk variants in the region harboring IL2 and IL21. *Nat Genet* 2007;39:827-9.
  51. Festen EA, Goyette P, Scott R, et al. Genetic variants in the region harbouring IL2/JL21 associated with ulcerative colitis. *Gut* 2009;58:799-804.
  52. Cooper JD, Smyth DJ, Smiles AM, et al. Meta-analysis of genome-wide association study data identifies additional type 1 diabetes risk loci. *Nat Genet* 2008;40:1399-401.
  53. Cua DJ, Sherlock J, Chen Y, et al. Interleukin-23 rather than interleukin-12 is the critical cytokine for autoimmune inflammation of the brain. *Nature* 2003;421:744-8.
  54. McGeachy MJ, Cua DJ. Th17 cell differentiation: the long and winding road. *Immunity* 2008;28:445-53.
  55. Duerr RH, Taylor KD, Brant SR, et al. A genome-wide association study identifies IL23R as an inflammatory bowel disease gene. *Science* 2006;314:1461-3.
  56. Cargill M, Schrodi SJ, Chang M, et al. A large-scale genetic association study confirms IL12B and leads to the identification of IL23R as psoriasis-risk genes. *Am J Hum Genet* 2007;80:273-90.
  57. Burton PR, Clayton DG, Cardon LR, et al. Association scan of 14,500 nonsynonymous SNPs in four diseases identifies autoimmunity variants. *Nat Genet* 2007;39:1329-37.
  58. Parham C, Chirica M, Timans J, et al. A receptor for the heterodimeric cytokine IL-23 is composed of IL-12Rbeta1 and a novel cytokine receptor subunit, IL-23R. *J Immunol* 2002;168:5699-708.
  59. Abraham C, Cho JH. Inflammatory bowel disease. *N Engl J Med* 2009;361:2066-78.
  60. Guarner F, Bourdet-Sicard R, Brandtzaeg P, et al. Mechanisms of disease: the hygiene hypothesis revisited. *Nat Clin Pract Gastroenterol Hepatol* 2006;3:275-84.
  61. Preidis GA, Versalovic J. Targeting the human microbiome with antibiotics, probiotics, and prebiotics: gastroenterology enters the metagenomics era. *Gastroenterology* 2009;136:2015-31.
  62. Wen L, Ley RE, Volchkov PY, et al. Innate immunity and intestinal microbiota in the development of type 1 diabetes. *Nature* 2008;455:1109-13.
  63. Hall JC, Rosen A. Type I interferons: crucial participants in disease amplification in autoimmunity. *Nat Rev Rheumatol* 2010;6:40-9.
  64. Graham RR, Kyogoku C, Sigurdsson S, et al. Three functional variants of IFN regulatory factor 5 (IRF5) define risk and protective haplotypes for human lupus. *Proc Natl Acad Sci U S A* 2007;104:6758-63.
  65. Hirschfield GM, Liu X, Han Y, et al. Variants at IRF5-TNPO3, 17q12-21 and MMEL1 are associated with primary biliary cirrhosis. *Nat Genet* 2010;42:655-7.
  66. Liu X, Invernizzi P, Lu Y, et al. Genome-wide meta-analyses identify three loci associated with primary biliary cirrhosis. *Nat Genet* 2010;42:658-60.
  67. Takaoka A, Yanai H, Kondo S, et al. Integral role of IRF-5 in the gene induction programme activated by Toll-like receptors. *Nature* 2005;434:243-9.
  68. Nejentsev S, Walker N, Riches D, Egholm M, Todd JA. Rare variants of IFIH1, a gene implicated in antiviral responses, protect against type 1 diabetes. *Science* 2009;324:387-9.
  69. Hyöty H, Taylor KW. The role of viruses in human diabetes. *Diabetologia* 2002;45:1353-61.
  70. Oliver JE, Silman AJ. What epidemiology has told us about risk factors and aetiopathogenesis in rheumatic diseases. *Arthritis Res Ther* 2009;11:223.
  71. Cooper GS, Wither J, Bernatsky S, et al. Occupational and environmental exposures and risk of systemic lupus erythematosus: silica, sunlight, solvents. *Rheumatology (Oxford)* 2010;49:2172-80.
  72. Klareskog L, Stolt P, Lundberg K, et al. A new model for an etiology of rheumatoid arthritis: smoking may trigger HLA-DR (shared epitope)-restricted immune reactions to autoantigens modified by citrullination. *Arthritis Rheum* 2006;54:38-46.
  73. Mahdi H, Fisher BA, Kallberg H, et al. Specific interaction between genotype, smoking and autoimmunity to citrullinated alpha-enolase in the etiology of rheumatoid arthritis. *Nat Genet* 2009;41:1319-24.
  74. Klareskog L, Catrina AI, Paget S. Rheumatoid arthritis. *Lancet* 2009;373:659-72.
  75. Cui J, Saevarsdottir S, Thomson B, et al. Rheumatoid arthritis risk allele PTPRC is also associated with response to anti-tumor necrosis factor alpha therapy. *Arthritis Rheum* 2010;62:1849-61.
  76. Arbuckle MR, McClain MT, Rubertone MV, et al. Development of autoantibodies before the clinical onset of systemic lupus erythematosus. *N Engl J Med* 2003;349:1526-33.
  77. Griffiths CE, Strober BE, van de Kerkhof P, et al. Comparison of ustekinumab and etanercept for moderate-to-severe psoriasis. *N Engl J Med* 2010;362:118-28.
  78. Perl A. Emerging new pathways of pathogenesis and targets for treatment in systemic lupus erythematosus and Sjogren's syndrome. *Curr Opin Rheumatol* 2009;21:443-7.
  79. Opar A. Kinase inhibitors attract attention as oral rheumatoid arthritis drugs. *Nat Rev Drug Discov* 2010;9:257-8.
  80. Eichler EE, Flint J, Gibson G, et al. Missing heritability and strategies for finding the underlying causes of complex disease. *Nat Rev Genet* 2010;11:446-50.
  81. Orrú V, Tsai SJ, Rueda B, et al. A loss-of-function variant of PTPN22 is associated with reduced risk of systemic lupus erythematosus. *Hum Mol Genet* 2009;18:569-79.
  82. Surolia I, Pirnie SP, Chellappa V, et al. Functionally defective germline variants of sialic acid acetyltransferase in autoimmunity. *Nature* 2010;466:243-7.
  83. Cohen JC, Kiss RS, Pertsemlidis A, Marcel YL, McPherson R, Hobbs HH. Multiple rare alleles contribute to low plasma levels of HDL cholesterol. *Science* 2004;305:869-72.

Copyright © 2011 Massachusetts Medical Society.