

# Genetic basis of autoimmunity

Alexander Marson,<sup>1,2</sup> William J. Housley,<sup>3</sup> and David A. Hafler<sup>3,4</sup>

<sup>1</sup>Diabetes Center and Division of Infectious Diseases, Department of Medicine, UCSF, San Francisco, California, USA. <sup>2</sup>Innovative Genomics Initiative, University of California, Berkeley, California, USA.

<sup>3</sup>Departments of Neurology and Immunobiology, Yale School of Medicine, New Haven, Connecticut, USA. <sup>4</sup>Broad Institute of MIT and Harvard, Cambridge, Massachusetts, USA.

**Autoimmune diseases affect up to approximately 10% of the population. While rare Mendelian autoimmunity syndromes can result from monogenic mutations disrupting essential mechanisms of central and peripheral tolerance, more common human autoimmune diseases are complex disorders that arise from the interaction between polygenic risk factors and environmental factors. Although the risk attributable to most individual nucleotide variants is modest, genome-wide association studies (GWAS) have the potential to provide an unbiased view of biological pathways that drive human autoimmune diseases. Interpretation of GWAS requires integration of multiple genomic datasets including dense genotyping, *cis*-regulatory maps of primary immune cells, and genotyped studies of gene expression in relevant cell types and cellular conditions. Improved understanding of the genetic basis of autoimmunity may lead to a more sophisticated understanding of underlying cellular phenotypes and, eventually, novel diagnostics and targeted therapies.**

## Introduction

Human autoimmune diseases are a major health issue, affecting up to approximately 10% of the population (1). Common human autoimmune diseases are complex disorders that arise from the interactions between polygenic risk factors and environmental factors (2). Intensive investigation of autoimmune disease genetics has the potential to offer an unbiased view of the underlying etiologies of these conditions and, perhaps, identify therapeutic targets. Our early understanding of the disease heritability derived from high rates of autoimmune disease concordance in twins (3, 4) or first-degree family members (5, 6) compared with nonbiological relatives with a shared environment (7). Despite this recognition that autoimmune disease risk is influenced by genetics, it has been challenging to identify the causal nucleotide variants and their functional effects.

The sequencing of the human genome and rapidly emerging genomic technologies are enabling comprehensive interrogation of genetic variants that contribute to autoimmune disease risk. Our understanding of the genetic basis of human autoimmune disease has expanded dramatically in the last 15 years. Here we review biological lessons from genetic studies of human autoimmune diseases. Rare monogenic autoimmune disease syndromes have revealed highly penetrant mutations that disrupt essential mechanisms of central and peripheral immune tolerance. Genome-wide association studies (GWAS) have provided insight into the more subtle immune dysregulation caused by common genetic variants that contribute risk of autoimmunity.

We focus on how new genome sequencing technologies are providing frameworks for the interpretation of GWAS. Human

genetic studies of autoimmune diseases benefit from the integration of multiple genomic datasets, including dense genotyping, epigenomic annotation of functional elements in primary immune cells, and large quantitative studies of gene expression in the relevant cell types and cellular conditions. An improved understanding of the genetic basis of autoimmunity will likely lead to a more sophisticated understanding of the cellular phenotypes underlying autoimmune diseases and, eventually, novel diagnostics and targeted therapies.

## Monogenic immune dysregulation syndromes

Rare autoimmune disease syndromes have provided insight into biological pathways necessary for the maintenance of immune homeostasis (8). Although far less prevalent than polygenic autoimmune diseases, patients with Mendelian immune dysregulation syndromes caused by monogenic mutations have been identified. Successful linkage analysis and positional cloning combined with mouse genetic models have identified causal mutations. Notably, studies of patients with immune dysregulation, polyendocrinopathy, enteropathy, X-linked (IPEX) syndrome have highlighted the crucial role of FOXP3 in Treg development and function (9–12). The study of patients with autoimmune polyendocrinopathy-candidiasis-ectodermal dystrophy (APECED; also known as autoimmune polyendocrine syndrome type 1) revealed the essential role of autoimmune regulator (AIRE) in thymic selection and central tolerance (13–18). These highly penetrant mutations in key protein-coding genes suggest pathways, relevant cell types, and mechanisms of tolerance that may undergo more subtle forms of dysregulation as a result of common genetic variation.

Mutations in the gene encoding *CTLA4* have recently been discovered in families with Mendelian multi-organ autoimmune disease syndromes (19–21). These rare mutations (including missense, nonsense, and splice variants) cause severe disease in heterozygous patients, albeit with incomplete penetrance. The mutations are associated with impairment of Treg suppressive function and

**Conflict of interest:** David A. Hafler has consulted for Allergan Pharmaceuticals, Bristol-Myers Squibb, EMD Serono, Genzyme Corporation, MedImmune, Mylan, Novartis Pharmaceuticals, Questcor, and Teva Pharmaceuticals and has received grant support from Bristol-Myers Squibb.

**Reference information:** *J Clin Invest.* 2015;125(6):2234–2241. doi:10.1172/JCI78086.

immune dysregulation. Common genetic variants in the *CTLA4* locus are associated with more modest increases in risk of autoimmune diseases including type 1 diabetes (T1D) and Graves' disease (22, 23). The functional consequences of a common risk allele (SNP rs3087243 located in the 3'-UTR of the *CTLA4*; A/G) were investigated using phosphorylated site-specific mAbs targeting components of TCR signaling in naive and memory T cells. The relative responsiveness to TCR stimulation, as assessed by phosphorylation levels of downstream signaling molecules, was altered in naive (CD4<sup>+</sup>CD45RA<sup>hi</sup>) and memory (CD4<sup>+</sup>CD45RA<sup>lo</sup>) T cells obtained from individuals with the disease susceptibility allele at *CTLA4*. This was among the early reports that allelic variation associated with autoimmune disease can alter the signaling threshold of CD4<sup>+</sup> T cells (24). Taken together, these findings are consistent with a spectrum of allelic variants at the same locus causing differing degrees of immune dysregulation and autoimmune disease risk.

### The era of GWAS

Identification and characterization of the common variants in autoimmune diseases has been more challenging because of their relatively small individual contributions to disease risk. Common variants have not been purified by negative selection, perhaps a result of their somewhat more modest biological effects as compared with gene disruptions seen in Mendelian diseases. A few notable loci, including *MHC* on chromosome 6p21 (25–27) and the *NOD2* locus (28–30), were initially linked to autoimmunity due to their relatively high odds ratios. However, identifying a larger set of loci that contribute to complex autoimmune disease risk has required the development of new genetic tools.

GWAS are large, case-control studies designed to detect variants that confer a modest risk of common diseases as opposed to rare diseases caused by highly penetrant mutations (31). The basic approach is to identify genetic variants that are preferentially associated with patients with a disease or trait relative to healthy individuals. GWAS using SNPs from the International HapMap Project (<http://hapmap.ncbi.nlm.nih.gov/>) have allowed an unbiased approach in scanning the whole genome and identifying disease-associated regions, with particular success in identifying a large number of loci associated with human autoimmune diseases.

Thousands of loci have been linked to hundreds of human diseases by GWAS (32, 33). Over 20 autoimmune diseases have been subjected to large GWAS that have successfully implicated risk loci of genome-wide significance. GWAS rapidly revealed shared genetic associations among the different autoimmune diseases (34–36). We recently reported that about two-thirds of loci associated with autoimmunity were shared risk factors for multiple autoimmune diseases (37). These findings are consistent with some shared pathological features of disparate autoimmune diseases and familial clustering of multiple autoimmune conditions.

Perhaps one of the more striking observations in applying genetics to autoimmune diseases is that, while the odds ratio to disease risk for each of the genetic variants is small (most have odds ratios less than 1.2), putative etiological pathways have begun to emerge based on the genes in associated genomic regions (31). Deeper understanding requires more thorough investigation of the specific mechanisms by which these genetic variants influence the biology of human disease.

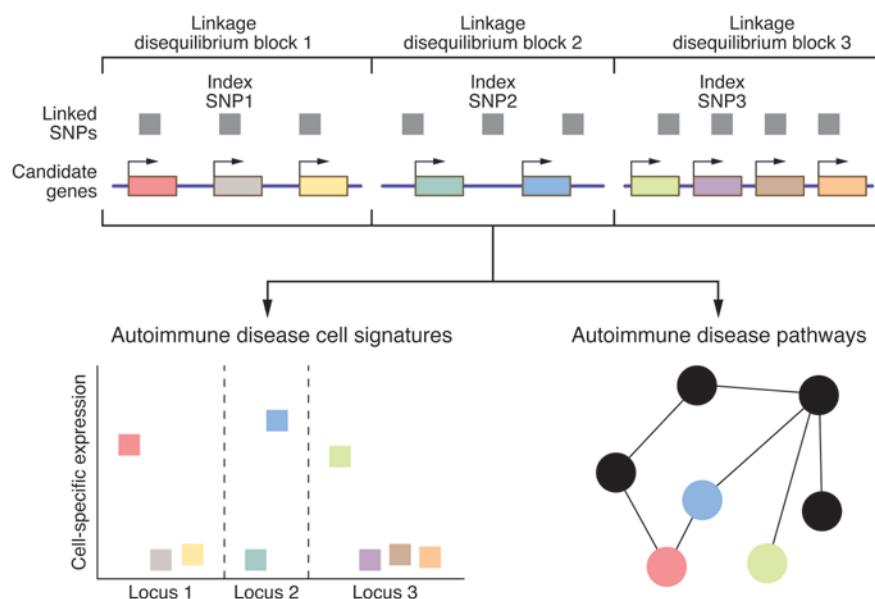
### Biological interpretation of GWAS loci

Biological insight from the thousands of loci implicated by GWAS depends on analytic tools for interpretation (Figure 1). A model whereby many common SNPs exert modest effects on a shared biological process predicts that disease-associated SNPs would exert their effect on a limited number of pathogenic cell types. Although common SNPs are inherited in the germ line, they could affect genes that are selectively expressed in particular cell types. Indeed, GWAS loci associated with autoimmune disease are enriched in genes that are preferentially expressed in particular immune subsets (38). A comparison of disease loci with gene expression in 223 cell types of the ImmGen dataset (39) revealed that SLE loci encode genes preferentially expressed in B cell subsets, whereas RA-associated loci are most strongly enriched in genes preferentially expressed in CD4<sup>+</sup> effector memory T cells. Furthermore, knowledge of these pathogenic cell types may suggest biologically relevant genes affected by GWAS variants that do not meet strict genome-wide significance thresholds; shared expression patterns could help to triage GWAS results. Subsequent analysis of human RNAseq data has confirmed enrichment of genes in loci associated with immune-mediated diseases that are preferentially expressed in CD4<sup>+</sup> T cell subsets (40).

Not only are candidate genes in GWAS loci coexpressed in pathogenic cell types, but they also encode proteins that participate in a disproportionate number of physical interactions that form biological pathways (41). In addition, integration of multiple annotation tools may help to extract enriched gene sets from loci linked to disease phenotypes (42, 43). Collectively, these efforts have two important goals: triage of candidate genes with linkage disequilibrium blocks and discovery of protein complexes and pathways that are affected by disease variants. Ultimately, discovery of such pathways will highlight the biology underlying complex autoimmune diseases and may direct drug discovery efforts. Protein interaction-based pathway analysis has suggested candidate causal genes within GWAS loci and has implicated key pathogenic pathways, including annotated pathways for JAK/STAT signaling and TCR signaling (44).

### The influence of disease variants on cellular phenotype

Pathway analysis has begun to reveal patterns among the loci linked to autoimmune diseases, but there is also a need to understand the functional mechanisms by which specific nucleotide variants contribute to disease risk. This has been challenging because the causal SNPs that contribute to autoimmune disease risk tend to be inherited along with neighboring neutral SNPs as a result of linkage disequilibrium. The index SNPs that are genotyped and associated with disease risk in GWAS implicate genomic loci comprised of multiple linked SNPs (Figure 1). Nonetheless, a number of studies have investigated the functional impact of polymorphisms. In one commonly used approach to correlate genotype with cellular phenotype, healthy subjects homozygous for either the risk-associated or protective haplotype are investigated. As the phenotypes are closely related to the genotypes, the numbers of subjects required are significantly less than what is required for elucidation of disease risk. However, in these phenotypic studies on a small number of subjects, the genetic



**Figure 1. Moving from GWAS loci to cellular pathways.** The causal SNPs that contribute to autoimmune disease risk are often inherited along with neighboring neutral SNPs as a result of linkage disequilibrium. The index SNPs that are genotyped and associated with disease risk in GWAS implicate genomic loci – linkage disequilibrium blocks – composed of multiple linked SNPs (gray boxes). GWAS loci associated with autoimmune disease are enriched in genes (rectangles) that are preferentially expressed in particular immune cell subsets (autoimmune disease cell signatures; bottom left) (38) and encode proteins (circles) that participate in a disproportionate number of direct and indirect physical interactions to form biological pathways (autoimmune disease pathways; bottom right) (41). Expression patterns and protein interaction network analysis have been used to triage candidate genes within linkage disequilibrium blocks. These analyses also suggest pathogenic cell types, protein complexes, and pathways that are affected by disease variants, which begins to elucidate the biology underlying complex autoimmune diseases and could direct drug discovery efforts. Adapted with permission from *American Journal of Human Genetics* (38) and *PLoS Genetics* (41).

background cannot be controlled, making it challenging to dissect the functional consequences of any individual nucleotide variant. Studies on human cells with naturally arising genetic variation have been complemented by genetic engineering of mice and human cell lines, although not all relevant phenotypes can be captured in these systems. Looking forward, targeted genome editing of disease variants (termed “SNP editing”) in primary human cell types associated with specific diseases is likely to open new possibilities in understanding pathogenesis of autoimmunity.

We discuss several examples (*PTPN22*, *CD6*, *TNFRSF1A*, *NFKB*, *CD25*) of progress and key remaining challenges in determining biologic function from GWAS signals, followed by a discussion of emerging strategies to develop more comprehensive elucidation of the cellular consequences of genetic variation.

***PTPN22: shared risk by regulation of multiple immune signals.*** One of the best-characterized genetic variants linked to multiple autoimmune diseases is a nonsynonymous coding SNP in the *PTPN22* gene, which encodes a critical immunoregulatory tyrosine phosphatase. In T cells, *PTPN22* negatively regulates TCR signaling by dephosphorylation of the TCR coreceptor CD3 and the tyrosine kinases LCK, FYN, and ZAP70 (45). In myeloid cells, *PTPN22* modulates TLR signaling by regulating phosphorylation and ubiquitination of TNF receptor-associated factor 3 and promoting type I IFN production (46). Despite strong evi-

dence implicating this specific nucleotide substitution in *PTPN22* as a causal autoimmunity variant, dissecting the specific cellular phenotype or phenotypes has remained challenging.

An 1858C>T polymorphism in *PTPN22* is strongly associated with many autoimmune diseases including T1D, SLE, Graves’ disease, and RA (47–51). Of note, the risk variant for these diseases is not associated with multiple sclerosis (MS) and may be a protective variant in Behcet’s disease and Crohn’s disease (52–56). The 1858C>T variant results in an arginine-to-tryptophan substitution at position 620 (R620W). The impact of the *PTPN22* risk variant on TCR signaling remains controversial, with studies suggesting both enhancement and suppression of TCR activation (57, 58). Functional interpretation is further complicated by the possibility of distinct signaling effects in thymic development versus peripheral stimulation and disparate effects on various T cell compartments including naive, memory, and effector subsets and Tregs (59). In B cells, carriers of the risk allele have altered B cell receptor (BCR) signaling and a defect in both central and peripheral B cell tolerance checkpoints, resulting in expansion of autoreactive B cells (60, 61). Through its effects in myeloid cells, the *PTPN22* R620W variant

may also contribute to autoimmunity by altering TLR signaling and impairing type I IFN induction (46).

***CD6: genetic modulation of immune receptor isoforms.*** *CD6* is a transmembrane glycoprotein expressed on both B and T cells. The ligand for *CD6*, activated leukocyte adhesion molecule (ALCAM), is expressed on activated APCs. Interaction of *CD6* and ALCAM results in phosphorylation of lymphocyte cytosolic protein 2 (SLP-76) and T cell costimulation (62). A GWAS meta-analysis of MS patients identified a variant (rs17824933) that falls within a broad linkage haplotype containing *CD6* (63). Homozygous carriers have diminished expression of the full-length *CD6* and an increase in a shortened *CD6Δ3* isoform. This isoform lacks the ALCAM binding site, and ligation of *CD6* with ALCAM results in decreased proliferation in *CD4<sup>+</sup>* T cells that are homozygous for the risk variant (rs17824933 GG) (64).

***TNFRSF1A/NF-κB: genetic modulation of an inflammatory pathway.*** The type 1 TNF receptor (*TNFR1*, encoded by *TNFRSF1A*) binds TNF- $\alpha$  and lymphotoxin- $\alpha$ 3 to initiate an NF- $\kappa$ B-mediated proinflammatory pathway or a caspase-3-mediated apoptotic pathway. GWAS identified the rs1800693 variant within *TNFRSF1A* as strongly associated with MS (63, 65, 66). The strong association signal and the low level of linkage disequilibrium in this region suggest that this variant is the most likely causative SNP. The SNP falls within a splice acceptor site upstream of exon 6, resulting in loss of

exon 6 and a premature stop codon. The result is a shortened splice variant in approximately 10% of transcripts. Full-length TNFR1 traffics normally to the cell surface, while the shortened variant accumulates within the cell. The C risk allele results in increased expression of IFN- $\gamma$ -inducible protein 10 (IP10; also known as CXCL10) from monocytes after stimulation with TNF- $\alpha$ , suggestive of an increased signaling response through TNFR1 (67). Gregory et al. suggested that the risk genotype leads to increased cleavage of TNFR1 from the cell surface, resulting in a soluble decoy receptor capable of blocking TNF- $\alpha$  (68); however, other studies have not observed changes in soluble TNFR1 (67, 69).

The transcription factor NF- $\kappa$ B is a central regulator of inflammation, and a significant number of variants within the NF- $\kappa$ B signaling cascade have been identified in MS and ulcerative colitis. We recently found that MS-associated variants proximal to *NFKB1* (rs228614) and in *TNFRSF1A* (*TNFR1*, rs1800693) are associated with increased NF- $\kappa$ B signaling after TNF- $\alpha$  stimulation (70). Both variants result in increased degradation of I $\kappa$ B $\alpha$ , a negative regulator of NF- $\kappa$ B and nuclear translocation of p65 NF- $\kappa$ B. The variant proximal to *NFKB1* controls signaling responses by altering expression of NF- $\kappa$ B itself, with the GG risk genotype expressing 20-fold more p50 NF- $\kappa$ B. Thus, genetic variants associated with risk of developing MS alter NF- $\kappa$ B signaling pathways and result in enhanced NF- $\kappa$ B activation and greater responsiveness to inflammatory stimuli.

**IL-2 receptor  $\alpha$  (*CD25*): complex effects on immune regulation.** Mechanistic characterization of causal variants at the IL-2 receptor  $\alpha$  (*IL2RA*; *CD25*) locus is notably complex, as multiple distinct variants have been associated with different autoimmune diseases. IL2R is a heterotrimeric receptor comprising the IL2R $\beta$  chain (CD122), the common  $\gamma$  chain (CD132), and the high-affinity IL-2Ra chain (CD25). CD25 is upregulated rapidly on T cells after stimulation and is constitutively expressed on Tregs. IL-2 signaling through CD25 stimulates T cell proliferation via phosphorylation of STAT5. Implicated variants at *CD25* appear to map to regulatory rather than coding sequences. Recent work suggests that this is a so-called “super-enhancer” locus, where a dense cluster of putative regulatory elements are active in T cells (71).

Multiple studies have attempted to elucidate how genetic variants in this locus affect *CD25*. For example, the rs2104286 risk allele in *CD25* is associated with both MS and T1D, suggesting that alterations in IL-2 signaling may represent a common mechanism across autoimmune diseases. Although the risk variant is associated with increased expression of *CD25* on effector T cells and Tregs (72–74), it is also associated with decreased STAT5 signaling downstream of IL-2 (73). Despite diminished IL-2 signaling through STAT5, the MS risk variant was associated with expression of GM-CSF after IL-2 stimulation (72). Recent investigation of five candidate loci within the 5' flanking region and the first intron of *CD25* showed marked effects on *CD25* expression, with the A allele associated with lower transcript levels and altered transcription factor binding (75). This suggests that variants within the *CD25* locus can affect transcription factor binding sites and alter enhancer activity.

### Novel approaches to determining the significance of genetic variation

A systematic approach to interrogate the function of genetic variants, including noncoding variants, is needed. This type of

approach requires identification of the causative SNPs in each haplotype. Use of newer Bayesian approaches to identify causative SNPs (discussed below) in combination with tools such as RNAseq, chromatin mapping, proteomics, and eventually SNP editing in relevant cell types may allow functional testing of the genetic variants found in individual patients to reveal disease-specific pathways for each subject.

### Genetic fine-mapping

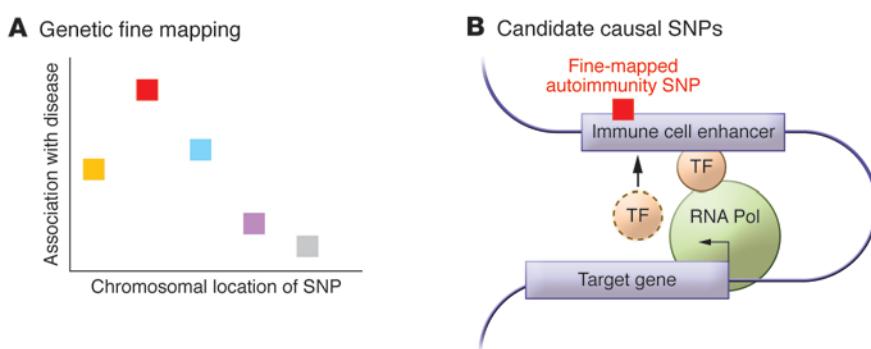
While GWAS have successfully identified broad risk loci, linkage disequilibrium limits their ability to pinpoint causal nucleotide variants within these loci. A risk locus associated with autoimmunity may encode as dozens of proteins as well as noncoding transcripts and regulatory elements. Likely candidate genes within these loci have been identified by systematic efforts to characterize the biological pathways affected by multiple autoimmunity risk variants. Now major efforts are underway to fine map the causal variants based on improved genetic data. Fine mapping requires significantly denser genotyping data than was provided by the initial GWAS studies.

The Immunochip was an international collaborative effort involving investigators of multiple immune-mediated diseases to create a platform to interrogate the autoimmune loci identified to date (76). Whereas initial GWAS genotyping was accomplished with tagging SNPs spread across the whole genome at low density, the Immunochip was an efficient genotyping platform designed to deeply interrogate 186 non-MHC loci with genome-wide significant associations with at least one autoimmune disease (104,425 SNPs, median coverage 486 SNPs/region). It also provided lighter coverage of other genomic regions with suggestive association evidence (49,198 SNPs). GWAS has been replicated for multiple autoimmune diseases, with dense genotyping data generated with large association studies on the Immunochip. This includes MS (66), celiac disease (77), autoimmune thyroid disease (78), primary biliary cirrhosis (PBC) (79), ankylosing spondylitis (80), atopic dermatitis (81), primary sclerosing cholangitis (82), juvenile idiopathic arthritis (JIA) (83), psoriasis (84), inflammatory bowel disease (85), and T1D (23).

Dense genotyping enables statistical analysis to identify a set of candidate causal or “credible” SNPs for a disease phenotype. A Bayesian analysis was pioneered to refine the association signal at 14 loci associated with three different diseases to a set of credible SNPs (86). We recently extended this work to develop a new algorithm for fine-mapping causal variants based on genetic evidence (37). Probabilistic identification of causal SNPs (PICS) is a Bayesian algorithm modeled on dense genotyping data from the recent Immunochip study of MS (66). We were able to assign explicit probabilities of causality to each SNP within GWAS loci based on strength of disease association and haplotype structure. Importantly, we were able to predict causal variants for GWAS data even when dense genotyping data were not available based on imputation to the 1000 Genomes Project (87). PICS is therefore broadly applicable in refining GWAS loci for a wide range of diseases and phenotypes to a tractable set of candidate causal SNPs.

### Functional interpretation of noncoding variants

Long-standing knowledge of the genetic code provides a powerful framework to understand deleterious effects of the variants that



**Figure 2. Integration of genetic and epigenetic fine mapping reveals candidate causal SNPs and pathogenic cell circuits.** (A) The goal of genetic fine mapping is to progress from multiple linked SNPs to one (or very few) candidate causal SNPs. Bayesian algorithms have been developed to fine map credible sets of candidate causal SNPs (37, 86) on the basis of dense genotyping data (e.g., Immunochip data [ref. 76]) or on the basis of imputation of sparser genotyping data to the 1000 Genomes Project (87). (B) Approximately 90% of causal variants associated with autoimmune diseases are noncoding. Genome-wide chromatin maps of active regulatory elements in primary immune cells in both resting and stimulated conditions serve as a powerful resource to identify functional noncoding elements that can be disrupted by disease-associated variants. Candidate causal disease variants (red box) map to regulatory elements (notably enhancers) active in specialized cell types, especially stimulated CD4<sup>+</sup> T cell subsets. Enhancers contain binding sites for transcription factors and regulate transcription through long-range interactions with RNA polymerase (RNA Pol) machinery (88). The locations of disease-associated SNPs highlight pathogenic cell types and cell conditions based on the activity patterns of affected enhancers. Quantitative trait studies are beginning to characterize the functional effects of causal autoimmune disease variants in modulating transcription factor (TF) binding (circles), chromatin state, target gene regulation, and cellular phenotype. Detailed studies might eventually diagnose specific gene regulatory defects caused by autoimmune disease variants and provide novel targets for therapeutic intervention. Adapted with permission from *Immunity* (116).

alter the amino acid code. However, approximately 90% of causal autoimmune disease variants appear to be noncoding (37). Until recently, the mechanisms by which noncoding DNA variation contributes to autoimmune disease have defied understanding.

Recent studies, including systematic efforts of large consortia, have revealed that intergenic regions are densely populated with hundreds of thousands of *cis*-regulatory elements, including enhancers that shape cell type-specific gene expression programs; mounting evidence suggests that noncoding variation that contributes to the risk of autoimmunity may affect these regulatory elements to cause disease (88). Enhancers are distal *cis*-regulatory elements (in contrast to promoters, which are *cis*-regulatory elements proximal to transcriptional start sites) that are essential for proper control of gene expression programs controlling cell identity and condition-specific cellular responses (89). Enhancers contain binding sites for transcription factors and regulate transcription through long-range interactions with RNA polymerase machinery. Their incomplete evolutionary conservation (90) and genomic distance from protein-coding genes have hampered efforts to identify these elements. However, recently stereotyped chromatin patterns of active and poised enhancers have been identified, enabling large-scale sequencing efforts to map regulatory elements throughout the human genome (88, 91, 92).

Genetic variation in humans might dysregulate transcriptional circuits to alter specialized cell functions and contribute to disease risk. This is in accordance with reports that direct targets of key transcriptional regulators in the immune system include genes that have been implicated by GWAS of autoimmunity

(93–95). Disease genetics could reveal key regulatory circuits in immune cells that fail in autoimmune diseases as a result of noncoding nucleotide variants, shedding new light on the underlying etiologies of human autoimmune diseases.

**Epigenomic annotation of noncoding variants.** Initial efforts have suggested that *cis*-regulatory maps may provide a useful interpretative framework for functional annotation of noncoding disease variants (71, 91, 96–99). However, two major challenges have limited efforts to identify and characterize noncoding variants: (a) GWAS studies do not readily distinguish causal variants from neighboring noncausal variants inherited in linkage disequilibrium and (b) epigenomes and transcriptomes must be mapped in diverse human cell types under multiple physiological conditions to identify the full range of regulatory elements and expression programs that may be affected by noncoding variants to cause disease.

Computational fine mapping approaches based on dense genotyping and imputation to the 1000 Genomes project provide a new opportunity to map causal autoimmune disease variants to functional noncoding elements in immune cells (refs. 23, 37, and A). To interpret fine-mapped autoimmune disease variants, we generated a large epigenomic resource of active regulatory elements in well-defined, primary immune cells in both resting and stimulated conditions. These epigenomic annotations allowed us to map disease variants to regulatory elements active in specialized cell types, especially stimulated CD4<sup>+</sup> T cell subsets (Figure 2B). Although the candidate causal SNPs tend to occur in immune enhancers — often at sites bound by multiple key transcription factors — in most cases, they do not directly disrupt or create recognizable transcription factor binding motifs. Instead, the genomic sites of disease SNPs strongly suggest the importance of noncanonical sequences with crucial roles in immune cell gene regulation (37).

We were able to cluster human diseases based on overlap strength between disease-associated SNPs and cell type-specific enhancer locations. Enhancer activity patterns displayed strict cell type (and even cell condition) restrictions. As a result, the locations of disease-associated SNPs highlighted pathogenic cell types and cell conditions based on the activity patterns of affected enhancers. Genetic data for more diseases and epigenomic data for more cell types under more conditions will continue to fill in a comprehensive map of human diseases by pathogenic cell signatures.

**Dysregulation of gene expression by disease-associated variants.** Genetic mapping of quantitative traits including gene expression is a complementary approach to epigenomic annotation of functional genomic elements. Considerable efforts have linked expression quantitative trait loci (eQTL) to the loci implicated

in disease risk (100, 101). Genetic factors affecting immune cell numbers may also contribute to autoimmune disease risk (102). We found that approximately 10% of fine-mapped variants contributing to autoimmune disease risk also contribute to alterations of transcript levels in a population-based study of the transcriptome in peripheral blood cells (37, 103). Despite the statistically significant overlap between disease risk and eQTL, the analysis also suggests that many of the regulatory effects of disease variants will only be observable in restricted cellular contexts. This is consistent with the finding of pervasive response eQTL (reQTL), where expression phenotypes are unmasked in response to selective stimulation conditions (104). Indeed, several recent studies have discovered sets of human reQTL variants that shape transcriptome responses to immune cell stimulation and polarization, demonstrating various degrees of overlap with the loci implicated by GWAS studies of immune-mediated diseases (75, 105–107). eQTL in CD4<sup>+</sup> T cells overlap with disease-associated SNPs for RA and MS, whereas SNPs associated with Alzheimer's and Parkinson's diseases preferentially overlap with eQTL in monocytes, underscoring the cell type-specific dysregulation of gene expression that likely contributes to disease phenotype (108).

Transcription factor binding studies in inbred mouse strains revealed that genetic variation can significantly alter the genome-wide binding landscape of key immune regulators (109). Quantitative trait studies in human cells suggest that genetic variation has detectable effects of transcription factor binding, gene expression, and chromatin state of regulatory elements (110). Although not statistically significant, recent work suggests that genetic variants associated with asthma risk can be found at Th2 cell enhancers that are differentially active between asthma patients and healthy controls (111). Such studies are beginning to characterize the functional effects of causal autoimmune disease variants in modulating transcription factor binding, chromatin state, gene regulation, and cellular phenotype. Detailed studies might eventually diagnose specific defects of autoimmune disease variants and provide novel targets for therapeutic intervention.

### Concluding remarks

Genetic studies of monogenic immune dysregulation syndromes have informed our understanding of essential mechanisms of immune tolerance. New genomic technologies developed over the past decade are now enabling systematic studies of complex autoimmune diseases. Hundreds of relevant loci have been implicated by autoimmune disease GWAS. One approach to understand pathological consequences of GWAS variation is phenotypic analysis of cells from a cohort of genotyped individuals. Here we also discussed systems-scale approaches that have allowed progress from GWAS hits to pathogenic pathways, transcriptional circuits, and immune cell types in specific cellular conditions.

Advances in genomic annotation and integrative analyses should accelerate efforts to connect the human genotypes associated with disease to specific cellular phenotypes. Extensive work has linked specific genetic variations to alterations in amino acid code, transcription, and splice regulation, especially

in genes important for cytokine signaling and TCR signaling pathways. There is a growing set of candidate causal SNPs to examine and improved information to inform hypotheses about which biological pathways will be affected. An improved understanding of pathogenic pathways and cell types may eventually help to reveal rare disease-associated variants, epistatic interactions among disease-associated SNPs, and gene/environment interactions, which have been challenging to identify and may contribute to heritability that remains unexplained by GWAS (112). Integrative analyses of autoimmune disease genetics have helped to prioritize likely causal genes in GWAS loci, suggesting not only key biological pathways, but perhaps even drug targets and candidate therapies (113). Looking forward, there is hope that such studies will uncover effects of genetic variants that contribute not only to disease risk, but also to heterogeneity in disease progression (114) and responses to therapy.

Even with rapid progress in genetic analysis, functional studies are essential to advance our understanding of disease genetics (115). Targeted genome editing of disease variants in human cells and animal models will provide opportunities for functional assessment of genetic effects on specific cellular pathways that contribute to autoimmune pathology. These studies will benefit from global analyses of cell types and conditions that are relevant for study; however, considerable work will be required to dissect the subtle and pleiotropic effects of disease variants, which may only be observable in restricted cellular contexts.

Improved analytic and experimental techniques also raise hopes for clinical applications of autoimmune genetics. A refined understanding of the causal variants contributing to autoimmune disease pathology could aid in the development of new diagnostic tests for disease risk and, perhaps, biomarkers for disease progression or therapeutic response. Genetic insight into key pathogenic circuits will focus future drug discovery efforts to correct the specific biochemical and epigenetic pathways that are dysregulated in human autoimmune diseases.

### Acknowledgments

We thank C.J. Ye, G.E. Haliburton, and K.K. Farh for suggestions on the manuscript. The review also benefitted from our conversations with all members of the Marson and Hafler laboratories. We are grateful for support from a UCSF Sandler Fellowship (to A. Marson); a gift from Jake Aronov (to A. Marson); a National MS Society Collaborative Research Centre Award (CA1061-A-18, to D.A. Hafler and A. Marson); a Life Sciences Research Foundation Pfizer Fellowship (to W.J. Housley); NIH grants P01 AI045757, U19 AI046130, U19 AI070352, and P01 AI039671 (to D.A. Hafler); the Penates Foundation (to D.A. Hafler); and the Nancy Taylor Foundation for Chronic Diseases Inc. (to D.A. Hafler).

Address correspondence to: Alexander Marson, University of California, San Francisco, 513 Parnassus Ave., Health Sciences West, Room 1053, San Francisco, California 94143, USA. Phone: 415.502.2611; E-mail: alexander.marson@ucsf.edu. Or to: David Hafler, Yale University School of Medicine, 300 George Street, New Haven, Connecticut 06511, USA. Phone: 203.737.4802; E-mail: david.hafler@yale.edu.

**REVIEW SERIES: AUTOIMMUNITY****The Journal of Clinical Investigation**

1. Cooper GS, Bynum ML, Somers EC. Recent insights in the epidemiology of autoimmune diseases: improved prevalence estimates and understanding of clustering of diseases. *J Autoimmun*. 2009;33(3-4):197-207.
2. Vyse TJ, Todd JA. Genetic analysis of autoimmune disease. *Cell*. 1996;85(3):311-318.
3. Sadovnick AD, et al. A population-based study of multiple sclerosis in twins: update. *Ann Neurol*. 1993;33(3):281-285.
4. Ebers GC, et al. A population-based study of multiple sclerosis in twins. *N Engl J Med*. 1986;315(26):1638-1642.
5. Sadovnick AD, Baird PA, Ward RH. Multiple sclerosis: updated risks for relatives. *Am J Med Genet*. 1988;29(3):533-541.
6. Sadovnick AD, Baird PA. The familial nature of multiple sclerosis: age-corrected empiric recurrence risks for children and siblings of patients. *Neurology*. 1988;38(6):990-991.
7. Ebers GC, Sadovnick AD, Risch NJ. A genetic basis for familial aggregation in multiple sclerosis. Canadian Collaborative Study Group. *Nature*. 1995;377(6545):150-151.
8. Cheng MH, Anderson MS. Monogenic autoimmunity. *Annu Rev Immunol*. 2012;30:393-427.
9. Bennett CL, et al. The immune dysregulation, polyendocrinopathy, enteropathy, X-linked syndrome (IPEX) is caused by mutations of FOXP3. *Nat Genet*. 2001;27(1):20-21.
10. Fontenot JD, Gavin MA, Rudensky AY. Foxp3 programs the development and function of CD4<sup>+</sup>CD25<sup>+</sup> regulatory T cells. *Nat Immunol*. 2003;4(4):330-336.
11. Hori S, Nomura T, Sakaguchi S. Control of regulatory T cell development by the transcription factor Foxp3. *Science*. 2003;299(5609):1057-1061.
12. Wildin RS, et al. X-linked neonatal diabetes mellitus, enteropathy and endocrinopathy syndrome is the human equivalent of mouse scurfy. *Nat Genet*. 2001;27(1):18-20.
13. Anderson MS, et al. Projection of an immunological self shadow within the thymus by the aire protein. *Science*. 2002;298(5597):1395-1401.
14. Nagamine K, et al. Positional cloning of the APECED gene. *Nat Genet*. 1997;17(4):393-398.
15. Klamp T, Sahin U, Kyewski B, Schwendemann J, Dhaene K, Tureci O. Expression profiling of autoimmune regulator AIRE mRNA in a comprehensive set of human normal and neoplastic tissues. *Immunol Lett*. 2006;106(2):172-179.
16. Heino M, et al. RNA and protein expression of the murine autoimmune regulator gene (Aire) in normal, RelB-deficient and in NOD mouse. *Eur J Immunol*. 2000;30(7):1884-1893.
17. Heino M, et al. Autoimmune regulator is expressed in the cells regulating immune tolerance in thymus medulla. *Biochem Biophys Res Commun*. 1999;257(3):821-825.
18. Altonen J, Björkes P, Sandkuyl L, Perheentupa J, Peltonen L. An autosomal locus causing autoimmune disease: autoimmune polyglandular disease type I assigned to chromosome 21. *Nat Genet*. 1994;8(1):83-87.
19. Schubert D, et al. Autosomal dominant immune dysregulation syndrome in humans with CTLA4 mutations. *Nat Med*. 2014;20(12):1410-1416.
20. Zeissig S, et al. Early-onset Crohn's disease autoimmunity associated with a variant in CTLA-4 [published online ahead of print November 3, 2014]. *Gut*. doi:10.1136/gutjnl-2014-308541.
21. Kuehn HS, et al. Immune dysregulation in human subjects with heterozygous germline mutations in CTLA4. *Science*. 2014;345(6204):1623-1627.
22. Chu X, et al. A genome-wide association study identifies two new risk loci for Graves' disease. *Nat Genet*. 2011;43(9):897-901.
23. Onengut-Gumuscu S, et al. Fine mapping of type 1 diabetes susceptibility loci evidence for colocalization of causal variants with lymphoid gene enhancers. *Nat Genet*. 2015;47(4):381-386.
24. Maier LM, Anderson DE, De Jager PL, Wicker LS, Hafler DA. Allelic variant in CTLA4 alters T cell phosphorylation patterns. *Proc Natl Acad Sci U S A*. 2007;104(47):18607-18612.
25. Brewerton DA, Caffrey M, Nicholls A, Walters D, James DC. Acute anterior uveitis and HL-A 27. *Lancet*. 1973;302(7836):994-996.
26. Brewerton DA, Caffrey M, Nicholls A, Walters D, Oates JK, James DC. Reiter's disease and HL-A 27. *Lancet*. 1973;302(7836):996-998.
27. Woodrow JC. HL-A 27 and Reiter's syndrome. *Lancet*. 1973;2(7830):671-672.
28. Hampe J, et al. Association between insertion mutation in NOD2 gene and Crohn's disease in German and British populations. *Lancet*. 2001;357(9272):1925-1928.
29. Hugot JP, et al. Association of NOD2 leucine-rich repeat variants with susceptibility to Crohn's disease. *Nature*. 2001;411(6837):599-603.
30. Ogura Y, et al. A frameshift mutation in NOD2 associated with susceptibility to Crohn's disease. *Nature*. 2001;411(6837):603-606.
31. Altshuler D, Daly MJ, Lander ES. Genetic mapping in human disease. *Science*. 2008;322(5903):881-888.
32. Hindorff LA, et al. Potential etiologic and functional implications of genome-wide association loci for human diseases and traits. *Proc Natl Acad Sci U S A*. 2009;106(23):9362-9367.
33. Welter D, et al. The NHGRI GWAS Catalog, a curated resource of SNP-trait associations. *Nucleic Acids Res*. 2014;42(Database issue):D1001-D1006.
34. Zhernakova A, van Diemen CC, Wijmenga C. Detecting shared pathogenesis from the shared genetics of immune-related diseases. *Nat Rev Genet*. 2009;10(1):43-55.
35. Cotsapas C, et al. Pervasive sharing of genetic effects in autoimmune disease. *PLoS Genet*. 2011;7(8):e1002254.
36. Cotsapas C, Hafler DA. Immune-mediated disease genetics: the shared basis of pathogenesis. *Trends Immunol*. 2013;34(1):22-26.
37. Farh KK, et al. Genetic and epigenetic fine mapping of causal autoimmune disease variants. *Nature*. 2015;518(7539):337-343.
38. Hu X, Kim H, Stahl E, Plenge R, Daly M, Raychaudhuri S. Integrating autoimmune risk loci with gene-expression data identifies specific pathogenic immune cell subsets. *Am J Hum Genet*. 2011;89(4):496-506.
39. Heng TS, Painter MW, Immunological Genome Project C. The Immunological Genome Project: networks of gene expression in immune cells. *Nat Immunol*. 2008;9(10):1091-1094.
40. Zhang W, et al. Effector CD4<sup>+</sup> T cell expression signatures and immune-mediated disease associated genes. *PLoS One*. 2012;7(6):e38510.
41. Rossin EJ, et al. Proteins encoded in genomic regions associated with immune-mediated disease physically interact and suggest underlying biology. *PLoS Genet*. 2011;7(1):e1001273.
42. Segre AV, et al. Common inherited variation in mitochondrial genes is not enriched for associations with type 2 diabetes or related glycemic traits. *PLoS Genet*. 2010;6(8):e1001058.
43. Franke L, van Bakel H, Fokkens L, de Jong ED, Egmont-Petersen M, Wijmenga C. Reconstruction of a functional human gene network, with an application for prioritizing positional candidate genes. *Am J Hum Genet*. 2006;78(6):1011-1025.
44. International Multiple Sclerosis Genetics Consortium. Network-based multiple sclerosis pathway analysis with GWAS data from 15,000 cases and 30,000 controls. *Am J Hum Genet*. 2013;92(6):854-865.
45. Stanford SM, Bottini N. PTPN22: the archetypal non-HLA autoimmunity gene. *Nat Rev Rheumatol*. 2014;10(10):602-611.
46. Wang Y, et al. The autoimmunity-associated gene PTPN22 potentiates toll-like receptor-driven, type 1 interferon-dependent immunity. *Immunity*. 2013;39(1):111-122.
47. International Consortium for Systemic Lupus Erythematosus Genetics, et al. Genome-wide association scan in women with systemic lupus erythematosus identifies susceptibility variants in ITGAM, PXK, KIAA1542 and other loci. *Nat Genet*. 2008;40(2):204-210.
48. Begovich AB, et al. A missense single-nucleotide polymorphism in a gene encoding a protein tyrosine phosphatase (PTPN22) is associated with rheumatoid arthritis. *Am J Hum Genet*. 2004;75(2):330-337.
49. Bottini N, et al. A functional variant of lymphoid tyrosine phosphatase is associated with type I diabetes. *Nat Genet*. 2004;36(4):337-338.
50. Kyogoku C, et al. Genetic association of the R620W polymorphism of protein tyrosine phosphatase PTPN22 with human SLE. *Am J Hum Genet*. 2004;75(3):504-507.
51. Velaga MR, et al. The codon 620 tryptophan allele of the lymphoid tyrosine phosphatase (LYP) gene is a major determinant of Graves' disease. *J Clin Endocrinol Metab*. 2004;89(11):5862-5865.
52. Baranathan V, et al. The association of the PTPN22 620W polymorphism with Behcet's disease. *Ann Rheum Dis*. 2007;66(11):1531-1533.
53. Begovich AB, et al. The R620W polymorphism of the protein tyrosine phosphatase PTPN22 is not associated with multiple sclerosis. *Am J Hum Genet*. 2005;76(1):184-187.
54. Criswell LA, et al. Analysis of families in the multiple autoimmune disease genetics consortium (MADGC) collection: the PTPN22 620W allele associates with multiple autoimmune phenotypes. *Am J Hum Genet*. 2005;76(4):561-571.
55. De Jager PL, et al. Evaluating the role of the 620W allele of protein tyrosine phosphatase PTPN22 in Crohn's disease and multiple sclerosis. *Eur J Hum Genet*. 2006;14(3):317-321.
56. Rivas MA, et al. Deep resequencing of GWAS loci identifies independent rare variants associated with inflammatory bowel disease. *Nat Genet*. 2011;43(11):1066-1073.

57. Vang T, et al. Autoimmune-associated lymphoid tyrosine phosphatase is a gain-of-function variant. *Nat Genet.* 2005;37(12):1317–1319.
58. Zhang J, et al. The autoimmune disease-associated PTPN22 variant promotes calpain-mediated Lyp/Pep degradation associated with lymphocyte and dendritic cell hyperresponsiveness. *Nat Genet.* 2011;43(9):902–907.
59. Rawlings DJ, Dai X, Buckner JH. The role of PTPN22 risk variant in the development of autoimmunity: finding common ground between mouse and human. *J Immunol.* 2015;194(7):2977–2984.
60. Menard L, et al. The PTPN22 allele encoding an R620W variant interferes with the removal of developing autoreactive B cells in humans. *J Clin Invest.* 2011;121(9):3635–3644.
61. Arechiga AF, et al. Cutting edge: the PTPN22 allelic variant associated with autoimmunity impairs B cell signaling. *J Immunol.* 2009;182(6):3343–3347.
62. Hassan NJ, et al. CD6 regulates T-cell responses through activation-dependent recruitment of the positive regulator SLP-76. *Mol Cell Biol.* 2006;26(17):6727–6738.
63. De Jager PL, et al. Meta-analysis of genome scans and replication identify CD6, IRF8 and TNFRSF1A as new multiple sclerosis susceptibility loci. *Nat Genet.* 2009;41(7):776–782.
64. Kofler DM, Severson CA, Mousessian N, De Jager PL, Hafler DA. The CD6 multiple sclerosis susceptibility allele is associated with alterations in CD4<sup>+</sup> T cell proliferation. *J Immunol.* 2011;187(6):3286–3291.
65. International Multiple Sclerosis Genetics Consortium Wellcome Trust Case Control Consortium 2, et al. Genetic risk and a primary role for cell-mediated immune mechanisms in multiple sclerosis. *Nature.* 2011;476(7359):214–219.
66. International Multiple Sclerosis Genetics Consortium, et al. Analysis of immune-related loci identifies 48 new susceptibility variants for multiple sclerosis. *Nat Genet.* 2013;45(11):1353–1360.
67. Ottoboni L, et al. Clinical relevance and functional consequences of the TNFRSF1A multiple sclerosis locus. *Neurology.* 2013;81(22):1891–1899.
68. Gregory AP, et al. TNF receptor 1 genetic risk mirrors outcome of anti-TNF therapy in multiple sclerosis. *Nature.* 2012;488(7412):508–511.
69. Comabella M, et al. TNFRSF1A polymorphisms rs1800693 and rs4149584 in patients with multiple sclerosis. *Neurology.* 2013;80(22):2010–2016.
70. Housley WJ, et al. Genetic variants associated with autoimmunity drive NFκB signaling and responses to inflammatory stimuli. *Sci Transl Med.* In Press.
71. Hnisz D, et al. Super-enhancers in the control of cell identity and disease. *Cell.* 2013;155(4):934–947.
72. Hartmann FJ, et al. Multiple sclerosis-associated IL2RA polymorphism controls GM-CSF production in human TH cells. *Nat Commun.* 2014;5:5056.
73. Cerosaletti K, et al. Multiple autoimmune-associated variants confer decreased IL-2R signaling in CD4<sup>+</sup> CD25(hi) T cells of type 1 diabetic and multiple sclerosis patients. *PLoS One.* 2013;8(12):e83811.
74. Dendrou CA, et al. Cell-specific protein phenotypes for the autoimmune locus IL2RA using a genotype-selectable human bioresource. *Nat Genet.* 2009;41(9):1011–1015.
75. Ye CJ, et al. Intersection of population variation and autoimmunity genetics in human T cell activation. *Science.* 2014;345(6202):1254665.
76. Cortes A, Brown MA. Promise and pitfalls of the Immunochip. *Arthritis Res Ther.* 2011;13(1):101.
77. Trynka G, et al. Dense genotyping identifies and localizes multiple common and rare variant association signals in celiac disease. *Nat Genet.* 2011;43(12):1193–1201.
78. Cooper JD, et al. Seven newly identified loci for autoimmune thyroid disease. *Hum Mol Genet.* 2012;21(23):5202–5208.
79. Liu JZ, et al. Dense fine-mapping study identifies new susceptibility loci for primary biliary cirrhosis. *Nat Genet.* 2012;44(10):1137–1141.
80. Cortes A, et al. Identification of multiple risk variants for ankylosing spondylitis through high-density genotyping of immune-related loci. *Nat Genet.* 2013;45(7):730–738.
81. Ellinghaus D, et al. High-density genotyping study identifies four new susceptibility loci for atopic dermatitis. *Nat Genet.* 2013;45(7):808–812.
82. Liu JZ, et al. Dense genotyping of immune-related disease regions identifies nine new risk loci for primary sclerosing cholangitis. *Nat Genet.* 2013;45(6):670–675.
83. Hinks A, et al. Dense genotyping of immune-related disease regions identifies 14 new susceptibility loci for juvenile idiopathic arthritis. *Nat Genet.* 2013;45(6):664–669.
84. Tsai LC, et al. Identification of 15 new psoriasis susceptibility loci highlights the role of innate immunity. *Nat Genet.* 2012;44(12):1341–1348.
85. Chen GB, et al. Estimation and partitioning of (co)heritability of inflammatory bowel disease from GWAS and immunochip data. *Hum Mol Genet.* 2014;23(17):4710–4720.
86. Wellcome Trust Case Control Consortium, et al. Bayesian refinement of association signals for 14 loci in 3 common diseases. *Nat Genet.* 2012;44(12):1294–1301.
87. 1000 Genomes Project Consortium, et al. An integrated map of genetic variation from 1,092 human genomes. *Nature.* 2012;491(7422):56–65.
88. Rivera CM, Ren B. Mapping human epigenomes. *Cell.* 2013;155(1):39–55.
89. Shlyueva D, Stampfel G, Stark A. Transcriptional enhancers: from properties to genome-wide predictions. *Nat Rev Genet.* 2014;15(4):272–286.
90. Villar D, Flieck P, Odom DT. Evolution of transcription factor binding in metazoans - mechanisms and functional implications. *Nat Rev Genet.* 2014;15(4):221–233.
91. Ernst J, et al. Mapping and analysis of chromatin state dynamics in nine human cell types. *Nature.* 2011;473(7345):43–49.
92. Heintzman ND, et al. Histone modifications at human enhancers reflect global cell-type-specific gene expression. *Nature.* 2009;459(7243):108–112.
93. Marson A, et al. Foxp3 occupancy and regulation of key target genes during T-cell stimulation. *Nature.* 2007;445(7130):931–935.
94. Samstein RM, et al. Foxp3 exploits a pre-existent enhancer landscape for regulatory T cell lineage specification. *Cell.* 2012;151(1):153–166.
95. Ciofani M, et al. A validated regulatory network for Th17 cell specification. *Cell.* 2012;151(2):289–303.
96. Maurano MT, et al. Systematic localization of common disease-associated variation in regulatory DNA. *Science.* 2012;337(6099):1190–1195.
97. Parker SC, et al. Chromatin stretch enhancer states drive cell-specific gene regulation and harbor human disease risk variants. *Proc Natl Acad Sci U S A.* 2013;110(44):17921–17926.
98. Trynka G, et al. Chromatin marks identify critical cell types for fine mapping complex trait variants. *Nat Genet.* 2013;45(2):124–130.
99. Hawkins RD, et al. Global chromatin state analysis reveals lineage-specific enhancers during the initiation of human T helper 1 and T helper 2 cell polarization. *Immunity.* 2013;38(6):1271–1284.
100. Gilad Y, Rifkin SA, Pritchard JK. Revealing the architecture of gene regulation: the promise of eQTL studies. *Trends Genet.* 2008;24(8):408–415.
101. Stranger BE, et al. Population genomics of human gene expression. *Nat Genet.* 2007;39(10):1217–1224.
102. Orru V, et al. Genetic variants regulating immune cell levels in health and disease. *Cell.* 2013;155(1):242–256.
103. Wright FA, et al. Heritability and genomics of gene expression in peripheral blood. *Nat Genet.* 2014;46(5):430–437.
104. Gat-Viks I, et al. Deciphering molecular circuits from genetic variation underlying transcriptional responsiveness to stimuli. *Nat Biotechnol.* 2013;31(4):342–349.
105. Fairfax BP, et al. Innate immune activity conditions the effect of regulatory variants upon monocyte gene expression. *Science.* 2014;343(6175):1246949.
106. Hu X, et al. Regulation of gene expression in autoimmune disease loci and the genetic basis of proliferation in CD4<sup>+</sup> effector memory T cells. *PLoS Genet.* 2014;10(6):e1004404.
107. Lee MN, et al. Common genetic variants modulate pathogen-sensing responses in human dendritic cells. *Science.* 2014;343(6175):1246980.
108. Raj T, et al. Polarization of the effects of autoimmune and neurodegenerative risk alleles in leukocytes. *Science.* 2014;344(6183):519–523.
109. Heinz S, et al. Effect of natural genetic variation on enhancer selection and function. *Nature.* 2013;503(7477):487–492.
110. McMicker G, et al. Identification of genetic variants that affect histone modifications in human cells. *Science.* 2013;342(6159):747–749.
111. Seumois G, et al. Epigenomic analysis of primary human T cells reveals enhancers associated with TH2 memory cell differentiation and asthma susceptibility. *Nat Immunol.* 2014;15(8):777–788.
112. Manolio TA, et al. Finding the missing heritability of complex diseases. *Nature.* 2009;461(7265):747–753.
113. Okada Y, et al. Genetics of rheumatoid arthritis contributes to biology and drug discovery. *Nature.* 2014;506(7488):376–381.
114. Lee JC, et al. Human SNP links differential outcomes in inflammatory and infectious disease to a FOXO3-regulated pathway. *Cell.* 2013;155(1):57–69.
115. Chakravarti A, Clark AG, Mootha VK. Distilling pathophysiology from complex disease genetics. *Cell.* 2013;155(1):21–26.
116. Fyre S, Worthington J. Take your PICS: moving from GWAS to immune function. *Immunity.* 2014;41(6):883–885.