Published in final edited form as:
Cell. 2010 March 19; 140(6): 791–797. doi:10.1016/j.cell.2010.03.003.

Unraveling the Genetics of Autoimmunity

Lauren A. Zenewicz[1], **Clara Abraham**[2], **Richard A. Flavell**[1,4], and **Judy H. Cho**[2,3,*]
[1]Department of Immunobiology, Yale University, New Haven, CT 06520, USA
[2]Department of Medicine, Yale University, New Haven, CT 06520, USA
[3]Department of Genetics, Yale University, New Haven, CT 06520, USA
[4]Howard Hughes Medical Institute, Yale University, New Haven, CT 06520, USA

Abstract

The chronic autoimmune diseases include multiple complex genetic disorders. Recently, genome-wide association studies (GWAS) have identified a large number of major loci, with many associations shared between various autoimmune diseases. These associations highlight key roles for lymphocyte activation and prioritize specific cytokine pathways and mechanisms of host-microbe recognition. Despite success in identifying loci, comprehensive models of disease pathogenesis are currently lacking. Future efforts comparing association patterns between autoimmune diseases may be particularly illustrative. New genomic technologies applied to classic genetic studies involving twins, early onset cases, and phenotypic extremes may provide key insights into developmental and gene-environment interactions in autoimmunity.

Autoimmune disorders are characterized by an inappropriate, ultimately excessive, inflammatory response against self, resulting in tissue destruction. Although many individuals affected by autoimmunity demonstrate multiorgan involvement, the primary end-organ target (e.g., autoimmune destruction of pancreatic islet cells in type 1 diabetes mellitus) typically drives the clinical presentation and disease definition. Evidence for both B and T lymphocyte hyper-reactivity is typically observed, with the presence of autoantibodies and genetic associations involving the major histocompatibility complex (MHC) providing the most significant association evidence for many autoimmune diseases.

Most cases of autoimmunity arise in the absence of a positive family history. However, evidence that genetic factors contribute to disease pathogenesis has been provided by familial clustering in some cases, which reflects shared genetic, developmental, and environmental factors. The contribution of genetic factors is established through twin studies demonstrating higher disease concordance in monozygotic compared to dizygotic twins. In contrast to single-gene, Mendelian disorders, complex genetic disorders such as many autoimmune diseases are associated with multiple genetic loci, conferring varying effects on disease susceptibility.

Within families, clustering of distinct autoimmune diseases has been reported, and this suggests the presence of shared pathogenic factors across autoimmunity. For example, a large, population-based survey demonstrated that families with a rheumatoid arthritis (RA) (Lin et al., 1998) or multiple sclerosis (MS) (Broadley et al., 2000) proband were more likely to also manifest other autoimmune disorders. Compared to families with a single member affected by MS, the frequency of other autoimmune diseases was higher in families

*Correspondence: judy.cho@yale.edu.

containing multiple members with MS, suggesting a cumulative enrichment of autoimmune susceptibility loci in these select cohorts.

The application of genome-wide association studies (GWAS) to autoimmune diseases has identified a growing number of disease-associated loci. GWAS involve the genotyping of several hundred thousand single-nucleotide polymorphisms (SNPs) throughout the genome in large case-control cohorts (Manolio et al., 2009). Because of the large number of statistical tests applied, stringent statistical thresholds are required (p value less than 5×10^{-8}) to establish genome-wide evidence for association. For many associated loci, the association signals do not directly implicate a single, protein-coding gene, and the causative role for candidate genes in the region can only be speculated. A striking number of major loci have been observed to demonstrate genome-wide evidence for association in multiple, distinct autoimmune disorders (Table 1) (Barrett et al., 2008, 2009; Festen et al., 2009; Franke et al., 2008; Gateva et al., 2009; Gregersen et al., 2009; Han et al., 2009; Hom et al., 2008; Raychaudhuri et al., 2009; van Heel et al., 2007; WTCCC, 2007; Zhernakova et al., 2007). Given the a priori epidemiological support for shared pathogenesis across autoimmune disorders, it may be argued that less stringent evidence for association would be required to establish contributions for those loci previously established in another, distinct autoimmune disease. Efforts to comprehensively genotype variation at all genome-wide significant autoimmune loci across all autoimmune diseases are ongoing. These studies will provide enormous insight into shared and distinct patterns of genetic associations across autoimmunity.

Notably, many inflammatory genes implicated in autoimmunity demonstrate broad expression patterns and pleio-tropic functions. However, integrative themes are emerging, implicating both established and new mechanisms of inflammation. Associated loci include a broad array of immune-associated genes involved in lymphocyte activation (receptor signaling pathways and costimulation), microbial recognition, and cytokines or cytokine receptors (Gregersen and Olsson, 2009). In the following sections, we highlight select genetic associations demonstrating the most significant evidence for associations in multiple autoimmune diseases.

Lymphocyte Activation in Autoimmunity

For many autoimmune disorders, the MHC represents the predominant association, highlighting its central and complex role in mediating host inflammatory responses to evolutionarily significant pathogens. The nature of antigenic responses to self- or non-self-antigens is shaped extensively by the unique coding and noncoding genetic variation of HLA alleles. Extensive allelic variation and linkage disequilibrium (nonrandom or correlated association of alleles) are observed throughout the MHC region. Recent high-density mapping in multiple autoimmune diseases demonstrated complex, multilocus effects that span the entire region, with evidence for shared and unique loci across diseases (Rioux et al., 2009).

After the MHC, one of the most common genetic associations across autoimmune disorders is observed at the protein tyrosine phosphatase gene *PTPN22*, expressed in lymphocytes. The minor tryptophan allele at Arg620Trp within *PTPN22* has been associated with numerous autoimmune diseases including type 1 diabetes mellitus, RA, autoimmune thyroiditis, and systemic lupus erythematosis (SLE) (Barrett et al., 2008; Bottini et al., 2004, WTCCC, 2007; Criswell et al., 2005). Interestingly, the more common arginine allele is associated with an inflammatory bowel disorder called Crohn's disease (Barrett et al., 2008). The tryptophan allele results in a gain of function by the phosphatase protein relative to the arginine allele, such that B and T cell activation is inhibited; tryptophan homozygotes are

characterized by a profound defect in lymphocyte receptor signaling (Rieck et al., 2007; Vang et al., 2005). It is possible that impaired T cell signaling increases autoimmune susceptibility by contributing to a failure to delete autoreactive T cells during thymic selection, impaired activity of regulatory T cell populations (Vang et al., 2005), or defects in clearance of microbes, such as viruses.

The plethora of associated loci containing candidate genes encoding molecules expressed by lymphocytes that modulate costimulatory functions (e.g., *CTLA-4*, *CD2*/*CD58*, *CD28*, *ICOSLG*, *TNFSF15*) (Cooper et al., 2008; De Jager et al., 2009; Gregersen et al., 2009; Gregersen and Olsson, 2009; Raychaudhuri et al., 2009) further highlights the centrality of lymphocyte activation in human autoimmunity. However, the nature (expression isoforms, cell-specific expression levels) of altered gene function resulting from disease-associated polymorphisms is largely undefined.

Cytokine Pathways

Numerous combinations of cytokines and cytokine receptors have been associated with autoimmunity, with common and distinct patterns of association evidence reported across disorders. For example, both type 1 diabetes mellitus and MS demonstrate association to interleukin 2 receptor alpha (*IL2RA*), and MS also demonstrates association to interleukin 7 receptor alpha (*IL7R*) (Barrett et al., 2009; Hafler et al., 2007). However, no evidence for association in either disease has been observed thus far for the shared interleukin 2 receptor gamma chain (*IL2RG*) required for interleukin 7 and interleukin 2 signaling. Although IL-7 has not been clearly associated with autoimmunity, associations in a gene region on chromosome 4q27 near the IL-2 and IL-21 genes have been reported in celiac disease, inflammatory bowel disease (IBD), RA, and type 1 diabetes (Barrett et al., 2009; Festen et al., 2009; van Heel et al., 2007; Zhernakova et al., 2007), but not thus far in MS (Table 1).

A more consistent example of cytokine-cytokine receptor involvement is provided by the IL-12/IL-23 pathway in IBD, ankylosing spondylitis, and psoriasis (A.B. Begovich et al., 2006, Am. J. Hum. Genet. abstract; Burton et al., 2007; Duerr et al., 2006). Both IBD and psoriasis demonstrate association to the interleukin 12 p40 subunit (*IL12B*) (Begovich et al., 2006; Nair et al., 2009), common to IL-12 and IL-23 cytokines, as well as to the interleukin 23 receptor (*IL23R*), which represents the IL-23-specific component of the heterodimeric cytokine receptor. The interleukin 23 pathway plays a central role in antimicrobial defenses and mucosal immunity, in part through expansion of Th17 lymphocytes. However, the role of Th17-lineage cytokines in the tissues associated with these respective diseases can lead to distinct outcomes. Recent data have shown that IL-23 mediates dermal inflammation through the Th17 cytokine IL-22 (Zheng et al., 2007). IL-22 stimulation of keratinocytes leads to hyperplasia of keratinocyte layers and induction of expression of antimicrobial peptides such as β-defensin (Zenewicz and Flavell, 2008). On the other hand, experimental mouse models of IBD suggest that IL-22 is protective in the intestine (Sugimoto et al., 2008; Zenewicz et al., 2008). IL-22 stimulation of colonic epithelial cells may help to maintain the barrier function of the intestine, as well as induce expression of antimicrobial peptides and mucins. The dual nature of this cytokine, protective versus inflammatory, likely depends on the inflammatory context and tissues involved. In addition to cytokine-cytokine receptor associations, IBD associations implicate downstream members of the interleukin 23 pathway including Janus-activated kinase 2 (JAK2) and signal transducer and activator of transcription 3 (STAT3) (Abraham and Cho, 2009).

The tumor necrosis factor (TNF) pathway represents a central therapeutic target across many autoimmune diseases. Multiple downstream components of TNF signaling have been associated in autoimmunity, most notably the tumor necrosis factor-inducible protein A20

(TNFAIP3), which terminates TNF- and pattern recognition receptor-induced responses of the transcription factor NF-κB. The gene region near *TNFAIP3* has been associated with SLE, psoriasis, and RA (Nair et al., 2009; Plenge et al., 2007a; Thomson et al., 2007). In addition, TNIP1 (TNFAIP3-interacting protein 1), which interacts with TNFAIP3, is also associated with psoriasis and SLE (Han et al., 2009; Nair et al., 2009). Additional downstream components of TNF signaling have associations to RA, for example, TNF receptor-associated factor 1 (*TRAF1*) (Plenge et al., 2007b) and Rel, a component of NF-κB signaling (Gregersen et al., 2009).

Microbial Responses

Although genetic variants altering lymphocyte activation and cytokine signaling modulate microbial responses, more specific examples of host genetic variation provide additional insight into mechanisms of autoimmunity. The complex relationships between viral infection and autoimmunity are highlighted through associations in genes encoding interferon regulatory factor 5 (IRF5) and interferon induced with helicase C domain 1 (IFIH1). The *IRF5* gene has been associated with SLE (Graham et al., 2006), with some evidence for association to other autoimmune diseases. IRF5 is downstream of pattern recognition receptor signaling and induces numerous cytokines, including type 1 interferons. The SLE-associated IRF5 association signals are complex, with some polymorphisms characterized by increased expression of IRF5 isoforms (Graham et al., 2006). Microarray studies of peripheral blood from SLE patients have demonstrated a strong type 1 interferon signature (Pascual et al., 2006).

More recently, uncommon polymorphisms in *IFIH1*, including a nonsense mutation, have been demonstrated to confer protection against developing type 1 diabetes mellitus (Nejentsev et al., 2009). IFIH1 recognizes RNA from picornaviruses and mediates immune activation. Importantly, infections with enteroviruses, members of the picornavirus family, are more frequent in newly diagnosed cases of type 1 diabetes mellitus and antedate the onset of disease-associated autoantibodies. Future studies testing whether wild-type IFIH1-mediated immune responses to enteroviruses induce autoreactive lymphocytes will provide key insight into its role in disease pathogenesis.

If the *IRF5* and *IFIH1* associations highlight pathogenic roles of increased and wild-type host microbial responses, respectively, equally important is the concept that impaired, initial microbial responses may also result ultimately in increased inflammation. IBD is comprised of two major subtypes, Crohn's disease and ulcerative colitis. Both subtypes are associated with multiple IL-23 pathway genes. However, only Crohn's disease is associated with loss-of-function polymorphisms in NOD2 (pattern recognition receptor for bacterial peptidoglycan, normally resulting in NF-κB activation) and in ATG16L1-mediated autophagy. IBD is believed to result from an inappropriate host response to commensal intestinal microbes, resulting ultimately in intestinal damage from excessive inflammation (Abraham and Cho, 2009). There is a functional link between pattern recognition receptors, including NOD2 and ATG16L1 (Cooney et al., 2009; Travassos et al., 2009). In addition, both NOD2 and ATG16L1 are expressed in a variety of innate cells, lymphocytes, and gut epithelial Paneth cells, which secrete potent antimicrobial peptides (Abraham and Cho, 2009). Therefore, in Crohn's disease, defects in first line mucosal clearance of microbes may ultimately contribute to excessive chronic inflammatory responses.

Beyond GWAS: Assessment and Challenges

Current GWAS have sampled common variation throughout the genome for disease associations (Table 2). Newer genotyping platforms that more comprehensively assay common variation will likely identify important new disease associations. GWAS in non-

European populations provide important comparative insight. In addition, combining larger case-control cohorts will likely identify genomic loci of modest effect, for which smaller studies were underpowered. Identification of additional genes will provide cumulative insight into disease pathways and complex functional networks. However, for all of these complex autoimmune disorders, presently identified loci account for only a minor fraction of the predicted heritability. The identification of additional loci requiring ever larger meta-analysis cohorts using similar methodology will likely not have a significant impact on accounting for overall heritability. This would indicate that genetic variation that is not well assayed through present approaches may yet provide a significant contribution to overall disease heritability (Clayton, 2009; Manolio et al., 2009).

More comprehensive cataloguing of genetic variation is ongoing, accelerated significantly by high-throughput sequencing technologies. Uncommon SNPs, and other not well assayed genetic variations, may provide significant additional sources of disease-associated genetic variation. Equally important, these sequencing technologies are providing data on tissue-specific transcriptome expression and epigenetic regulation, which may be particularly relevant to autoimmune diseases. Given the central role of host-microbe interactions, sequence-based analyses of microbial communities will also provide important insight. The plethora of data that will be generated through these emerging approaches presents challenges and opportunities to define the underlying rules governing immune system responses and perturbations resulting in autoimmunity.

Beyond GWAS: Uncommon and Non-SNP-Based Variation

In many cases, the association signals identified by GWAS will be largely driven by relatively common genetic variation of modest effect, with increased risks of 1.1 to 1.5 per associated allele typically observed (Altshuler et al., 2008). However, functional genetic variation having a greater effect on disease susceptibility would be maintained at relatively low frequencies within populations due to purifying selection. It is anticipated that uncommon genetic variation, in the form of SNPs, copy number polymorphisms, and other insertions/deletions throughout the genome, may contribute significantly to the genetics of autoimmune diseases. A comprehensive cataloguing of these less common variants is being completed through the 1000 Genomes Project (http://www.1000genomes.org). By definition, identifying less common genetic variation will require deeper interrogation of larger sequencing cohorts from diverse populations.

There are both practical and methodological challenges to establishing disease associations for uncommon alleles. Practical considerations include increased technical difficulties associated with definitively establishing and validating the presence of uncommon alleles. In addition, there is the potential for less accurate genotype assignments, which are often based on clustering algorithms. Methodological challenges include the decreased power to establish disease associations for uncommon alleles compared to common alleles. The plethora of independent, uncommon variants compared with the more limited subset of common variants increases the potential number of statistical tests applied in a comprehensive genotyping approach for uncommon variants. Because of these factors, more intensive surveys for uncommon variation have focused on protein-coding genes (Ng et al., 2009) and developing gene-based tests of association that incorporate multiple uncommon alleles simultaneously. With the advent of whole-genome, exon-capture techniques followed by high-throughput sequencing, studies testing for association to uncommon, coding region variants have now been reported (Choi et al., 2009).

Identifying Functional Variation within Disease-Associated Loci

The presence of only one gene within an association signal typically implicates it in disease pathogenesis, with the strong caveat that functional proof of disease causation will typically not be established for complex disease gene associations. As opposed to Mendelian diseases, in complex, multigenic diseases, single-gene alterations by themselves are insufficient to drive disease expression. Because of the highly correlated nature of common variants in the human genome, association signals often encompass multiple, protein-coding genes demonstrating equivalent statistical evidence for association. In these cases, dissection of which gene is generating the association signal may be difficult to establish. At the other end of the spectrum, many GWAS-identified association signals are confined to gene deserts containing no protein-coding genes.

Disease-associated missense mutations at conserved amino acid residues represent strong candidates meriting further functional analyses. In many cases, however, functional variants may exert their major effects through more subtle regulatory effects on gene expression. Genome-wide linkage and association studies mapping DNA variation with variable mRNA expression through expression quantitative trait loci (eQTL) mapping have generated a rich source of likely functional polymorphisms (Dixon et al., 2007). However, these studies do not generally confirm that the DNA polymorphisms regulate variable protein expression, and posttranscriptional mechanisms of gene regulation have not been comprehensively evaluated. Equally important, eQTL studies confined to a single cell type likely represent only a subset of DNA polymorphisms regulating gene expression. In this regard, studies utilizing primary human cells or relevant tissues (Emilsson et al., 2008) to map DNA variation with variable RNA expression may be particularly useful in autoimmune diseases. These approaches have now been applied to various diseases. For example, a study using primary human peripheral blood mononuclear cells demonstrated that a DNA haplotype in a region including *IL2RA* (CD25) that confers protection against type 1 diabetes mellitus and MS is associated with increased CD25 expression in $CD4^+$ memory T cells. Key to the successful completion of such well-powered studies is the availability of large bioresources of healthy controls that allow for efficient recall of individuals stratified on genotypes of interest (Dendrou et al., 2009).

Much of the impetus toward defining functional consequences of disease-associated polymorphisms has focused on protein-coding genes. However, although protein-coding genes represent a small minority of the genome, it is important to note that the majority of the genome is actively transcribed, resulting in a complex network of short and long noncoding RNAs that have complex, and currently incompletely understood, effects on gene regulation and cellular function (Encode Project Consortium, 2007). The pervasiveness of these non-coding RNAs would suggest that at least some of the association signals presently observed in “gene deserts,” as well as within regions containing the more familiar protein-coding genes, may be driven by altered functional effects of noncoding RNAs.

Developing Predictive Disease Models

Establishing a more complete catalog of common and uncommon genetic associations, as well as more precisely defining the functional variation at each associated locus, represent important, intermediate goals in the genetics of autoimmunity. However, a major goal moving forward is the development of improved predictive models for disease. Such models, applied to both human disease and model systems, ideally would encompass an improved capacity to predict disease risk and model disease progression, improved disease classification across and within autoimmune diseases based on pathophysiological mechanisms, and improved prediction of therapeutic responses, both in the development of

new therapeutic agents and in the application to individual patients. The rapid advances and decreasing costs of high-throughput sequencing approaches provide a powerful tool to address some of these challenges. Applying recent high-throughput sequencing technologies to classic genetic approaches of examining phenotypic extremes, early onset cases, and twin studies may provide important insight. Although the significance of GWAS rests with definitive DNA-to-disease mappings, defining intermediate mappings between DNA sequence, chromatin modifications, and RNA expression will provide insight into molecular networks and predictive behavior that may form the basis for more rational therapeutic development (Schadt, 2009). For example, twin studies provide an important means of examining the interplay between genetic, environmental, and development factors contributing to autoimmunity.

Twin Studies in Genetics and Autoimmunity

The finding of significantly higher monozygotic compared to dizygotic twin concordance for a given disease is used as proof that genetic factors play a significant role in disease pathogenesis, and that genetic mapping studies are merited. Twin concordance studies form a basis for composite disease heritability estimates. However, the composite cohort of twin pairs for a given disease encompasses a wide range of genetic backgrounds, with likely very different genetic propensities for disease and subphenotype expression.

Perhaps one of the more tractable environmental covariates modulating autoimmune disease involves the bidirectional crosstalk between the host and intestinal microbiomes. A recent study demonstrated that the human gut microbiome demonstrates familial similarity, with comparable degrees of covariance between adult dizygotic and monozygotic twins; a nonsignificant trend toward increased similarity between monozygotic twin pairs was observed (Turnbaugh et al., 2009). However, a clear role for host genetic background in altering distal gut microbiome composition was demonstrated in MyD88-deficient mice compared to their littermate controls, with an altered propensity to develop autoimmune diabetes (Wen et al., 2008).

Developmental factors also significantly modulate gene expression; this is clearly illustrated in twin studies comparing methylation status assessed by microarray analysis of CpG islands. Overall, monozygotic twin correlation of methylation status was higher than that observed in dizygotic twins. However, within monozygotic twins, correlations were lower for dichorionic compared to monochorionic monozygotic twins, indicating that factors other than DNA sequence modulate epigenetic regulation (Kaminsky et al., 2009). The potential contribution of the underlying DNA sequence in chromatin modification can be systematically defined through chromatin immunoprecipitation (ChIP-seq) or through formaldehyde-assisted isolation of regulatory elements (FAIRE-seq), followed by high-throughput sequencing. Although extensive work has focused on epigenetic alterations in neoplasia, the development of genome-wide maps of chromatin states in pluripotent and lineage-committed cells can now be applied to the genetics and genomics of autoimmunity (Meissner et al., 2008).

Epigenetics in Autoimmunity

Genome-wide maps of epigenetic regulation will provide enormous amounts of information, along with the challenge of defining the underlying rules that govern gene expression and cellular differentiation. For example, key transcription factors and lineage-specific cytokines in $CD4^+$ T cell subsets demonstrate a broad spectrum of epigenetic states, contributing to a useful framework for understanding specificity and plasticity of $CD4^+$ T cell subsets. H3K4me3 and H3K27me3 histone modifications are enriched in active and inactive chromatin regions, respectively, and effectively discriminate between genes that are

expressed, poised for expression, or stably repressed (Meissner et al., 2008). An improved understanding of epigenetics and autoimmune diseases will likely provide insight into the pathophysiology of specific autoimmune diseases and may assist in improved diagnosis and treatment.

Model Development in Autoimmunity

The application of genetic and genomic approaches to build useful predictive models of autoimmune disease ideally would encompass development of murine models of disease that integrate primary factors in human disease pathogenesis and predict therapeutic responses in humans. Predictive models in humans ideally would include defining individuals at risk for disease development, as well as predicting disease prognosis and informing therapeutic management. The continuum between monogenic and polygenic human diseases illustrates the challenges and opportunities. For example, the identification of rare Mendelian autoimmune disorders in humans (e.g., mutations in AIRE; Mathis and Benoist, 2009) has provided enormous insight into basic mechanisms of autoimmunity, enhanced by similar phenotypic manifestations in corresponding murine genetic models. Conversely, the development of single-gene, murine models of disease has provided significant insight; the identification of correlative disease associations in a given gene through GWAS in humans would prioritize that gene and pathway as playing a more significant role in human disease. However, complete knockdown of single-gene expression in murine models will result in very different downstream consequences compared to the more subtle effects on gene function resulting from the human SNP associations that have typically been identified through GWAS. Reporter constructs tagging immune cell subsets integrated with single-gene and spontaneous autoimmune disease models provide important refinement of the ability to study key immune cell subsets. For example, a double-transgenic mouse model distinguishing T cells expressing $FoxP3^+$ and “exFoxP3” established that the exFoxP3 subset of cells expressing both IL-17 and IFN-γ led to the rapid development of diabetes, underscoring the potential significance of cellular plasticity in autoimmunity (Zhou et al., 2009).

More broadly, modeling complex autoimmune diseases in mice has limitations due to inherent differences between the murine and human immune systems (Mestas and Hughes, 2004). To better recapitulate the human immune response, researchers have developed “humanized” mice. Humanized mice encompass two distinct models: (1) transgenic mice expressing human genes and (2) immunodeficient mice engrafted with human peripheral blood mononuclear cells or hematopoietic cells to generate a “human” immune system. These systems, alone or combined together, allow us to model human immune responses within the confines of mice, with all of their research advantages. The optimization of engraftment of human stem cells into immunodeficient mice is an area of active research. Mice expressing human cytokines or cytokine receptors will be potentially superior hosts for more closely understanding the development and activation of human immune responses in vivo (Huntington et al., 2009). Preclinical studies testing new therapies in increasingly refined murine models of autoimmunity may enhance predictive capacity for new treatments in human disease.

Human Studies of Autoimmunity

Ultimately, methodologies to improve human studies, in parallel with refined model systems, will be required for future advances. Recent genetic association studies have prioritized shared and unique mechanisms of lymphocyte activation and cytokine signaling between various autoimmune disorders. Comparative studies between various autoimmune disorders, integrating genetic and expression data, as well as immune system function, will

be important in prioritizing new therapeutic approaches, categorizing diseases into subsets and delineating individuals with distinct immune system functions, and modulating inflammatory responses with therapeutic interventions.

Clinical applications of expression analyses have been very successfully applied to some autoimmune disorders, notably SLE. SLE affects many different organs, including the heart, kidneys, skin, joints, and nervous system (Rahman and Isenberg, 2008), and the disease course is often unpredictable. Due to the heterogeneity of SLE, the use of biomarker signatures for the diagnosis and tracking of disease progression would be a valuable tool. Modular analysis, in which transcription of a group of genes is assessed, has recently been applied to SLE (Chaussabel et al., 2008). SLE patient transcriptional profiles were followed over time and found to correlate with disease activity. These transcriptional analyses have been applied to diagnosis, monitoring of the disease course, and therapeutic interventions. It may be argued that peripheral blood immune system responses may not reflect relevant, organ-specific immune responses; the success of these approaches in SLE may reflect its systemic nature. However, alterations in systemic immune system function have been identified across autoimmune diseases, and similar approaches in other autoimmune disorders will likely identify important comparative insight. More generally, the development of innovative and uniform methods for immune monitoring in health and disease and in the context of clinical trials represents an extremely high priority. Given the extensive genetic, functional, and therapeutic overlap between various autoimmune disorders, it is anticipated that the most rapid and significant advances will be attained through comparative approaches. The mosaic of disease-associated autoimmune loci identified by GWAS highlights key functional polymorphisms, which together with developmental and environmental factors result in an increased propensity for developing inflammatory disease in humans.

Acknowledgments

This work was supported by grants from the National Institutes of Health: DK P30-34989 (J.H.C. and C.A.), U19AI082713 (J.H.C. and C.A.), R01DK077905 (C.A.), R01DK077903, U01 DK62429, U01 DK062422, UL1 RR024139, RCDK086800 (J.H.C.); the Bohmfalk Funds for Medical Research (J.H.C.); and the Crohn's and Colitis Foundation of America (J.H.C. and C.A.).

References

Abraham C, Cho JH. N Engl J Med. 2009; 361:2066–2078. [PubMed: 19923578]

Altshuler D, Daly MJ, Lander ES. Science. 2008; 322:881–888. [PubMed: 18988837]

Barrett JC, Hansoul S, Nicolae DL, Cho JH, Duerr RH, Rioux JD, Brant SR, Silverberg MS, Taylor KD, Barmada MM, et al. Nat Genet. 2008; 40:955–962. [PubMed: 18587394]

Barrett JC, Clayton DG, Concannon P, Akolkar B, Cooper JD, Erlich HA, Julier C, Morahan G, Nerup J, Nierras C, et al. Nat Genet. 2009 Published online May 10, 2009. 10.1038/ng.381

Bottini N, Musumeci L, Alonso A, Rahmouni S, Nika K, Rostamkhani M, MacMurray J, Meloni GF, Lucarelli P, Pellecchia M, et al. Nat Genet. 2004; 36:337–338. [PubMed: 15004560]

Broadley SA, Deans J, Sawcer SJ, Clayton D, Compston DA. Brain. 2000; 123:1102–1111. [PubMed: 10825350]

Burton PR, Clayton DG, Cardon LR, Craddock N, Deloukas P, Duncanson A, Kwiatkowski DP, McCarthy MI, Ouwehand WH, Samani NJ, et al. Nat Genet. 2007; 39:1329–1337. [PubMed: 17952073]

Chaussabel D, Quinn C, Shen J, Patel P, Glaser C, Baldwin N, Stichweh D, Blankenship D, Li L, Munagala I, et al. Immunity. 2008; 29:150–164. [PubMed: 18631455]

Choi M, Scholl UI, Ji W, Liu T, Tikhonova IR, Zumbo P, Nayir A, Bakkaloglu A, Ozen S, Sanjad S, et al. Proc Natl Acad Sci USA. 2009; 106:19096–19101. [PubMed: 19861545]

Clayton DG. PLoS genetics. 2009; 5:e1000540. [PubMed: 19584936]

Cooney R, Baker J, Brain O, Danis B, Pichulik T, Allan P, Ferguson DJ, Campbell BJ, Jewell D, Simmons A. Nat Med. 2009; 16:90–97. [PubMed: 19966812]

Cooper JD, Smyth DJ, Smiles AM, Plagnol V, Walker NM, Allen JE, Downes K, Barrett JC, Healy BC, Mychaleckyj JC, et al. Nat Genet. 2008; 40:1399–1401. [PubMed: 18978792]

Criswell LA, Pfeiffer KA, Lum RF, Gonzales B, Novitzke J, Kern M, Moser KL, Begovich AB, Carlton VE, Li W, et al. Am J Hum Genet. 2005; 76:561–571. [PubMed: 15719322]

De Jager PL, Jia X, Wang J, de Bakker PI, Ottoboni L, Aggarwal NT, Piccio L, Raychaudhuri S, Tran D, Aubin C, et al. Nat Genet. 2009; 41:776–782. [PubMed: 19525953]

Dendrou CA, Plagnol V, Fung E, Yang JH, Downes K, Cooper JD, Nutland S, Coleman G, Himsworth M, Hardy M, et al. Nat Genet. 2009; 41:1011–1015. [PubMed: 19701192]

Dixon AL, Liang L, Moffatt MF, Chen W, Heath S, Wong KC, Taylor J, Burnett E, Gut I, Farrall M, et al. Nat Genet. 2007; 39:1202–1207. [PubMed: 17873877]

Duerr RH, Taylor KD, Brant SR, Rioux JD, Silverberg MS, Daly MJ, Steinhart AH, Abraham C, Regueiro M, Griffiths A, et al. Science. 2006; 314:1461–1463. [PubMed: 17068223]

Emilsson V, Thorleifsson G, Zhang B, Leonardson AS, Zink F, Zhu J, Carlson S, Helgason A, Walters GB, Gunnarsdottir S, et al. Nature. 2008; 452:423–428. [PubMed: 18344981]

Encode Project Consortium. Nature. 2007; 447:799–816. [PubMed: 17571346]

Festen EA, Goyette P, Scott R, Annese V, Zhernakova A, Lian J, Lefebvre C, Brant SR, Cho JH, Silverberg MS, et al. Gut. 2009; 58:799–804. [PubMed: 19201773]

Franke A, Balschun T, Karlsen TH, Sventoraityte J, Nikolaus S, Mayr G, Domingues FS, Albrecht M, Nothnagel M, Ellinghaus D, et al. Nat Genet. 2008; 40:1319–1323. [PubMed: 18836448]

Gateva V, Sandling JK, Hom G, Taylor KE, Chung SA, Sun X, Ortmann W, Kosoy R, Ferreira RC, Nordmark G, et al. Nat Genet. 2009; 41:1228–1233. [PubMed: 19838195]

Graham RR, Kozyrev SV, Baechler EC, Reddy MV, Plenge RM, Bauer JW, Ortmann WA, Koeuth T, Gonzalez Escribano MF, Pons-Estel B, et al. Nat Genet. 2006; 38:550–555. [PubMed: 16642019]

Gregersen PK, Amos CI, Lee AT, Lu Y, Remmers EF, Kastner DL, Seldin MF, Criswell LA, Plenge RM, Holers VM, et al. Nat Genet. 2009; 41:820–823. [PubMed: 19503088]

Gregersen PK, Olsson LM. Annu Rev Immunol. 2009; 27:363–391. [PubMed: 19302045]

Hafler DA, Compston A, Sawcer S, Lander ES, Daly MJ, De Jager PL, de Bakker PI, Gabriel SB, Mirel DB, Ivinson AJ, et al. N Engl J Med. 2007; 357:851–862. [PubMed: 17660530]

Han JW, Zheng HF, Cui Y, Sun LD, Ye DQ, Hu Z, Xu JH, Cai ZM, Huang W, Zhao GP, et al. Nat Genet. 2009; 41:1234–1237. [PubMed: 19838193]

Hom G, Graham RR, Modrek B, Taylor KE, Ortmann W, Garnier S, Lee AT, Chung SA, Ferreira RC, Pant PV, et al. N Engl J Med. 2008; 358:900–909. [PubMed: 18204098]

Huntington ND, Legrand N, Alves NL, Jaron B, Weijer K, Plet A, Corcuff E, Mortier E, Jacques Y, Spits H, et al. J Exp Med. 2009; 206:25–34. [PubMed: 19103877]

Kaminsky ZA, Tang T, Wang SC, Ptak C, Oh GH, Wong AH, Feldcamp LA, Virtanen C, Halfvarson J, Tysk C, et al. Nat Genet. 2009; 41:240–245. [PubMed: 19151718]

Lin JP, Cash JM, Doyle SZ, Peden S, Kanik K, Amos CI, Bale SJ, Wilder RL. Hum Genet. 1998; 103:475–482. [PubMed: 9856493]

Manolio TA, Collins FS, Cox NJ, Goldstein DB, Hindorff LA, Hunter DJ, McCarthy MI, Ramos EM, Cardon LR, Chakravarti A, et al. Nature. 2009; 461:747–753. [PubMed: 19812666]

Mathis D, Benoist C. Annu Rev Immunol. 2009; 27:287–312. [PubMed: 19302042]

Meissner A, Mikkelsen TS, Gu H, Wernig M, Hanna J, Sivachenko A, Zhang X, Bernstein BE, Nusbaum C, Jaffe DB, et al. Nature. 2008; 454:766–770. [PubMed: 18600261]

Mestas J, Hughes CC. J Immunol. 2004; 172:2731–2738. [PubMed: 14978070]

Nair RP, Duffin KC, Helms C, Ding J, Stuart PE, Goldgar D, Gudjonsson JE, Li Y, Tejasvi T, Feng BJ, et al. Nat Genet. 2009; 41:199–204. [PubMed: 19169254]

Nejentsev S, Walker N, Riches D, Egholm M, Todd JA. Science. 2009; 324:387–389. [PubMed: 19264985]

Ng SB, Turner EH, Robertson PD, Flygare SD, Bigham AW, Lee C, Shaffer T, Wong M, Bhattacharjee A, Eichler EE, et al. Nature. 2009; 461:272–276. [PubMed: 19684571]

Pascual V, Farkas L, Banchereau J. Curr Opin Immunol. 2006; 18:676–682. [PubMed: 17011763]

Plenge RM, Cotsapas C, Davies L, Price AL, de Bakker PI, Maller J, Pe'er I, Burtt NP, Blumenstiel B, DeFelice M, et al. Nat Genet. 2007a; 39:1477–1482. [PubMed: 17982456]

Plenge RM, Seielstad M, Padyukov L, Lee AT, Remmers EF, Ding B, Liew A, Khalili H, Chandrasekaran A, Davies LR, et al. N Engl J Med. 2007b; 357:1199–1209. [PubMed: 17804836]

Rahman A, Isenberg DA. N Engl J Med. 2008; 358:929–939. [PubMed: 18305268]

Raychaudhuri S, Thomson BP, Remmers EF, Eyre S, Hinks A, Guiducci C, Catanese JJ, Xie G, Stahl EA, Chen R, et al. Nat Genet. 2009; 41:1313–1318. [PubMed: 19898481]

Rieck M, Arechiga A, Onengut-Gumuscu S, Greenbaum C, Concannon P, Buckner JH. J Immunol. 2007; 179:4704–4710. [PubMed: 17878369]

Rioux JD, Goyette P, Vyse TJ, Hammarstrom L, Fernando MM, Green T, De Jager PL, Foisy S, Wang J, de Bakker PI, et al. Proc Natl Acad Sci USA. 2009; 106:18680–18685. [PubMed: 19846760]

Schadt EE. Nature. 2009; 461:218–223. [PubMed: 19741703]

Sugimoto K, Ogawa A, Mizoguchi E, Shimomura Y, Andoh A, Bhan AK, Blumberg RS, Xavier RJ, Mizoguchi A. J Clin Invest. 2008; 118:534–544. [PubMed: 18172556]

Thomson W, Barton A, Ke X, Eyre S, Hinks A, Bowes J, Donn R, Symmons D, Hider S, Bruce IN, et al. Nat Genet. 2007; 39:1431–1433. [PubMed: 17982455]

Travassos LH, Carneiro LA, Ramjeet M, Hussey S, Kim YG, Magalhaes JG, Yuan L, Soares F, Chea E, Le Bourhis L, et al. Nat Immunol. 2009; 11:55–62. [PubMed: 19898471]

Turnbaugh PJ, Hamady M, Yatsunenko T, Cantarel BL, Duncan A, Ley RE, Sogin ML, Jones WJ, Roe BA, Affourtit JP, et al. Nature. 2009; 457:480–484. [PubMed: 19043404]

van Heel DA, Franke L, Hunt KA, Gwilliam R, Zhernakova A, Inouye M, Wapenaar MC, Barnardo MC, Bethel G, Holmes GK, et al. Nat Genet. 2007; 39:827–829. [PubMed: 17558408]

Vang T, Congia M, Macis MD, Musumeci L, Orru V, Zavattari P, Nika K, Tautz L, Tasken K, Cucca F, et al. Nat Genet. 2005; 37:1317–1319. [PubMed: 16273109]

Wen L, Ley RE, Volchkov PY, Stranges PB, Avanesyan L, Stonebraker AC, Hu C, Wong FS, Szot GL, Bluestone JA, et al. Nature. 2008; 455:1109–1113. [PubMed: 18806780]

WTCCC. Nature. 2007; 447:661–678. [PubMed: 17554300]

Zenewicz LA, Flavell RA. Eur J Immunol. 2008; 38:3265–3268. [PubMed: 19016525]

Zenewicz LA, Yancopoulos GD, Valenzuela DM, Murphy AJ, Stevens S, Flavell RA. Immunity. 2008; 29:947–957. [PubMed: 19100701]

Zheng Y, Danilenko DM, Valdez P, Kasman I, Eastham-Anderson J, Wu J, Ouyang W. Nature. 2007; 445:648–651. [PubMed: 17187052]

Zhernakova A, Alizadeh BZ, Bevova M, van Leeuwen MA, Coenen MJ, Franke B, Franke L, Posthumus MD, van Heel DA, van der Steege G, et al. Am J Hum Genet. 2007; 81:1284–1288. [PubMed: 17999365]

Zhou X, Bailey-Bucktrout SL, Jeker LT, Penaranda C, Martinez-Llordella M, Ashby M, Nakayama M, Rosenthal W, Bluestone JA. Nat Immunol. 2009; 10:1000–1007. [PubMed: 19633673]

Table 1

Association of Genomic Loci with Autoimmune Diseases

Chromosome Region	Genes of Interest[a]	Function	Diseases[b]
1p13	*PTPN22*	T and B cell receptor signaling	RA, T1D, CD
2q33	*CTLA4*	Transmits inhibitory signals to T cells	T1D, RA
6p21	*MHC*	Major histocompatibility complex	Most autoimmune disorders
1p13	*CD2/CD58*	Activation of T lymphocytes	RA, MS
1p31	*IL23R*	Unique component of the heterodimeric IL-23 receptor	IBD, PS, AS
1q32	*IL10*	Downregulates immune responses, including cytokines, MHC class II and costimulatory molecules	IBD, SLE, T1D
4q26	*IL2/IL21*	T cell trophic growth factors	CeD, IBD, RA, T1D
5q33	*IL12B*	p40 subunit common to IL-12 and IL-23	IBD, PS
10p15	*IL2RA*	IL-2 receptor α chain	MS, T1D
6q23	*TNFAIP3*	Induced by TNF and pattern recognition receptor activation; inhibits NF-κB signaling	RA, SLE, PS
5q33	*TNIP1*	Interacts with TNFAIP3	SLE, PS
6q21	*PRDM1*	Transcriptional repressor of IFN-β; induces B cell maturation	RA, SLE
8p23	*BLK*	B lymphoid tyrosine kinase	SLE, RA
18p11	*PTPN2*	T cell protein tyrosine phosphatase	IBD, T1D

Genome-wide significant association defined as p value $< 5 \times 10^{-8}$.

[a] Association regions often encompass either no genes or multiple genes, with the precise causal gene often not definitively established.

[b] AS, ankylosing spondylitis; CeD, celiac disease; IBD, inflammatory bowel disease; MS, multiple sclerosis; PS, psoriasis; RA, rheumatoid arthritis; T1D, type 1 diabetes mellitus.

Table 2

Emerging Approaches and Potential Advances

Approach	Established and Potential Advances
Genome-wide association studies	Have identified a large number of definitive associations across autoimmunity, with many shared across autoimmune diseases
Search for uncommon DNA variants	May identify more penetrant alleles with larger functional effects
Transcriptome sequencing	Will identify tissue-specific alternative isoforms, noncoding RNAs
Expression quantitative trait loci mapping	Mapping DNA polymorphisms to variable RNA expression
Epigenetic analysis: chromatin modifications	More comprehensive maps of DNA sequences modulating transcriptional regulation
Sequence analysis of the intestinal microbiome	Potentially tractable environment covariate modulating intestinal and systemic immune responses
Humanized mice	Incorporates key human immune response components in model systems
Human immune analyses	Prioritize new therapies, identify disease subtypes, and follow disease course