Published in final edited form as:

Curr Opin Immunol. 2014 December ; 31: 51–57. doi:10.1016/j.coi.2014.09.007.

Genetics of immune-mediated disorders: from genome-wide association to molecular mechanism

Vinod Kumar^{1,*}, Cisca Wijmenga¹, and Ramnik J. Xavier²

¹University of Groningen, University Medical Centre Groningen, Department of Genetics, Groningen, the Netherlands ²Broad Institute of MIT and Harvard, Cambridge, USA; Centre for Computational and Integrative Biology, Massachusetts General Hospital, Harvard Medical School, Boston, Massachusetts 02114, USA

Abstract

Genetic association studies have identified not only hundreds of susceptibility loci to immune-mediated diseases but also pinpointed causal amino-acid variants of *HLA* genes that contribute to many autoimmune reactions. Majority of non-*HLA* genetic variants are located within non-coding regulatory region. Expression QTL studies have shown that these variants affect disease mainly by regulating gene expression. We discuss recent findings on shared genetic loci between infectious and immune-mediated diseases and provide potential clues to explore genetic associations in the context of these infectious agents. We propose that the interdisciplinary studies (genetics-genomics-immunology-infection-bioinformatics) are the future post-GWAS approaches to advance our understanding of the pathogenesis of immune-mediated diseases.

Introduction

The innate and adaptive immune processes are the two vital defence mechanisms of the host's immune system. An inappropriate response of the human immune system against self-antigens leads to immune-mediated conditions such as autoimmune diseases and inflammatory disorders. Although immune-mediated diseases occur in individuals due to complex interactions between environmental, microbial, and genetic factors, twin studies have suggested that genetic variation is an important contributor to differential susceptibility [1,2]. Based on this notion, several genetic studies (mainly candidate-gene analyses and family-based linkage studies) have identified causal genes for immune-mediated diseases with the aim of understanding associated pathogenic mechanisms. However, these studies were unable to identify novel loci quickly, since candidate gene studies rely on previous knowledge of the genes and their molecular connection to disease pathology. In contrast, genome-wide association studies (GWAS) have the significant advantage of finding disease genes in a hypothesis-free manner and not biased by previous knowledge of the disease aetiology. Thus, using single nucleotide polymorphisms (SNPs) as molecular markers, GWAS can reveal novel and unexpected insights into the biology of disease. Indeed, GWAS

*Corresponding author: v.kumar@umcg.nl.

of immune-mediated diseases have accelerated the process of finding hundreds of susceptibility loci [3].

Although GWAS have been highly successful in identifying risk loci, a direct consequence of extensive linkage disequilibrium (LD) is that many regions identified by GWAS are relatively large and contain multiple genes. Fine-mapping of the GWAS-associated autoimmune loci was therefore performed with the immunochip - a dedicated genotyping array designed to refine association signals in autoimmune diseases. The consortium informing the design of the immunochip initially targeted 12 immune-mediated diseases: ankylosing spondylitis, autoimmune thyroid disease, coeliac disease (CeD), Crohn's disease (CD), immunoglobulin A deficiency, multiple sclerosis, primary biliary cirrhosis, psoriasis, rheumatoid arthritis (RA), systemic lupus erythematosus (SLE), type 1 diabetes, and ulcerative colitis. A review by Ricaño-Ponce and Wijmenga [4] has already summarized the genetics of immune-mediated diseases revealed by GWAS studies. Several other studies have also outlined potential post-GWAS approaches that could be taken to prioritize functional genes and variants from GWAS loci, such as expression quantitative trait loci (eQTL) mapping, pathway analyses, and functional annotation of SNPs using regulatory information [5–7]. In this review, we discuss findings from other genomic approaches to show how immune-disease associated SNPs may cause disease and we consider important observations made from GWAS and immunochip studies.

Genetics of immune-mediated diseases revealed by GWAS and Immunochip

The human leukocyte antigen (*HLA*) alleles within the major histocompatibility complex (*MHC*) on chromosome 6p21 have been implicated as susceptibility alleles for several immune-mediated diseases by many pre-GWAS candidate gene association studies [4]. GWAS confirmed the prominent and strong genetic association of many immune-mediated diseases to particular *HLA* alleles. This was anticipated as the *MHC* region is enriched for genes with immune functions, and many *HLA* alleles are required to present self-antigens to T cells, which then induce autoimmune reactions [8]. This work led to post-GWAS efforts, not only to confirm the associations to *HLA* alleles, but also to pinpoint causal amino-acid variants of *HLA* genes that contribute to autoimmune reactions. For example, *HLA-DRB1* alleles have historically been implicated as RA risk alleles. Extending these observations, Raychaudhuri et al. [9] performed haplotype and conditional analyses using RA GWAS data and identified five amino-acid polymorphisms (three in *HLA-DRB1*, one in *HLA-B* and one in *HLA-DPB1*) that all localize to peptide-binding grooves, thereby suggesting effects on presentation of auto-antigens. Building on these results, another study showed that one of the amino-acid variations within *HLA-DRB1* is associated with a differential risk between anti-citrullinated-protein-autoantibody (ACPA)-positive RA and ACPA-negative RA [10]. This clearly suggested roles for both *HLA* variation and distinct auto-antigens in contributing to clinical heterogeneity. These analyses expanded our understanding of the molecular connection between *HLA* alleles and auto-antigens. For example, major *HLA* factors (*HLA-DQ2* and/or *HLA-DQ8*) and an environmental trigger (gliadin peptides) have been well characterized in CeD [11–13], while structural studies have identified dominant

gliadin peptides that interact with HLA-DQ2 and HLA-DQ8 specifically [14,15]. A recent study by Petersen et al. [16] showed that the T cell receptors specific for HLA-DQ2 interact differently than other T cell receptors to these gliadin peptides, suggesting the involvement of a biased T cell repertoire selection in CeD. It is therefore important to test whether this concept is also true in other HLA-associated immune-mediated diseases, as it may provide a potential explanation for the variable inflammatory response seen within a group of patients and might point to an antigen-specific immunotherapy.

In addition to increased understanding of the HLA-associations, GWAS of immune-mediated diseases have led to two additional observations crucial for conceptualizing pathogenic mechanisms of these diseases. The first was that the considerable number of association signals mapping to non-*HLA* loci were shared by different immune-mediated diseases. Hrdlickova et al. (Manuscript submitted) queried the 284 loci associated to eight immune-mediated diseases by Immunochip studies (data available per 1st June 2013; autoimmune thyroid disease [17], CeD [18], inflammatory bowel disease (IBD) [19], juvenile idiopathic arthritis [20], primary biliary cirrhosis [21], psoriasis [22], primary sclerosing cholangitis [23] and RA [24]), and found that more than 41% (119 loci) of the 284 loci were shared by two or more immune-mediated diseases. This further strengthens the idea that there may be common causal mechanisms that lead to most immune-mediated diseases [25].

The second important observation was that more than 90% of SNPs associated with immune-mediated disease are located within the non-coding regions of the genome [4], suggesting that these variants may be regulatory in function. Functional annotation of GWAS SNPs associated with immune-mediated diseases using Encyclopaedia of DNA Elements (ENCODE) data revealed that 7.6% of the variants map to promoter histone marks, 18.8% map to enhancer histone marks, 32.1% are in DNase-hypersensitive sites, 42.3% are SNPs that change motifs, and 14.4% map to protein-binding regions [4]. This observation is also true for variants identified in non-immune disease GWAS, in which more than 93% of SNPs overlap with gene regulatory regions [26]. This regulatory role for most SNPs associated with immune-mediated diseases clearly implies that they may cause disease by affecting expression levels of genes. eQTL mapping studies have further confirmed this, as discussed in more detail below. Following the discovery of functional regulatory SNPs, finding the causal genes becomes vital to understanding the GWAS results and to clarifying how genes are associated with immune-mediated diseases. We believe that this can be efficiently explored by testing whether SNPs associated with immune-mediated diseases are also linked to infectious diseases (Figure 1); we discuss below why this method is beneficial for understanding immune-mediated diseases.

Susceptibility genes revealed by eQTL mapping

Correlating gene expression levels with the genotypes of any genetic variant is called eQTL mapping. If an affected gene is located in close proximity to a SNP (i.e. less than one megabase from the SNP), common usage refers to this as a *cis*-eQTL, whereas a *trans*-eQTL effect occurs when an affected gene is located at a greater distance from the SNP (i.e. more than five megabases to the SNP or on a different chromosome than the SNP). eQTL

mapping has been demonstrated to be a powerful approach for prioritizing disease susceptibility genes from GWAS loci [27]. For example, 39 of 71 (55%) CD loci shown an eQTL effect [28], as do 14 of 26 (54%) CeD loci [29], and 32 of 53 (60%) type 1 diabetes loci [30]. Recently, Ricaño-Ponce and Wijmenga [4] systematically mapped eQTL for all the GWAS SNPs associated with 12 different immune-mediated diseases and found that 54% of these SNPs were *cis*-eQTLs. More insight into underlying disease pathways can be obtained by mapping *trans*-eQTL. For example, the SLE-associated SNP rs4917014 was found to affect the *IKZF1* gene in *cis* [31]. *IKZF1* encodes a transcription factor and hence regulates many other genes. It is therefore conceivable that the SLE SNP could also be associated with expression levels of *IKZF1* targets. Indeed, the same SLE SNP was also shown to be associated with the expression levels of eight different genes (*CLEC10A*, *CLEC4C*, *C1QB*, *IFI6*, *HERC5*, *IFIT1*, *TNFRSF21*, and *MX1*) in *trans* that are involved in type 1 interferon (IFN) response or complement activation [31]. This finding reveals the mechanism of how a GWAS-identified SNP for SLE regulates the network of genes responsible for increased type 1 IFN response and decreased complement expression, which are hallmark features of SLE [32–34].

Linking the genetics of infectious disease and immune-mediated diseases

Among several environmental factors [35], infections have been largely implicated in the onset and/or promotion of immune-mediated diseases [36,37]. Several mechanisms have been proposed to explain how infectious agents, such as viruses, bacteria, fungi and parasites, can trigger immune-mediated diseases [35]. However, it has not been clear how to mechanistically link SNPs associated with immune-mediated diseases to explain the role of an infectious triggers in provoking disease. Here, we provide examples of genetic studies to highlight how these approaches may help us understand the functional basis for susceptibility loci shared by both infectious and immune-mediated diseases. The first genetic insight arose from a GWAS study of leprosy, in which leprosy susceptibility loci were identified that overlap with 5 IBD risk loci [38]. A recent immunochip study of IBD extended this overlap by implicating seven of the eight leprosy loci as IBD loci [19]. Furthermore, they found a large overlap of IBD loci with genes involved in primary immune-deficiencies resulting in severe infections, including a Mendelian susceptibility to mycobacterial disease [19]. Taking virus infection as an example, an enrichment of shared HLA susceptibility alleles, including a variant associated with high surface expression of HLA-C, was found for HIV-1 infection and psoriasis [39]. Later, it was shown that higher HLA-C expression levels are associated with protection in HIV infection, but also with increased risk of CD [40]. In another study, Lee et al. [41] analysed CD GWAS data and identified SNP (rs12212067) in the *FOXO3A* locus as being associated with prognosis of CD and RA. On lipopolysaccharide (LPS) stimulation of monocytes, the G allele at rs12212067 was associated with the up-regulation of *FOXO3A* and differential levels of pro- and anti-inflammatory cytokines. Furthermore, the authors demonstrate that the same variant (rs12212067) is associated with an increased risk of severe malaria.

Interestingly, this genetic overlap is not restricted to bacterial, viral, and parasitic infections. We recently performed an immunochip-wide association study to identify susceptibility genes to *Candida* infection (Nat. Communications, In press). We found three loci associated

with immune-mediated diseases (psoriasis, CeD and multiple sclerosis) to be also significantly associated with candidemia with very high effect sizes (odds ratios of 2.96 to 4.68). Further, it is interesting to note that, in general, SNPs associated with infectious diseases show very high effect sizes (odds ratios of > 2) [42], in contrast to immune-mediated disease SNPs. It is therefore plausible that the genetic architecture of infectious diseases could be driven by multiple, individual, rare genetic variants with incomplete penetrance, but with a strong impact on disease. These findings strongly support the hypothesis that infectious pressures on the immune system select genetic variants that result in a strong immune responses towards eliminating pathogens, which in turn lead to a greater susceptibility to inflammatory or autoimmune diseases [43]. One of the earliest studies to support this idea focused on the *SH2B3* locus, which is associated with several autoimmune diseases including CeD. *SH2B3* encodes a protein with a structure similar to proteins that inhibit cytokine-signalling. The *SH2B3* SNP (rs3184504) is a non-synonymous variant (R262W) associated with enhanced cytokine responses after immune cells have been stimulated with bacterial antigens [44]. Although this SNP is very common in Europeans, it is rare in Asians and Africans. Hence, the study estimated that a wave of positive selection in Europe may have occurred 1,200–1,700 years ago, as a protective variant against the Justinian plague (*Yersinia pestis*) [44].

Fortunately, current genomic technologies allow us to explore these findings in more detail to understand the functional consequences of the positively selected variants. For example, an eQTL study using steady-state gene expression levels in blood cells mapped eQTL to GWAS SNPs [31]. This study not only confirmed the *cis*-eQTL effect of the SNP on *SH2B3* but also revealed its *trans*-eQTL effect on 29 other genes (Figure 2). Interestingly, the risk allele for CeD (allele A at rs3184504) was shown to up-regulate 11 genes that are enriched for toll-like receptor signalling and response to IFN-gamma (Figure 2). It is therefore conceivable that the positively selected *SH2B3* allele is capable of inducing a strong innate immune response as a defence mechanism against pathogens, but enhances risk for autoimmune disease due to an enhanced inflammatory response. These eQTL studies provided valuable insights to advance the understanding of the associated disease mechanisms. However, there are some limitations with these studies, because the role of genotype-environment interactions, cell-type-specific gene expression patterns, and non-coding RNAs have not been considered. Below we discuss some recent studies to highlight the relevance of these factors in gaining a much deeper insight and for future work.

Linking infectious triggers to GWAS immune-mediated disease loci

The majority of the eQTL analyses (discussed above) have been performed in peripheral blood mononuclear cells using steady-state gene expression levels. However, the effect of genetic variants on gene expression often dependent on cell type [45]. In addition, the complex interaction between genetic factors, infections, and autoimmune diseases highlights some pathogenic relationships between different infectious agents and immune-mediated diseases. Thus, mapping eQTL using gene expression levels obtained from immune cell types that have been stimulated with infectious agents can yield valuable insights. Indeed, recent studies have highlighted a complex pattern of gene-regulation through SNPs associated with immune-mediated diseases, which is highly dependent on cell-type,

stimulation, and duration of stimulation [46–48]. Fairfax et al. [46] mapped eQTL in primary CD14⁺ human monocytes upon activation with IFN-gamma or with LPS for 2 or 24 hours. They found that more than 50% of the *cis*-eQTL that overlapped GWAS loci were observed only in cells stimulated with IFN-gamma. Several LPS-stimulation-specific eQTL were also identified, many of which depended on the duration of stimulation. For example, the multiple sclerosis associated SNP rs17445836 was identified as a strong regulator of the *IRF8* gene expression, specifically in cells treated with LPS for 2 hours. In another study, Lee et al. [47] performed eQTL analysis in dendritic cells treated with three innate antigens (LPS, influenza virus, and IFN- β). This study found 38 SNPs associated with immune-mediated diseases to be significant eQTL only in the stimulation state. Examples include *TRAF1* for RA and CeD, *CREM* for CD and ulcerative colitis, and *IRF7* for SLE. Results from these studies provide novel links between SNPs associated with immune-mediated disease and specific pathways in innate immune cells.

To date, many eQTL studies have relied on microarray technology to measure gene expression levels and have therefore only correlated the expression of protein-coding genes with disease-associated SNPs. However, the advent of next generation sequencing has led to the identification of thousands of non-protein coding genes, mainly long non-coding RNAs (lncRNA) [49,50]. lncRNAs are critical regulators of gene expression and thus play a significant role in many important biological aspects, including cell differentiation, cancer and the immune system [51–53]. lncRNAs are also involved in regulating specific immune processes such as NF- κ B signalling, the anti-viral response, CD4⁺ and CD8⁺ T cell differentiation, and the inflammatory response [54–57]. In the context of immune-mediated diseases, Ricaño-Ponce and Wijmenga [4] have shown that approximately 10% of SNPs associated with immune-mediated diseases map to lncRNAs. To gain insight into the role of lncRNAs in immune-mediated diseases, Hrdlickova et al. (Manuscript submitted) tested the expression of lncRNAs and protein-coding genes located in loci associated with eight immune-mediated diseases, across 11 distinct immune cell types (granulocytes, monocytes, NK cells, B cells, memory T cells, naïve CD4⁺ T cells, and naïve CD8⁺ T cells, and four different CD4⁺ T-helper cell populations). This analysis revealed that the lncRNAs mapping to immune-mediated diseases are significantly enriched in immune cell types compared to lncRNAs from the whole genome. Since human disease-associated SNPs can influence the expression levels of lncRNAs [58], future eQTL studies should test whether SNPs associated with immune-mediated disease affect lncRNA expression by using RNAseq data from immune cell types, in the context of external factors, particularly infectious agents (Figure 1).

Conclusions

Understanding the biological pathways involved in immune-mediated disease is crucial for identifying appropriate targets for therapy. Genetic studies have contributed significantly towards identifying important susceptibility loci and have shown that disease is caused by many individual genetic variations, with a major contribution from environmental factors. To date, eQTL studies have provided a functional basis for approximately 50% of the SNPs associated with immune-mediated disease. Future eQTL studies should consider genetic variations in the context of external factors, particularly infectious agents, to provide a

comprehensive functional basis to GWAS associations. Thus, future post-GWAS studies should focus on integrative approaches [59] to test whether SNPs associated with immune-mediated disease have other effects, including changes to cytokine responses to stimulation, epigenetic markers, metabolites, or the microbiome component. This strategy will not only provide a mechanistic explanation for the immune-disease associations, but also help us to choose relevant therapeutic targets.

Acknowledgments

This work was supported by the European Research Council Advanced Grant (ERC-671274 to CW), the Dutch Digestive Diseases Foundation (MLDS WO11-30 to CW and VK), the European Union's Seventh Framework Programme (EU FP7) TANDEM project (HEALTH-F3-2012-305279 to CW and VK) and CCFA, Helmsley Charitable Trust and National Institutes of Health (DK43351 to RJX). We thank Jackie Senior and Kate McIntyre for editing this paper.

References and recommended reading

Papers of particular interest, published within the period of review, have been highlighted as:

- of special interest
- of outstanding interest

1. Amur S, Parekh A, Mummaneni P. Sex differences and genomics in autoimmune diseases. *J Autoimmun.* 2012; 38:J254–265. [PubMed: 22204900]
2. Wandstrat A, Wakeland E. The genetics of complex autoimmune diseases: non-MHC susceptibility genes. *Nat Immunol.* 2001; 2:802–809. [PubMed: 11526390]
3. Welter D, MacArthur J, Morales J, Burdett T, Hall P, Junkins H, Klemm A, Flicek P, Manolio T, Hindorf L, et al. The NHGRI GWAS Catalog, a curated resource of SNP-trait associations. *Nucleic Acids Res.* 2014; 42:D1001–1006. [PubMed: 24316577]
- 4••. Ricano-Ponce I, Wijmenga C. Mapping of Immune-Mediated Disease Genes. *Annu Rev Genomics Hum Genet.* 2013 In this review, the authors have performed a systematic annotation of all autoimmune disease associated loci to provide a complete picture of genetics of immune-mediated diseases.
5. Hou L, Zhao H. A review of post-GWAS prioritization approaches. *Front Genet.* 2013; 4:280. [PubMed: 24367376]
6. Zhang X, Bailey SD, Lupien M. Laying a solid foundation for Manhattan - 'setting the functional basis for the post-GWAS era'. *Trends Genet.* 2014; 30:140–149. [PubMed: 24661571]
7. Kumar V, Wijmenga C, Withoff S. From genome-wide association studies to disease mechanisms: celiac disease as a model for autoimmune diseases. *Semin Immunopathol.* 2012; 34:567–580. [PubMed: 22580835]
8. Horton R, Wilming L, Rand V, Lovering RC, Bruford EA, Khodiyar VK, Lush MJ, Povey S, Talbot CC Jr, Wright MW, et al. Gene map of the extended human MHC. *Nat Rev Genet.* 2004; 5:889–899. [PubMed: 15573121]
- 9•. Raychaudhuri S, Sandor C, Stahl EA, Freudenberg J, Lee HS, Jia X, Alfredsson L, Padyukov L, Klareskog L, Worthington J, et al. Five amino acids in three HLA proteins explain most of the association between MHC and seropositive rheumatoid arthritis. *Nat Genet.* 2012; 44:291–296. This is the first study to pinpoint the exact HLA amino acids involved in rheumatoid arthritis using genetic data. [PubMed: 22286218]
10. Han B, Diogo D, Eyre S, Kallberg H, Zhernakova A, Bowes J, Padyukov L, Okada Y, Gonzalez-Gay MA, Rantapaa-Dahlqvist S, et al. Fine Mapping Seronegative and Seropositive Rheumatoid Arthritis to Shared and Distinct HLA Alleles by Adjusting for the Effects of Heterogeneity. *Am J Hum Genet.* 2014; 94:522–532. [PubMed: 24656864]

- 11•. Abadie V, Sollid LM, Barreiro LB, Jabri B. Integration of genetic and immunological insights into a model of celiac disease pathogenesis. *Annu Rev Immunol.* 2011; 29:493–525. An excellent review which not only provides interesting immunological insights into celiac disease but also provides a framework to explore genetic association results to understand other immune-mediated diseases. [PubMed: 21219178]
12. Jabri B, Sollid LM. Tissue-mediated control of immunopathology in coeliac disease. *Nat Rev Immunol.* 2009; 9:858–870. [PubMed: 19935805]
- 13•. Sollid LM, Jabri B. Triggers and drivers of autoimmunity: lessons from coeliac disease. *Nat Rev Immunol.* 2013; 13:294–302. The authors in this review highlight the role of infectious agents in causing the autoimmune phenotype in celiac disease. [PubMed: 23493116]
14. Henderson KN, Reid HH, Borg NA, Broughton SE, Huyton T, Anderson RP, McCluskey J, Rossjohn J. The production and crystallization of the human leukocyte antigen class II molecules HLA-DQ2 and HLA-DQ8 complexed with deamidated gliadin peptides implicated in coeliac disease. *Acta Crystallogr Sect F Struct Biol Cryst Commun.* 2007; 63:1021–1025.
15. Kim CY, Quarsten H, Bergseng E, Khosla C, Sollid LM. Structural basis for HLA-DQ2-mediated presentation of gluten epitopes in celiac disease. *Proc Natl Acad Sci U S A.* 2004; 101:4175–4179. [PubMed: 15020763]
16. Petersen J, Montserrat V, Mujico JR, Loh KL, Beringer DX, van Lummel M, Thompson A, Mearin ML, Schweizer J, Kooy-Winkelaar Y, et al. T-cell receptor recognition of HLA-DQ2-gliadin complexes associated with celiac disease. *Nat Struct Mol Biol.* 2014; 21:480–488. [PubMed: 24777060]
17. Cooper JD, Simmonds MJ, Walker NM, Burren O, Brand OJ, Guo H, Wallace C, Stevens H, Coleman G, Franklyn JA, et al. Seven newly identified loci for autoimmune thyroid disease. *Hum Mol Genet.* 2012; 21:5202–5208. [PubMed: 22922229]
18. Trynka G, Hunt KA, Bockett NA, Romanos J, Mistry V, Szperl A, Bakker SF, Bardella MT, Bhaw-Rosun L, Castillejo G, et al. Dense genotyping identifies and localizes multiple common and rare variant association signals in celiac disease. *Nat Genet.* 2011; 43:1193–1201. [PubMed: 22057235]
- 19•. Jostins L, Ripke S, Weersma RK, Duerr RH, McGovern DP, Hui KY, Lee JC, Schumm LP, Sharma Y, Anderson CA, et al. Host-microbe interactions have shaped the genetic architecture of inflammatory bowel disease. *Nature.* 2012; 491:119–124. A large case-control study that highlighted the significant contribution of microbes in causing the inflammatory bowel disease. [PubMed: 23128233]
20. Hinks A, Cobb J, Marion MC, Prahalad S, Sudman M, Bowes J, Martin P, Comeau ME, Sajuthi S, Andrews R, et al. Dense genotyping of immune-related disease regions identifies 14 new susceptibility loci for juvenile idiopathic arthritis. *Nat Genet.* 2013; 45:664–669. [PubMed: 23603761]
21. Liu JZ, Almarri MA, Gaffney DJ, Mells GF, Jostins L, Cordell HJ, Ducker SJ, Day DB, Heneghan MA, Neuberger JM, et al. Dense fine-mapping study identifies new susceptibility loci for primary biliary cirrhosis. *Nat Genet.* 2012; 44:1137–1141. [PubMed: 22961000]
22. Tsoi LC, Spain SL, Knight J, Ellinghaus E, Stuart PE, Capon F, Ding J, Li Y, Tejasvi T, Gudjonsson JE, et al. Identification of 15 new psoriasis susceptibility loci highlights the role of innate immunity. *Nat Genet.* 2012; 44:1341–1348. [PubMed: 23143594]
23. Liu JZ, Hov JR, Folseraas T, Ellinghaus E, Rushbrook SM, Doncheva NT, Andreassen OA, Weersma RK, Weismuller TJ, Eksteen B, et al. Dense genotyping of immune-related disease regions identifies nine new risk loci for primary sclerosing cholangitis. *Nat Genet.* 2013; 45:670–675. [PubMed: 23603763]
24. Eyre S, Bowes J, Diogo D, Lee A, Barton A, Martin P, Zhernakova A, Stahl E, Viatte S, McAllister K, et al. High-density genetic mapping identifies new susceptibility loci for rheumatoid arthritis. *Nat Genet.* 2012; 44:1336–1340. [PubMed: 23143596]
- 25•. Zhernakova A, van Diemen CC, Wijmenga C. Detecting shared pathogenesis from the shared genetics of immune-related diseases. *Nat Rev Genet.* 2009; 10:43–55. One of the first studies to provide evidence for shared genes across many immune-mediated diseases. [PubMed: 19092835]
- 26•. Maurano MT, Humbert R, Rynes E, Thurman RE, Haugen E, Wang H, Reynolds AP, Sandstrom R, Qu H, Brody J, et al. Systematic localization of common disease-associated variation in

regulatory DNA. *Science*. 2012; 337:1190–1195. This study clearly shows that majority of disease-associated common polymorphisms are located within regulatory regions, suggesting their role of gene-expression regulation. [PubMed: 22955828]

27. Westra HJ, Franke L. From genome to function by studying eQTLs. *Biochim Biophys Acta*. 2014
28. Franke A, McGovern DP, Barrett JC, Wang K, Radford-Smith GL, Ahmad T, Lees CW, Balschun T, Lee J, Roberts R, et al. Genome-wide meta-analysis increases to 71 the number of confirmed Crohn's disease susceptibility loci. *Nat Genet*. 2010; 42:1118–1125. [PubMed: 21102463]
29. Dubois PC, Trynka G, Franke L, Hunt KA, Romanos J, Curtotti A, Zhernakova A, Heap GA, Adany R, Aromaa A, et al. Multiple common variants for celiac disease influencing immune gene expression. *Nat Genet*. 2010; 42:295–302. [PubMed: 20190752]
30. Barrett JC, Clayton DG, Concannon P, Akolkar B, Cooper JD, Erlich HA, Julier C, Morahan G, Nerup J, Nierras C, et al. Genome-wide association study and meta-analysis find that over 40 loci affect risk of type 1 diabetes. *Nat Genet*. 2009; 41:703–707. [PubMed: 19430480]
31. Westra HJ, Peters MJ, Esko T, Yaghootkar H, Schurmann C, Kettunen J, Christiansen MW, Fairfax BP, Schramm K, Powell JE, et al. Systematic identification of trans eQTLs as putative drivers of known disease associations. *Nat Genet*. 2013; 45:1238–1243. By using the largest number of samples for eQTL mapping to date, the authors provide clear connection between disease associated SNPs and the affected pathways. [PubMed: 24013639]
32. Baechler EC, Batliwalla FM, Karypis G, Gaffney PM, Ortmann WA, Espe KJ, Shark KB, Grande WJ, Hughes KM, Kapur V, et al. Interferon-inducible gene expression signature in peripheral blood cells of patients with severe lupus. *Proc Natl Acad Sci U S A*. 2003; 100:2610–2615. [PubMed: 12604793]
33. Bennett L, Palucka AK, Arce E, Cantrell V, Borvak J, Banchereau J, Pascual V. Interferon and granulopoiesis signatures in systemic lupus erythematosus blood. *J Exp Med*. 2003; 197:711–723. [PubMed: 12642603]
34. Botto M, Dell'Agnola C, Bygrave AE, Thompson EM, Cook HT, Petry F, Loos M, Pandolfi PP, Walport MJ. Homozygous C1q deficiency causes glomerulonephritis associated with multiple apoptotic bodies. *Nat Genet*. 1998; 19:56–59. [PubMed: 9590289]
35. Vojdani A. A Potential Link between Environmental Triggers and Autoimmunity. *Autoimmune Dis*. 2014; 2014:437231. [PubMed: 24688790]
36. Munz C, Lunemann JD, Getts MT, Miller SD. Antiviral immune responses: triggers of or triggered by autoimmunity? *Nat Rev Immunol*. 2009; 9:246–258. [PubMed: 19319143]
37. Kivity S, Agmon-Levin N, Blank M, Shoenfeld Y. Infections and autoimmunity--friends or foes? *Trends Immunol*. 2009; 30:409–414. [PubMed: 19643667]
38. Zhang F, Liu H, Chen S, Low H, Sun L, Cui Y, Chu T, Li Y, Fu X, Yu Y, et al. Identification of two new loci at IL23R and RAB32 that influence susceptibility to leprosy. *Nat Genet*. 2011; 43:1247–1251. [PubMed: 22019778]
39. Chen H, Hayashi G, Lai OY, Dilthey A, Kuebler PJ, Wong TV, Martin MP, Fernandez Vina MA, McVean G, Wabl M, et al. Psoriasis patients are enriched for genetic variants that protect against HIV-1 disease. *PLoS Genet*. 2012; 8:e1002514. [PubMed: 22577363]
40. Apps R, Qi Y, Carlson JM, Chen H, Gao X, Thomas R, Yuki Y, Del Prete GQ, Goulder P, Brumme ZL, et al. Influence of HLA-C expression level on HIV control. *Science*. 2013; 340:87–91. [PubMed: 23559252]
41. Lee JC, Espeli M, Anderson CA, Linterman MA, Pocock JM, Williams NJ, Roberts R, Viatte S, Fu B, Peshu N, et al. Human SNP links differential outcomes in inflammatory and infectious disease to a FOXO3-regulated pathway. *Cell*. 2013; 155:57–69. Among the first examples of how a inflammatory disease associated SNP can also influence infectious disease. [PubMed: 24035192]
42. Hill AV. Evolution, revolution and heresy in the genetics of infectious disease susceptibility. *Philos Trans R Soc Lond B Biol Sci*. 2012; 367:840–849. [PubMed: 22312051]
43. Netea MG, Wijmenga C, O'Neill LA. Genetic variation in Toll-like receptors and disease susceptibility. *Nat Immunol*. 2012; 13:535–542. [PubMed: 22610250]
44. Zhernakova A, Elbers CC, Ferwerda B, Romanos J, Trynka G, Dubois PC, de Kovel CG, Franke L, Oosting M, Barisani D, et al. Evolutionary and functional analysis of celiac risk loci reveals

- SH2B3 as a protective factor against bacterial infection. *Am J Hum Genet.* 2010; 86:970–977. [PubMed: 20560212]
45. Su AI, Cooke MP, Ching KA, Hakak Y, Walker JR, Wiltshire T, Orth AP, Vega RG, Sapinoso LM, Moqrich A, et al. Large-scale analysis of the human and mouse transcriptomes. *Proc Natl Acad Sci U S A.* 2002; 99:4465–4470. [PubMed: 11904358]
  46. Fairfax BP, Humburg P, Makino S, Naranbhai V, Wong D, Lau E, Jostins L, Plant K, Andrews R, McGee C, et al. Innate immune activity conditions the effect of regulatory variants upon monocyte gene expression. *Science.* 2014; 343:1246949. One of the first studies that indicated that some of the immune-mediated SNPs regulate gene-expression in a stimulation and time dependent manner. [PubMed: 24604202]
  47. Lee MN, Ye C, Villani AC, Raj T, Li W, Eisenhaure TM, Imboywa SH, Chipendo PI, Ran FA, Slowikowski K, et al. Common genetic variants modulate pathogen-sensing responses in human dendritic cells. *Science.* 2014; 343:1246980. One of the first studies that indicated that some of the immune-mediated SNPs regulate gene-expression in a stimulation and time dependent manner. [PubMed: 24604203]
  48. Raj T, Rothamel K, Mostafavi S, Ye C, Lee MN, Replogle JM, Feng T, Lee M, Asinovski N, Frohlich I, et al. Polarization of the effects of autoimmune and neurodegenerative risk alleles in leukocytes. *Science.* 2014; 344:519–523. [PubMed: 24786080]
  49. Guttman M, Amit I, Garber M, French C, Lin MF, Feldser D, Huarte M, Zuk O, Carey BW, Cassady JP, et al. Chromatin signature reveals over a thousand highly conserved large non-coding RNAs in mammals. *Nature.* 2009; 458:223–227. [PubMed: 19182780]
  50. Mercer TR, Dinger ME, Mattick JS. Long non-coding RNAs: insights into functions. *Nat Rev Genet.* 2009; 10:155–159. [PubMed: 19188922]
  51. Rinn JL, Chang HY. Genome regulation by long noncoding RNAs. *Annu Rev Biochem.* 2012; 81:145–166. [PubMed: 22663078]
  52. Hu W, Alvarez-Dominguez JR, Lodish HF. Regulation of mammalian cell differentiation by long non-coding RNAs. *EMBO Rep.* 2012; 13:971–983. [PubMed: 23070366]
  53. Fitzgerald KA, Caffrey DR. Long noncoding RNAs in innate and adaptive immunity. *Curr Opin Immunol.* 2014; 26:140–146. [PubMed: 24556411]
  54. Peng X, Gralinski L, Armour CD, Ferris MT, Thomas MJ, Proll S, Bradel-Tretheway BG, Korth MJ, Castle JC, Biery MC, et al. Unique signatures of long noncoding RNA expression in response to virus infection and altered innate immune signaling. *MBio.* 2010;1.
  55. Collier SP, Collins PL, Williams CL, Boothby MR, Aune TM. Cutting edge: influence of Tmevpg1, a long intergenic noncoding RNA, on the expression of Ifng by Th1 cells. *J Immunol.* 2012; 189:2084–2088. [PubMed: 22851706]
  56. Carpenter S, Aiello D, Atianand MK, Ricci EP, Gandhi P, Hall LL, Byron M, Monks B, Henry-Bezy M, Lawrence JB, et al. A long noncoding RNA mediates both activation and repression of immune response genes. *Science.* 2013; 341:789–792. One of the first studies to provide mechanistic insights into the role of long non-coding RNAs in regulating immune response. [PubMed: 23907535]
  57. Pang KC, Dinger ME, Mercer TR, Malquori L, Grimmond SM, Chen W, Mattick JS. Genome-wide identification of long noncoding RNAs in CD8+ T cells. *J Immunol.* 2009; 182:7738–7748. [PubMed: 19494298]
  58. Kumar V, Westra HJ, Karjalainen J, Zhernakova DV, Esko T, Hrdlickova B, Almeida R, Zhernakova A, Reinmaa E, Vosa U, et al. Human disease-associated genetic variation impacts large intergenic non-coding RNA expression. *PLoS Genet.* 2013; 9:e1003201. [PubMed: 23341781]
  59. Smeeckens SP, Ng A, Kumar V, Johnson MD, Plantinga TS, van Diemen C, Arts P, Verwiel ET, Gresnigt MS, Fransen K, et al. Functional genomics identifies type I interferon pathway as central for host defense against *Candida albicans*. *Nat Commun.* 2013; 4:1342. [PubMed: 23299892]

Highlights

1. Majority of the SNPs associated with immune-mediated disease impact gene expression.
2. Growing evidence on shared genetics between infectious and immune-mediated diseases.
3. Several autoimmune SNPs are stimulation or condition specific eQTLs.
4. Integrative approaches are powerful post-GWAS strategies.

Figure 1.

Shown is a flowchart outlining the steps to identify relevant triggers of autoimmunity. Analyzing the intersection between SNPs associated with immune-mediated diseases and the genetics of infectious diseases and other endophenotypes helps us to prioritize microbial and environmental triggers. These triggers can then be used as stimuli to activate immune cells to obtain transcriptional responses by RNA sequencing. *Cis*- and *trans*-eQTL mapping can then identify both the causal genes and pathways. This information will yield insight into disease mechanisms and in turn inform the choice of relevant therapeutic targets.

Figure 2.

SNP rs3184504 on human chromosome 12 is associated with autoimmune diseases. eQTL mapping showed that the autoimmune risk allele (A allele) up-regulates *SH2B3* gene expression (*cis*-eQTL) and also affects 29 other genes on different chromosomes (*trans*-eQTL). Pathway analysis showed enrichment of genes for innate immunity among the up-regulated genes (green dotted arrows), and enrichment of genes for chemokine signalling pathway among the down-regulated genes (red dotted arrows).