Emerging patterns of genetic overlap across autoimmune disorders.

Permalink
https://escholarship.org/uc/item/5x43v70s

Journal
Genome medicine, 4(1)

ISSN
1756-994X

Authors
Richard-Miceli, Corinne
Criswell, Lindsey A

Publication Date
2012-01-27

DOI
10.1186/gm305

Peer reviewed
Emerging patterns of genetic overlap across autoimmune disorders

Corinne Richard-Miceli and Lindsey A Criswell*

Abstract

Most of the recently identified autoimmunity loci are shared among multiple autoimmune diseases. The pattern of genetic association with autoimmune phenotypes varies, suggesting that certain subgroups of autoimmune diseases are likely to share etiological similarities and underlying mechanisms of disease. In this review, we summarize the major findings from recent studies that have sought to refine genotype-phenotype associations in autoimmune disease by identifying both shared and distinct autoimmunity loci. More specifically, we focus on information from recent genome-wide association studies of rheumatoid arthritis, ankylosing spondylitis, celiac disease, multiple sclerosis, systemic lupus erythematosus, type 1 diabetes and inflammatory bowel disease. Additional work in this area is warranted given both the opportunity it provides to elucidate pathogenic mechanisms in autoimmunity and its potential to inform the development of improved diagnostic and therapeutic tools for this group of complex human disorders.

Genetic factors shared among diverse autoimmune disorders

Autoimmune disorders (AIDs), which as a group affect approximately 8.5% of individuals worldwide [1], are responsible for a substantial amount of disability and morbidity. Some AIDs are organ specific (for example, type 1 diabetes (T1D) targets the pancreas, autoimmune thyroid disease (AITD) attacks the thyroid gland), whereas others can affect multiple organs and/or be associated with systemic manifestations. Systemic lupus erythematosus (SLE) is the prototypic systemic AID that can affect multiple organs and can also be associated with significant systemic manifestations, morbidity and early mortality [2]. Most AIDs, including rheumatoid arthritis (RA), ankylosing spondylitis (AS), inflammatory bowel disease (IBD) and multiple sclerosis (MS) have a predilection for specific organs (for example, the synovial joints in RA and the gastrointestinal tract in IBD) but are also associated with manifestations outside the primary target organ. Reasons for the diverse manifestations exhibited by different AIDs remain unclear, but recent progress in elucidating genetic susceptibility loci for this group of disorders promises to shed light on this important issue.

Although AIDs encompass a broad range of phenotypic manifestations and severity, several features suggest that they share common etiologic factors. For example, most AIDs are characterized by female predominance, and many are associated with the production of autoantibodies (for example, anti-citrullinated-peptide antibodies are observed among 70 to 80% of RA patients). These shared disease features, in conjunction with epidemiologic evidence that demonstrates the clustering of multiple AIDs within individuals and families, strongly implicate shared etiologic factors, including shared genetic loci.

Familial clustering of autoimmune disorders has been long recognized and supports a role for shared genetic predisposition. For example, family studies have documented the clustering of certain autoimmune diseases among the relatives of individuals who have RA, MS, SLE, T1D and other diseases [3-9]. One of the earliest autoimmune disease clusters to be described involved RA, T1D and AITD, and this cluster was recently supported by a systematic review of studies describing clustering of RA, T1D, AITD and MS [3]. Interestingly, that systematic review found evidence of inverse clustering of RA and MS, suggesting that MS and RA might be less closely related than some other AIDs [3]. This conclusion was reinforced by a recent study comparing the genetic variation profiles of six AIDs [10]. That study, by Sirota et al. [10], was based on an analysis of a large number of genetic variants examined in recent genome-wide association studies (GWAS). It found that RA and AS appeared to represent one AID cluster that is distinct from another represented by MS and AITD, with T1D...
showing similarity to both groups and Crohn's disease to neither.

Early candidate gene studies, particularly those focusing on genes within the human leukocyte antigen (HLA) region [11], also supported the notion of shared ‘autoimmunity’ loci. Strong support for genetic loci that are shared across autoimmune disorders and located outside the HLA region has been demonstrated for several loci encoding proteins that have immune-mediating functions, including cytokotic T-lymphocyte antigen 4 (CTLA4; a member of the immunoglobulin superfamily that is expressed on the surface of helper T cells and transmits an inhibitory signal to T cells), protein tyrosine phosphatase non-receptor type 22 (PTPN22; which is expressed primarily in lymphoid tissue and plays a role in the regulation of T-cell receptor signaling pathways), and tumor necrosis factor (TNF) alpha-induced protein 3 (TNFAIP3; which inhibits NF-kappa B activation as well as TNF-mediated apoptosis) [12-14]. Many of the recently identified AID loci involve pathways related to B-cell or T-cell activation and differentiation, innate immunity, and regulation of cytokine signaling [15,16]. Certain loci, however, appear to be associated with specific autoimmune diseases. For example, variants in NOD2 (nucleotide-binding oligomerization domain containing 2) and ATG16L1 (ATG16 autophagy-related-16-like 1) have been associated with defective autophagy in dendritic cells from Crohn's disease patients [17].

A theme emerging from recent genetic studies of AIDs relates to the surprising degree of overlap between genetic loci for this diverse group of disorders, given the phenotypic diversity. Several recent reviews have summarized emerging work that identifies both genetic loci that are shared across the spectrum of autoimmune disease and the biologic pathways whose involvement is implicated by these loci [15,16,18,19]. For example, Zhernakova et al. [18] completed a detailed review of 16 genome-wide association (GWA) or non-synonymous SNP scans for 11 immune-related disorders that were published in 2007 or 2008. Their analysis underscores the extensive sharing of genetic risk loci across this spectrum of disorders, and the fact that most of these loci can be mapped to a few shared biologic pathways, including those related to innate immunity, immune signaling, T-cell differentiation, cytokines and chemokines.

The analysis by Zhernakova et al. [18] also suggests that the degree to which each of these disorders is characterized by shared (rather than unique) susceptibility loci varies substantially, from all loci shared (for RA) to 50% or more shared for celiac disease, psoriasis, MS, SLE, T1D, AS and AITD [18]. The two types of IBD examined, Crohn's disease and ulcerative colitis (UC), shared substantial numbers of loci between them but relatively few with the other AIDs studied. The extent to which the T-cell differentiation, immune cell signaling, innate immunity and TNF signaling, or other pathways are implicated for each of these disorders varies, but overall the analysis by Zhernakova et al. [18] suggests that most of these pathways contribute (to a variable degree) to most of these disorders.

In this review we focus on recent studies that have sought to refine genotype-phenotype associations by comparing susceptibility loci between specific AIDs. We concentrate on RA, AS, celiac disease, MS, SLE, T1D and IBD. Table 1 summarizes these AIDs in terms of their prevalence in the population and major phenotypic features. In particular, we focus on comparative studies that use GWA results to distinguish genetic variants that are specific to individual AIDs from those that are shared among multiple AIDs. We also summarize the results of a recently published cross-phenotype meta-analysis that uses genetic association results to highlight four main AID clusters. A detailed understanding of these shared and distinct genetic loci provides insight into fundamental etiologic mechanisms in autoimmune disease. It has the potential to inform the choice of current therapies and the development of novel targeted therapies and other interventions that could improve our ability to manage these complex human disorders.

Comparative studies of GWA data to identify shared and distinct genetic loci for AIDs

GWA and other recent genetic studies of AIDs have been remarkably successful in terms of the number of genetic loci that have been identified. For example, more than 70 genetic loci have now been firmly established as susceptibility factors for Crohn's disease [20], and more than 30 loci that contribute to the risk of RA and/or SLE have been identified [21,22]. Following the completion of many large GWAS of individual AIDs, a number of studies have sought to refine the specificity of loci that are associated with AIDs. More specifically, these recent studies have examined AID risk loci identified for one AID in case-control collections that have been developed for another AID to distinguish between those risk loci that are shared and those that are distinct for the AIDs being compared. In Table 2, the specific disease comparison studies that are discussed in this review are listed alongside both the major shared (and unique) loci that have been identified and the biologic pathways or mechanisms implicated by these analyses. Table 3 and Figure 1 further highlight the patterns of shared AID risk loci and their associated pathways revealed in these studies. In the next section, we summarize very recent work by Cotsapas et al. [23] that addresses these relationships in an analytically powerful way. Specifically, they utilized GWA data generated for seven AIDs and performed a cross-phenotype meta-analysis (CPMA) to distinguish between
Table 1. Prevalence and major phenotypic features of autoimmune diseases

| Autoimmune disease         | Frequency in the general population | Major phenotypic features                                                                 |
|-----------------------------|--------------------------------------|------------------------------------------------------------------------------------------|
| Celiac disease [39]         | 750 per 100,000 (United States)      | Break in immune tolerance to gluten, with malabsorptive diarrhea, and villous atrophy of intestinal mucosae |
| Type 1 diabetes [40]        | 50 to 200 per 100,000 (United States) | Hyperglycemia with macro- and microvascular complications                                  |
| Inflammatory bowel disease  | 200 per 100,000 (United States)      | Ulcerative colitis: superficial continuous ulceration of the large bowel. Crohn’s disease: transmural lesions that can occur throughout the gastrointestinal tract |
| Rheumatoid arthritis [41]  | 500 to 1,100 per 100,000 (United States and northern Europe) | Chronic inflammatory joint disease with potential joint destruction                       |
| Juvenile idiopathic arthritis [42] | 12 per 100,000 (United States) | Clinically heterogeneous group of disorders characterized by chronic inflammatory arthritis in children |
| Multiple sclerosis [43]     | 22 to 177 per 100,000 (United States) | Inflammatory disorder of the central nervous system with a wide range of neurological symptoms resulting from white matter lesions |
| Systemic lupus erythematosus [44] | 150 per 100,000 (United States) | Wide range of clinical manifestations with multi-organ involvement                        |
| Spondyloarthropathies [45]  | 1,300 per 100,000 (United States)    | Inflammatory rheumatic disorders characterized by axial and or peripheral arthritis, associated with enthesitis, dactylitis and potential extra-articular manifestations such as uveitis and skin rash |

 genetic variants that are common to all of these seven AIDs from variants that are common to some but not all of these AIDs and variants that are specific for one AID.

Type 1 diabetes

Several recent studies have compared the pattern of genetic association between T1D and other AIDs, including celiac disease [24], RA [25] and IBD [26]. On the basis of the association of both T1D and celiac disease with HLA class II loci, Smyth et al. [24] evaluated both the association of eight non-HLA celiac disease risk loci with T1D and of 18 T1D loci with risk of celiac disease in very large samples of patients, controls and families. Their analysis revealed that seven loci were common to these two AIDs, including RGS1 on chromosome 1q31, IL18RAP on chromosome 2q12, TAGAP on chromosome 6q25, a 32-bp insertion-deletion variant on chromosome 3p21, PTPN2 on chromosome 18p11, CTLA4 on chromosome 2q33, and SH2B3 on chromosome 12q24. Further, the associated alleles for IL18RAP and TAGAP confer risk of celiac disease but protection against T1D. Non-shared loci include the T1D risk loci INS (chromosome 11p15), IL2RA (chromosome 10p15), and PTPN22 (chromosome 1p13) and the celiac disease risk loci IL12A (chromosome 3q25) and LPP (chromosome 3q28).

Eyre et al. [25] extended this work by investigating genetic overlap between T1D and celiac disease risk loci and RA. They studied eight celiac disease risk loci and six T1D risk loci in a large sample of RA patients and control individuals. Although they found significant evidence for association of the TAGAP locus (which is associated with both celiac disease and T1D but with opposing effects) with RA and modest evidence of association between the CIQTNF6 T1D risk locus and RA, overall their investigation revealed little evidence of association between celiac disease and T1D risk loci and RA, suggesting that RA might be more genetically distinct.

Finally, Wang et al. [26] studied GWA data from large collections of IBD patients (Crohn’s disease and UC), T1D patients and control individuals of European ancestry to identify shared susceptibility loci. Although they identified a number of overlapping susceptibility loci among these diseases, their results were notable for the frequency with which risk alleles for one disease appear to provide protection against another. They interpret these data as indicating that many AID risk loci could be under balancing selection and that variants that have opposing effects on different AIDs might contribute to the maintenance of common susceptibility alleles in human populations.

Inflammatory bowel disease

Although the two types of IBD, Crohn’s disease and UC, differ in several important ways, such as the depth and location of inflammation in the gastrointestinal tract (Table 1), the clustering of these diseases within certain families and their overlapping risk loci support their etiologic relationship. Thus, these diseases have often been considered together in GWA and other genetic studies.

In addition to the aforementioned investigation of IBD and T1D [26], other recent work has investigated genetic overlap between IBD and other AIDs, including AS, celiac disease, psoriasis, SLE, RA and MS [20,27-29]. As mentioned previously, a large number of loci that
**Table 2. Summary of studies included in this review and the shared autoimmune disease risk loci that they identified**

| Autoimmune diseases compared | Study design | Shared risk loci (see Table 3 for other gene definitions) |
|------------------------------|-------------|---------------------------------------------------------|
| T1D and CeD [24]            | 8,064 T1D; 3,064 parent-child trios 2,560 CeD 9,339 controls | RGS1, IL18RAP, TAGAP, CCR3, PTNP2, CTLA4, SH2B3 |
| RA and SLE [30]             | 3,962 RA and 9,275 controls Screening of 11 SLE-associated SNPs Case control study and meta-analysis | BLK (B lymphoid tyrosine kinase), UBE2L3 (Ubiquitin-conjugating enzyme E2L 3) |
| RA and SLE [31]             | 1,635 RA and 1,906 controls Screening of nine SLE-associated SNPs Case control study and meta-analysis | No shared SNPs except the BLK locus in the subgroup of patients with sicca syndrome |
| RA and CeD [32]             | 1,368 RA, 795 CeD and 1,683 controls Screening of 11 RA and 11 CeD loci Case control study and meta-analysis | SH2B3, TNFAIP3, IL2/IL21, SH2B3, LPP, MMEL1/TNFSF14, PFKB3/PFKCQ |
| RA, T1D and CeD [25]        | 3,962 RA and 3,531 controls Screening of 8 CeD- and 6 T1D-associated SNPs Case control study | TAGAP (CeD- and RA- associated SNP) CeD: minor allele at-risk RA: minor allele protective |
| JIA and AID loci [33]       | Exploratory cohort: 809 JIA and 3,535 controls Replication cohort: 1,015 JIA and 1,560 controls Screening of 519 AID-associated SNPs Case-control study and meta-analysis | PTPN22, PTPN2, IL2/IL21, STAT4 (Signal transducer and activator of transcription 4), TNFAIP3, COG6 (Component of oligomeric Golgi complex 6), ANGST1 |
| AS and Crohn’s [27]         | Screening of 39 Crohn’s disease-associated SNPs 182 AS patients | SH2B3-associated locus in the subgroup of patients with sicca syndrome |
| AS and Crohn’s [28]         | Screening of 53 Crohn’s disease-associated SNPs 2,773 AS patients and 2,215 controls | Chr 1q32 (near KIF21B), STAT3, IL23R |
| Crohn’s and CeD [29]        | Meta-analysis of GWAS data from Crohn’s disease and CeD | ORMDL3 |
| IBD and T1D [26]            | Screening of 81 non-MHC loci associated with various AIDs from previous GWAS 1,689 Crohn’s disease, 777 UC, 989 T1D and 6,197 controls | PTPN2, IL18RAP, TAGAP, PUS10 (Pseudoubinylate synthase 1) |
| MS and AID loci [36]        | 2,864 MS and 2,930 controls Screening of 20 SNPs associated with various AIDs from previous GWAS Case control study and meta-analysis | SH2B3 (SH2B adaptropin protein), associated with RA, SLE and T1D, KIF5A (Kinesin family member 5A) and CD226 associated with RA and T1D |
| SLE and AID loci [34]       | 1,500 SLE and 5,706 controls Screening of 446 variants associated with 17 AIDs from previous GWAS Replication study with 2,085 SLE and 2,854 controls Case control study and meta-analysis | JIA-associated locus: VTCN1 Crohn’s disease-associated loci: PLC1, 1, VEGFA, ZGAP MS-associated loci: IL12A, RPL19P8, CD40 CeD-associated loci: IL12A, TRAF1 UC-associated loci: VEGFA, ZGAP RA-associated loci: TRAF1, CD40 T1D-associated locus: CCAR7 |

Contribute to risk of IBD have been identified through GWA and related studies; some of these loci are shared by Crohn’s disease and UC, whereas others are specific to one of these forms of IBD. This work has recently been reviewed by Lees et al. [20] and is briefly discussed here. Among the shared loci, components of the IL-23 pathway are of particular interest because of the development of therapies directed against IL12p40, the protein subunit shared by both IL-12 and IL-23. Other AIDs that are associated with variants of IL-23R (that is, psoriasis and AS) are therefore candidates for the assessment of treatments that target the IL-23 pathway. By contrast, loci that are specific for Crohn’s disease highlight the role of bacterial component clearance by infected cells through autophagy, mechanisms that might not be involved to the same extent in other AIDs. The spectrum of loci shared by IBD and other AIDs is more limited, involving innate immunity (IRF5), T-cell activation (PTPN22, IL2, IL2RA and IL21) or the activation of the unfolded protein response (ORMDL3). Interestingly, some variants affect the risk of different associated AIDs in opposite directions. For example, the PTPN22 allele, which is associated with increased risk of RA and SLE, reduces the risk of IBD.

**Rheumatoid arthritis and juvenile idiopathic arthritis**

Several recent studies have investigated genetic risk loci that are shared between RA or juvenile idiopathic arthritis (JIA) and other AIDs, including SLE, T1D, celiac disease, IBD and MS [25,30–33]. Studies investigating
Table 3. Shared autoimmune disease loci and biologic pathways that are implicated according to the NCBI database (entrez Gene summary) or UniProtKB/Swiss-Prot

| Shared loci                                                                 | Biologic pathways                                                                                   |
|----------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------|
| ANGPT1 (Angiopoietin 1)                                                    | Roles in vascular development and angiogenesis                                                      |
| BLK (B lymphoid tyrosine kinase)                                           | B-cell receptor signaling and B-cell development                                                    |
| CCNY (Cyclin Y)                                                            | Control of cell division cycles and regulation of cyclin-dependent kinases                        |
| CCR3 (Chemokine receptor 3)                                                | Binds to eotaxin, eotaxin-3, MCP-3, MCP-4, RANTES and MIP-1                                      |
| CCR7 (Chemokine receptor 7)                                                | Receptor for the MIP-3-beta chemokine                                                             |
| CD40                                                                       | Member of the TNF-receptor superfamily; receptor for CD40/L                                      |
| CD226                                                                      | Receptor involved in intercellular adhesion, lymphocyte signaling, cytotoxicity and lymphokine secretion mediated by cytotoxic T-lymphocyte (CTL) and NK cells |
| CDG6 (Component of oligomeric Golgi complex 6)                            | Required for normal Golgi function                                                                |
| CTLA4 (Cytotoxic T-lymphocyte-associated protein 4)                       | Negative regulator of T-cell responses                                                             |
| HEC2 (Hect domain and RLD 2)                                              | E3 ubiquitin-protein ligase                                                                      |
| FCGR2A (Fc fragment of IgG)                                                | Binds to the Fc region of immunoglobulins gamma                                                    |
| KCOSLG (Inducible T-cell co-stimulator ligand)                             | Co-stimulatory signal for T-cell proliferation and cytokine secretion. Ligand for the T-cell-specific cell surface receptor ICOS |
| KIF5A (Kinesin family member SA)                                          | Transcriptional regulator of hematopoietic cell differentiation                                     |
| IL10                                                                       | Inhibits the synthesis of a number of cytokines, including interferon-gamma, IL-2, IL-3, TNF and GM-CSF, produced by activated macrophages and by helper T-cells |
| IL12A                                                                      | Cytokine that can act as a growth factor for activated T and NK cells                              |
| IL18RAP                                                                    | NFkB and JNK activation (IL-18 dependent)                                                        |
| IL2/IL21                                                                   | Cytokines required for T-cell or B-cell proliferation                                            |
| IL23R                                                                      | Activates STAT1 and STAT3, MAPK1/3 (ERK1/2), JUN and AKT                                          |
| IL26                                                                       | Broad functions in adaptive immunity                                                              |
| IL77                                                                       | Plays a negative regulatory role in cells of the immune system                                    |
| IRF5 (Interferon regulatory factor 5)                                      | Might play a role in the innate immune response by regulating autophagy formation in response to intracellular pathogens |
| IRF8 (Interferon regulatory factor 8)                                      | Microtubule-dependent motor required for intracellular protein transport                          |
| IRGM (Immunity-related GTPase family, M)                                   | Role in cell shape and motility                                                                  |
| KIF5A (Kinesin family member SA)                                          | Metalloprotease involved in sperm function                                                        |
| LPP (LIM-domain-containing preferred translocation partner in IpaNA)       | Activates NFkB and stimulates T-cell proliferation                                                |
| MMEL1 (Membrane metallo-endopeptidase-like 1)                             | Might indirectly regulate endoplasmic reticulum-mediated Ca++ signaling                           |
| TNFSF14                                                                    | Involved in an inositol phospholipid-based intracellular signaling cascade                        |
| ORMOL (ORM1-like 3)                                                       | TCR-mediated T-cell activation                                                                   |
| PLCL1 (Phospholipase C-like 1)                                            | Lymphocyte cell signaling                                                                        |
| PTPN2 (Protein tyrosine phosphatase, non-receptor type 2)                 | Involved in the TCR signaling pathway                                                            |
| PTPN22                                                                     | Post-transcriptional nucleotide modification of RNAs                                             |
| PUS10 (Pseudouridylate synthase 10)                                       | Might be involved in the regulation of B cell activation and proliferation                        |
| RGS5 (Regulator of G-protein signaling 1)                                  | Pseudogene                                                                                        |
| RPL19P8 (Ribosomal protein L19 pseudogene 8)                              | T-cell receptor activation signaling                                                             |
| SH2B3 (SH2B adaptor protein 3)                                            | Th17 differentiation pathway                                                                     |
| STAT3 (Signal transducer and activator of transcription 3)               | T-cell differentiation (interferon-gamma expression)                                             |
| STAT4 (Signal transducer and activator of transcription 4)               | T-cell activation                                                                                 |
| TAGAP (T-cell activation RhoGTPase activating protein)                    | Negative regulator of the TNF- or LPS-mediated activation of NF-kappa-B                           |
| TNFAP3                                                                    | Mediates activation of NF-kappa-B                                                                |
| TNFSF15                                                                   | Adapter molecule that regulates the activation of NF-kappa-B and JNK                              |
| Traf1 (TNF receptor-associated factor 1)                                   | Ubiquitin of intracellular component                                                              |
| UBE2L3 (Ubiquitin-conjugating enzyme E2L 3)                                | Signaling protein involved in the regulation of angiogenesis and vasculogenesis                  |
| VEGFA (Vascular endothelial growth factor A)                              | Negatively regulates T-cell-mediated immune response by inhibiting T-cell activation, proliferation, cytokine production and development of cytotoxicity |
| VTCN1 (V-set domain containing T cell activation inhibitor 1)             | Negatively regulates expression of epidermal growth factor receptor (EGFR)                      |

GM-CSF, granulocyte-macrophage colony stimulating factor; LPS, lipopolysaccharide; NK, natural killer; TCR, T-cell receptor.
genetic overlap between RA and SLE have revealed both shared and distinct risk loci. For example, Orozco et al. [30] studied 11 SLE risk loci (TNFSF4, BANK1, TNIPI, PTG1, UHRF1BPI, ATG5, JAZF1, BLK, KIAA1542, ITGAM and UBE2L3) for evidence of association with RA. They found that RA cases were enriched, overall, for SLE genetic risk alleles. Variants within the BLK and UBE2L3 loci were most strongly associated with RA. Similarly, Suarez-Gestal et al. [31] studied nine variants at the following SLE risk loci for evidence of association with RA among cases and controls from Spain: ITGAM, C8orf13-BLK, TYK2, 1q25-1, PKX, KIAA1542, MECP2, BANK1, and LY9. Previously established shared risk loci, including HLA, PTPN22, STAT4 and 6q23, were not studied. None of the nine SLE risk variants studied was significantly associated with RA, suggesting that the genetic contribution to these two AIDs is relatively distinct, although it is also possible that this study was not powered sufficiently to identify shared risk loci.

Coenen et al. [32] investigated the extent of genetic overlap between RA and celiac disease. Specifically, they evaluated 11 RA and 11 celiac disease risk loci among Dutch RA patients, celiac disease patients and control individuals. Their analyses revealed six risk loci that were shared by RA and celiac disease, which included the TNFAIP3, IL2/IL21, SH2B3, LPP, MMEL1/TNFRSF14 and PFKFB3/PRKCOQ loci. Overall, the shared loci supported the importance of both adaptive and innate immunity in susceptibility to these two disorders.

Thompson et al. [33] used GWAS to investigate more than 400 genetic variants that had been associated previously with one or more AIDs to determine whether they are also associated with the risk of JIA. Using a two-stage design (discovery and replication cohorts), they found strong evidence of association for seven distinct loci, including PTPN22, PTPN2, ADA-1-IL2-IL21, STAT4, C12orf30, COG6 and ANGPT1. These loci have been previously associated with RA, T1D, Crohn's disease and/
or MS, supporting genetic overlap between JIA and a broad spectrum of AIDs. Given the phenotypic heterogeneity of JIA, which includes several distinct clinical and serologic subgroups, however, further work will be needed to clarify whether patterns of shared risk loci differ among specific JIA subsets.

**Systemic lupus erythematosus**

In addition to the studies mentioned above that have compared the overlap between SLE and RA risk loci [30,31], recent work by Ramos et al. [34] suggests only modest overlap between SLE and other AID risk loci. More specifically, these authors evaluated 446 genetic variants that had previously been associated with one or more of 17 AIDs to determine which loci were significantly associated with SLE susceptibility. A number of AID loci, including FCGR2A, IL10, IRGM, TNFAIP3, IKZF1, IRF5, BLK, IRF8, and UBE2L3, were associated with SLE and one or more other AIDs. However, many SLE loci, including ITGAM, TNFSF4, PTTG1, PHRF1, WDFy4 and BANK1, were associated with other AIDs only weakly, if at all.

**Multiple sclerosis (MS)**

Multiple sclerosis is characterized by very strong associations with HLA class II variants, but some of the other genes that are strongly associated with multiple AIDs, such as PTPN22, do not appear to contribute substantially to the risk of MS. Nonetheless, emerging evidence from GWAS supports overlap of MS-associated genes with genes that have been linked to a broad spectrum of AIDs [35]. For example, work by Alcina et al. [36] in which 12 genetic variants previously associated with other AIDs were studied in a large collection of Spanish MS cases and control individuals identified three shared susceptibility loci, including KIF5A, SH2B3 and CD226, that also influence risk of RA, T1D and SLE (SH2B3). More recently, a collaborative GWAS involving almost 10,000 MS cases recruited from 15 different countries has identified a large number of susceptibility loci, most of which map to regions containing immunologically relevant genes [37]. Particularly over-represented are loci implicated in T-helper-cell differentiation. Further, just over one-third of the MS risk loci identified overlap with regions previously identified in GWAS of other AIDs. Most of these shared risk loci have been associated with celiac disease, T1D, RA and/or IBD [37].

As mentioned previously, recent work by Sirota et al. [10] highlights the fact that certain variants that are associated with increased risk for some AIDs appear to be protective for others. More specifically, Sirota et al. [10] studied six AIDs (T1D, MS, AS, RA, Crohn’s disease and AITD) and found that AS and RA formed one group, and MS and AITD formed another group (with T1D showing similarity to both groups and Crohn’s disease to neither). Further, susceptibility variants that are associated with the first class of AIDs generally had a protective effect in relation to the second class of AIDs. As an example, TAP2, which is involved in the transport of antigens from the cytoplasm to the endoplasmic reticulum for association with MHC class I molecules, was found to be a susceptibility factor for AS, RA and T1D, but a protective factor for MS and AITD.

**Results of CPMA to identify shared and distinct genetic loci in AID**

Cotsapas et al. [23] have recently completed a CPMA that significantly extends our understanding of shared and distinct AID loci. More specifically, they studied 107 SNPs associated in recent GWAS with one or more of the following AIDs: celiac disease, Crohn’s disease, MS, psoriasis, RA, SLE and T1D. Their study indicates that almost half of these loci (47/107, 44%) are associated with multiple AIDs; many of these variants were not previously known to be shared across AIDs. Nine of these 47 variants had opposing effects in different AIDs. Cotsapas and colleagues also examined patterns of disease association for the 47 shared loci, and found that just one locus, a variant in an exon of SH2B3 (rs3185404) was significantly associated with all seven of the AIDs examined. The remaining 46 variants were associated with subsets of the seven AIDs.

The authors extended their analysis of these variants to try to elucidate the molecular pathways underlying these subgroups of AID. Four clusters were revealed on the basis of the patterns of AID associations. The first cluster, represented by variants in IL23R, IL12B, PTGER4, JAK2, KIF21B, STAT3 and other genes, was most strongly associated with Crohn’s disease, psoriasis and MS. A second cluster, represented by variants in STAT4, IRF5, TNFAIP3, RGS1, CCR1, IL18RAP, IL2-IL21 and UBE2L3, was most strongly associated with celiac disease, RA and SLE. A third cluster, represented by variants in ORMDL3, CLEC16A, IL2RA, PRKCQ, CYP27B1, IKZF1 and ETS1, was most strongly associated with T1D, MS and RA. A fourth cluster, represented by variants in SH2B3, PTPN2, PTPN22, PRKCQ, CTLA4, UBAH3A, IL10, IFI1H1, IL2, BACH2, IL27, CD226 and other genes, was most strongly associated with T1D, RA, celiac disease, Crohn’s disease and SLE. Further, an analysis of protein-protein interactions revealed that the proteins encoded by variants within groups were more likely to interact with each other (either directly or via intermediates) than with proteins encoded by variants in other groups, underscoring the biologic relevance of the AID relationships defined by this CPMA.

The delineation of genes and pathways that relate more specifically to certain AIDs than to others provides
valuable information that can be used to target autoimmune phenotypes with interventions that are relevant to those pathways. The highlighted biologic pathways then provide a focus for more fundamental research, aimed at elucidating the underlying disease mechanisms in autoimmunity, and they could inform the development of novel therapies. The success of anti-TNF targeted therapies for a diverse group of autoimmune disorders, including RA, IBD, psoriasis, AS and others [38], nicely illustrates the potential value of this information. Similarly, the aforementioned collaborative GWAS of MS [37] highlights loci that are related to MS therapies, including VCAM1 (natalizumab) and IL2RA (daclizumab).

As recently summarized in a review by Rai and Wakeland [16], despite the dramatic increase both in the number of risk loci recently identified for human AIDs and in information about patterns of shared risk and biologic pathways, the current literature does not provide a complete mechanistic understanding of biologic pathways that explain the pattern of AID susceptibility in human populations. Additional work will be required to refine genotype-phenotype relationships in autoimmune disease more completely. This research should include larger case-control studies in diverse population groups and the application of new technologies, such as next-generating sequencing, to define all of the relevant genetic variation. Given the ‘missing heritability’ of human AIDs, and the fact that current GWAS have captured primarily common genetic SNP variants, it is likely that rare or structural variants explain much of the missing heritability, the identification of which will require new and emerging technologies. Finally, once the complete genetic architecture underlying human AIDs has been characterized, additional methods will be required to define the functional mechanisms that explain these genetic associations.

Summary and conclusions

Owing to the rapid pace of identification of AID-associated genes during the past 5 years, primarily as a result of GWAS, there is now a wealth of information available that allows for a more thorough delineation of the extent of genetic overlap across this broad group of disorders. Loci that are shared between various AIDs and involved in a wide range of immune pathways (for example, T-cell activation, B-cell activation, cytokine signaling) might help explain common pathogenic features and inform the development of novel therapies. Further, the lack of overlap for other loci and pathways (for example, IL23R and STAT3 in IBD or spondyloarthritits) also suggests distinct pathogenic mechanisms that could explain, at least in part, the phenotypic diversity across the spectrum of autoimmune disease. It is important to keep in mind, however, that current studies are likely insufficiently powered to characterize fully the genetic architecture of AIDs, including shared and distinct loci and biologic pathways. Thus, the ongoing generation and analysis of data emerging from GWAS and other genetic studies is warranted in order to better define genotype-phenotype associations in human AIDs and to clarify which pathways and specific targets are most relevant to the diseases within this diverse group of human disorders.

Abbreviations

AID, autoimmune disorder; AITD, autoimmune thyroid disease; AS, ankylosing spondylitis; CPMA, cross-phenotype meta-analysis; CTLA4, cytotoxic T-lymphocyte antigen 4; GWA, genome-wide association; GWAS, GWA studies; HLA, human leukocyte antigen; IBD, inflammatory bowel disease; JIA, juvenile idiopathic arthritis; PTPN22, protein tyrosine phosphatase non-receptor type 22; MS, multiple sclerosis; RA, rheumatoid arthritis; SLE, systemic lupus erythematosus; T1D, type 1 diabetes; TNF, tumor necrosis factor; TNFAIP3, TNF alpha-induced protein 3; ULC, ulcerative colitis.

Competing interests

The authors declare that they have no competing interests.

Author details

1. Université Paris-Sud 11, Hôpital Bicêtre, Assistance Publique-Hôpitaux de Paris (AP-HP), Le Kremlin Bicêtre, France. 2. University of California San Francisco, Rosalind Russell Medical Research Center for Arthritis, Department of Medicine, Parnassus Avenue, San Francisco, CA 94143, USA.

Published: 27 January 2012

References

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, 33:197-207.
2. Rahman A, Isenberg DA. Systemic lupus erythematosus. N Engl J Med 2008, 358:929-939.
3. Somers EC, Thomas SL, Smeeth L, Hall AI. Autoimmune diseases co-occurring within individuals and within families: a systematic review. Epidemiology 2006, 17:202-217.
4. Michou L, Rat AC, Lasbleiz S, Bardin T, Cornells F. Prevalence and distribution of autoimmune diseases in 368 rheumatoid arthritis families. J Rheumatol 2008, 35:790-796.
5. Barcellos LF, Kamdar BB, Ramsay PP, DeLoa C, Lincoln RR, Callies S, Schmidt S, Haines JL, Pericak-Vance MA, Okensgen JR, Hauser SL. Clustering of autoimmune diseases in families with a high-risk for multiple sclerosis: a descriptive study. Lancet Neurol 2006, 5:924-931.
6. Alarcon-Segovia D, Alarcon-Riquelme ME, Cardiel MH, Caireo F, Massardo L, Villa AR, Pons-Estel BA. Familial aggregation of systemic lupus erythematosus, rheumatoid arthritis, and other autoimmune diseases in 1,177 lupus patients from the GLADEL cohort. Arthritis Rheum 2010, 52:1138-1147.
7. Anaya JM, Castiblanco J, Tobon GJ, Garcia J, Abad V, Cuervo H, Velasquez A, Angel ID, Vega P, Arango A. Familial clustering of autoimmune diseases in patients with type 1 diabetes mellitus. J Autoimmun 2006, 26:208-214.
8. Anaya JM, Tobon GJ, Vega P, Castiblanco J. Autoimmune disease aggregation in families with primary Sjogren’s syndrome. J Rheumatol 2006, 33:2227-2234.
9. Prahallad S, Shear ES, Thompson SD, Giannini EH, Glass DN. Increased prevalence of familial autoimmunity in simplex and multiplex families with juvenile rheumatoid arthritis. Arthritis Rheum 2002, 46:1851-1856.
10. Sireota M, Schaub MA, Batsoglou S, Robinson WH, Butte AJ. Autoimmune disease classification by inverse association with SNP alleles. PLoS Genet 2009, 5:e1000792.
11. Fernando MM, Stevens CR, Walsh EC, De Jager PL, Goyette P, Penenge RM, Vyse TJ, Rioux JD. Defining the role of the MHC in autoimmunity: a review and pooled analysis. PLoS Genet 2008, 4:e1000024.
12. Kristiansen OP, Larsen ZM, Pociot F. CTLA-4 in autoimmune diseases - a general susceptibility gene to autoimmunity? Genes Immun 2000, 1:170-184.
13. Chung SA, Criswell LA: PTPN22: its role in SLE and autoimmunity. Autoimmunity 2007, 40:582-590.

14. Vereecke L, Beyaert R, van Loo G: The ubiquitin-editing enzyme A20 (TNFAIP3) is a central regulator of immunopathology. Trends Immunol 2009, 30:383-391.

15. Geergeren PK, Olsson LM: Recent advances in the genetics of autoimmune disease. Annu Rev Immunol 2009, 27:363-391.

16. Rai E, Wakeland EK: Genetic predisposition to autoimmunity - what have we learned? Semin Immunol 2011, 23:67-83.

17. Homer CR, Richmond AL, Rebert NA, Achkar JP, McDonald C: ATG16L1 and NOD2 interact in an autophagy-dependent antibacterial pathway implicated in Crohn's disease pathogenesis. Gastroenterology 2010, 139:1630-1641, 1641.e1-2.

18. Zhernakova A, van Diemen CC, Wijmenga C: Detecting shared pathogenicity from the shared genetics of immune-related diseases. Nat Rev Genet 2009, 10:443-55.

19. Lettre G, Rioux JD: Autoimmune diseases: insights from genome-wide association studies. Hum Mol Genet 2008, 17(R1):16-121.

20. Lees CW, Barrett JC, Parkes M, Satsangi J: New IBD genetics: common pathways with other diseases. Gut 2011, 60:1739-1753.

21. Raychaudhuri S: Recent advances in the genetics of rheumatoid arthritis. Rev Exp Clin Immunol 2010, 4.461-479.

22. Cottapas C, Voght BF, Rossin E, Lage K, Neale BM, Wallace C, Abecasis GR, Barrett JC, Behrens T, Cho J, De Jager PL, Elder JT, Graham RR, Geogresen P, Klafkis L, Simovitch KA, van Heel DA, Wijmenga C, Worthington J, Todd JA, Haffner DA, Rich SS, Daly MJ, FOCIS Network of Consortia: Pervasive sharing of genetic effects in autoimmune disease. PLoS Genet 2011, 7:e1002155.

23. Smyth DJ, Plagnol V, Walker NM, Cooper JD, Downes K, Yang JH, Howson JM, Stevens H, McManus R, Wijmenga C, Heap GA, Dubois PC, Clayton DG, Hunt KA, van Heel DA, Todd JA: Shared and distinct genetic variants in type 1 diabetes and celiac disease. N Engl J Med 2008, 359:2767-2777.

24. Eyre S, Hincks A, Bowes J, Flynn E, Martin P, Wilson AG, Emery P, Steer S, Hocking LJ, Reid DM, Harrison P, Wordsworth PB; Yorkshire Early Arthritis Consortium; Biologics in RA Control Consortium, Thomson W, Worthington J, Barton A: Overlapping genetic susceptibility variants between three autoimmune disorders: rheumatoid arthritis, type 1 diabetes and coeliac disease. Arthritis Res Ther 2010, 12:R175.

25. Wang K, Baldassano R, Zhang H, Qu H, Imelinskas M, Kugathasan S, Anness V, Dubinsky M, Rotter JI, Russell RK, Bradfield JP, Siegelman PM, Gie arens J, Walters R, Houlston R, McEvoy L, Cattaneo A, van Heel DA, Wijmenga C, Worthington J, Todd JA, Haffner DA, Rich SS, Daly MJ, FOCIS Network of Consortia: Autoimmune diseases: insights from genome-wide association studies. Hum Mol Genet 2008, 17(R1):16-121.