Available online at www.sciencedirect.com
ScienceDirect
Current Opinion in Immunology

The genetics of autoimmune diseases: a networked perspective

Sergio E Baranzini

Modern tools for genetic analysis are producing a large impact on our understanding of autoimmunity. More than 30 genome-wide association studies (GWAS) have been published to date in several autoimmune diseases (AID) and hundreds of common variants have been identified that confer risk or protection. While statistical adjustments are essential to refine the list of potential associations with each disease, valuable information can be extracted by the systematic collection of moderately significant variants present in more than one trait. In this article, a compilation of all GWAS published to date in seven common AID is provided and a network-based analysis of shared susceptibility genes at different levels of significance is presented. While involvement of the MHC region in chromosome 6p21 is not in question for most AID, the complex genetic architecture of this locus poses a significant analytical challenge. On the other hand, by considering the contribution of non-MHC-related genes, similarities and differences among AID can be readily computed thus gaining insights into possible pathogenic mechanisms. Statistically significant excess sharing of non-MHC genes was found between type I diabetes (T1D) and all other AID studied, a result also seen for RA. A smaller but significant degree of sharing was observed for multiple sclerosis (MS), Celiac disease (CeD) and Crohn's disease (CD). The availability of GWAS data allows for a systematic analysis of similarities and differences among several AID. Using this class of approaches the unique genetic landscape for each autoimmune disease can start to be defined.

Address
Department of Neurology, School of Medicine, University of California San Francisco, 513 Parnassus Ave. Room S-256, San Francisco, CA 94143-0435, United States

Corresponding author: Baranzini, Sergio E (sebaran@cgl.ucsf.edu)

Current Opinion in Immunology 2009, **21**:1–10

This review comes from a themed issue on
Autoimmunity
Edited by Jeffrey Bluestone and Vijay Kuchroo

0952-7915/$ – see front matter
Published by Elsevier Ltd.

DOI 10.1016/j.coi.2009.09.014

Autoimmune disorders arise when physiological tolerance to ''self'' antigens is lost. Although several mechanisms may be involved in this pathogenic process, dysregulation of T-cell and B-cell activation and of pathways leading to inflammation are logical candidates. Susceptibility to autoimmune diseases has been associated with multiple factors including genetics, epigenetics, and the environment. While the modest concordance rate in monozygotic twins suggests that environmental factors are major players in most autoimmune diseases, increased heritability within families and the decrease in risk with the degree of relatedness all argue in favor of genetic factors. With the advent of high-throughput genomics, massive amounts of genetic data are being produced and reported on a monthly basis. Although considerable insight has been gained from each of these individual studies, a detailed comparative analysis will likely identify both unique and common pathways operating in autoimmunity. This kind of analysis may set the basis for more targeted and rational therapeutic approaches.

Certain autoimmune disorders co-occur significantly in a single individual or within nuclear families more often than expected suggesting the presence of genetic variants that predispose to or protect against autoimmunity [1–4]. In a recent analysis Rzhetzky et al. reviewed 1.5 million medical records involving 161 diseases and computed pairwise correlations of disease co-occurrences [5•]. Indeed, several autoimmune disorders co-occurred in the same individuals more often than expected by chance. T1D most often correlated not only with the presence of type 2 diabetes mellitus (T2D), but also with RA, and psoriasis. Similarly MS correlated with systemic lupus erythematosus (SLE), T1D, T2D and psoriasis, while RA strongly correlated with SLE, ankylosing spondilitis, T1D, T2D, Sjogren's, and Psoriasis. Although these data were derived from medical records and not from genetic analysis, the results suggest that common genetic mechanisms may be at play in different AID.

Genetic polymorphisms are heritable sequence alterations in the genome that contribute to phenotypic variability, and can modulate the expression and/or function of genes thus affecting the behavior of biological pathways, potentially determining susceptibility to diseases. With the advent of genomic tools that made available miniaturization and automation of genotyping platforms, more than 200 genome-wide association (GWA) studies have been performed in different diseases to date [6••,7] including 31 studies in 7 common AID. In this review I will summarize the findings of these studies, elaborate hypotheses about the possible pathogenic mechanisms implicated in each disorder, and provide a global view of shared and specific genes that characterize them.

The aim of GWA studies is to characterize the genetic architecture of complex genetic traits through the identification of disease variants against the background of

random variation seen in a population as a whole. In a typical study, hundreds of thousands of markers covering a significant portion of the common variation in the population are tested simultaneously in cases and controls and the allelic frequencies of each marker are compared between the two groups. The large number of common genetic variation in the human population means that prior odds that a randomly chosen marker is relevant for a given disease are extremely low (estimated at 10^{-5} for MS [8]). For this reason, and in the absence of other prior knowledge, only highly significant markers should be taken into consideration.

A survey of all studies reported in the GWAS Catalog [7] (as of July 2009) in 7 common AID (CeD, CD, MS, Ps, RA, SLE, and T1D) and T2D revealed that variants in 45 genes (7 MHC-related) are associated with disease susceptibility (Table 1). Only markers with highly significant associations ($P < 10^{-10}$) or identified at $P < 10^{-7}$ in two or more studies are shown in this list. However, several truly associated variants may never reach this significance due to the limited power of most studies. Thus, exploratory analyses using lower significance thresholds may uncover important associations, particularly, if they occur in candidate genes or pathways. Although by simply tabulating data, genes associated with more than one disease can be easily identified (e.g. *PTPN22* with RA and T1D, *IL2RA* with MS and T1D, *IL12B* in Ps and CD), this task becomes more difficult at a lower significance cut-off as the number of genes increases considerably. Unfortunately, the GWAS Catalog only lists associations at 10^{-6} or lower, thus preventing any analysis using a more liberal significance threshold.

Table 1

Top associated genes in seven common autoimmune diseases

Disease	Gene	Reference
CeD	IL21	[40,41]
	RGS1	[40]
	HLA-DQA1	[41]
MS	HLA-DRB1	[16•,42–44]
	METTL1, CYP27B1	[42]
	CD58	[16•,42]
	HLA-B	[16•]
	TNFRSF1A	[16•]
	IL2RA	[16•,44]
Psoriasis	HLA-C	[45–47]
	IL12B	[45,48]
	TNIP1	[45]
	IL13	[45]
	TNFAIP3	[45]
	LCE3D, LCE3A	[48]
Crohn's	IL23R	[35,38,49–51]
	ATG16L1	[35,38,50]
	PTGER4	[38]
	NOD2	[35,38,49]
	ZNF365	[38]
	PTPN2	[35,38,52]
	NKX2-3	[38,52]
	IRGM	[38,52]
	IL12B	[38]
	MST1	[38,52]
	CCR6	[38]
	STAT3	[38]
	LRRK2, MUC19	[38]
	TNFSF15	[38]
	CDKAL1	[38]
	BSN, MST1	[35]
	CARD15	[50]
RA	PTPN22	[32–35]
	REL	[32]
	OLIG3, TNFIP3	[33,53]
	HLA-DRB1	[33–35]
	HLA-DQA1, HLA-DQA2	[35,54]
	TRAF1-C5	[34]
SLE	TNFAIP3	[55]
	STAT4	[37,55]
	HLA-DQA1	
	IRF5, TNPO3	[37]
	ITGAM, ITGAX	[37]
	C8orf13, BLK	[37]
	BANK1	[56]
T1D	MHC	[35,38,57,58]
	PTPN22	[35,38,57–59]
	INS	[38,58,59]
	C10orf59	[38]
	SH2B3	[35,38]
	ERBB3	[35,38,57,59,60]
	CLEC16A	[38,57]
	CTLA4	[38,57]
	PTPN2	[38,57,59]
	IL2RA	[38]
	IL27	[38]
	C6orf173	[38]
	IL2	[38]
	ORMDL3	[38]
	GLIS3	[38]
	CD69	[38]
	UBASH3A	[38,61]
	IFIH1	[38,59]
	BACH2	[38,57]
	CTSH	[38,57]
	PRKCQ	[38,57]
	C1QTNF6	[38]
	C12orf30	[57]
	C1QTNF6	[57]
	KIAA0350	[35,58,59]
	C12orf30	[59]

A straightforward solution to the problem of low powered datasets is to conduct larger studies. This is the rationale behind the second phase of the Wellcome Trust Case Control Consortium (WTCCC2), a massive collaborative project that is genotyping 120 000 samples in 13 diseases (including ankylosing spondylitis, MS and Ps) and two quantitative phenotypes [9]. Another strategy to increase the prior odds of finding a true significant marker is to incorporate prior biological knowledge. A reasonable approach to accomplish this is through the integrated

Please cite this article in press as: Baranzini SE. The genetics of autoimmune diseases: a networked perspective, Curr Opin Immunol (2009), doi:10.1016/j.coi.2009.09.014

analysis of susceptibility alleles into biological pathways [10–12]. This, however, requires the inclusion of variants with only nominal evidence of genetic association that are typically filtered out in most studies in order to minimize type I error (false positives). A recent article by our group used this approach to identify novel susceptibility pathways in MS, and revealed genetic overlaps between MS and Alzheimer's disease and between MS and bipolar disorder. In addition, the presence of common variants in the MHC region between MS, RA and T1D but not T2D or CD with MHC alleles was highlighted [13••].

Analytic and computational approaches that integrate results from multiple GWA datasets represents an alternative strategy that may strengthen previous conclusions, suggest novel loci or pathways, and refine the localization of association signals [14,15]. For example a recent meta-analysis of three MS studies identified *CD6*, *TNFRSF1A*, and *IRF8*, three non-MHC-related genes not found in any of the previous GWAS in this disease [16•]. In a larger study, Johnson and O'Donnell collected and catalogued results from 118 GWAS published through March 1, 2008, all of which tested trait associations with >50 000 markers [17•]. This study listed all the *P*-values as provided by the authors in the original publications, in some cases as relaxed as $P < 0.05$. Although several autoimmune diseases were included in that set comparing genes across traits was out of the scope of that work. We then extracted all moderately significant ($P < 10^{-4}$) associations from each study (plus those in T2D) to analyze and compare the genetic contribution to these autoimmune disorders (Table 2). When multiple studies for the same disorder reported on the same gene, *P*-values were combined using the Fisher's method [18]. Altogether, 1201 genes with modest evidence for association in at least one of these autoimmune disorders were identified.

Recently, Goh *et al.* integrated all available genetic data from the Online Mendelian Inheritance in Man (OMIM) database using a bipartite network-based visualization approach [19•]. Since OMIM focuses primarily on Mendelian disorders, genetic data on complex disorders were derived from literature mining and thus, less accurately represented in this analysis. Nevertheless, this strategy identified groups of diseases that shared susceptibility genes and grouped them together, thus creating a disease landscape based on genetic similarity. To address whether genes involved in one AID also confer susceptibility to another we carried out a similar approach to that used by Goh using evidence from GWAS (Figure 1).

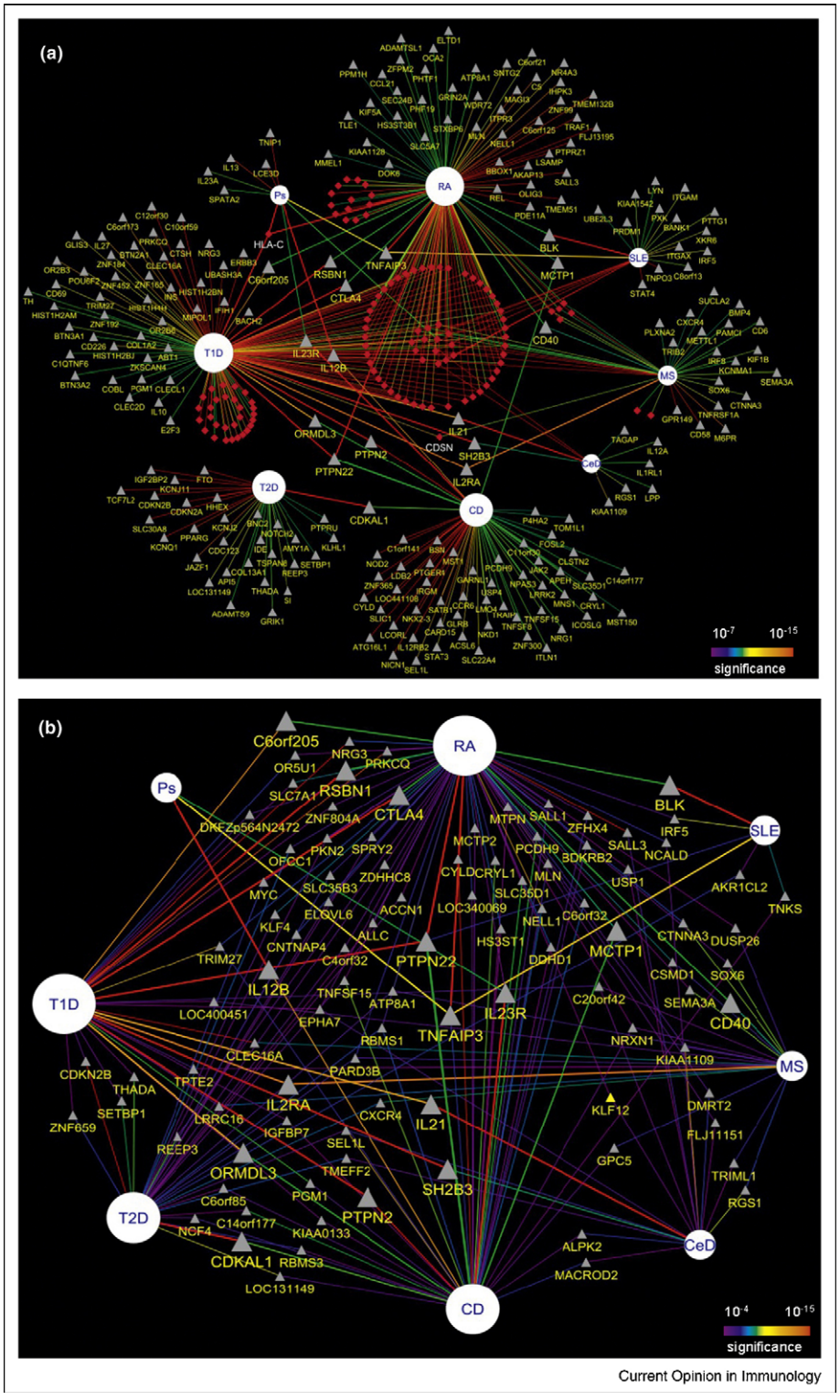
Table 2

GWAS Studies used for network analysis

Phenotype	Cases	Controls	Analyzed SNPs	#SNPs reported (criteria)	Reference
CeD	778	1422	310 605	50 ($P < 10^{-2}$)	[41]
CD	382 (trios)	–	164 279	62 ($P < 10^{-3}$)	[49]
CD	393	399	92 387	139 ($P < 10^{-2}$)	[62]
CD	547	548	308 332	6 (top)	[63]
CD	547	928	302 451	1 (top)	[51]
CD	946	977	304 413	23 ($P < 10^{-4}$)	[50]
CD	94	752	72 738	4 ($P < 10^{-3}$)	[64]
CD	2000	3000	469 557	502 ($P < 10^{-3}$)	[35]
MS	931	2431	334 923	114 ($P < 10^{-3}$)	[44]
MS	978	883	551 642	44 ($P < 10^{-3}$)	[65]
Ps	318	288	313 830	3 ($P < 10^{-4}$)	[66]
RA	1522	1850	297 086	193 ($P < 10^{-2}$)	[34]
RA	625	558	203 269	14 ($P < 10^{-2}$)	[67]
RA (CCP+)	397	1211	79 853	205 ($P < 10^{-3}$)	[53]
RA	2000	3000	469 557	380 ($P < 10^{-3}$)	[35]
SLE	94	538	52 608	1 (top)	[68]
SLE	51	54	262 264	5 (top)	[69]
SLE	720	2337	265 648	35 ($P < 10^{-2}$)	[36]
T1D	1028	1143	534 071	88 (top)	[58]
T1D	2000	3000	469 557	102 ($P < 10^{-2}$)	[35]
T2D	1464	1467	386 371	102 ($P < 10^{-2}$)	[70]
T2D	105	102	115 352	72 ($P < 10^{-2}$)	[71]
T2D	640	674	80 044	89 (top)	[72]
T2D	124	295	82 485	125 (top)	[73]
T2D	500	497	315 917	7 (top)	[74]
T2D	1161	1174	315 635	97 ($P < 10^{-2}$)	[75]
T2D	661	614	392 935	50 ($P < 10^{-2}$)	[76]
T2D	1399	5275	313 179	48 (top)	[77]
T2D	3757	5346	393 453	65 ($P < 10^{-3}$)	[78]
T2D	307 (trios)	–	66 543	45 ($P < 10^{-3}$)	[79]
T2D	91	1083	70 987	5 (top)	[80]

Please cite this article in press as: Baranzini SE. The genetics of autoimmune diseases: a networked perspective, Curr Opin Immunol (2009), doi:10.1016/j.coi.2009.09.014

Figure 1



Disease-gene network **(A)** Top genetic associations in 7 autoimmune diseases and T2D. The most significant SNP per gene was selected. Only associations with the significance of at least $P < 10^{-7}$ are visualized. If a given gene was identified in more than one disease, multiple lines connecting it with each disease were drawn. Lines are colored using a ''heat'' scheme according to the evidence for association. Thus ''hot'' edges (e.g. red, orange) represent more significant associations than ''cold'' edges (e.g. purple, blue). Diseases are depicted by circles of size proportional to the number of associated genes, non-MHC genes by grey triangles, and genes in the MHC region are shown as red diamonds. **(B)** Similar to (A), but with the threshold of significance lowered to $P < 10^{-4}$. To aid visualization, only genes shared by at least two diseases are shown.

Please cite this article in press as: Baranzini SE. The genetics of autoimmune diseases: a networked perspective, Curr Opin Immunol (2009), doi:10.1016/j.coi.2009.09.014

Table 3

Genes shared by at least three diseases at (aggregate) $P < 10^{-4}$

Gene	Chr	Description	Crohn's	RA	T1D	Celiac	MS	SLE	Ps
PTPN22	1	Protein tyrosine phosphatase, non-receptor type	10^{-8}	10^{-90}	10^{-226}			10^{-5}	
IL23R	1	Interleukin 23 receptor	10^{-102}	10^{-4}					10^{-7}
NRXN1	2	Neurexin 1 isoform beta precursor	10^{-4}	10^{-4}			10^{-4}		
KIAA1109	4	Hypothetical protein LOC84162		10^{-5}		10^{-12}	10^{-5}		
EPHA7	6	Ephrin receptor EphA7		10^{-4}	10^{-4}	10^{-5}			
TRIM27	6	Tripartite motif-containing 27	10^{-5}	10^{-6}	10^{-11}				
TNFAIP3	6	Tumor necrosis factor, alpha-induced protein 3	10^{-5}	10^{-20}				10^{-11}	10^{-11}
TNKS	8	Tankyrase	10^{-4}				10^{-4}	10^{-6}	
C20orf42	20	Fermitin family homolog 1			10^{-5}	10^{-4}	10^{-4}		

When the network is visualized with only those genes exceeding the genome-wide significance level of $P < 10^{-7}$, a large connected core of genes and diseases can be observed (Figure 1A). In this visualization a strong cluster of MHC-associated genes is readily identified, not only for RA and T1D, but also for MS and SLE. This is shown by the prominent circles of red diamonds in the center of the figure.

All of the MHC-related genes associated with SLE and CeD were also shared by either RA, T1D, or MS. However, this observation may be a consequence of the strong linkage disequilibrium operating in that region of the genome. The only two diseases showing no MHC associations were Crohn's and T2D. This finding is likely a consequence of the absence of GWAS signals in chromosome 6 reported in several studies in CD and the fact that T2D is primarily a metabolic disease, included in this analysis for comparison [20]. Despite this observation, CD is still connected to the main core by sharing *MCTP1* with RA, *IL23* and *IL12B* with Ps, *ORMDL3*, *PTPN2* and *PTPN22* with T1D, and *CDKAL1* with T2D.

While the connection among diseases through MHC-associated genes is illuminating, the extensive LD in this region obscures the identification of additional shared genes. If the MHC locus is ignored, 16 genes are still associated with more than one disease at this significance level (displayed towards the center of Figure 1A). This select list of potentially ''general'' autoimmunity genes includes *PTPN22*, a tyrosine phosphatase strongly associated with T1D (aggregate $P < 10^{-226}$), RA (aggregate $P < 10^{-90}$), and to a lesser extent with CD (aggregate $P < 10^{-8}$). The risk allele of this non-synonymous SNP (R620W), disrupts the P1 proline-rich motif that is important for interaction with cytoplasmic tyrosine kinase (CSK), potentially altering these protein's normal function as a negative regulator of T cell activation. *PTPN22* has also been associated with other autoimmune diseases including Addison's disease [21] and Graves' thyroiditis [22]. *TNFAIP3* is also highly associated with RA (aggregate $P < 10^{-20}$), SLE (aggregate $P < 10^{-11}$), Ps (aggregate $P < 10^{-11}$), and moderately associated with CD (aggregate $P < 10^{-5}$). This TNFa-induced gene is essential for limiting inflammation by terminating NF-kappa B responses.

If a more relaxed threshold ($P < 10^{-4}$) is used to visualize the reported associations, 71 non-MHC genes are identified as shared by at least two diseases (Figure 1B), seven by three diseases, and only 2 by four diseases (Table 3). In addition to *PTPN22* and *TNFAIP3* described above, *IL23R* and the *KIAA1109* locus appear as additional general autoimmunity genes. IL23R is also a key component of the T cell activation pathway and plays a critical role in differentiation, expansion and stabilization of proinflammatory TH17 cells [23]. KIAA1109 is located within a region of high linkage disequilibrium in chromosome 4 that also encompasses the genes for ADAD1, IL2, and IL21. The role of IL2 in T and B cell proliferation and its potential implications in autoimmunity have been extensively documented [24,25]. Meanwhile, IL-21 acts as a co-stimulator of proliferation, enhances memory response, and modulates homeostasis. Within the innate immune system IL-21 has a role in the terminal differentiation of NK cells, enhancing cytotoxic function while also decreasing cellular viability. These immune maturation and stimulating functions have resulted in IL-21 being tested in a variety of models of immunity [26–28]. Finally, another gene with seemingly general autoimmune properties is *CTLA4*, although in this analysis it only reached genome-wide significance (aggregate $P < 10^{-7}$) in T1D and RA. When engaged by its receptors CD80 or CD86, CTLA4 initiate signals resulting in the inhibition of T cell activation. Several additional reports relating *CTLA4* with multiple autoimmune diseases exist [29], but most of these followed a candidate gene approach, and thus were not included here. Altogether, the data presented here suggest that genes involved in activation, proliferation, and homeostasis of cells involved in adaptive immune responses are more likely to represent general autoimmunity genes. This is further supported by the observation that a large proportion of these genes physically interact among each other (S. Baranzini, unpublished observation), thus possibly taking part in the same or highly overlapping biological

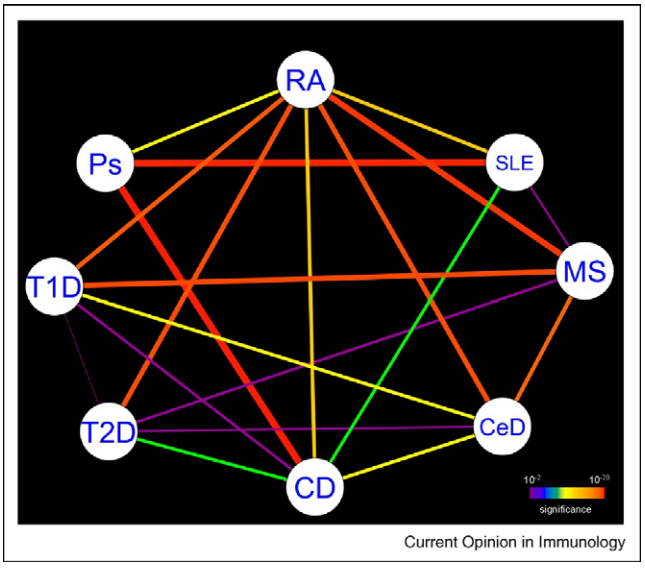
Please cite this article in press as: Baranzini SE. The genetics of autoimmune diseases: a networked perspective, Curr Opin Immunol (2009), doi:10.1016/j.coi.2009.09.014

pathways. A corollary to this observation indicates that associations unique to each disease would be responsible for attracting immune responses to specific tissues, although more functional studies in animal models will be needed to confirm this.

While at $P < 10^{-4}$ several of these discoveries may represent false positives, potential associations with autoimmunity genes may be discovered if genes are shared by two or more diseases at a higher proportion than what it would be expected by chance alone. To this end we counted all shared association between all pairs of diseases and assessed enrichment by computing a chi-squared test (Table 4). We found that T1D shared significantly more genes than expected with all other diseases studied. For example, it shared 12 genes with MS (1.49 expected, $P < 10^{-18}$), 12 with T1D (2.13 expected, $P < 10^{-12}$), and 18 with T2D (3.6 expected, $P < 10^{-14}$). On the other hand, genes shared between T2D and most other diseases (except RA), barely exceeded the expected number. The network in Figure 2 summarizes the relationships among these autoimmune diseases, and Supplementary Table 1 lists each gene with their corresponding *P*-value. Of note, none of these multiple associations is inflated by LD, as most of these genes are encoded in different chromosomes.

The finding that susceptibility alleles for one disease can be protective in another represents an additional level of complexity in the genetics of AID. For example, the common allele (T) of SNP rs11594656 in *IL2RA* (CD25) confers susceptibility to T1D, but protection

Figure 2



Disease similarity network Diseases are connected by nodes of a color that is proportional to the *P*-value of the chi-squared test for excess shared susceptibility genes.

Table 4

Number of non-MHC genes shared across diseases

	CD			MS			Ps			RA			SLE			T1D			T2D		
	O	E	*P*	O	E	*P*	O	E	*P*	O	E	*P*	O	E	*P*	O	E	*P*	O	E	*P*
CeD	4	0.48	4.0E−07	3	0.2	8.0E−10	0	0.01	0.89	7	0.7	1.7E−14	0	0.06	0.8	3	0.3	6.7E−7	2	0.5	0.03
CD				3	1.05	0.06	3	0.09	1.2E−22	4	0.44	5.2E−8	3	0.3	1E−6	6	1.5	2.3E−4	10	2.54	2.3E−6
MS							0	0.04	0.84	12	1.49	3.4E−18	1	0.13	0.02	7	0.64	1.75E−15	5	1.08	1.5E−4
Ps										2	0.13	1.1E−7	1	0.01	7.4E−21	0	0.05	0.81	0	0.09	0.76
RA													4	0.43	5.2E−8	12	2.13	9E−12	18	3.6	1.4E−14
SLE																1	0.2	0.06	1	0.32	0.22
T1D																			4	1.56	0.05

O = observed; E = expected; *P* = chi-squared *P* value.

Please cite this article in press as: Baranzini SE. The genetics of autoimmune diseases: a networked perspective, Curr Opin Immunol (2009), doi:10.1016/j.coi.2009.09.014

to multiple sclerosis [30,31]. Similarly, while the minor allele of the R620W polymorphism in PTPN22 has been associated with susceptibility to T1D, RA [32–35], and SLE [36,37], it appears to confer protection to CD [38]. Functional studies aimed at detecting tissue (or cell) specific variation in the expression or function of the target gene may contribute to elucidate the pathogenic mechanisms operating in each AID [31].

Noteworthy, all of the reported associations are derived from microarray-based studies, where only common variants (minor allele frequency > 10%) are interrogated in GWAS. Thus, the almost certain influence of evolutionarily younger (rare) variants has not been adequately evaluated. With the advent of next generation sequencing, more data are expected to be gathered in the near future to address this important question. A recent report found that rare variants in the interferon-inducible gene IFIH1 are protective against T1D [39••]. While this study re-sequenced exons and splice junctions of a few candidate genes, more rare variants are likely to be found when sequencing of entire genomes becomes mainstream. The 1000 genomes project (www.1000genomes.org), an international research consortium formed to create the most detailed and medically useful picture to date of human genetic variation, is currently producing whole-genome high-quality sequences of 1000 individuals of different ethnic backgrounds and health conditions. Data emerging from this effort will undoubtedly uncover a multitude of private variants giving rise to a new catalog of human variation, an invaluable resource in the search for novel disease associations.

While the identification of the precise pathways involved in susceptibility to AID will clearly require additional time and effort, integration of data from multiple diseases represents the logical next step in discovering similarities and differences among them. While similarities will shed light into the general mechanism behind autoimmunity, genetic features private to a given disease will help pinpoint the basis for tissue specificity.

Concluding remarks

The genetic bases of autoimmunity are just starting to be uncovered. While several bona-fide susceptibility genes have been identified in most common autoimmune diseases, technological advances in high-throughput genotyping platforms and more affordable next generation sequencing methods will contribute to significantly expand these lists. In addition, structural variants (insertion/deletion polymorphisms, copy number variations, etc.) are also likely to play a significant role in determining susceptibility to AID. Taking into account the known susceptibility loci for each trait, disease-specific custom genotyping chips will be designed so as to cover a wide spectrum of variants in larger cohorts of individuals. At the same time, deep sequencing of candidate regions will be carried out to identify rare (private) mutations and structural variants that affect only a few individuals. Together, these approaches will eventually discover most if not all of the genetic contribution to these diseases and allow for the systematic search of similarity and differences among them. One obvious benefit of this new knowledge would be the cross-utilization of drugs for diseases with similar genetic fingerprint. Ultimately, this high-resolution genetic disease landscape will contribute to more accurate models of pathogenesis setting the bases for the development of more rational therapeutic approaches.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.coi.2009.09.014.

Acknowledgements

SEB is a Harry Weaver Neuroscience Scholar of the U.S. National Multiple Sclerosis Society (NMSS). This work was supported by grant RG4051A2/1 from the NMSS.

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. Lin JP, Cash JM, Doyle SZ, Peden S, Kanik K, Amos CI, Bale SJ, Wilder RL: **Familial clustering of rheumatoid arthritis with other autoimmune diseases**. *Hum Genet* 1998, **103**:475-482.

2. Tait KF, Marshall T, Berman J, Carr-Smith J, Rowe B, Todd JA, Bain SC, Barnett AH, Gough SC: **Clustering of autoimmune disease in parents of siblings from the Type 1 diabetes Warren repository**. *Diabet Med* 2004, **21**:358-362.

3. Barcellos LF, Kamdar BB, Ramsay PP, DeLoa C, Lincoln RR, Caillier S, Schmidt S, Haines JL, Pericak-Vance MA, Oksenberg JR *et al.*: **Clustering of autoimmune diseases in families with a high-risk for multiple sclerosis: a descriptive study**. *Lancet Neurol* 2006, **5**:924-931.

4. IMSGC, **The expanding genetic overlap between multiple sclerosis and type I diabetes**. *Genes Immun* 2009, 10:11–14.

5. • Rzhetsky A, Wajngurt D, Park N, Zheng T: **Probing genetic overlap among complex human phenotypes**. *Proc Natl Acad Sci U S A* 2007, **104**:11694-11699.
Authors reviewed 1.5 million medical records and created models of co-morbidity among more than 160 diseases.

6. •• Hindorff LA, Sethupathy P, Junkins HA, Ramos EM, Mehta JP, Collins FS, Manolio TA: **Potential etiologic and functional implications of genome-wide association loci for human diseases and traits**. *Proc Natl Acad Sci U S A* 2009, **106**:9362-9367.
This paper reports the creation of a searchable catalog of all published GWAS.

7. A Catalog of Published Genome-Wide Association Studies on World Wide Web URL, http://www.genome.gov/26525384.

8. Oksenberg JR, Baranzini SE, Sawcer S, Hauser SL: **The genetics of multiple sclerosis: SNPs to pathways to pathogenesis**. *Nat Rev Genet* 2008, **9**:516-526.

9. Wellcome Trust Case Control Consortium 2 on World Wide Web URL, https://www.wtccc.org.uk/ccc2/index.shtml.

Please cite this article in press as: Baranzini SE. The genetics of autoimmune diseases: a networked perspective, Curr Opin Immunol (2009), doi:10.1016/j.coi.2009.09.014

10. Torkamani A, Topol EJ, Schork NJ: **Pathway analysis of seven common diseases assessed by genome-wide association**. *Genomics* 2008, **92**:265-272.

11. Lesnick TG, Papapetropoulos S, Mash DC, Ffrench-Mullen J, Shehadeh L, de Andrade M, Henley JR, Rocca WA, Ahlskog JE, Maraganore DM: **A genomic pathway approach to a complex disease: axon guidance and Parkinson disease**. *PLoS Genet* 2007, **3**:e98.

12. Wang K, Li M, Bucan M: **Pathway-based approaches for analysis of genomewide association studies**. *Am J Hum Genet* 2007:81.

13. •• Baranzini SE, Galwey NW, Wang J, Khankhanian P, Lindberg R, Pelletier D, Wu W, Uitdehaag BM, Kappos L, Polman CH *et al.*: **Pathway and network-based analysis of genome-wide association studies in multiple sclerosis**. *Hum Mol Genet* 2009, **18**:2078-2090.
In this article, authors overlap evidence of genetic association with evidence of physical interaction to identify new candidates involved in MS susceptibility.

14. Chen GK, Witte JS: **Enriching the analysis of genomewide association studies with hierarchical modeling**. *Am J Hum Genet* 2007, **81**:397-404.

15. Forabosco P, Bouzigon E, Ng MY, Hermanowski J, Fisher SA, Criswell LA, Lewis CM: **Meta-analysis of genome-wide linkage studies across autoimmune diseases**. *Eur J Hum Genet* 2009, **17**:236-243.

16. • De Jager PL, Jia X, Wang J, de Bakker PI, Ottoboni L, Aggarwal NT, Piccio L, Raychaudhuri S, Tran D, Aubin C *et al.*: **Meta-analysis of genome scans and replication identify CD6 IRF8 and TNFRSF1A as new multiple sclerosis susceptibility loci**. *Nat Genet* 2009, **41**:776-782.
In this article, authors describe a meta-analysis across genotyping platforms using linkage disequilibrium-based allele imputation. Novel candidates for MS are reported.

17. • Johnson AD, O'Donnell CJ: **An open access database of genome-wide association results**. *BMC Med Genet* 2009, **10**:6.
This article describes a compilation of *P*-values from 118 GWAS.

18. Fisher RA: **Combining independent tests of significance**. *American Statistician* 1948, **2**:30.

19. • Goh KI, Cusick ME, Valle D, Childs B, Vidal M, Barabasi AL: **The human disease network**. *Proc Natl Acad Sci U S A* 2007, **104**:8685-8690.
Authors create a network of genes and diseases using information from OMIM database.

20. Kumar V, Abbas AK, Fausto N, Robbins SL, Cotran RS: *Robbins and Cotran pathologic basis of disease*. 7th. Philadelphia: Elsevier Saunders; 2005.

21. Husebye E, Lovas K: **Pathogenesis of primary adrenal insufficiency**. *Best Pract Res Clin Endocrinol Metab* 2009, **23**:147-157.

22. Smyth D, Cooper JD, Collins JE, Heward JM, Franklyn JA, Howson JM, Vella A, Nutland S, Rance HE, Maier L *et al.*: **Replication of an association between the lymphoid tyrosine phosphatase locus (LYP/PTPN22) with type 1 diabetes, and evidence for its role as a general autoimmunity locus**. *Diabetes* 2004, **53**:3020-3023.

23. Abraham C, Cho JH: **IL-23 and autoimmunity: new insights into the pathogenesis of inflammatory bowel disease**. *Annu Rev Med* 2009, **60**:97-110.

24. Nelson BH: **IL-2, regulatory T cells, and tolerance**. *J Immunol* 2004, **172**:3983-3988.

25. Malek TR, Bayer AL: **Tolerance, not immunity, crucially depends on IL-2**. *Nat Rev Immunol* 2004, **4**:665-674.

26. Jones JL, Phuah CL, Cox AL, Thompson SA, Ban M, Shawcross J, Walton A, Sawcer SJ, Compston A, Coles AJ: **IL-21 drives secondary autoimmunity in patients with multiple sclerosis, following therapeutic lymphocyte depletion with alemtuzumab (Campath-1H)**. *J Clin Invest* 2009, **119**:2052-2061.

27. Ettinger R, Kuchen S, Lipsky PE: **The role of IL-21 in regulating B-cell function in health and disease**. *Immunol Rev* 2008, **223**:60-86.

28. Ettinger R, Kuchen S, Lipsky PE: **Interleukin 21 as a target of intervention in autoimmune disease**. *Ann Rheum Dis* 2008, **67(Suppl 3)**:iii83-iii86.

29. Scalapino KJ, Daikh DI: **CTLA-4: a key regulatory point in the control of autoimmune disease**. *Immunol Rev* 2008, **223**:143-155.

30. Maier LM, Lowe CE, Cooper J, Downes K, Anderson DE, Severson C, Clark PM, Healy B, Walker N, Aubin C *et al.*: **IL2RA genetic heterogeneity in multiple sclerosis and type 1 diabetes susceptibility and soluble interleukin-2 receptor production**. *PLoS Genet* 2009, **5**:e1000322.

31. Dendrou CA, Plagnol V, Fung E, Yang JH, Downes K, Cooper JD, Nutland S, Coleman G, Himsworth M, Hardy M *et al.*: **Cell-specific protein phenotypes for the autoimmune locus IL2RA using a genotype-selectable human bioresource**. *Nat Genet* 2009, **41**:1011-1015.

32. Gregersen PK, Amos CI, Lee AT, Lu Y, Remmers EF, Kastner DL, Seldin MF, Criswell LA, Plenge RM, Holers VM *et al.*: **REL, encoding a member of the NF-kappaB family of transcription factors, is a newly defined risk locus for rheumatoid arthritis**. *Nat Genet* 2009, **41**:820-823.

33. Raychaudhuri S, Remmers EF, Lee AT, Hackett R, Guiducci C, Burtt NP, Gianniny L, Korman BD, Padyukov L, Kurreeman FA *et al.*: **Common variants at CD40 and other loci confer risk of rheumatoid arthritis**. *Nat Genet* 2008, **40**:1216-1223.

34. Plenge RM, Seielstad M, Padyukov L, Lee AT, Remmers EF, Ding B, Liew A, Khalili H, Chandrasekaran A, Davies LR *et al.*: **TRAF1-C5 as a risk locus for rheumatoid arthritis—a genomewide study**. *N Engl J Med* 2007, **357**:1199-1209.

35. WTCCC, Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. *Nature* 2007, 447:661–678.

36. Harley JB, Alarcon-Riquelme ME, Criswell LA, Jacob CO, Kimberly RP, Moser KL, Tsao BP, Vyse TJ, Langefeld CD, Nath SK *et al.*: **Genome-wide association scan in women with systemic lupus erythematosus identifies susceptibility variants in ITGAM, PXK, KIAA1542 and other loci**. *Nat Genet* 2008, **40**:204-210.

37. Hom G, Graham RR, Modrek B, Taylor KE, Ortmann W, Garnier S, Lee AT, Chung SA, Ferreira RC, Pant PV *et al.*: **Association of systemic lupus erythematosus with C8orf13-BLK and ITGAM-ITGAX**. *N Engl J Med* 2008, **358**:900-909.

38. Barrett JC, Hansoul S, Nicolae DL, Cho JH, Duerr RH, Rioux JD, Brant SR, Silverberg MS, Taylor KD, Barmada MM *et al.*: **Genome-wide association defines more than 30 distinct susceptibility loci for Crohn's disease**. *Nat Genet* 2008, **40**:955-962.

39. •• Nejentsev S, Walker N, Riches D, Egholm M, Todd JA: **Rare variants of IFIH1, a gene implicated in antiviral responses, protect against type 1 diabetes**. *Science* 2009, **324**:387-389.
This article describes the use of next generation sequencing to identify rare variants associated with type I diabetes

40. Hunt KA, Zhernakova A, Turner G, Heap GA, Franke L, Bruinenberg M, Romanos J, Dinesen LC, Ryan AW, Panesar D *et al.*: **Newly identified genetic risk variants for celiac disease related to the immune response**. *Nat Genet* 2008, **40**:395-402.

41. van Heel DA, Franke L, Hunt KA, Gwilliam R, Zhernakova A, Inouye M, Wapenaar MC, Barnardo MC, Bethel G, Holmes GK *et al.*: **A genome-wide association study for celiac disease identifies risk variants in the region harboring IL2 and IL21**. *Nat Genet* 2007, **39**:827-829.

42. ANZgene: **Genome-wide association study identifies new multiple sclerosis susceptibility loci on chromosomes 12 and 20**. *Nat Genet* 2009, **41**:824-828.

43. Comabella M, Craig DW, Camina-Tato M, Morcillo C, Lopez C, Navarro A, Rio J, Montalban X, Martin R: **Identification of a novel risk locus for multiple sclerosis at 13q31.3 by a pooled genome-wide scan of 500,000 single nucleotide polymorphisms**. *PLoS One* 2008, **3**:e3490.

Please cite this article in press as: Baranzini SE. The genetics of autoimmune diseases: a networked perspective, Curr Opin Immunol (2009), doi:10.1016/j.coi.2009.09.014

44. Hafler DA, Compston A, Sawcer S, Lander ES, Daly MJ, De Jager PL, de Bakker PI, Gabriel SB, Mirel DB, Ivinson AJ *et al.*: **Risk alleles for multiple sclerosis identified by a genomewide study**. *N Engl J Med* 2007, **357**:851-862.

45. Nair RP, Duffin KC, Helms C, Ding J, Stuart PE, Goldgar D, Gudjonsson JE, Li Y, Tejasvi T, Feng BJ *et al.*: **Genome-wide scan reveals association of psoriasis with IL-23 and NF-kappaB pathways**. *Nat Genet* 2009, **41**:199-204.

46. Liu Y, Helms C, Liao W, Zaba LC, Duan S, Gardner J, Wise C, Miner A, Malloy MJ, Pullinger CR *et al.*: **A genome-wide association study of psoriasis and psoriatic arthritis identifies new disease loci**. *PLoS Genet* 2008, **4**:e1000041.

47. Capon F, Bijlmakers MJ, Wolf N, Quaranta M, Huffmeier U, Allen M, Timms K, Abkevich V, Gutin A, Smith R *et al.*: **Identification of ZNF313/RNF114 as a novel psoriasis susceptibility gene**. *Hum Mol Genet* 2008, **17**:1938-1945.

48. Zhang XJ, Huang W, Yang S, Sun LD, Zhang FY, Zhu QX, Zhang FR, Zhang C, Du WH, Pu XM *et al.*: **Psoriasis genome-wide association study identifies susceptibility variants within LCE gene cluster at 1q21**. *Nat Genet* 2009, **41**:205-210.

49. Raelson JV, Little RD, Ruether A, Fournier H, Paquin B, Van Eerdewegh P, Bradley WE, Croteau P, Nguyen-Huu Q, Segal J *et al.*: **Genome-wide association study for Crohn's disease in the Quebec Founder Population identifies multiple validated disease loci**. *Proc Natl Acad Sci U S A* 2007, **104**:14747-14752.

50. Rioux JD, Xavier RJ, Taylor KD, Silverberg MS, Goyette P, Huett A, Green T, Kuballa P, Barmada MM, Datta LW *et al.*: **Genome-wide association study identifies new susceptibility loci for Crohn disease and implicates autophagy in disease pathogenesis**. *Nat Genet* 2007, **39**:596-604.

51. Libioulle C, Louis E, Hansoul S, Sandor C, Farnir F, Franchimont D, Vermeire S, Dewit O, de Vos M, Dixon A *et al.*: **Novel Crohn disease locus identified by genome-wide association maps to a gene desert on 5p13.1 and modulates expression of PTGER4**. *PLoS Genet* 2007, **3**:e58.

52. Parkes M, Barrett JC, Prescott NJ, Tremelling M, Anderson CA, Fisher SA, Roberts RG, Nimmo ER, Cummings FR, Soars D *et al.*: **Sequence variants in the autophagy gene IRGM and multiple other replicating loci contribute to Crohn's disease susceptibility**. *Nat Genet* 2007, **39**:830-832.

53. Plenge RM, Cotsapas C, Davies L, Price AL, de Bakker PI, Maller J, Pe'er I, Burtt NP, Blumenstiel B, DeFelice M *et al.*: **Two independent alleles at 6q23 associated with risk of rheumatoid arthritis**. *Nat Genet* 2007, **39**:1477-1482.

54. Julia A, Ballina J, Canete JD, Balsa A, Tornero-Molina J, Naranjo A, Alperi-Lopez M, Erra A, Pascual-Salcedo D, Barcelo P *et al.*: **Genome-wide association study of rheumatoid arthritis in the Spanish population: KLF12 as a risk locus for rheumatoid arthritis susceptibility**. *Arthritis Rheum* 2008, **58**:2275-2286.

55. Graham RR, Cotsapas C, Davies L, Hackett R, Lessard CJ, Leon JM, Burtt NP, Guiducci C, Parkin M, Gates C *et al.*: **Genetic variants near TNFAIP3 on 6q23 are associated with systemic lupus erythematosus**. *Nat Genet* 2008, **40**:1059-1061.

56. Kozyrev SV, Abelson AK, Wojcik J, Zaghlool A, Linga Reddy MV, Sanchez E, Gunnarsson I, Svenungsson E, Sturfelt G, Jonsen A *et al.*: **Functional variants in the B-cell gene BANK1 are associated with systemic lupus erythematosus**. *Nat Genet* 2008, **40**:211-216.

57. Cooper JD, Smyth DJ, Smiles AM, Plagnol V, Walker NM, Allen JE, Downes K, Barrett JC, Healy BC, Mychaleckyj JC *et al.*: **Meta-analysis of genome-wide association study data identifies additional type 1 diabetes risk loci**. *Nat Genet* 2008, **40**:1399-1401.

58. Hakonarson H, Grant SF, Bradfield JP, Marchand L, Kim CE, Glessner JT, Grabs R, Casalunovo T, Taback SP, Frackelton EC *et al.*: **A genome-wide association study identifies KIAA0350 as a type 1 diabetes gene**. *Nature* 2007, **448**:591-594.

59. Todd JA, Walker NM, Cooper JD, Smyth DJ, Downes K, Plagnol V, Bailey R, Nejentsev S, Field SF, Payne F *et al.*: **Robust associations of four new chromosome regions from genome-wide analyses of type 1 diabetes**. *Nat Genet* 2007, **39**:857-864.

60. Hakonarson H, Qu HQ, Bradfield JP, Marchand L, Kim CE, Glessner JT, Grabs R, Casalunovo T, Taback SP, Frackelton EC *et al.*: **A novel susceptibility locus for type 1 diabetes on Chr12q13 identified by a genome-wide association study**. *Diabetes* 2008, **57**:1143-1146.

61. Grant SF, Qu HQ, Bradfield JP, Marchand L, Kim CE, Glessner JT, Grabs R, Taback SP, Frackelton EC, Eckert AW *et al.*: **Follow-up analysis of genome-wide association data identifies novel loci for type 1 diabetes**. *Diabetes* 2009, **58**:290-295.

62. Franke A, Hampe J, Rosenstiel P, Becker C, Wagner F, Hasler R, Little RD, Huse K, Ruether A, Balschun T *et al.*: **Systematic association mapping identifies NELL1 as a novel IBD disease gene**. *PLoS One* 2007, **2**:e691.

63. Duerr RH, Taylor KD, Brant SR, Rioux JD, Silverberg MS, Daly MJ, Steinhart AH, Abraham C, Regueiro M, Griffiths A *et al.*: **A genome-wide association study identifies IL23R as an inflammatory bowel disease gene**. *Science* 2006, **314**:1461-1463.

64. Yamazaki K, McGovern D, Ragoussis J, Paolucci M, Butler H, Jewell D, Cardon L, Takazoe M, Tanaka T, Ichimori T *et al.*: **Single nucleotide polymorphisms in TNFSF15 confer susceptibility to Crohn's disease**. *Hum Mol Genet* 2005, **14**:3499-3506.

65. Baranzini SE, Wang J, Gibson RA, Galwey N, Naegelin Y, Barkhof F, Radue EW, Lindberg RL, Uitdehaag BM, Johnson MR *et al.*: **Genome-wide association analysis of susceptibility and clinical phenotype in multiple sclerosis**. *Hum Mol Genet* 2009, **18**:767-778.

66. Capon F, Di Meglio P, Szaub J, Prescott NJ, Dunster C, Baumber L, Timms K, Gutin A, Abkevic V, Burden AD *et al.*: **Sequence variants in the genes for the interleukin-23 receptor (IL23R) and its ligand (IL12B) confer protection against psoriasis**. *Hum Genet* 2007, **122**:201-206.

67. Steer S, Abkevich V, Gutin A, Cordell HJ, Gendall KL, Merriman ME, Rodger RA, Rowley KA, Chapman P, Gow P *et al.*: **Genomic DNA pooling for whole-genome association scans in complex disease: empirical demonstration of efficacy in rheumatoid arthritis**. *Genes Immun* 2007, **8**:57-68.

68. Kamatani Y, Matsuda K, Ohishi T, Ohtsubo S, Yamazaki K, Iida A, Hosono N, Kubo M, Yumura W, Nitta K *et al.*: **Identification of a significant association of a single nucleotide polymorphism in TNXB with systemic lupus erythematosus in a Japanese population**. *J Hum Genet* 2008, **53**:64-73.

69. Cervino AC, Tsinoremas NF, Hoffman RW: **A genome-wide study of lupus: preliminary analysis and data release**. *Ann N Y Acad Sci* 2007, **1110**:131-139.

70. Saxena R, Voight BF, Lyssenko V, Burtt NP, de Bakker PI, Chen H, Roix JJ, Kathiresan S, Hirschhorn JN, Daly MJ *et al.*: **Genome-wide association analysis identifies loci for type 2 diabetes and triglyceride levels**. *Science* 2007, **316**:1331-1336.

71. Hanson RL, Bogardus C, Duggan D, Kobes S, Knowlton M, Infante AM, Marovich L, Benitez D, Baier LJ, Knowler WC: **A search for variants associated with young-onset type 2 diabetes in American Indians in a 100K genotyping array**. *Diabetes* 2007, **56**:3045-3052.

72. Hayes MG, Pluzhnikov A, Miyake K, Sun Y, Ng MC, Roe CA, Below JE, Nicolae RI, Konkashbaev A, Bell GI *et al.*: **Identification of type 2 diabetes genes in Mexican Americans through genome-wide association studies**. *Diabetes* 2007, **56**:3033-3044.

73. Rampersaud E, Damcott CM, Fu M, Shen H, McArdle P, Shi X, Shelton J, Yin J, Chang YP, Ott SH *et al.*: **Identification of novel candidate genes for type 2 diabetes from a genome-wide association scan in the Old Order Amish: evidence for replication from diabetes-related quantitative traits and from independent populations**. *Diabetes* 2007, **56**:3053-3062.

74. Salonen JT, Uimari P, Aalto JM, Pirskanen M, Kaikkonen J, Todorova B, Hypponen J, Korhonen VP, Asikainen J, Devine C *et al.*: **Type 2 diabetes whole-genome association study in four populations: the DiaGen consortium**. *Am J Hum Genet* 2007, **81**:338-345.

Please cite this article in press as: Baranzini SE. The genetics of autoimmune diseases: a networked perspective, Curr Opin Immunol (2009), doi:10.1016/j.coi.2009.09.014

75. Scott LJ, Mohlke KL, Bonnycastle LL, Willer CJ, Li Y, Duren WL, Erdos MR, Stringham HM, Chines PS, Jackson AU *et al.*: **A genome-wide association study of type 2 diabetes in Finns detects multiple susceptibility variants**. *Science* 2007, **316**:1341-1345.

76. Sladek R, Rocheleau G, Rung J, Dina C, Shen L, Serre D, Boutin P, Vincent D, Belisle A, Hadjadj S *et al.*: **A genome-wide association study identifies novel risk loci for type 2 diabetes**. *Nature* 2007, **445**:881-885.

77. Steinthorsdottir V, Thorleifsson G, Reynisdottir I, Benediktsson R, Jonsdottir T, Walters GB, Styrkarsdottir U, Gretarsdottir S, Emilsson V, Ghosh S *et al.*: **A variant in CDKAL1 influences insulin response and risk of type 2 diabetes**. *Nat Genet* 2007, **39**:770-775.

78. Zeggini E, Weedon MN, Lindgren CM, Frayling TM, Elliott KS, Lango H, Timpson NJ, Perry JR, Rayner NW, Freathy RM *et al.*: **Replication of genome-wide association signals in UK samples reveals risk loci for type 2 diabetes**. *Science* 2007, **316**:1336-1341.

79. Florez JC, Manning AK, Dupuis J, McAteer J, Irenze K, Gianniny L, Mirel DB, Fox CS, Cupples LA, Meigs JB: **A 100 K genome-wide association scan for diabetes and related traits in the Framingham Heart Study: replication and integration with other genome-wide datasets**. *Diabetes* 2007, **56**:3063-3074.

80. Meigs JB, Manning AK, Fox CS, Florez JC, Liu C, Cupples LA, Dupuis J: **Genome-wide association with diabetes-related traits in the Framingham Heart Study**. *BMC Med Genet* 2007, **8(Suppl 1)**:S16.

Please cite this article in press as: Baranzini SE. The genetics of autoimmune diseases: a networked perspective, Curr Opin Immunol (2009), doi:10.1016/j.coi.2009.09.014