

Detecting shared pathogenesis from the shared genetics of immune-related diseases

Alexandra Zhernakova*[§], Cleo C. van Diemen*[§] and Cisca Wijmenga*[†]

Abstract | Recent genetic studies have revealed shared immunological mechanisms in several immune-related disorders that further our understanding of the development and concomitance of these diseases. Our Review focuses on these shared aspects, using the novel findings of recently performed genome-wide association studies and non-synonymous SNP scans as a starting point. We discuss how identifying new genes that are associated with more than one autoimmune or chronic inflammatory disorder could explain the genetic basis of the shared pathogenesis of immune-related diseases. This analysis helps to highlight the key molecular pathways that are involved in these disorders and the potential roles of novel genes in immune-related diseases.

Innate immune system

An immediate nonspecific immune response to foreign infectious agents. It includes chemical defence mechanisms (for example, mucus and complement production), as well as cellular functions, such as phagocytosis by macrophages and neutrophils.

Adaptive immune system

A specific immune response initiated by highly specific receptors that are present on B and T cells, which recognize antigens. There is extensive crosstalk between the innate and adaptive immune responses.

*Department of Medical Genetics, University Medical Centre Utrecht, Utrecht 3508 AB, the Netherlands.

[†]Department of Genetics, University Medical Centre Groningen, University of Groningen, Groningen 9700 RB, the Netherlands.

[§]These authors contributed equally to this work.

Correspondence to C.W. e-mail:

c.wijmenga@umcutrecht.nl
doi:10.1038/nrg2489

The immune system plays an essential part in the host defence against harmful antigens and in the balance between tolerance and immunity to other antigens. Dysregulation of the immune system can result in disorders of both the innate immune system and adaptive immune system, and in inflammation. Immune-related disorders consist of a range of diseases, which include autoimmune and inflammatory disorders, such as [asthma](#), [rheumatoid arthritis](#), [inflammatory bowel disease](#) and [type 1 diabetes](#) (T1D). Autoimmune disorders arise from inappropriate destruction of normal tissue by the immune system owing to a failure of the immune self-tolerance mechanism, which is the ability of the immune system to discriminate between 'self' and 'non-self' antigens¹. By contrast, inflammatory disorders result from an excessive inflammatory response that is more harmful to host tissue than to exogenous antigens, such as pathogens and other environmental stimuli^{2,3}.

Until recently, autoimmune disorders were thought to develop by different mechanisms from those that provoke inflammatory disorders, but recent insights suggest that similar pathogenetic effector pathways operate in both types of disorder. In autoimmune and inflammatory disorders, the key elements that regulate the attack or destruction of tissues or organ systems by an inappropriate or excessive immune response include: known genetic factors, such as the human leukocyte antigen (HLA) genes; innate and adaptive immune regulation processes; and triggering by environmental factors, such as infections or dietary components. The notion of a

common pathophysiological mechanism is further supported by the observation of parallel inflammatory and autoimmune disorders in some patients and families. For example, co-occurrence of T1D, autoimmune thyroiditis ([Graves' disease](#) and [Hashimoto's disease](#)), rheumatoid arthritis, [coeliac disease](#) and [multiple sclerosis](#) has been observed^{4,5}. However, most studies investigating co-morbidity have been based on selected disease cohorts rather than population data, and suffer from many shortcomings that include inadequate sample sizes, ascertainment bias and heterogeneity of outcome.

It is known that genetic factors play an important part in the development of immune-related disorders, and the heritable component of these disorders is evident from the high concordance observed in monozygotic twin pairs and increased familial clustering. The high heritability of these diseases has prompted much research to identify the underlying disease susceptibility genes. Nonetheless, elucidating the genetics of these disorders is severely hampered by the existence of genetic heterogeneity, the low penetrance of individual disease alleles, and the potential for gene-gene and gene-environment interactions. Hence, it is likely that many different susceptibility alleles contribute to each disease, each of which have only a modest effect. Genome-wide association studies (GWA studies) have recently led to the identification of tens of loci in the human genome that are associated with susceptibility to immune-related disorders⁶. These studies also indicate that certain loci and genes seem to predispose to multiple

Ascertainment bias

A false conclusion that is made as a result of nonrandom sampling.

Heritability

The proportion of the total phenotypic variation for a given characteristic in a population that can be attributed to genetic variance among individuals.

Genetic heterogeneity

(Also called locus heterogeneity). A situation in which variation in different genes might cause identical or similar forms of the disease in different individuals.

Penetrance

The probability of observing a specific phenotype in individuals who carry a particular genotype. If this probability is less than one for all genotypes of a variant, then the variant has incomplete penetrance.

Genome-wide association study

(GWA study). A large-scale genotyping analysis of markers across the human genome, which is designed to identify genetic association with diseases or observable traits.

Genome-wide non-synonymous SNP scan

Genome-wide scan for disease association that includes only non-synonymous SNPs (nsSNPs).

Linkage analysis

The process of mapping genes by typing genetic markers in families to identify chromosome regions that are associated with disease or trait values within pedigrees more often than would be expected by chance. Such linked regions are more likely to contain a causal genetic variant.

Functional candidate gene

A gene that might be involved in a particular disease because of its biological relevance.

Identical by descent

Describes multiple alleles that are identical because they arose from the same allele in an earlier generation.

immune-related disorders, which confirm the suggestion that shared molecular mechanisms that are due to common genetic variants contribute to a spectrum of diseases.

This Review focuses on the shared aspects of immune-related disorders by taking the new genetic findings from recent GWA studies and genome-wide non-synonymous SNP scans (nsSNP scans) as a starting point. We have brought together the evidence from GWA studies and nsSNP scans from 11 immune-related diseases (BOX 1; TABLE 1) that show extensive co-morbidity (TABLE 2). This analysis highlights the fact that the majority of newly identified genes can be mapped to a few shared immunological molecular pathways. We discuss the common pathways underlying these immune-related diseases and how genetic analysis might further our understanding of immune disease pathology.

Genetics of immune-related diseases

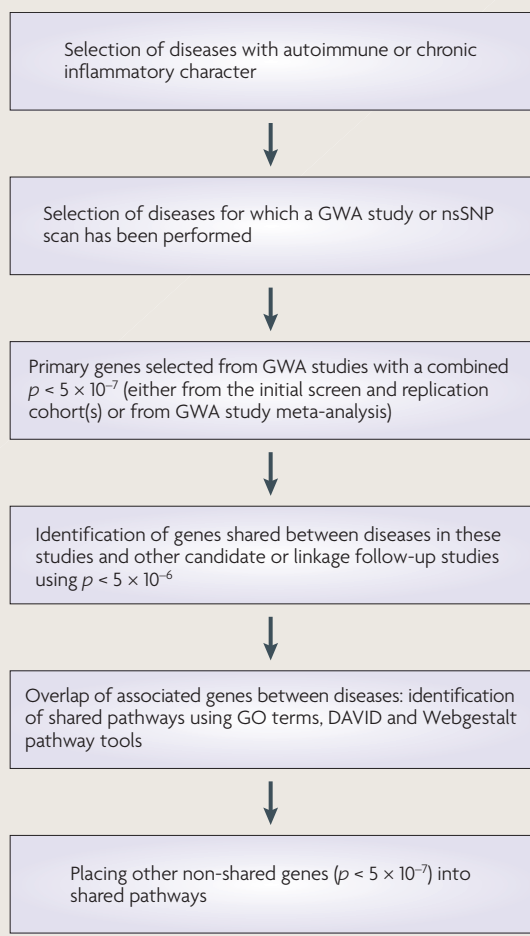
Until recently, the search for genetic risk factors for complex disorders was limited to genome-wide linkage analysis and association studies using single functional candidate genes. Most of the association studies in immune-related diseases were based on genes involved in T-cell signalling and the adaptive immune response. Hundreds

of associations between the HLA system and autoimmune disorders have been established (TABLE 1) and the first associations were discovered in the 1970s, which was long before molecular techniques were introduced⁷. Despite a large number of non-HLA candidate gene association studies, only two genes have consistently been found to be associated with multiple autoimmune disorders: *CTLA4* (cytotoxic T-lymphocyte-associated 4) and *PTPN22* (protein tyrosine phosphatase, non-receptor type 22 (lymphoid))^{8,9}.

The linkage studies that have been performed for many immune-related diseases have shown that a large number of regions in the genome co-segregate with disease in families, or that shared alleles that are identical by descent are present in affected sibling pairs. These genetic studies indicated that certain chromosomal regions seem to confer predisposition to multiple immune-related disorders, thereby supporting the suggestion that a shared group of genes contribute to the spectrum of immune diseases¹⁰⁻¹². Initially, the identification and independent replication of true disease genes from these linkage regions was disappointing, except for a few genes, such as the *NOD2* gene (nucleotide-binding oligomerization domain containing 2; also known as *CARD15*), which is involved in inflammatory bowel

Box 1 | Selection of studies and mapping of shared pathways

We chose to focus on 11 immune-related diseases (TABLE 1) for which at least one genome-wide association (GWA) study or non-synonymous SNP (nsSNP) scan was available. From the GWA studies and nsSNP scans, we selected all the genes or loci that were associated with these diseases with a combined $p < 5 \times 10^{-7}$ (either from the GWA studies, nsSNP scan or meta-analysis). All associated genes were replicated in an independent cohort either within the original study or in a follow-up paper (we also included the cytotoxic T-lymphocyte-associated 4 (*CTLA4*) gene, which was associated in the original screen with $p < 5 \times 10^{-5}$, but not included for replication, as it was a confirmed type 1 diabetes gene⁹). This analysis revealed 88 associated loci. To investigate which of these loci overlapped with two or more immune diseases, we reconsidered the GWA studies, nsSNP scans and GWA meta-analysis studies, as well as independent candidate gene studies, replication papers, and linkage fine-mapping studies, using a less stringent cut-off for the p value of less than 5×10^{-6} (Supplementary information S1 (table)). The resulting list of 23 genes, which were encoded by 20 loci (Supplementary information S3 (figure)) overlapping two or more diseases, was analysed using three pathway analysis tools: the [Database for Annotation, Visualization and Integrated Discovery \(DAVID\)](#), [Webgestalt](#) and [Gene Ontology \(GO\)](#) (detailed results are given in [Supplementary information S5 \(table\)](#)). This approach enabled us to link many of the 23 genes to shared immunological pathways or to related categories (TABLE 3). After identifying the major shared pathways, we investigated whether the 68 remaining non-shared associated loci (with $p < 5 \times 10^{-7}$) (Supplementary information S4 (table)) could be placed in the same pathways. This selection strategy is presented in the flowchart. The ascertainment bias in our selection of genes is discussed in BOX 2.



Population attributable risk (PAR). Calculated using the following formula, where *f* is the allele frequency in the population, and *RR* is the relative risk:

$$PAR = \frac{f(RR-1)}{1+f(RR-1)}$$

disease. Despite the enormous efforts that were invested in linkage and candidate gene studies, no clear picture emerged regarding the pathways that might be causally involved in susceptibility to immune-related diseases. This situation changed dramatically with the arrival of GWA studies.

GWA studies aim to cover the majority of the common variants in the genome by using knowledge of the linkage disequilibrium (LD) relationships between the hundreds of thousands of SNPs generated by the HapMap project¹³. GWA studies now provide an unbiased approach for identifying the association of genes

Table 1 | **Chronic inflammatory and autoimmune disorders**

Disease	Clinical manifestation	Disease pathogenesis	Strength of association* to HLA [‡] ; PAR [§]	Prevalence ^{84,85}	Refs
Ankylosing spondylitis (OMIM 106300)	Chronic, degenerative inflammatory arthritis that primarily affects the spine and sacroiliac joints, causing eventual fusion of the spine	Specific auto-antibodies cannot be detected. Possibly cross-reaction with antigens of the <i>Klebsiella</i> spp. or other bacterial strain	Strong (PAR = 0.43)	0.025%	27
Asthma (OMIM 600807)	Chronic condition in the respiratory system in which the airways occasionally constrict, become inflamed, and are lined with excessive amounts of mucus	Inflammation in response to exposure to an environmental stimulant, such as an allergen, smoke or perfume, which is mediated by a T _H 2-type immune response and includes mast cells, eosinophil infiltrates and IgE antibodies	Weak (PAR = 0.03)	6–8%	86
Autoimmune thyroid diseases (Graves' disease and Hashimoto's disease) (OMIM 27500 and OMIM 140300)	Graves' disease: chronic hyperthyroidism Hashimoto's disease: chronic hypothyroidism	Graves' disease: autoimmune reaction to the receptor for thyroid-stimulating hormone Hashimoto's disease: autoimmune response against thyroid peroxidase and thyroglobulin	Moderate (PAR = 0.14)	0.8–1.2%	27, 87
Coeliac disease (OMIM 212750)	Chronic inflammation of the intestine and flattening of the mucosa	Exposure to gluten peptides elicits T cell-mediated autoimmune reaction against tissue transglutaminase, and less often to other auto-antigens	Strong (PAR = 0.43)	0.5–2%	30, 38
Crohn's disease (OMIM 266600)	Chronic, episodic, inflammatory bowel disease, which primarily causes ulceration of the small and large intestines but can affect any region of the digestive system	Unknown; involves an inappropriate immune response to commensal bacteria	Weak (PAR = 0.03)	0.25–0.65% (inflammatory bowel disease)	7,18
Multiple sclerosis (OMIM 126200)	Autoimmune attack of the central nervous system, which leads to demyelination of neurons, causing potentially debilitating physical and mental symptoms	After infection in the brain, trapped T cells initiate an autoimmune response to foreign myelin, thereby triggering inflammatory processes, stimulating other immune cells, cytokines and antibodies	Moderate (PAR = 0.19)	0.02%	88, 89
Psoriasis (OMIM 177900)	Inflammation and the rapid growth and reproduction of skin cells	Unknown trigger; T cells become active, migrate to the dermis and activate the release of cytokines (particularly TNF α), leading to inflammation	Moderate–strong (PAR = 0.21)	0.5–1% worldwide, 2% of Europeans	90
Rheumatoid arthritis (OMIM 180300)	Chronic inflammation of synovial joints	Autoimmune reaction against connective tissue components. Presence of rheumatoid factor and ACPA	Strong (PAR = 0.5)	0.5–1%	49, 91
Systemic lupus erythematosus (OMIM 152700)	Chronic inflammation, can affect any part of the body, but often the heart, joints, skin, lungs, blood vessels, liver, kidneys and nervous system	Autoimmune reaction against nuclear proteins, which leads to the formation of immune complexes	Moderate (PAR = 0.13)	0.04–0.12%	37
Type 1 diabetes (OMIM 222100)	Destruction of pancreatic β -cells, which leads to insufficient release of insulin from the pancreas	T cell-mediated autoimmune response, and production of auto-antibodies against islet cells, insulin, glutamic acid decarboxylase and protein tyrosine phosphatase	Strong (PAR = 0.49)	0.2–0.3%	26
Ulcerative colitis (OMIM 191390)	Chronic inflammation and ulcers in the top layer of the lining of the large intestine	Unknown; characterized by abnormal activation of the immune system	Weak–moderate (PAR = 0.16)	ND	7,20, 92

*Only human leukocyte antigen (HLA) association in Caucasians has been reported. [‡]HLA association is annotated as strong if the odds ratio for reported alleles is above 4; moderate if the odds ratio is between 2 and 4; and weak if the odds ratio is below 2. [§]Population attributable risk (PAR) was calculated based on the HLA SNP that showed the strongest association to the disease in genome-wide association (GWA) studies, with the exception of the asthma study⁸⁶. An overview of the associated HLA alleles is presented in [Supplementary information S2](#) (table). ACPA, anti-citrulline-peptide antibody; IgE, immunoglobulin E; ND, not determined; OMIM, Online Mendelian Inheritance in Man; T_H, T helper; TNF α , tumour necrosis factor α .

Table 2 | Concomitance of chronic inflammatory and autoimmune diseases

Disease	Other diseases seen in patients	Familial clustering	Refs
Ankylosing spondylitis	Ulcerative colitis (5%), Crohn's disease (3%), psoriasis (16%)	IBD* (7%), psoriasis (10%)	94
Asthma	Crohn's disease, ulcerative colitis	ND	95
AIT disease (including Graves' and Hashimoto's disease)	Rheumatoid arthritis (1.4–17.6%), T1D (3–15%), coeliac disease (5.4%)	ND	4,96
Coeliac disease	Asthma (24.6%), SLE (2.4%), AIT (5%), T1D (5%), psoriasis	ND	97–99
Crohn's disease	Rheumatoid arthritis (1.7–1.6%), asthma (7.1–13.0%), psoriasis (1.7–1.9%), coeliac disease (19%), multiple sclerosis (0.4%)	Psoriasis (10%)	95,100–102
Multiple sclerosis	AIT (0.5–3.9%), psoriasis (6%), IBD (3%)*, rheumatoid arthritis (0.35–2.4%), T1D (0–2.6%)	AIT (10%), psoriasis (6%), IBD (3%)*, rheumatoid arthritis (2%)	4
Psoriasis	Crohn's disease, ulcerative colitis, coeliac disease	ND	95
Rheumatoid arthritis	Asthma (10.0%), T1D (0.3–6%), AIT (0.5–9.8%)	AIT (2.1%), T1D (0.44%), SLE (0.36%), multiple sclerosis (0.29%)	4,85,99
SLE	Diabetes mellitus (11.6%)	ND	103
T1D	Coeliac disease (4–9%), AIT (0–24%), asthma (5%), rheumatoid arthritis (2%)	Coeliac disease (6%), AIT (8%) (2.7%), rheumatoid arthritis (0.97%)	4,99,104, 105
Ulcerative colitis	Rheumatoid arthritis (1.1–1.6%), asthma (7.9–12.0%), psoriasis (1.7–1.8%), multiple sclerosis (0.4–0.54%)	ND	95,102

*No discrimination made between Crohn's disease and ulcerative colitis. AIT, autoimmune thyroid disease; IBD, inflammatory bowel disease; ND, not determined; SLE, systemic lupus erythematosus; T1D, type 1 diabetes.

Linkage disequilibrium (LD). The nonrandom association of genetic marker alleles. Two markers are in LD when some combinations of alleles in a population occur more or less frequently than would be expected if random assortment occurred.

Major histocompatibility complex (MHC). A 4-Mb region of human chromosome 6 that contains many genes with immunological functions. It is encoded by the human leukocyte antigen (HLA) locus.

Deep sequence
Massive parallel sequencing of the same DNA target with new-generation sequencing platforms, such as the Roche 454 FLX system, the Illumina Genome Analyzer and the Applied Biosystems SOLiD system.

Meta-analysis
An approach that combines the results of several studies that address a set of related research hypotheses to overcome the problem of reduced statistical power in studies with small sample sizes.

with diseases. In the past 2 years, ~300 loci for ~120 diseases and traits have been identified using GWA studies. Of the 187 studies reported in the [Catalogue of Published Genome-wide Association Studies](#) before 15 October 2008, 22 studies were performed on the 11 immune-mediated diseases that we have focused on and that are analysed in this Review (BOX 1; [Supplementary information S1](#) (table)). These GWA studies successfully identified novel genetic risk factors in each of the diseases discussed. However, it is important to realise that GWA studies also have limitations (BOX 2). An alternative to GWA studies is genome-wide nsSNP scans, in which only nsSNPs are tested. The two methods are similar because both have the advantage of being a hypothesis-free method. However, the lower numbers of SNPs tested in genome-wide nsSNP scans substantially limits the coverage of the human genome in these scans. Moreover, in contrast to monogenic disorders, many of the variants identified so far for complex diseases are located in non-coding regions, which are not covered by nsSNP scans.

The main genetic region linked to immune disorders before the advent of GWA studies was the HLA region, which is encoded by the major histocompatibility complex (MHC) locus (TABLE 1; [Supplementary information S2](#) (table))¹⁴. The MHC locus spans approximately 4 Mb and contains ~250 genes, of which ~60% have immune-related functions. The MHC region is characterized by extended LD blocks (up to 3 Mb), and by a strong and complicated LD pattern between the blocks¹⁵. These features make it difficult to pinpoint the exact location of the associated signal. Despite the fact that association of the MHC region to many immune diseases has been confirmed using GWA studies with many SNPs, these studies have been unable to pinpoint the causal genes. One approach to solve this problem would be to fine map

and deep sequence a large cohort to define which genes in the MHC region are causal, and thus confirm the association of these genes with susceptibility to certain immune-related diseases. This approach was recently used to define the causal genes for T1D¹⁶. Another approach would be to identify the epitope that elicits the immune response that occurs in disease to find the exact causal HLA molecule. For example, in coeliac disease, the gluten epitope was found to bind to the HLA-DQ2 molecule, indicating that the *HLA-DQA1*0501* and *HLA-DQB1*0201* alleles were causal variants.

Implications of a shared genetic background

In this Review, we have brought together data from GWA studies and nsSNP scans to identify shared genetic components of different immune disorders and map them to common pathogenetic pathways. We have analysed data from 22 primary GWA studies, 6 nsSNP scans and 28 replication and meta-analysis papers for the 11 immune-related diseases we have focused on (information about these studies can be found in [Supplementary information S1](#) (table)). To search for the common genetic pathways in immune-related disorders, we used a stepwise approach (BOX 1). By applying stringent significance criteria, we identified 23 genes that are shared by two or more diseases ([Supplementary information S3](#) (figure)). Using pathway analysis tools, we were able to allocate most of the shared genes to a few key pathways. In summary, the genes shared between the 11 immune-related diseases were involved in three major immunological pathways (T-cell differentiation, immune-cell signalling and the innate immune response) or were members of two functional groups —one group consisted of genes that are shared between ulcerative colitis and Crohn's disease (inflammatory bowel disease), and the other group encoded cytokines and chemokines (TABLE 3).

Box 2 | Ascertainment bias

For a number of reasons, our selection of genes (described in BOX 1) was subject to ascertainment bias, partly owing to some general drawbacks of genome-wide association (GWA) studies:

- GWA studies generally identify only common genetic variants. Hence, the disease susceptibility alleles that have been identified so far are common in the general population (minor allele frequencies are usually > 5%), and show low penetrance and modest effect sizes (odds ratios are usually < 1.5), resulting in low predictive values (the common disease–common variant hypothesis). However, the architecture of complex diseases is expected to involve not only common variants with low penetrance, but also low-frequency variants ('rare variants') with high penetrance, as well as structural variants (insertion or deletion (indels) polymorphisms and copy number variants).
- Our gene selection from disease-associated regions that contain multiple genes in linkage disequilibrium (LD) with each other might be biased. The selection was driven by three factors: the function of the genes was already known; genetical genomics, for example, *ORMDL3* in asthma and Crohn's disease, *IL18RAP* (interleukin-18 receptor accessory protein) in coeliac disease, *SLC22A5* (solute carrier family 22, member 5) in the *IBD5* (inflammatory bowel disease 5) locus in Crohn's disease; or the presence of a non-synonymous SNP (nsSNP) that is known to result in altered gene or protein function.
- The non-random coverage of scans means that a subset of genomic regions are poorly covered in GWA studies — this is mostly true for copy number variable regions.
- Limited follow-up with respect to a number of SNPs might mean that the significance of some SNPs is not well characterized, but might also mean that more extended follow-up and meta-analyses could identify additional shared genes.
- Despite the large size of GWA studies, the results might not have sufficient power to detect moderate effects, as we observed for most of the associated genes.
- For some diseases (for example, asthma, ankylosing spondylitis and autoimmune thyroid disease), only a single GWA study or nsSNP scan has been performed so far, or only a limited number of genes showed association.
- Nearly all the studies we analysed were performed in Caucasian populations and therefore represent a minority of the human population.
- Many of the loci identified by GWA studies and nsSNP scans cannot be connected to certain pathways, partly because the gene function is unknown, or the LD structure is too dense to locate the causal gene, or the associated signal is located in a gene desert. Extensive genetic and functional studies are required to fully characterize the associated loci.

Common disease–common variant hypothesis

This states that many genetic variants that underlie complex diseases are common and are therefore susceptible to detection by population association studies. An alternative possibility is that the genetic contributions to complex diseases arise from many variants, all of which are rare.

Genetical genomics

An approach that brings together genetic analysis and gene expression studies by directly characterizing the genetic influence of gene expression.

 T_H1 cells

A subset of T-helper cells that produce interferon- γ (and other cytokines) and that activate macrophages.

 T_H17 cells

A subset of CD4⁺ T-helper cells that produce interleukin 17 (IL-17) and that are thought to be important in inflammatory and autoimmune diseases.

Regulatory T cells

(T_{reg} cells). A subset of CD4⁺ T-helper cells that suppresses or regulates effector T cells and other immune cells. The absence or presence of dysfunctional T_{reg} cells are associated with severe autoimmunity.

Many of the disease-specific genes identified in the studies that we analysed also belong to the three immunological pathways outlined above. Strikingly, at least one associated gene was involved in each pathway for almost all of the immune diseases that were studied, which highlights the overlap in these immunological pathways between diseases (FIG. 1). However, some of the disease-specific genes cannot be mapped to a shared pathway and these genes might indicate new disease-specific pathways. Further information on the shared genes and the categorization of the pathways is given in [Supplementary information S4](#) (table). In the next sections, we focus on the three major immunological pathways that we identified as being shared between different immune-related diseases and highlight the role of these pathways in disease pathogenesis.

Genes involved in T-cell differentiation

T-helper cells. We observed that many genes identified by GWA studies were involved in T-cell differentiation. The T-cell differentiation pathway contains the largest number of shared immune-related genes that were identified in this study. Recent immunological studies have also shown that many immune-related diseases are not only characterized by high numbers of T cells, but also by an imbalance in T-cell subsets. This imbalance might be explained to some extent by a genetic effect on T-cell differentiation. Most of the genes that were identified are specifically involved in the differentiation of T_H1 cells, T_H17 cells and regulatory T cells (T_{reg} cells).

T_H1 cells have long been implicated in many inflammatory diseases and the involvement of T_H1 genes in disease is now emerging, as seen by the association of *IL18RAP* (interleukin-18 receptor accessory protein), *IL12*, *IL10*, *STAT3* (signal transducer and activator of transcription 3) and *STAT4* (FIG. 2) with almost all of the immune-related diseases we analysed. T_H17 cells were identified as a separate T-helper lineage, but they are now thought to be more important than T_H1 cells as mediators in immune-related diseases¹⁷. Chronically inflamed tissues (for example, in Crohn's disease, rheumatoid arthritis, asthma and psoriasis) are infiltrated with highly differentiated T_H17 cells and, in particular, genes associated with T_H17 cells (for example, *IL23R* and *IL21*) ([Supplementary information S3](#) (figure)) are associated with nearly all immune-related diseases. Intriguingly, one allele of *IL23R* has been associated with Crohn's disease and ulcerative colitis, and the other allele has been associated with psoriasis^{18–20}. This association might be due to the effect of different causal variants on different haplotypes or a biological difference in the function of IL-23 and IL-23R signalling in these different diseases. For example, IL-23 stimulates survival and proliferation of T_H17 cells, and thus serves as a key master cytokine regulator; the causal genetic variant of *IL23R* can cause up or downregulation of IL-23 signalling and thus affect T_H17 cells. The effect of the *IL23R* variant on T_H17 cells can be either beneficial or harmful, depending on the disease. However, this model is speculative until the causal variant is

identified. Thus far, the GWA data have indicated that the dysregulation of T_H17-cell differentiation in patients with immune disorders can be influenced by genetics.

Regulatory T cells. There is increasing evidence that T_{reg} cells are less active in chronic immune-related diseases^{21,22}. *IL-2* and its receptor (encoded by *IL2RA*, *IL2RB* and *IL2RG*) are crucial in the activation and

Table 3 | Immune-disease associated genes grouped by pathways and categories

Disease	T-cell differentiation	Immune-cell activation, signalling	Innate immunity and TNF signalling	Other categories			Refs
				Other cytokines or chemokines	Inflammatory bowel disease shared	Disease-specific or shared but with unknown function	
Ankylosing spondylitis	<i>IL23R*</i>	None	None	None	None	<i>ERAP1</i>	27
Asthma	None	None	<i>TRAF1–C5*</i>	None	None	<i>ORMDL3**</i>	106,107
Autoimmune thyroid disease	None	<i>CTLA4*</i>	None	None	None	<i>TSHR</i>	9,27,39
Coeliac disease	<i>IL2–IL21*</i> , <i>IL18RAP*</i> , <i>IL12A</i>	<i>SH2B3*</i> , <i>TAGAP</i>	<i>TNFAIP3*</i>	<i>CCR1–3</i>	None	<i>LPP</i> , <i>RGS1</i>	30,38, G. Trynka, A.Z. and C.W., unpublished observations
Crohn's disease	<i>IL12B*</i> , <i>IL18RAP*</i> , <i>IL23R*</i> , <i>STAT3</i>	<i>PTPN2*</i> , <i>PTPN22*</i> , <i>ICOSLG</i> , <i>JAK2</i>	<i>TNFRSF6B*</i> , <i>IRF5</i> , <i>TNFSF15</i> , <i>IRGM</i> , <i>LRRK2–MUC19</i> , <i>NOD2</i> , <i>ATG16L1</i>	<i>CCR6</i>	<i>MST1–BSN*</i> , <i>NKX2–3*</i> , <i>PSMG1*</i>	<i>ORMDL3**</i> , <i>CCNY</i> , <i>CDKAL1</i> , <i>CUL2</i> , <i>GCKR</i> , <i>IBD5 (SLC22A5)</i> , <i>ITLN1</i> , <i>KIF21B</i> , <i>SBNO2</i> , <i>ZNF365</i> , <i>1q24.3</i> , <i>1q31.2</i> , <i>5p13 (PTGER4)</i> , <i>5q23</i> , <i>6q21</i> , <i>7p12</i> , <i>8q24</i> , <i>11q13</i> , <i>13q14</i> , <i>15q21.3</i> , <i>17q11.1</i> , <i>17q23.2</i> , <i>21q21</i>	18,26,108–119
Multiple sclerosis	<i>IL2RA*</i>	None	None	<i>IL7R*</i>	None	<i>CLEC16A**</i>	27,76
Psoriasis	<i>IL12B*</i> , <i>IL23R*</i>	None	β-defensin	None	None	<i>ZNF313</i>	19,90,120,121
Rheumatoid arthritis	<i>STAT4*</i>	<i>PTPN22*</i> , <i>CD40</i> , <i>PRKCCQ</i> , <i>PIP4K2C</i>	<i>IRF5*</i> , <i>TNFAIP3*</i> , <i>TRAF1–C5*</i> , <i>TNFRSF14</i>	<i>CCL21</i>	None	None	26,49,122–129
Systemic lupus erythematosus	<i>STAT4*</i>	<i>PTPN22*</i> , <i>TNFSF4</i> , B-cell specific: <i>BANK1</i> , <i>BLK</i> , <i>LYN</i>	<i>IRF5*</i> , <i>TNFAIP3*</i> , <i>ITGAM</i> , <i>UBE2L3</i> , <i>ATG5</i>	None	None	<i>ICA1</i> , <i>PHRF1</i> , <i>NMNAT2</i> , <i>PXK</i> , <i>SCUBE1</i> , <i>1q25.1</i> , <i>5q33.3</i>	37,123,130–135
Type 1 diabetes	<i>IL2–IL21*</i> , <i>IL2RA*</i>	<i>CTLA4*</i> , <i>PTPN2*</i> , <i>PTPN22*</i> , <i>SH2B3*</i> , <i>CD226</i>	<i>IFIH1</i>	<i>IL7R*</i>	None	<i>CLEC16A**</i> , <i>COL1A2</i> , <i>ERBB3</i> , <i>INS</i> , <i>LPHN2</i> , <i>NRP1</i>	9,26,28,29, 31,136–138
Ulcerative colitis	<i>IL10</i>	None	<i>IRF5*</i> , <i>TNFRSF6B*</i>	None	<i>MST1–BSN*</i> , <i>NKX2–3*</i> , <i>PSMG1*</i>	<i>ECM1</i>	20,92,110, 118,119,140

Genes identified by genome-wide association (GWA) studies and nonsynonymous SNP (nsSNP) scans as being associated with immune-related diseases were selected by the method described in BOX 1. Pathway analyses grouped some of these genes into three immune pathways: T-cell differentiation; immune-cell activation and signalling; innate immunity and tumour necrosis factor (TNF) signalling. Additional genes were placed in other categories: other cytokines or chemokines; inflammatory bowel disease shared; and disease-specific or shared but with unknown function. *The 23 genes that are directly shared between the diseases that were used to identify the shared pathways. This grouping of genes is also illustrated in FIG. 1. †These genes overlap in two diseases, but little is known about their function so they could not be placed in shared pathways. *ATG*, autophagy-related; *BANK1*, B-cell scaffold with ankyrin repeats; *BLK*, B-lymphoid tyrosine kinase; *BSN*, bassoon; *C5*, complement component 5; *CCL21*, chemokine (C-C motif) ligand 21; *CCNY*, cyclin Y; *CCR*, chemokine (C-C motif) receptor; *CDKAL1*, cyclin-dependent kinase 5 regulatory subunit associated 1-like 1; *CLEC16A*, C-type lectin domain family 16; *COL1A2*, collagen 1α2; *CTLA4*, cytotoxic T-lymphocyte-associated protein 4; *CUL2*, cullin 2; *ECM1*, extracellular matrix 1; *ERAP1*, endoplasmic reticulum aminopeptidase 1; *GCKR*, glucokinase regulator; *IBD5*, inflammatory bowel disease 5; *ICA1*, islet-cell autoantigen 1; *ICOSLG*, inducible T-cell co-stimulator ligand; *IFIH1*, interferon-induced with helicase C domain; *IL18RAP*, interleukin 18 receptor accessory protein; *INS*, insulin; *IRF5*, interferon regulatory factor 5; *IRGM*, immunity-related GTPase; *ITGAM*, integrin-αM; *ITLN1*, intelectin 1; *JAK2*, Janus kinase 2; *KIF21B*, kinesin family member 21B; *LPHN2*, latrophilin 2; *LPP*, LIM-domain-containing preferred translocation partner in lipoma; *LRRK2*, leucine-rich repeat kinase 2; *MST1*, macrophage stimulating 1; *MUC19*, mucin 19; *NKX2-3*, NK2 transcription factor-related, locus 3; *NMNAT2*, nicotinamide nucleotide adenyltransferase 2; *NOD2*, nucleotide-binding oligomerization domain containing 2; *NRP1*, neuropilin 1; *PHRF1*, PHD and ring finger domains 1; *PIP4K2C*, phosphatidylinositol-5-phosphate 4-kinase type 2γ; *PRKCCQ*, protein kinase Cγ; *PSMG1*, proteasome assembly chaperone 1; *PTGER4*, prostaglandin E receptor 4; *PTPN*, protein tyrosine phosphatase non-receptor; *PXK*, PX domain-containing kinase; *RGS1*, regulator of G-protein signalling 1; *SBNO2*, strawberry notch homologue 2; *SCUBE1*, signal peptide, CUB domain, EGF-like 1; *STAT*, signal transducer and activator of transcription; *TAGAP*, T-cell activation RhoGTPase activating; *TNFAIP3*, tumour necrosis factor (TNF), α-induced 3; *TNFRSF6B*, TNF receptor superfamily, member 6b; *TNFSF*, TNF superfamily; *TRAF1*, TNF receptor-associated factor 1; *TSHR*, thyroid stimulating hormone receptor; *UBE2L3*, ubiquitin-conjugating enzyme E2L3; *ZNF*, zinc finger.

function of T_{reg} cells, and in the prevention of development of autoimmunity. *IL2*^{-/-} mice develop autoimmunity, which is most prominently seen as ulcerative colitis, and a combined *IL2RA* and *CTLA4* deficiency in mice causes severe autoimmune disease^{23,24}. Similarly, *CTLA4* blockade or IL-2 neutralization for a short period elicits T cell-mediated autoimmune reactions in mice. Genetic *IL2* variants in mice were shown to alter the functions of T_{reg} cells and predispose to autoimmune disease²⁵. Two genes in the T_{reg} activation cascade have now been associated with multiple autoimmune diseases: *IL2RA* (also known as *CD25*) and the locus that includes *IL2* and *IL21* (REFS 26–31). Although the presence of strong LD makes it difficult to locate the true associated gene, both *IL2* and *IL21* are plausible candidates, as they promote differentiation and activation of T_{reg} and T_H17 cells, respectively. A deficiency in IL-2 signalling might result in a lower number of T_{reg} cells or less activated T_{reg} cells, which might lead to loss of tolerance and onset of autoimmunity. High or dysregulated IL-21 levels might potentially lead to autoimmune disorders that result from antibody overproduction. For example, high levels of IL-21 have been found in the guts of patients with coeliac disease and Crohn's disease^{32–34}.

The association of T-cell differentiation pathway genes with multiple inflammatory diseases has added to the knowledge of the functional roles of T_H1, T_H17 and T_{reg} molecules in autoimmunity. The combination of genetic associations and the molecular profiles in inflamed tissues should help to highlight targets for anti-inflammatory drug development.

Immune-cell signalling

Genetic links between T-cell signalling and immune disorders, including T1D, coeliac disease and systemic lupus erythematosus (SLE), were known before GWA studies were performed, because associations with HLA, *CTLA4* and *PTPN22* had been previously described. The functional role of signal transduction molecules in the development of autoimmunity had also been previously demonstrated in animal models^{35,36}. GWA studies have now enabled the identification of new, shared genetic components in immune disorders.

Four genes that are primarily involved in immunological signalling (that encode the co-stimulation molecule *CTLA4*, the protein tyrosine phosphatases *PTPN22* and *PTPN2*, and the adaptor protein *SH2B3*) are shared between two or more autoimmune diseases^{9,18,29,37–39}. In addition, several disease-specific associated genes are also involved in the same immunological pathway, these genes include: *TAGAP* (T cell-specific Rho-GTPase activation protein); co-stimulation molecules *ICOSLG* (inducible T-cell co-stimulator ligand), *TNFSF4* (tumour necrosis factor superfamily member 4) and integrin signalling molecule *CD226*; and other molecules involved in immune cell signalling, such as *PRKCCQ* (protein kinase C τ), *PIP4K2C* (phosphatidylinositol-5-phosphate 4-kinase type-2 γ), *JAK2* (Janus kinase 2) and *TNFSF5* (TNF superfamily member 5; also known as *CD40*) (TABLE 3).

T-cell signalling. The role of two genes that are shared between immune diseases, *PTPN22* and *CTLA4*, had already been extensively studied using candidate gene approaches before their role in autoimmunity was confirmed by GWA studies^{40,41}. Another gene that is shared between immune disorders, *PTPN2*, plays an important part in the negative regulation of the inflammatory response in T cells⁴². The *SH2B3* (also known as *LNK*) gene is another interesting locus that is shared between immune diseases, as it regulates T-cell receptor signalling, and growth factor and cytokine receptor-mediated signalling (including the TNF α pathway), which are implicated in leukocyte and myeloid-cell homeostasis^{43,44}. The same *SH2B3* Arg262Trp allele is associated with both coeliac disease and T1D^{29,30}. However, understanding the role of shared genes in immune signalling pathways is complicated by the discovery of association of different alleles of the same gene to different immune-related disorders. A well-known example is *PTPN22*, for which the functional Arg602Trp allele is a strong risk factor for T1D and rheumatoid arthritis, but it also confers protection against Crohn's disease. The *PTPN22* molecule is involved in the activation of both naive and effector T cells. Although the biology underlying this observation is unknown, insufficient stimulation of naive T cells might lead to breakdown of immunological tolerance, but overactivation of effector T cells might increase the inflammatory response.

B-cell signalling. Interestingly, a disease-specific association with the signalling molecules that are primarily expressed on B cells was observed for SLE. The subset

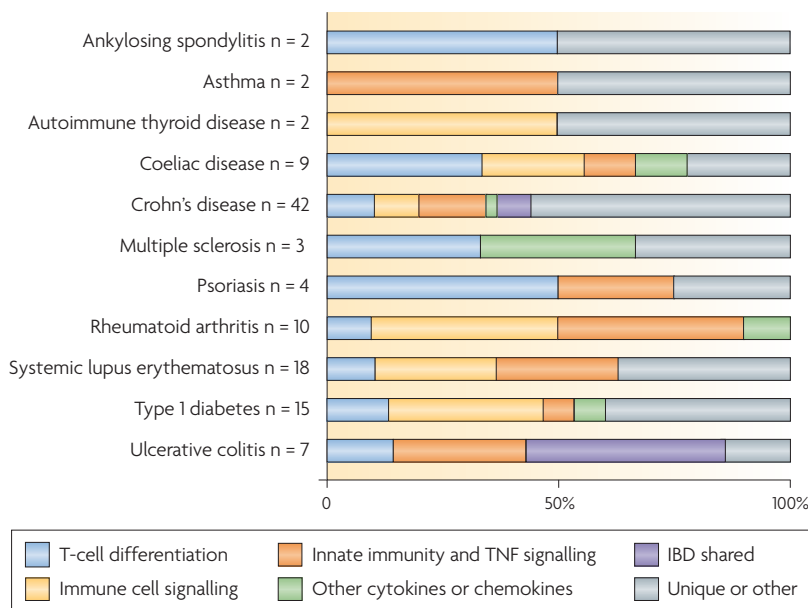


Figure 1 | The proportion of disease-associated genes in different pathways or categories. The relative proportion of genes associated with each immune-related disease that map to each of the pathways or categories shown in TABLE 3 is illustrated schematically in this figure. The colour of each block refers to a pathway or category, and the size of the block reflects the proportion of genes that belong to that pathway. IBD, inflammatory bowel disease; TNF, tumour necrosis factor.

of associated genes that is specific for B-cell activation includes the B-cell kinases LYN and BLK, and the scaffold protein BANK1 (B-cell scaffold protein with ankyrin repeats 1). This association could reflect the primary

pathogenic role of auto-antibodies in lupus development⁴⁵. Auto-antibodies to dsDNA are highly specific for lupus (they are present in 70% of patients) and correlate with disease activity⁴⁶. The auto-antibodies found

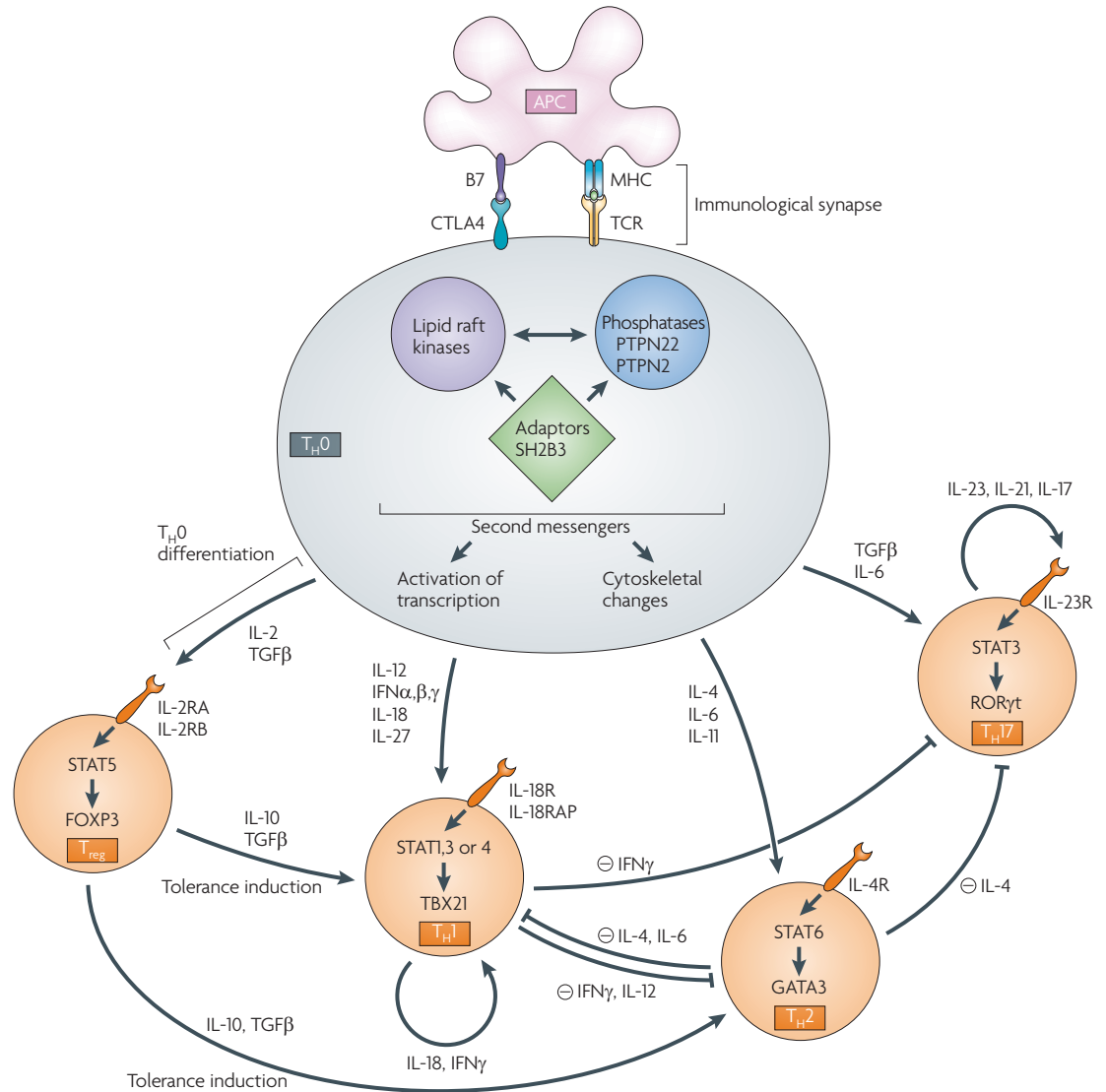


Figure 2 | T-cell differentiation and signalling. T-cell signalling requires the formation of a contact interface between an antigen-presenting cell (APC) and a T cell, known as the immunological synapse. The key event in T-cell stimulation is the activation of the T-cell receptor (TCR) by antigenic peptides presented by major histocompatibility complex (MHC) molecules on the APC. Other factors required for the formation of the stable immunological synapse and the proper activation of T cells include: the activation of co-stimulatory molecules, such as CTLA4 (cytotoxic T-lymphocyte-associated protein 4) by their ligands (B7); activation of integrin cascades; and co-stimulation by various cytokines via cytokine receptors (not shown). The sequence of further T-cell activation involves multiple kinases, phosphatases, adaptor proteins and effector enzymes, and leads to the synthesis and activation of transcription factors, and cytoskeleton reorganization. These processes maintain the effector functions of T cells, such as cytokine synthesis, cellular proliferation and differentiation. The type of antigen that elicits the immune response (and thus the desired type of immune response) and the type of cytokines secreted by APCs determine whether T helper (T_H) 0 cells differentiate into effector T_H1 , T_H2 , T_H17 or regulatory T (T_{reg}) cells. Effector T_H cells determine the type of immune response by the production of specific cytokines, which can also downregulate other types of effector T_H cells. The process of T-cell activation is illustrated using T_H0 cells, but this process occurs similarly in other T-cell types. FOXP3, forkhead box P3; GATA3, GATA-binding factor 3; IFN, interferon; IL-2RA, interleukin 2 receptor α -subunit; IL-2RB, interleukin 2 receptor β -subunit; IL-18RAP, interleukin 18 receptor-associated protein; PTPN, protein tyrosine phosphatase; ROR γ t, retanoid-related orphan receptor- γ ; STAT, signal transducer and activator of transcription; TBX21, T-box transcription factor 21; TGF β , transforming growth factor- β .

in lupus occur in less than 0.5% of healthy individuals. The presence of auto-antibodies in other autoimmune diseases (such as auto-antibodies to glutamic acid decarboxylase (GAD) and insulinoma-associated 2 (IA2) in T1D and anti-tissue transglutaminase (TTG) antibodies in coeliac disease) occurs more frequently in healthy individuals and might occur as a consequence of the disease rather than being the cause⁴⁷. A role for genetic factors in auto-antibody seropositivity has now started to emerge. For example, anti-citrulline-peptide antibody (ACPA)-positive patients with rheumatoid arthritis have a strong association with HLA-DR4, but ACPA-negative patients show association to HLA-DR3 (REF 48). Consistent with these associations, TNFSF5 has been mainly associated with an auto-antibody-positive subgroup of patients with rheumatoid arthritis⁴⁹. TNFSF5 is expressed primarily on B cells and has a role in B-cell activation, proliferation and antibody secretion^{50,51}. In some autoimmune diseases, auto-antibodies are present years before clinical symptoms⁵². Hence, auto-antibody positivity might represent sub-phenotypes within autoimmune diseases, and differential genetic associations in auto-antibody-positive and -negative groups could help us to understand differences in pathogenesis and disease development.

Innate immunity and TNF signalling

The functional role of innate immunity in autoimmune disorders is widely recognized. However, before GWA studies, the only convincing evidence for genetic association of the innate immunity pathway with inflammatory disorders was *NOD2* in Crohn's disease and *IRF5* (interferon regulatory factor 5) in SLE. This picture has changed dramatically with GWA results that have highlighted an assortment of shared and specific innate genes that are involved in the pathogenesis of immune-related diseases. So far, associations to more than 10 genes involved in different key mechanisms of innate immunity have been established (TABLE 3). Nearly all of the described diseases are associated with TNF-receptor and ligand genes, and two of these genes, the *C5-TRAF1* locus on chromosome 9q34 and the *OLIG3-TNFAIP3* locus on chromosome 6q23.3, are directly shared by two or more diseases. The two other important innate pathways — barrier function and autophagy — seem to show more disease-specific associations. The roles of some of the genes involved in these two pathways and in TNF signalling, and the insights they give into the underlying disease processes, are described below.

Barrier function. The epithelial layer forms the first line of defence against microorganisms by producing antimicrobial peptides (defensins and cryptocidins), forming a mucus layer that protects the intestinal epithelium from injury. Genetic findings that are involved in this barrier function include the association of the *NOD2* gene and the *MUC19* (mucin 19)-containing locus to Crohn's disease^{3,18}. The involvement of mucins in inflammatory bowel disease is further supported by evidence that deficiency in another component of the mucosal mucus layer, *MUC2*, leads to colonic inflammation

and predisposition to experimental colitis in mice^{53,54,55}. Interestingly, there is evidence that suggests there is an association of tight junction-complex genes, which are involved in intestinal permeability, with coeliac disease and ulcerative colitis⁵⁶. Another barrier risk factor is copy number variation (CNV) in defensin genes, and candidate gene studies have associated defensin CNV with psoriasis and a colonic form of Crohn's disease^{57,58}. These results have not yet been supported by GWA studies, probably owing to the poor coverage of the defensin gene cluster by current genotyping platforms. Overall, genetic defects in barrier molecules seem to be specific for barrier diseases^{12,59}, such as Crohn's disease and psoriasis, rather than for classical autoimmune diseases.

Autophagy. Three Crohn's disease genes, *ATG16L1* (autophagy 16 related-like 1), *IRGM* (immunity-related GTPase M) and *LRRK2* (leucine-rich repeat kinase 2), are involved in autophagy, which is unsurprising because this mechanism is known to be involved in Crohn's disease. This genetic association underscores the role of intracellular processing of bacteria in disease pathogenesis^{3,60}. More interestingly, SLE is associated to another key autophagy molecule, *ATG5* (REF. 37). The functional importance of this finding might lie in the essential part that *ATG5* plays in the autophagy process in plasmacytoid dendritic cells (pDCs)⁶¹. pDCs are involved in the recognition of viral pathogens, and pDCs from *ATG5* knockout mice produce less interferon- α (IFN α) in response to viral infections⁶¹. In addition, *IRF5* and *IFIH1* (interferon induced with helicase C domain; also known as *MDA5*) are also involved in the IFN α pathway, and are associated with several autoimmune and inflammatory diseases. Type 1 interferons mediate the early innate immune response to viral infections and have already been clinically implicated in SLE in two ways: first, by a correlation between IFN α serum levels and disease activity, and second, by the development of lupus symptoms in individuals who were treated with IFN α for other disorders⁶²⁻⁶⁴. The shared genetic associations of the IFN α and autophagy pathways underscore a role for viral and bacterial triggers in immune-related diseases⁶⁵.

TNF signalling. Interestingly, GWA studies now provide evidence that genes that are either involved in TNF signalling or that are activated via TNF have a role in immune-related diseases. As TNF signalling links the innate and adaptive immune systems, it could help to explain the involvement of both systems in autoimmune diseases. For example, the cytokine encoded by *TNFSF15*, which is associated with Crohn's disease, is activated via stimulation by lipopolysaccharides. It was therefore suggested that *TNFSF15* plays a part in the innate immune response⁶⁶; *TNFSF15* has also been reported to enhance both T_H1- and T_H17-effector functions⁶⁷. *TNFAIP3*, which is associated with SLE, rheumatoid arthritis and coeliac disease, encodes the A20 protein, which is required for termination of the nuclear factor- κ B (NF- κ B) signal that is mediated by innate immune receptors⁶⁸. Genetic deficiency of A20 in mice leads to persistent activation of NF- κ B by Toll-like

Auto-antibody seropositivity
The presence of antibodies that are directed against one or more of an individual's own proteins.

Autophagy
A cellular process of degradation of cellular components that occurs by transporting the components to lysosomes. This process maintains a balance between the synthesis and degradation of cellular products, and is also involved in the degradation of intracellular pathogens.

receptors, resulting in multi-organ inflammation and neonatal lethality^{69,70}. The loss of A20 decreases the tolerance of the innate immune system to commensal intestinal microflora⁷¹.

GWA findings have thus greatly expanded our view of the role of innate pathways in autoimmunity. It has become clear that the genes that control innate immunity are not only associated with inflammatory barrier disorders, such as psoriasis and Crohn's disease, but also with classical autoimmune diseases. These findings provide genetic support for the hypothesis that microbial and viral pathogens are triggers of autoimmune diseases^{72–75}.

Other associated genes

Genes that were not grouped to immunological pathways shared between immune disorders include: genes shared between Crohn's disease and ulcerative colitis; genes that encode several chemokines and their receptors; shared genes with unknown function; and disease-specific genes.

Several genes were only shared between Crohn's disease and ulcerative colitis, which are the two main types of inflammatory bowel disease. These diseases have many overlapping clinical features and co-occur in families; it is therefore conceivable that the genes that are shared between these two entities of inflammatory bowel diseases reflect shared pathogenesis.

Associations with genes that encode several cytokines and chemokines and their receptors have been reported for a few immune-related diseases: *CCL21* (chemokine (C-C motif) ligand 21) is associated with rheumatoid arthritis, *CCR6* (chemokine receptor 6) is associated with Crohn's disease, and the *CCR1–CCR3* locus is associated with coeliac disease. The major role of chemokines is to regulate the immune response and recruit effector immune cells to sites of inflammation. Chemokine genes usually form a cluster with strong LD between genes; for example, SNP rs6441961, which is associated with coeliac disease, is in LD with several chemokine genes, including *CCR1*, *CCR2*, *CCRL2*, *CCR3*, *CCR5* and *CCXCR1*. This clustering, together with the broad and overlapping functions of various cytokines, makes it difficult to determine the exact nature of the association of different chemokines with various autoimmune diseases. The *IL7R* gene, which is associated with T1D and multiple sclerosis, has an important role in regulating the adaptive immune system by maintaining sufficient numbers of effector and memory T cells^{29,76,77}.

CLEC16A (C-type lectin domain 16; also known as *KIAA0350*) and *ORMDL3* are two genes that are shared between immune disorders and have not been studied in detail, therefore their functional role in disease is unknown. *CLEC16A* is associated with T1D and multiple sclerosis and encodes a C-type lectin receptor, which might play a part in antigen sampling by dendritic cells⁷⁸. The *ORMDL3* gene is associated with asthma and Crohn's disease, is located in an extended LD block, and was prioritized for study owing to the strong *cis*-correlation between *ORMDL3* expression and its associated genotype. Further studies are required to establish the role of these shared genes in immunity.

The GWA studies and nsSNP scans also identified genes that are specifically associated with just one of the immune diseases analysed. Further characterization of these genes could help us to understand why individuals develop a specific immune disease. Two well-described disease-specific genes that have now been confirmed by GWA studies are *INS* (insulin) in T1D and *TSHR* (thyroid-stimulating hormone receptor) in autoimmune thyroid disease. Both genes are involved in pathways that are highly specific to T1D and autoimmune thyroid disease: *INS* modulates insulin expression in the thymus⁷⁹, and *TSHR* encodes a specific receptor for thyroid-stimulating hormone. New disease-specific genes that are identified by GWA studies could explain undiscovered aspects of disease pathogenesis. For example, *LPP* (LIM domain-containing preferred translocation partner in lipoma), which is associated with coeliac disease, might have a structural role in the intestine, and the association of *ERAP1* (endoplasmic reticulum aminopeptidase 1; also known as *ARTS1*) with ankylosing spondylitis implicates peptide processing during presentation to MHC class I molecules²⁷ in disease pathogenesis.

Discussion and future perspectives

The genes for immune-related diseases that have been identified by GWA studies thus far show that the general immune pathways of T-cell differentiation, T-cell signalling and the innate immune response are shared between some of these diseases. These genetic results confirm the clinical observations regarding the co-morbidity of immune-related diseases, but they also provide further information, as they now indicate why these immune diseases form clusters. Immunity — both the adaptive and innate pathways — lies at the heart of these diseases. Innate immunity is particularly interesting from a clinical perspective, as it provides links to environmental triggers of disease and might provide new tools for disease prevention. Vaccination is one such potential tool, as associations with bacterial and viral infections have been suggested for most of the diseases discussed in this Review. For example, a positive correlation between the presence of multiple sclerosis or SLE and high titres of auto-antibodies against Epstein–Barr virus has been reported^{65,73}. An increased frequency of rotaviral infections is associated with coeliac disease⁸⁰, and recent studies in mice strongly suggest that intestinal microbiota play a part in the development of T1D⁷⁵. Defining the profile of the genetic susceptibility pathway, together with knowledge of environmental triggers, might help prevent immune-related diseases through the vaccination of high-risk individuals or families.

GWA studies have highlighted the importance of previously unrecognized pathways, such as autophagy in Crohn's disease and SLE, or genes from the TNF α pathway in multiple immune-related diseases. Novel pathways could lead to the development of new strategies for treatment of immune-related disorders. For example, anti-TNF medicines are currently used for treating rheumatoid arthritis and Crohn's disease, but this treatment is invasive and its effectiveness varies between individuals. Moreover, a subset of patients with

LD block

A segment of DNA with markers that are in linkage disequilibrium (LD) with each other.

Prospective epidemiological studies

A research study that, over a period of time, follows groups of individuals who are alike in many ways but differ by certain characteristics, and compares them for a particular outcome.

rheumatoid arthritis that was treated with anti-TNF antibodies develops anti-dsDNA antibodies and SLE⁸¹. Genetic variances in the TNF α profiles of patients might explain these phenomena and thus help to design more effective treatments. A further example is provided by a new compound, STA-5326, that downregulates the p35 subunit of IL-12 and the p40 subunit of IL-12 and IL-23 at the transcriptional level, and inhibits the production of both IL-12 and IL-23 cytokines. STA-5326 is currently being investigated as a therapy in Crohn's disease and rheumatoid arthritis⁸².

An important question that needs to be raised is: how do we proceed from here? Prospective epidemiological studies based on large population-based biobanks allow us to follow individuals with certain genetic risks over a long period and to monitor their environmental exposures and disease outcome. These prospective studies will be instrumental for assessing the contributions of both genetic and environmental factors to disease risk, and for determining the relationships between these factors. In particular, prospective studies could help to establish which environmental triggers elicit disease. From such studies, we could also learn about the order and timing of the development of multiple immune disorders in the same individual. Another important aspect of immune disease that the epidemiological studies could help to elucidate is whether the presence of auto-antibodies predisposes to autoimmune disease, or whether auto-antibody production is a consequence of disease. Auto-antibodies are produced early in some immune diseases before the clinical symptoms appear; however, auto-antibodies are also found in healthy individuals⁴⁶. Epidemiological studies will also be instrumental in relating the presence of auto-antibodies to predisposing alleles⁵².

A crucial aspect of assessing the value of recent genetic findings is examining the possible benefits to patients. Can we use the knowledge of shared genetic pathways for predicting disease development, and can these findings be useful in clinical practice? The use of genetic findings for predicting progression rather than onset of disease

has been reported for Crohn's disease⁸³, although the current prediction models are based on disease populations, which overestimate the risk. Prospective cohort studies might solve this problem. Moreover, using information from shared pathways might help predict a general immune-related profile, rather than a disease-specific profile. Predictions relating to a specific immune pathway could provide a practical tool for planning therapeutic intervention. The accuracy of genetic profiling might be improved further as more genetic factors are characterized and the gene-environmental contribution to disease becomes clearer.

After we understand the different molecular pathways that underlie immune-related diseases, these pathways might provide targets for therapeutic intervention. If the genetic profile of a patient can be correlated to the effectiveness of a specific therapy, this would open up new avenues in clinical trials. The genetic predisposition to impairment of a certain pathway might help to define clinical subgroups of disease and prioritize patient groups for a specific therapy. If the pathways rather than the genes need to be targeted, we will soon have excellent starting points for the development of therapeutic targets for many immune diseases and the molecular phenotype of the patient will then become more relevant than the clinical phenotype. In addition, patients with different diseases but overlapping pathways might benefit from the same treatment, which could lead to the development of new and shared therapeutic prospects.

The genetic study of immune-related diseases has only revealed the tip of the iceberg thus far, as more genes need to be found and the causal variants need to be identified. Nevertheless, the notion of shared genetic pathways creates new and powerful approaches for discovering the full repertoire of susceptibility genes because genetic resources can be shared instead of focusing on single diseases. For example, performing meta-analyses across multiple immune-related diseases might be a powerful method for discovering more of the common genetic variation that contributes to the pathways that are shared in immune disorders.

- Davies, A. J. Immunological tolerance and the autoimmune response. *Autoimmun. Rev.* **7**, 538–543 (2008).
- Broide, D. New perspectives on mechanisms underlying chronic allergic inflammation and asthma in 2007. *J. Allergy Clin. Immunol.* **122**, 475–480 (2008).
- Cho, J. H. The genetics and immunopathogenesis of inflammatory bowel disease. *Nature Rev. Immunol.* **8**, 458–466 (2008).
This paper gives an up-to-date summary of genetic findings in Crohn's disease, and discusses their biological relevance and applications.
- Somers, E. C., Thomas, S. L., Smeeth, L. & Hall, A. J. Autoimmune diseases co-occurring within individuals and within families: a systematic review. *Epidemiology* **17**, 202–217 (2006).
- Barera, G. *et al.* Occurrence of celiac disease after onset of type 1 diabetes: a 6-year prospective longitudinal study. *Pediatrics* **109**, 833–838 (2002).
- Xavier, R. J. & Rioux, J. D. Genome-wide association studies: a new window into immune-mediated diseases. *Nature Rev. Immunol.* **8**, 631–643 (2008).
This review discusses the results of GWA studies in several immune-related diseases, their impact on understanding disease pathogenesis and future prospectives.
- Fernando, M. M. *et al.* Defining the role of the MHC in autoimmunity: a review and pooled analysis. *PLoS Genet.* **4**, e1000024 (2008).
This paper gives a comprehensive summary of the literature on HLA association with six immune-mediated diseases.
- Duffy, D. L. Genetic determinants of diabetes are similarly associated with other immune-mediated diseases. *Curr. Opin. Allergy Clin. Immunol.* **7**, 468–474 (2007).
- Ueda, H. *et al.* Association of the T-cell regulatory gene *CTLA4* with susceptibility to autoimmune disease. *Nature* **423**, 506–511 (2003).
- Becker, K. G. *et al.* Clustering of non-major histocompatibility complex susceptibility candidate loci in human autoimmune diseases. *Proc. Natl Acad. Sci. USA* **95**, 9979–9984 (1998).
- Becker, K. G. The common variants/multiple disease hypothesis of common complex genetic disorders. *Med. Hypotheses* **62**, 309–317 (2004).
- Schreiber, S., Rosenstiel, P., Albrecht, M., Hampe, J. & Krawczak, M. Genetics of Crohn disease, an archetypal inflammatory barrier disease. *Nature Rev. Genet.* **6**, 376–388 (2005).
- Frazer, K. A. *et al.* A second generation human haplotype map of over 3.1 million SNPs. *Nature* **449**, 851–861 (2007).
- This paper describes the current knowledge of human genome variations (the HapMap). This knowledge is essential for designing commercial genotyping platforms and understanding their advantages and limitations.**
- Manolio, T. A., Brooks, L. D. & Collins, F. S. A HapMap harvest of insights into the genetics of common disease. *J. Clin. Invest* **118**, 1590–1605 (2008).
This paper describes the achievements of the HapMap project and the practical implications for complex diseases, including the summary results of GWA studies for ~40 diseases and traits.
- Alper, C. A. *et al.* The haplotype structure of the human major histocompatibility complex. *Hum. Immunol.* **67**, 73–84 (2006).
- Nejentsev, S. *et al.* Localization of type 1 diabetes susceptibility to the MHC class I genes *HLA-B* and *HLA-A*. *Nature* **450**, 887–892 (2007).
This study presents the comprehensive fine-mapping of the MHC region in type 1 diabetes.
- Bettelli, E., Korn, T., Oukka, M. & Kuchroo, V. K. Induction and effector functions of T_H17 cells. *Nature* **453**, 1051–1057 (2008).
This is an excellent review of the newly described T_H17 cell lineage, and the role of T_H17 cells in autoimmunity and inflammation.

18. Barrett, J. C. *et al.* Genome-wide association defines more than 30 distinct susceptibility loci for Crohn's disease. *Nature Genet.* **40**, 955–962 (2008).
19. Cargill, M. *et al.* A large-scale genetic association study confirms IL12B and leads to the identification of IL23R as psoriasis-risk genes. *Am. J. Hum. Genet.* **80**, 273–290 (2007).
20. Franke, A. *et al.* Sequence variants in *IL10*, *ARPC2* and multiple other loci contribute to ulcerative colitis susceptibility. *Nature Genet.* **40**, 1319–1323 (2008).
21. Brusko, T. M., Putnam, A. L. & Bluestone, J. A. Human regulatory T cells: role in autoimmune disease and therapeutic opportunities. *Immunol. Rev.* **223**, 371–390 (2008).
22. Sakaguchi, S., Yamaguchi, T., Nomura, T. & Ono, M. Regulatory T cells and immune tolerance. *Cell* **133**, 775–787 (2008).
23. Sadlack, B. *et al.* Generalized autoimmune disease in interleukin-2-deficient mice is triggered by an uncontrolled activation and proliferation of CD4+ T cells. *Eur. J. Immunol.* **25**, 3053–3059 (1995).
24. Sadlack, B. *et al.* Ulcerative colitis-like disease in mice with a disrupted interleukin-2 gene. *Cell* **75**, 253–261 (1993).
25. Yamanouchi, J. *et al.* Interleukin-2 gene variation impairs regulatory T cell function and causes autoimmunity. *Nature Genet.* **39**, 329–337 (2007).
26. The Wellcome Trust Case Control Consortium. Genome-wide association study of 14,000 cases of seven common diseases and 3000 shared controls. *Nature* **447**, 661–678 (2007).
This paper describes the most extended genome-wide scan performed so far, which includes 17,000 individuals and 7 common diseases. This publication describes several key methodological aspects of GWA studies.
27. Burton, P. R. *et al.* Association scan of 14,500 nonsynonymous SNPs in four diseases identifies autoimmunity variants. *Nature Genet.* **39**, 1329–1337 (2007).
28. Lowe, C. E. *et al.* Large-scale genetic fine mapping and genotype-phenotype associations implicate polymorphism in the IL2RA region in type 1 diabetes. *Nature Genet.* **39**, 1074–1082 (2007).
29. Todd, J. A. *et al.* Robust associations of four new chromosome regions from genome-wide analyses of type 1 diabetes. *Nature Genet.* **39**, 857–864 (2007).
30. van Heel, D. A. *et al.* A genome-wide association study for celiac disease identifies risk variants in the region harboring IL2 and IL21. *Nature Genet.* **39**, 827–829 (2007).
31. Vella, A. *et al.* Localization of a type 1 diabetes locus in the IL2RA/CD25 region by use of tag single-nucleotide polymorphisms. *Am. J. Hum. Genet.* **76**, 773–779 (2005).
32. Fina, D. *et al.* Interleukin-21 contributes to the mucosal T helper cell type 1 response in celiac disease. *Gut* **57**, 887–892 (2008).
33. Fina, D., Caruso, R., Pallone, F. & Monteleone, G. Interleukin-21 (IL-21) controls inflammatory pathways in the gut. *Endocr. Metab. Immune. Disord. Drug Targets.* **7**, 288–291 (2007).
34. Monteleone, G., Pallone, F. & MacDonald, T. T. Interleukin-21: a critical regulator of the balance between effector and regulatory T-cell responses. *Trends Immunol.* **29**, 290–294 (2008).
35. Jiang, H. & Chess, L. Regulation of immune responses by T cells. *N. Engl. J. Med.* **354**, 1166–1176 (2006).
36. Ohashi, P. S. T-cell signalling and autoimmunity: molecular mechanisms of disease. *Nature Rev. Immunol.* **2**, 427–438 (2002).
37. Harley, J. B. *et al.* Genome-wide association scan in women with systemic lupus erythematosus identifies susceptibility variants in *ITGAM*, *PXK*, *KIAA1542* and other loci. *Nature Genet.* **40**, 204–210 (2008).
38. Hunt, K. A. *et al.* Newly identified genetic risk variants for celiac disease related to the immune response. *Nature Genet.* **40**, 395–402 (2008).
39. Kavvoura, F. K. *et al.* Cytotoxic T-lymphocyte associated antigen 4 gene polymorphisms and autoimmune thyroid disease: a meta-analysis. *J. Clin. Endocrinol. Metab.* **92**, 3162–3170 (2007).
40. Anjos, S. & Polychronakos, C. Mechanisms of genetic susceptibility to type 1 diabetes: beyond HLA. *Mol. Genet. Metab.* **81**, 187–195 (2004).
41. Vang, T., Miletic, A. V., Bottini, N. & Mustelin, T. Protein tyrosine phosphatase PTPN22 in human autoimmunity. *Autoimmunity* **40**, 453–461 (2007).
42. Vang, T. *et al.* Protein tyrosine phosphatases in autoimmunity. *Annu. Rev. Immunol.* **26**, 29–55 (2008).
43. Fitau, J., Boulday, G., Coulon, F., Quillard, T. & Charreau, B. The adaptor molecule Lnk negatively regulates tumor necrosis factor- α -dependent VCAM-1 expression in endothelial cells through inhibition of the ERK1 and -2 pathways. *J. Biol. Chem.* **281**, 20148–20159 (2006).
44. Li, Y., He, X., Schembri-King, J., Jakes, S. & Hayashi, J. Cloning and characterization of human Lnk, an adaptor protein with pleckstrin homology and Src homology 2 domains that can inhibit T cell activation. *J. Immunol.* **164**, 5199–5206 (2000).
45. Isenberg, D. A., Manson, J. J., Ehrenstein, M. R. & Rahman, A. Fifty years of anti-DNA antibodies: are we approaching journey's end? *Rheumatology* **46**, 1052–1056 (2007).
46. Rahman, A. & Isenberg, D. A. Systemic lupus erythematosus. *N. Engl. J. Med.* **358**, 929–939 (2008).
47. Vojdani, A. Antibodies as predictors of complex autoimmune diseases. *Int. J. Immunopathol. Pharmacol.* **21**, 267–278 (2008).
48. van der Helm-van Mil A. H., Huizinga, T. W., de Vries, R. R. & Toes, R. E. Emerging patterns of risk factor make-up enable subclassification of rheumatoid arthritis. *Arthritis Rheum.* **56**, 1728–1735 (2007).
49. Raychaudhuri, S. *et al.* Common variants at *CD40* and other loci confer risk of rheumatoid arthritis. *Nature Genet.* **40**, 1216–1223 (2008).
50. Armitage, R. J. *et al.* CD40L: a multi-functional ligand. *Semin. Immunol.* **5**, 401–412 (1993).
51. Arpin, C. *et al.* Generation of memory B cells and plasma cells *in vitro*. *Science* **268**, 720–722 (1995).
52. Arbuckle, M. R. *et al.* Development of autoantibodies before the clinical onset of systemic lupus erythematosus. *N. Engl. J. Med.* **349**, 1526–1533 (2003).
53. Moehle, C. *et al.* Aberrant intestinal expression and allelic variants of mucin genes associated with inflammatory bowel disease. *J. Mol. Med.* **84**, 1055–1066 (2006).
54. Einerhand, A. W. *et al.* Role of mucins in inflammatory bowel disease: important lessons from experimental models. *Eur. J. Gastroenterol. Hepatol.* **14**, 757–765 (2002).
55. van der Sluis, S. M. *et al.* Muc2-deficient mice spontaneously develop colitis, indicating that MUC2 is critical for colonic protection. *Gastroenterology* **131**, 117–129 (2006).
56. Wapenaar, M. C. *et al.* Associations with tight junction genes *PAR3* and *MAGI2* in Dutch patients point to a common barrier defect for coeliac disease and ulcerative colitis. *Gut* **57**, 463–467 (2008).
57. Fellermann, K. *et al.* A chromosome 8 gene-cluster polymorphism with low human beta-defensin 2 gene copy number predisposes to Crohn disease of the colon. *Am. J. Hum. Genet.* **79**, 439–448 (2006).
58. Hollox, E. J. *et al.* Psoriasis is associated with increased β -defensin genomic copy number. *Nature Genet.* **40**, 23–25 (2008).
59. Bowcock, A. M. & Krueger, J. G. Getting under the skin: the immunogenetics of psoriasis. *Nature Rev. Immunol.* **5**, 699–711 (2005).
60. Levine, B. & Deretic, V. Unveiling the roles of autophagy in innate and adaptive immunity. *Nature Rev. Immunol.* **7**, 767–777 (2007).
61. Lee, H. K., Lund, J. M., Ramanathan, B., Mizushima, N. & Iwasaki, A. Autophagy-dependent viral recognition by plasmacytoid dendritic cells. *Science* **315**, 1398–1401 (2007).
62. Niewold, T. B. & Swedler, W. I. Systemic lupus erythematosus arising during interferon- α therapy for cryoglobulinemic vasculitis associated with hepatitis C. *Clin. Rheumatol.* **24**, 178–181 (2005).
63. Schmidt, K. N. & Ouyang, W. Targeting interferon- α : a promising approach for systemic lupus erythematosus therapy. *Lupus* **13**, 348–352 (2004).
64. Kim, T. *et al.* Serum levels of interferons in patients with systemic lupus erythematosus. *Clin. Exp. Immunol.* **70**, 562–569 (1987).
65. James, J. A. *et al.* An increased prevalence of Epstein-Barr virus infection in young patients suggests a possible etiology for systemic lupus erythematosus. *J. Clin. Invest* **100**, 3019–3026 (1997).
66. Yamazaki, K. *et al.* Single nucleotide polymorphisms in *TNFSF15* confer susceptibility to Crohn's disease. *Hum. Mol. Genet.* **14**, 3499–3506 (2005).
67. Takedatsu, H. *et al.* TL1A (TNFSF15) regulates the development of chronic colitis by modulating both T-helper 1 and T-helper 17 activation. *Gastroenterology* **135**, 552–567 (2008).
68. Sun, S. C. Deubiquitylation and regulation of the immune response. *Nature Rev. Immunol.* **8**, 501–511 (2008).
69. Boone, D. L. *et al.* The ubiquitin-modifying enzyme A20 is required for termination of Toll-like receptor responses. *Nature Immunol.* **5**, 1052–1060 (2004).
70. Lee, E. G. *et al.* Failure to regulate TNF-induced NF- κ B and cell death responses in A20-deficient mice. *Science* **289**, 2350–2354 (2000).
71. Turer, E. E. *et al.* Homeostatic MyD88-dependent signals cause lethal inflammation in the absence of A20. *J. Exp. Med.* **205**, 451–464 (2008).
72. Murphy, P. M. Viral exploitation and subversion of the immune system through chemokine mimicry. *Nature Immunol.* **2**, 116–122 (2001).
73. Ascherio, A. *et al.* Epstein-Barr virus antibodies and risk of multiple sclerosis: a prospective study. *JAMA* **286**, 3083–3088 (2001).
74. Lipton, H. L., Liang, Z., Hertzler, S. & Son, K. N. A specific viral cause of multiple sclerosis: one virus, one disease. *Ann. Neurol.* **61**, 514–523 (2007).
75. Wen, L. *et al.* Innate immunity and intestinal microbiota in the development of Type 1 diabetes. *Nature* **455**, 1109–1113 (2008).
76. Gregory, S. G. *et al.* Interleukin 7 receptor α chain (IL7R) shows allelic and functional association with multiple sclerosis. *Nature Genet.* **39**, 1083–1091 (2007).
77. Kallies, A. Distinct regulation of effector and memory T-cell differentiation. *Immunol. Cell Biol.* **86**, 325–332 (2008).
78. Geijtenbeek, T. B., van Vliet, S. J., Engering, A., 't Hart, B. A. & van Kooyk, Y. Self- and non self-recognition by C-type lectins on dendritic cells. *Annu. Rev. Immunol.* **22**, 33–54 (2004).
79. Vafiadis, P. *et al.* Insulin expression in human thymus is modulated by *INS* VNTR alleles at the *IDDM2* locus. *Nature Genet.* **15**, 289–292 (1997).
80. Stene, L. C. *et al.* Rotavirus infection frequency and risk of celiac disease autoimmunity in early childhood: a longitudinal study. *Am. J. Gastroenterol.* **101**, 2333–2340 (2006).
81. Shakoob, N., Michalska, M., Harris, C. A. & Block, J. A. Drug-induced systemic lupus erythematosus associated with etanercept therapy. *Lancet* **359**, 579–580 (2002).
82. Wada, Y. *et al.* Selective abrogation of Th1 response by STA-5326, a potent IL-12/IL-23 inhibitor. *Blood* **109**, 1156–1164 (2007).
83. Weersma, R. K. *et al.* Molecular prediction of disease risk and severity in a large Dutch Crohn's disease cohort. *Gut* 29 Sep 2008 (doi:10.1136/gut.2007.144865).
84. Barcellos, L. F. *et al.* Clustering of autoimmune diseases in families with a high-risk for multiple sclerosis: a descriptive study. *Lancet Neurol.* **5**, 924–931 (2006).
85. Michou, L., Rat, A. C., Lasbleiz, S., Bardin, T. & Cornelis, F. Prevalence and distribution of autoimmune diseases in 368 rheumatoid arthritis families. *J. Rheumatol.* **35**, 790–796 (2008).
86. Munthe-Kaas, M. C. *et al.* HLA Dr-Dq haplotypes and the TNFA308 polymorphism: associations with asthma and allergy. *Allergy* **62**, 991–998 (2007).
87. Jacobson, E. M., Huber, A. & Tomer, Y. The HLA gene complex in thyroid autoimmunity: from epidemiology to etiology. *J. Autoimmun.* **30**, 58–62 (2008).
88. Fernando, M. M. *et al.* Identification of two independent risk factors for lupus within the MHC in United Kingdom families. *PLoS Genet.* **3**, e192 (2007).
89. Hafler, D. A. *et al.* Risk alleles for multiple sclerosis identified by a genomewide study. *N. Engl. J. Med.* **357**, 851–862 (2007).
90. Liu, Y. *et al.* A genome-wide association study of psoriasis and psoriatic arthritis identifies new disease loci. *PLoS Genet.* **4**, e1000041 (2008).
91. Bowes, J. & Barton, A. Recent advances in the genetics of RA susceptibility. *Rheumatology* **47**, 399–402 (2008).
92. Fisher, S. A. *et al.* Genetic determinants of ulcerative colitis include the *ECM1* locus and five loci implicated in Crohn's disease. *Nature Genet.* **40**, 710–712 (2008).
93. Schlesselman, J. J. *Case-Control Studies: Design, Conduct, Analysis* (Oxford Univ. Press, Oxford, 1992).
94. Brophy, S. *et al.* Inflammatory eye, skin, and bowel disease in spondyloarthritis: genetic, phenotypic, and environmental factors. *J. Rheumatol.* **28**, 2667–2673 (2001).
95. Bernstein, C. N., Wajda, A. & Blanchard, J. F. The clustering of other chronic inflammatory diseases in inflammatory bowel disease: a population-based study. *Gastroenterology* **129**, 827–836 (2005).

96. Spadaccino, A. C. *et al.* Celiac disease in north Italian patients with autoimmune thyroid diseases. *Autoimmunity* **41**, 116–121 (2008).
97. Freeman, H. J. Adult celiac disease followed by onset of systemic lupus erythematosus. *J. Clin. Gastroenterol.* **42**, 252–255 (2008).
98. Guariso, G. *et al.* Clinical, subclinical and potential autoimmune diseases in an Italian population of children with coeliac disease. *Aliment. Pharmacol. Ther.* **26**, 1409–1417 (2007).
99. Kero, J., Gissler, M., Hemminki, E. & Isolauri, E. Could TH1 and TH2 diseases coexist? Evaluation of asthma incidence in children with coeliac disease, type 1 diabetes, or rheumatoid arthritis: a register study. *J. Allergy Clin. Immunol.* **108**, 781–783 (2001).
100. Ch'ng, C. L., Jones, M. K. & Kingham, J. G. Celiac disease and autoimmune thyroid disease. *Clin. Med. Res.* **5**, 184–192 (2007).
101. Lee, F. I., Bellary, S. V. & Francis, C. Increased occurrence of psoriasis in patients with Crohn's disease and their relatives. *Am. J. Gastroenterol.* **85**, 962–963 (1990).
102. Weng, X., Liu, L., Barcellos, L. F., Allison, J. E. & Herrinton, L. J. Clustering of inflammatory bowel disease with immune mediated diseases among members of a northern California-managed care organization. *Am. J. Gastroenterol.* **102**, 1429–1435 (2007).
103. Molina, M. J. *et al.* Prevalence of systemic lupus erythematosus and associated comorbidities in Puerto Rico. *J. Clin. Rheumatol.* **13**, 202–204 (2007).
104. Barker, J. M. Clinical review: type 1 diabetes-associated autoimmunity: natural history, genetic associations, and screening. *J. Clin. Endocrinol. Metab* **91**, 1210–1217 (2006).
105. Cronin, C. C. *et al.* High prevalence of celiac disease among patients with insulin-dependent (type I) diabetes mellitus. *Am. J. Gastroenterol.* **92**, 2210–2212 (1997).
106. Hasegawa, K. *et al.* Variations in the C3, C3a receptor, and C5 genes affect susceptibility to bronchial asthma. *Hum. Genet.* **115**, 295–301 (2004).
107. Moffatt, M. F. *et al.* Genetic variants regulating *ORMDL3* expression contribute to the risk of childhood asthma. *Nature* **448**, 470–473 (2007).
108. Duerr, R. H. *et al.* A genome-wide association study identifies *IL23R* as an inflammatory bowel disease gene. *Science* **314**, 1461–1463 (2006).
109. Franke, A. *et al.* Systematic association mapping identifies *NELL1* as a novel IBD disease gene. *PLoS ONE* **2**, e691 (2007).
110. Franke, A. *et al.* Replication of signals from recent studies of Crohn's disease identifies previously unknown disease loci for ulcerative colitis. *Nature Genet.* **40**, 713–715 (2008).
111. Hampe, J. *et al.* A genome-wide association scan of nonsynonymous SNPs identifies a susceptibility variant for Crohn disease in *ATG16L1*. *Nature Genet.* **39**, 207–211 (2007).
112. Kugathasan, S. *et al.* Loci on 20q13 and 21q22 are associated with pediatric-onset inflammatory bowel disease. *Nature Genet.* **40**, 1211–1215 (2008).
113. Libioulle, C. *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.* **3**, e58 (2007).
114. Parkes, M. *et al.* Sequence variants in the autophagy gene *IRGM* and multiple other replicating loci contribute to Crohn's disease susceptibility. *Nature Genet.* **39**, 830–832 (2007).
115. Raelson, J. V. *et al.* Genome-wide association study for Crohn's disease in the Quebec founder population identifies multiple validated disease loci. *Proc. Natl Acad. Sci. USA* **104**, 14747–14752 (2007).
116. Rioux, J. D. *et al.* Genome-wide association study identifies new susceptibility loci for Crohn disease and implicates autophagy in disease pathogenesis. *Nature Genet.* **39**, 596–604 (2007).
117. Weersma, R. K. *et al.* *ATG16L1* and *IL23R* are associated with inflammatory bowel diseases but not with celiac disease in the Netherlands. *Am. J. Gastroenterol.* **103**, 621–627 (2008).
118. Zhernakova, A. *et al.* Genetic analysis of innate immunity in Crohn's disease and ulcerative colitis identifies two susceptibility loci harboring *CARD9* and *IL18RAP*. *Am. J. Hum. Genet.* **82**, 1202–1210 (2008).
119. Dideberg, V. *et al.* An insertion-deletion polymorphism in the interferon regulatory factor 5 (*IRF5*) gene confers risk of inflammatory bowel diseases. *Hum. Mol. Genet.* **16**, 3008–3016 (2007).
120. Capon, F. *et al.* Identification of *ZNF313/RNF114* as a novel psoriasis susceptibility gene. *Hum. Mol. Genet.* **17**, 1938–1945 (2008).
121. Hollox, E. J. *et al.* Psoriasis is associated with increased β -defensin genomic copy number. *Nature Genet.* **40**, 23–25 (2008).
122. Zhernakova, A. *et al.* Novel association in chromosome 4q27 region with rheumatoid arthritis and confirmation of type 1 diabetes point to a general risk locus for autoimmune diseases. *Am. J. Hum. Genet.* **81**, 1284–1288 (2007).
123. Remmers, E. F. *et al.* *STAT4* and the risk of rheumatoid arthritis and systemic lupus erythematosus. *N. Engl. J. Med.* **357**, 977–986 (2007).
124. Plenge, R. M. *et al.* *TRAF1-C5* as a risk locus for rheumatoid arthritis—a genomewide study. *N. Engl. J. Med.* **357**, 1199–1209 (2007).
125. Kochi, Y. *et al.* A functional variant in *FCRL3*, encoding Fc receptor-like 3, is associated with rheumatoid arthritis and several autoimmunities. *Nature Genet.* **37**, 478–485 (2005).
126. Thomson, W. *et al.* Rheumatoid arthritis association at 6q23. *Nature Genet.* **39**, 1431–1433 (2007).
127. Lee, C. C. *et al.* Interleukin-18 gene polymorphism, but not interleukin-2 gene polymorphism, is associated with rheumatoid arthritis. *Immunogenetics* **59**, 433–439 (2007).
128. Plenge, R. M. *et al.* Two independent alleles at 6q23 associated with risk of rheumatoid arthritis. *Nature Genet.* **39**, 1477–1482 (2007).
129. Dieguez-Gonzalez, R. *et al.* Association of interferon regulatory factor 5 haplotypes, similar to that found in systemic lupus erythematosus, in a large subgroup of patients with rheumatoid arthritis. *Arthritis Rheum.* **58**, 1264–1274 (2008).
130. Graham, D. S. *et al.* Association of LY9 in UK and Canadian SLE families. *Genes Immun.* **9**, 93–102 (2008).
131. Hom, G. *et al.* Association of systemic lupus erythematosus with C8orf13-BLK and ITGAM-ITGAX. *N. Engl. J. Med.* **358**, 900–909 (2008).
132. Kozyrev, S. V. *et al.* Functional variants in the B-cell gene *BANK1* are associated with systemic lupus erythematosus. *Nature Genet.* **40**, 211–216 (2008).
133. Nath, S. K. *et al.* A nonsynonymous functional variant in integrin- α_M (encoded by *ITGAM*) is associated with systemic lupus erythematosus. *Nature Genet.* **40**, 152–154 (2008).
134. Sigurdsson, S. *et al.* A common *STAT4* risk haplotype for systemic lupus erythematosus is over-expressed, correlates with anti-dsDNA production and shows additive effects with two *IRF5* risk alleles. *Hum. Mol. Genet.* **17**, 2868–2876 (2008).
135. Graham, D. S. *et al.* Polymorphism at the TNF superfamily gene *TNFSF4* confers susceptibility to systemic lupus erythematosus. *Nature Genet.* **40**, 83–89 (2008).
136. Hakonarson, H. *et al.* A genome-wide association study identifies *KIAA0350* as a type 1 diabetes gene. *Nature* **448**, 591–594 (2007).
137. Hakonarson, H. *et al.* A novel susceptibility locus for type 1 diabetes on chr 12q13 identified by a genome-wide association study. *Diabetes* **57**, 1143–1146 (2008).
138. Smyth, D. J. *et al.* A genome-wide association study of nonsynonymous SNPs identifies a type 1 diabetes locus in the interferon-induced helicase (*IFIH1*) region. *Nature Genet.* **38**, 617–619 (2006).
139. Dideberg, V. *et al.* An insertion-deletion polymorphism in the interferon regulatory factor 5 (*IRF5*) gene confers risk of inflammatory bowel diseases. *Hum. Mol. Genet.* **16**, 3008–3016 (2007).
140. Dubois, P. C. & van Heel, D. A. New susceptibility genes for ulcerative colitis. *Nature Genet.* **40**, 686–688 (2008).

Acknowledgements

We thank J. Senior and M. Wapenaar for their help preparing this manuscript. Our work is supported by grants from the Celiac Disease Consortium (an innovative cluster approved by the Netherlands Genomics Initiative and partly funded by the Dutch Government, grant BSIK03009 to C.W.) and the Netherlands Organization for Scientific Research (NWO-VICI grant 918.66.620 to C.W.).

DATABASES

Entrez Gene: <http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=gene>
[CTLA4](http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=OMIM) | [IL2RA](http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=OMIM) | [IL10](http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=OMIM) | [IL18RAP](http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=OMIM) | [IL21](http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=OMIM) | [IL23R](http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=OMIM) | [NOD2](http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=OMIM) | [PTPN22](http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=OMIM) | [STAT3](http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=OMIM) | [STAT4](http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=OMIM)
OMIM: <http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=OMIM>
[asthma](http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=OMIM) | [coeliac disease](http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=OMIM) | [Graves' disease](http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=OMIM) | [Hashimoto's disease](http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=OMIM) | [multiple sclerosis](http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=OMIM) | [rheumatoid arthritis](http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=OMIM) | [type 1 diabetes](http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=OMIM)
UniProtKB: <http://www.uniprot.org>
[IL-2](http://www.uniprot.org) | [PTPN2](http://www.uniprot.org) | [SH2B3](http://www.uniprot.org)

FURTHER INFORMATION

Cisca Wijmenga's homepage:
<http://www.rug.nl/umcg/faculteit/disciplinegroepen/medischegenetica/research/CVCiscaWijmenga>
HapMap: <http://www.hapmap.org>
Database for Annotation, Visualization and Integrated Discovery (DAVID) Bioinformatics Resources:
<http://david.abcc.ncifcrf.gov/home.jsp>
Web-based Gene Set Analysis Toolkit (Webgestalt):
<http://bioinfo.vanderbilt.edu/webgestalt>
Gene Ontology: <http://www.geneontology.org>
Catalogue of Published Genome-wide Association Studies:
<http://www.genome.gov/GWAstudies>

SUPPLEMENTARY INFORMATION

See online article: S1 (table) | S2 (table) | S3 (figure) | S4 (table) | S5 (Excel file)

ALL LINKS ARE ACTIVE IN THE ONLINE PDF