Genetics of autoimmune diseases — disorders of immune homeostasis

Peter K. Gregersen* and Timothy W. Behrens‡

Abstract | In the past few years, our extensive knowledge of the mammalian immune system and our increasing ability to understand the genetic causes of complex human disease have opened a window onto the pathways that lead to autoimmune disorders. In addition to the well-established role of genetic variation that affects the major histocompatibility complex, a number of rare and common variants that affect a range of immunological pathways are now known to have important influences on the phenotypic diversity that is seen among autoimmune diseases. Recent studies have also highlighted a previously unanticipated interplay between the innate and adaptive immune system, providing a new direction for research in this field.

Innate immune system
Nonspecific and phylogenetically ancient mechanisms that form the first line of defence against infection. Innate immune defence is inborn and does not involve memory; it uses a limited set of molecules that generally recognize common molecular patterns found in microorganisms.

Adaptive immune system
A flexible and specific immune response that can adjust to new structures and that retains a memory of prior exposure to these structures. A large and diverse set of recognition molecules — antibodies (produced by B cells) and T-cell receptors — mediate adaptive immune recognition.

*Robert S. Boas Center for Genomics and Human Genetics, Feinstein Institute for Medical Research, North Shore Long Island Jewish Health System, Manhasset, New York 11030, USA. ‡Genentech Inc., South San Francisco, California 94080, USA. Correspondence to P.K.G. e-mail: peterg@nshs.edu

The organism possesses certain contrivances by means of which the immunity reaction, so easily produced by all kinds of cells, is prevented from working against the organism's own elements. Paul Ehrlich

The destructive power and complexity of the human immune system necessitates the presence of sophisticated mechanisms to regulate its activity. The need to prevent the immune system from turning against the self was an early concept in immunology, reflected in the term 'horror autotoxicus' coined by immunologist Paul Ehrlich in the late nineteenth century[1]. The issue of self–non-self discrimination and regulation of autoimmunity has subsequently emerged as a central problem in modern immunology. Approximately 3% of the human population is affected by a recognized autoimmune disorder[2–4], and unrecognized autoimmune mechanisms might also contribute to other common disorders[5]. Understanding the genetic factors that contribute to these diseases is therefore likely to have a significant effect on public health. In addition, autoimmunity can be present without any obvious clinical manifestations; autoantibodies are common in normal subjects, and in some cases their presence is a risk factor for the future development of overt autoimmune disease[6–8]. Genetic factors also have a role in determining progression to clinical disease, and identifying these factors offers a fertile area for the development of targeted preventive therapies.

Most recent advances in understanding immune regulation have resulted from mouse studies, which have revealed that regulation by 'certain contrivances' occurs at several levels, with the involvement of many cellular and biochemical pathways. Broadly, immune defence is mediated by two complementary systems, the innate immune system and the adaptive immune system. Innate immunity is phylogenetically ancient and is generally directed towards immediate responses to commonly encountered threats in the environment. By contrast, adaptive immunity is concerned primarily with the development of long-lasting defence and memory for threats that might be encountered repeatedly. Until recently, defects in adaptive immunity were a predominant focus in autoimmune disease research, for example driving investigation into the origin and specificity of autoantibodies. However, it is increasingly evident that the innate and adaptive immune systems interact to form an efficient overall system of immune defence (FIG. 1); the pathogenesis of autoimmunity, and its genetic basis, must be understood in this context.

Rodent disease models have provided essential tools for testing concepts in autoimmunity as they might apply to disease phenotypes. However, whereas various 'knockout' and 'knock-in' genetic effects can be explored in animal models to examine disease pathways, it is often difficult to assess their relevance to the human conditions. The study of human genetics, difficult as it is, presents a more direct line to understanding the molecular basis of autoimmunity in human populations. Traditional linkage analysis approaches have generally not been productive for autoimmune diseases, with a few exceptions[9,10]. This is thought to be mainly because most autoimmune diseases have a complex, multifactorial basis, which linkage methods lack sufficient statistical power to dissect. However, in the past several years there has been increasing success
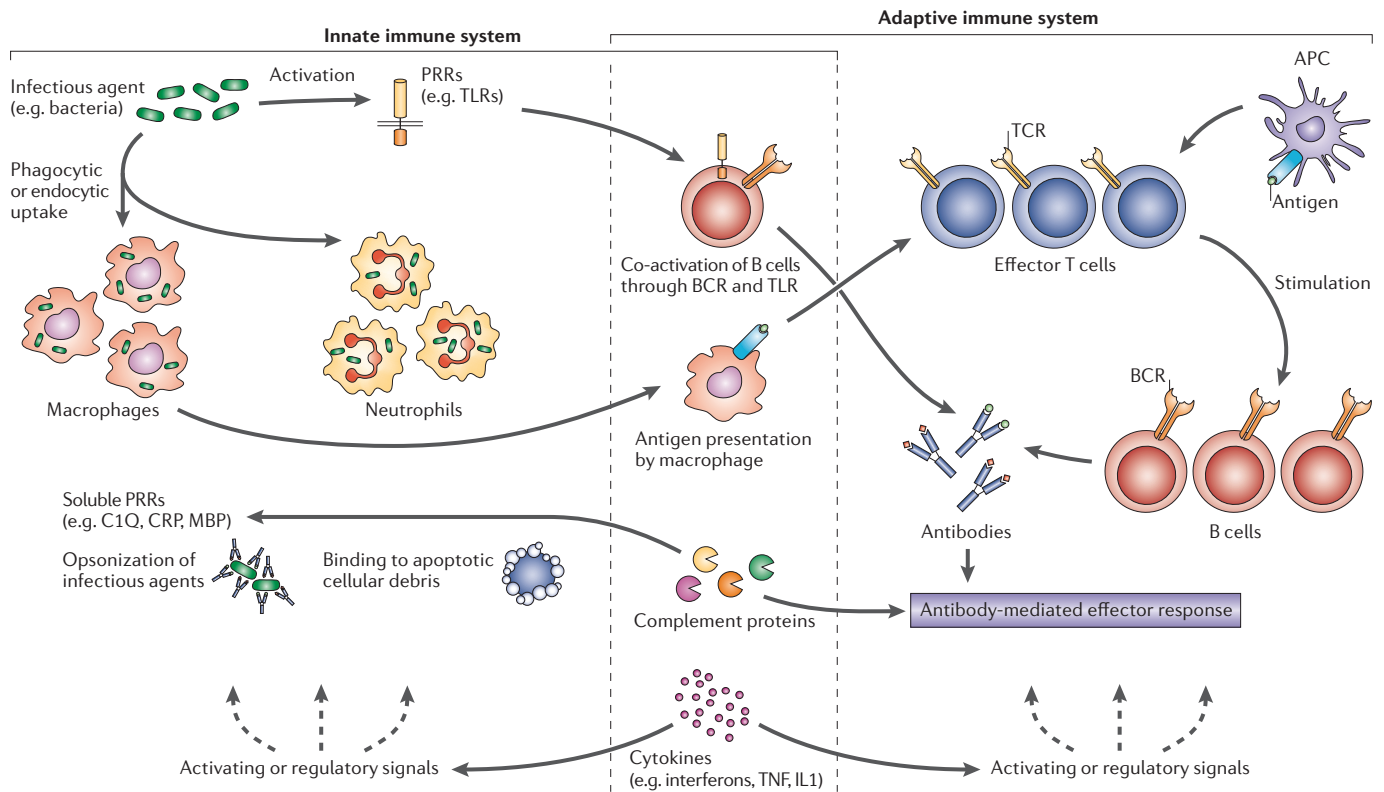
Figure 1 | **The innate and adaptive immune systems and the overlap between them.** Innate immune mechanisms generally involve immediate, nonspecific responses to foreign infectious agents. These include cellular functions such as phagocytosis and endocytosis by macrophages and neutrophils. Some of these activities are dependent on pattern-recognition receptors (PRRs), such as Toll-like receptors (TLRs) and NOD-like receptors, which recognize pathogen-associated molecular patterns (PAMPs) present on a variety of microorganisms. In addition, a variety of soluble PRRs, such as complement proteins (C1Q), mannose-binding protein (MBP), and acute phase reactants, such as C-reactive protein (CRP), have a role in innate immunity by opsonizing microrganisms and binding to apoptotic cellular debris in a nonspecific manner. Adaptive immune mechanisms involve the engagement of receptors that are selected for reactivity with specific antigens (T-cell receptors (TCRs) and immunoglobulin receptors on B cells). The full development of these responses requires the expansion and differentiation of the specific responder cells, which establishes a memory for the specific antigen response. The innate and adaptive immune systems are interrelated in ways that have not been fully established. For example, antigens that are phagocytosed or endocytosed in a nonspecific manner by macrophages are presented to T cells, generating a highly specific T-cell response. In addition, co-stimulation of B cells through TLRs (such as TLR9) can result in the production of specific antibodies to self antigens. Cytokines such as interferons, tumour necrosis factor (TNF), and interleukin 1 (IL1) might stimulate activity of both the innate and adaptive immune response. Complement proteins also mediate the effector responses induced by antibodies (not shown), and therefore have a role in both innate and adaptive immune functions. APC, antigen-presenting cell; BCR, B-cell receptor.

Linkage analysis
A method for tracking the transmission of genetic information across generations to identify the map location of genetic loci on the basis of co-inheritance of genetic markers and discernable phenotypes.

Gene association study
A study in which a genetic variant is genotyped in a population for which phenotypic information (such as disease occurrence, or a range of different trait values) is available. If a correlation is observed between the genotype and phenotype, there is said to be an association between the variant and the disease or trait.

in identifying causal genes for these disorders by taking advantage of more powerful gene association studies, using a candidate gene approach that is based on rapidly accumulating knowledge about the cellular and molecular biology of the immune response.

This Review brings together the emerging genetic data that have resulted from such studies. Recent findings support the view that some autoimmune disorders might share a common genetic and mechanistic basis. At the same time, genetic factors also seem to contribute to the phenotypic diversity that is seen within a given disorder. Genetic studies have also revealed a striking diversity of molecular pathways to disease, including unexpectedly important contributions of genetic variation in innate immune mechanisms for some forms of autoimmunity.

Diverse phenotypes, overlapping pathways

Human autoimmune diseases are phenotypically extremely heterogeneous. From a clinical perspective it is convenient to classify autoimmunity into 'systemic' versus 'organ-specific' diseases. However, this division does not necessarily reflect a fundamental mechanistic or genetic distinction between these groups. All have a significant genetic component, and there is clearly genetic overlap between organ-specific and systemic autoimmune disease. This is reflected in the overlapping associations of human leukocyte antigen (HLA) alleles with these disorders (TABLE 1). HLA molecules are encoded within the major histocompatibility complex (MHC) on chromosome 6 and are essential for the normal function of the adaptive immune system,

Table 1 | **Sibling risk and HLA associations for selected autoimmune disorders**

Disease	Relative risk*	HLA–DRB1 associations‡
'Organ-specific' autoimmune disorders		
Type 1 diabetes	15	DR3, DR4
Graves disease	15	DR3
Multiple sclerosis	20	DR2
'Systemic' autoimmune disorders		
Systemic lupus erythematosus	20	DR2, DR3
Rheumatoid arthritis	8	DR1, DR4

*The relative risk to siblings of affected individuals, compared with the risk in the general population; data taken from REF. 122. ‡The HLA associations that are listed represent broad families of alleles at the HLA–DRB1 locus. The specific allelic associations are known in considerable detail, although there is still controversy over the exact alleles or loci involved in some cases[123–127].

Candidate gene
A gene for which there is evidence of its possible role in the trait or disease that is under study.

T cells
Lymphocytes that have important roles in the primary immune response. Effector T cells fall into two classes — CD8+ killer or cytotoxic T cells, which destroy infected cells, and CD4+ or helper T cells, which regulate the function of other lymphocytes. A third class, regulatory T cells, regulate the self reactivity of effector T cells in the periphery.

Systemic lupus erythematosus
The prototypical autoimmune disease in which antibodies are produced to DNA and a variety of other self constituents. The disease has highly variable manifestations and can be mild or lead to widespread inflammation and tissue destruction of important organ systems.

Tolerance
A specific unresponsiveness of the adaptive immune system to particular antigens in an otherwise immunocompetent organism. In the absence of autoimmunity, tolerance to self antigens is characteristic of the normal immune system.

as they are central to the presentation of antigens for recognition by T cells. The allelic diversity of HLA molecules exceeds that of any other protein family, and this diversity is a major contributor to individual differences in immune responsiveness. Hundreds of associations between HLA alleles and autoimmune disorders, as well as other diseases, have been established over the last three decades[11]. Although some HLA alleles predispose to multiple diseases, in general there are specific HLA allelic associations with each disorder. An evolving concept, which we expand on below, posits that particular HLA alleles, in combination with other genetic variants, determine the specific phenotypic expression of an underlying autoimmune trait that is common to many of these disorders[12].

Even within a single autoimmune disease, there might be considerable phenotypic heterogeneity. Systemic lupus erythematosus (SLE) is the classical example, in which nearly every organ system can be affected, with substantial variability in clinical manifestations even within familial disease[13]. As discussed below, dysregulation of interferon pathways seems to be a central feature of SLE. However, HLA alleles might regulate the production of particular autoantibodies, which in turn predispose to specific clinical manifestations. Therefore, it can be challenging to distinguish genetic effects that relate to a single disorder, cross multiple autoimmune disorders, or might be specific for a particular clinical manifestation of a disease.

Defects in adaptive immune mechanisms
Central and peripheral mechanisms maintain T-cell tolerance. Many autoimmune diseases seem to result from a failure to maintain T-cell tolerance. There are several levels at which T cells are prevented from uncontrolled self reactivity (FIG. 2). The first of these involves thymic selection (see REFS 14,15 for reviews). Pre-T-cells (thymocytes) in the thymus undergo a stringent process of selection, during which 95% are eliminated[14]. Thymocytes with excessive reactivity to self antigens, which are presented by endogenous MHC molecules and activate thymocytes through interactions with T-cell receptors (TCRs), die through apoptosis. Surviving thymocytes undergo an additional process of positive

selection and emerge from the thymus as mature T cells. Most of these T cells are various types of effector T cell, generally either CD4+ or CD8+.

A second major population of T cells — regulatory T cells (T_{Reg} cells) — are also produced during thymic selection[16]. These cells are CD4+ and CD25+, and also express FOXP3, a member of the forkhead family of transcription factors. T_{Reg} cells have a key role in controlling the self reactivity of effector T cells in the periphery, because thymic selection does not completely eliminate autoreactive T cells[17]. T_{Reg} cells can also develop *de novo* in the periphery from precursors that might not be directly derived from T_{Reg} cells that emerge from the thymus[18,19]. The expression of FOXP3 is essential for the development of T_{Reg} cells, regardless of whether they arise in the thymus or the periphery[20,21]. In addition, T_{Reg} cells are critically dependent on the T-cell-derived cytokine interleukin 2 (IL2) for their maintenance and survival[22]. T_{Reg} cells regulate the activation and differentiation of effector T cells at many levels. This regulatory activity is complex and incompletely understood but seems to involve effects on antigen-presenting cells (APCs), T cells and B cells, with variable requirements for direct cell contact[23].

Genetic defects in central tolerance mechanisms. A number of autoimmune diseases can be ascribed to genetic abnormalities in thymic selection[14,15] (TABLE 2). Both quantitative differences in the expression of the selecting self antigens and in the responsiveness of the thymocytes to these antigens can have a major effect on the selection process.

Defective thymic deletion can be caused by reduced or altered expression of MHC molecules. As discussed above, thymocytes, like mature T cells, require MHC molecules for presentation of (self) antigens to the TCR. If the relevant MHC molecule is absent or expressed at a low level in the thymus, thymocytes will not be activated and will not undergo apoptosis. Indeed, this process is exquisitely sensitive both to the level of MHC molecule expression, and to the specific MHC alleles that are expressed in the thymus. This is likely to be a major mechanism by which HLA molecules regulate the peripheral T-cell repertoire and overall T-cell immune responsiveness in a given individual. Although this has been demonstrated in rodents, only indirect evidence of this mechanism can be obtained in humans, owing to technical and ethical barriers. Nevertheless, altered thymic selection remains a leading explanation for the large number of associations between HLA alleles and human diseases.

Low or inadequate expression of self antigens in the thymus can also cause or predispose to autoimmune disease. The poster child for genes that are involved in this type of disease mechanism is autoimmune regulator (*AIRE*)[24,25]. *AIRE* encodes a transcription factor that regulates the thymic expression of a variety of self antigens[25]. In humans, mutations in *AIRE* cause the rare Mendelian autosomal recessive disorder autoimmune polyendocrine syndrome 1 (APS1), in which autoimmunity is focused largely on endocrine organs[26,27].

Recently it has been shown that a remarkably large number of organ-specific self antigens are expressed in the thymus[14]. The absence of this expression can result in an escape from thymic deletion and the release of self-reactive T cells into the periphery[25]. In the absence
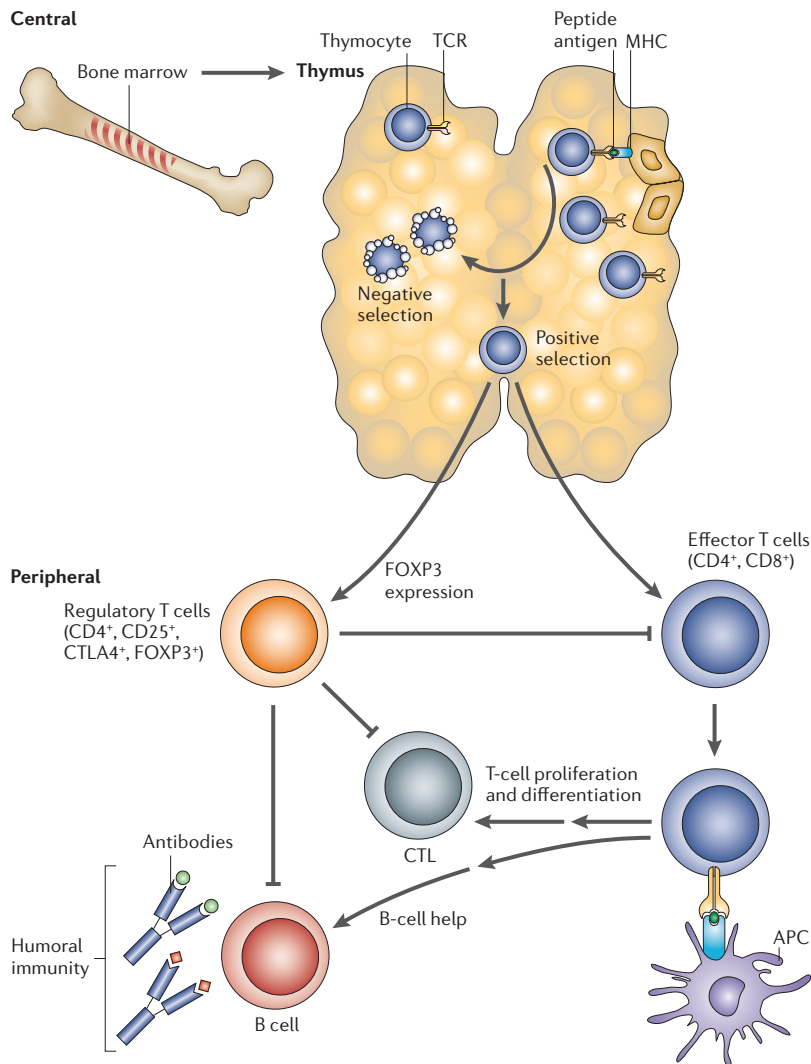

Figure 2 | Central and peripheral tolerance mechanisms. The top panel depicts events involved in central tolerance, which takes place in the thymus. Thymocytes undergo a maturation and selection process in which strongly self-reactive thymocytes — as determined by interactions with major histocompatibility complex (MHC) proteins in combination with self peptides — are deleted. Similarly, non-functional thymocytes undergo apoptosis. Only thymocytes that are activated by self peptide and MHC below a certain threshold are positively selected and migrate into the periphery as mature T cells. Most of these thymic emigrants develop into effector CD4[+] and CD8[+] T cells, and mediate both humoral (antibody-mediated) and cellular immune responses. A small percentage of T cells that emigrate from the thymus express the transcription factor FOXP3 and develop into CD4[+], CD25[+] and CTLA4[+] regulatory T cells (natural T_{Reg} cells). Once in the periphery, these cells are key mediators of peripheral tolerance. The mechanism of action of T_{Reg} cells is incompletely understood, but includes actions at many levels of the effector immune response. T_{Reg} cells might inhibit T-cell activation by antigen-presenting cells (APCs) and inhibit T-cell differentiation into cytotoxic effector cells, as well as preventing T cells from providing help to B cells in the production of antibodies. FOXP3[+] T_{Reg} cells can also be generated from peripheral T cells (not shown). CTL, cytotoxic lympmhocyte; TCR, T-cell receptor.

of *AIRE*, these antigens are not available to be presented by MHC molecules to the developing thymocytes, and the T cells that are reactive to these antigens escape into the periphery[26,27].

A large number of self antigens are expressed at highly variable levels in the thymic epithelium[14,28], and this might be a common underlying risk factor for a variety of complex autoimmune diseases, given the quantitative nature of the thymic selection process[15]. For example, regulatory polymorphisms of the insulin gene are a risk factor for type 1 diabetes[29–31]. The disease-associated alleles show reduced expression in the thymic epithelium, but increased expression in pancreatic islet cells, a situation that is predicted to predispose to a lack of central T-cell tolerance to insulin[32]. The relevance of this is further supported by the presence of insulin-specific T cells in the draining pancreatic lymph nodes of individuals with type 1 diabetes[33]. A recent study has suggested a similar mechanism for the acetylcholine receptor in myasthenia gravis[34], although this remains to be confirmed. The association of Graves disease with polymorphisms of the thyroid-stimulating hormone receptor (*TSHR*)[35], a target autoantigen in this disease, is also of interest in this context. This provides a rationale for a more detailed genetic analysis of promoter polymorphisms in the autoantigens involved in many other autoimmune disorders.

In addition to defects in antigen expression and presentation, if thymocytes are unable to respond adequately to signals delivered through the TCR they will not be properly selected. For example, in the mouse it has been shown that a mutation in the gene that encodes ZAP70, a key TCR signalling molecule, results in reduced negative selection and subsequent escape of self-reactive T cells into the periphery[36]. Interestingly, these animals develop an autoimmune arthritis that is similar to human rheumatoid arthritis, providing a compelling example of how a specific defect in central tolerance can lead to a particular autoimmune phenotype. A comprehensive analysis of *ZAP70* polymorphisms and related T-cell signalling molecules has not yet been carried out for human autoimmune diseases, although this is likely to be achieved soon in the context of whole-genome association studies or case–control studies that are based on the comprehensive genetic analysis of candidate pathways.

Genetic defects in T_{Reg} cell functions. A role for T_{Reg} cells in human autoimmunity is demonstrated by the rare X-linked recessive disorder IPEX (immune dysregulation, polyendocrinopathy, enteropathy, X-linked syndrome). This syndrome is accompanied by type 1 diabetes, thyroiditis, inflammatory bowel disease, atopic dermatitis and fatal infections[37], and is caused by mutations in *FOXP3* (REF. 38). The scurfy mouse is homologous to human IPEX[39], and in both the human and mouse forms of the disease a severe deficiency of T_{Reg} cells has a causal role[40,41].

T_{Reg} cells are critically dependent on IL2 for their proliferation and function[22] and a lack of IL2 or its receptors can lead to autoimmunity, as demonstrated in several mouse models[42,43]. On first consideration, this might

Table 2 | **Non-HLA susceptibility genes for autoimmune disease**

Affected gene or genetic defect	Diseases	Mode of inheritance or odds ratio	Mechanism	Refs
Mendelian disorders				
AIRE	Autoimmune polyendocrine syndrome type I	Autosomal recessive	Adaptive immunity affected; defective thymic expression of self antigens	24
FOXP3	Immune dysregulation, polyendocrinopathy, enteropathy, X-linked syndrome (homologue of scurfy mouse)	X-linked recessive	Adaptive immunity affected; defective production of T_{Reg} cells	38
Polygenic disorders				
Insulin gene	Type 1 diabetes	OR ≈ 2	Adaptive immunity affected; reduced thymic expression of insulin; reduced negative selection of T cells	29–31
PTPN22	Type 1 diabetes, rheumatoid arthritis, autoimmune thyroid disease	OR = 1.5–2	Adaptive immunity affected; reduced thymocyte responsiveness to negative selection; reduced T_{Reg} function; other mechanisms?	65–67
CTLA4	Type 1 diabetes, autoimmune thyroid disease, rheumatoid arthritis, others?	OR = 1.15 (type 1 diabetes) OR = 1.5 (autoimmune thyroid disease)	Adaptive immunity affected. Reduced T_{Reg} function; inadequate negative signalling in effector T cells; other mechanisms?	54, 55
PD1	Systemic lupus erythematosus, type 1 diabetes?, rheumatoid arthritis?	OR = 1.6 (systemic lupus erythematosus)	Adaptive immunity affected; inadequate negative signalling in effector T cells; other adaptive mechanisms?	128
Unknown	Autoimmune polyendocrine syndrome type II	Some Mendelian, others unknown	Adaptive and innate immunity affected; reduced T_{Reg} function; other mechanisms?	
IRF5	Systemic lupus erythematosus	OR ≈ 1.5	Innate immunity affected? Altered signalling thresholds in TLR pathways	96
Complement deficiencies *C1Q, C2, C4*	Systemic lupus erythematosus	Most rare, highly penetrant, others common with modest risk (C4)	Adaptive and innate immunity affected; altered pathogen recognition; altered effector functions (immune complex clearance, others)	99

AIRE, autoimmune regulator; *CTLA4*, cytotoxic T-lymphocyte-associated protein 4; *FOXP3*, a member of the forkhead/winged-helix family of transcriptional regulators; *IRF5*, interferon regulatory factor 5; OR, odds ratio; *PD1*, programmed death 1 receptor; *PTPN22*, protein tyrosine kinase, non-receptor type 22 (also known as *LYP*); TLR, Toll-like receptor; T_{Reg} cells, regulatory T cells.

seem paradoxical because IL2 also functions in effector T-cell growth and differentiation. However, recent studies indicate that effector T cells are less completely dependent on IL2 than T_{Reg} cells, suggesting that partial loss of IL2 function might lead to a relative deficiency of peripheral T_{Reg} cell function[44]. A recently identified association between an IL2 receptor locus and type 1 diabetes[45] might reflect a mechanism involving T_{Reg} cells, although the causative gene has not been definitively identified.

Genetic defects in T-cell co-stimulatory molecules. In addition to the part played by T_{Reg} cells, the activation of peripheral T cells is also controlled by an array of co-stimulatory molecules on the T-cell surface, which modulate activation of these cells through the TCR (FIG. 3). CD28, the inducible T-cell co-stimulator (ICOS) and the cytotoxic T-lymphocyte-associated protein 4 (CTLA4) are members of the immunoglobulin superfamily that are expressed on T cells[46] and bind homologous ligands on APCs. CD28 and ICOS provide positive signals[47], whereas CTLA4 is generally a negative regulator of T-cell activation[48]. However, the direction and magnitude of these effects can vary depending on the state of activation and differentiation of the cell. For example, although CTLA4 is generally inhibitory, it seems to be activating for T_{Reg} cells. *CTLA4* knockout is associated with florid lymphoproliferation[49,50].

CTLA4 polymorphisms have been associated with a variety of human autoimmune diseases, with odds ratios between 1.1 and 1.5 (REF. 51). Such associations were initially reported with Graves disease[52], and subsequently with type 1 diabetes[53], and most of these early studies focused on an A/G SNP in exon 1. *CTLA4* lies on chromosome 2q in a region that also encodes two other co-stimulatory molecules, CD28 and ICOS, and it was initially suggested that the associations with *CTLA4* might reflect linkage disequilibrium with other candidate genes in the region[51]. However, an extensive SNP analysis has now narrowed down a major risk region to a 6.1-kb segment in the 3′ region of *CTLA4*, including possible regulatory polymorphisms in the 3′ UTR[54]. In addition to type 1 diabetes and Graves disease, SNPs in this region have also been reported to be associated with other autoimmune disorders[51] such as Addison disease, coeliac disease and rheumatoid arthritis[55]. Therefore, *CTLA4* seems to be a general susceptibility locus for autoimmunity.

Several mechanisms have been proposed for the involvement of *CTLA4* in the loss of tolerance and autoimmunity. Interestingly, the human risk haplotype is associated with lower levels of a splice variant encoding a soluble form of CTLA4. Because soluble CTLA4 can interact with its ligands CD80 and CD86, and block their interaction with co-stimulatory molecules such as

Type I diabetes
An autoimmune disorder in which insulin-producing islet cells are destroyed by the immune system, leading to insulin deficiency. Onset is usually in childhood or adolescence.

Myasthenia gravis
An autoimmune disorder in which antibodies to the acetylcholine receptor disrupt neuromuscular transmission leading to symptoms of weakness.

Rheumatoid arthritis
A chronic inflammatory disorder of synovial joints and surrounding tissues. It often leads to bone and cartilage destruction if left untreated.
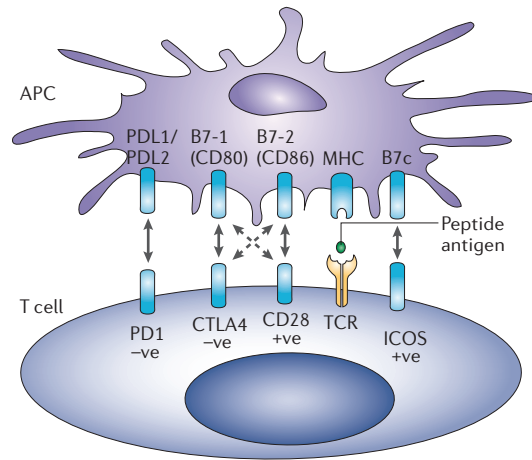
Figure 3 | **Co-stimulatory and co-inhibitory molecules set thresholds for T-cell activation.** T cells are activated through stimulation of the T-cell receptor (TCR). Co-stimulatory and co-inhibitory molecules that are also present at the T-cell surface set thresholds for this activation, and thereby regulate the degree of activation as well as the development of peripheral T-cell tolerance. CD28 and ICOS are activating receptors that bind members of the B7 family of ligands on antigen-presenting cells (APCs) such as macrophages and dendritic cells. The cytotoxic T-lymphocyte-associated protein 4 (CTLA4) is inhibitory on most T cells (with the exception of regulatory T cells), and, like CD28, can bind both CD80 and CD86, but with varying affinities (indicated by solid and dashed interaction lines). This can result in cross competition for CD28 co-stimulation. The programmed cell death 1 (PD1) protein is also inhibitory for T-cell activation, and can bind two related ligands, PDL1 and PDL2.

Whole-genome association studies
Studies in which associations between genetic variation and a phenotype of interest (for example, disease) are identified by genotyping cases and controls for a set of polymorphic markers that capture genetic variation across the entire genome.

Florid lymphoproliferation
Extensive proliferation of lymphoid elements, usually either T cells or B cells.

Odds ratio
A measure of relative risk that is usually estimated from case–control studies.

Addison disease
A condition resulting from inadequate production of cortisol by the adrenal glands owing to adrenal destruction by autoimmune mechanisms.

Coeliac disease
An autoimmune disorder induced by exposure to gluten and characterized by damage to all or part of the villi lining the small intestine, leading to malabsorption.

Macrophage
A specialized type of white blood cell that can engulf foreign particles and microorganisms.

Dendritic cells
These cells present antigen to T cells, and stimulate cell proliferation and the immune response.

Memory T cells
T lymphocytes that have been previously exposed to specific antigen and are primed to have an increased response following re-exposure to the same antigen.
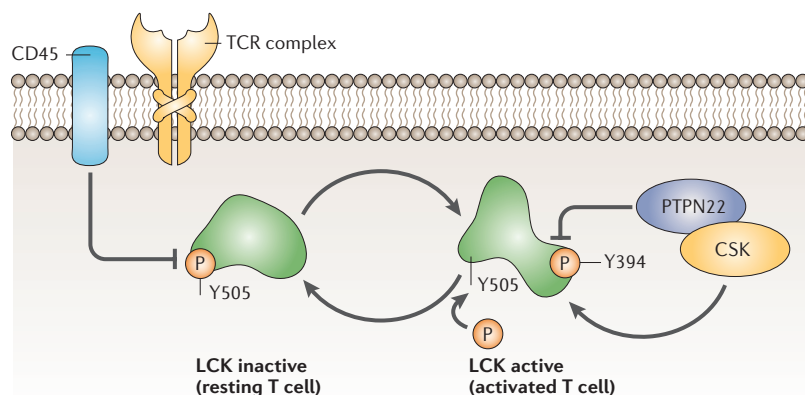
CD28 (FIG. 3), this might indicate a relative lack of local blocking of effector T-cell activity in subjects who are at risk for autoimmunity. However, investigations of this possibility at the protein level have been inconclusive[51]. CTLA4 is also constitutively expressed on T_{Reg} cells, where it functions as an activating co-stimulatory molecule on these cells. This suggests another potential mechanism for disease associations by virtue of diminished T_{Reg} cell activity. Several other mechanisms have also been proposed[51].

Similar to CTLA4, the programmed cell death 1 (PD1) protein is an important co-inhibitor of T-cell activation. Compared with *CTLA4*, *Pd1* knockout in the mouse leads to milder and more varied forms of autoimmunity with strain-dependent manifestations such as lupus-like disease and autoimmune cardiomyopathy[56,57]. *PD1* polymorphisms have recently been associated with human SLE[58], and might predispose to other autoimmune disorders (TABLE 2). One ligand for PD1 is PDL1, which is expressed both on APCs (FIG. 3) and on a variety of non-haematopoietic cells, and is induced by interferon-γ (IFN-γ) (REF. 59). A second ligand, PDL2, is primarily expressed on macrophages and dendritic cells. A recent study indicates that PDL1 expression by the placenta might have a role in maintaining immune

tolerance to the fetus during pregnancy[60]. Therefore, a continuing balance between stimulatory and inhibitory co-receptors seems to have a regulatory role during different phases of an immune response, including at the site of tissue inflammation[61]. Several additional members of this receptor family have been identified recently[48,59,62], and polymorphisms in some of these molecules provide other candidates for association with autoimmunity.

PTPN22 — setting thresholds for TCR signalling. *PTPN22* is an intracellular tyrosine phosphatase expressed in many haematopoietic cell types; it has a variety of substrates, not all of which have been characterized[63,64]. In the past few years, a functional polymorphism in *PTPN22* has been identified as a major common risk factor for several autoimmune disorders including type 1 diabetes[65], rheumatoid arthritis[66], autoimmune thyroid disease[12,67], SLE[63,64,68] and myasthenia gravis[69]. Similar to *CTLA4*, *PTPN22* therefore seems to be a general susceptibility locus for autoimmunity, although the strength of disease associations for *PTPN22* is generally higher than for *CTLA4*. Odds ratios in the range of 1.5–2.0 are typical for *PTPN22*, whereas these are in the range of 1.1–1.5 for most *CTLA4* associations.

PTPN22 turns off TCR signalling by functioning together with the intracellular kinase CSK. This combined activity downregulates the activation of LCK, a tyrosine kinase that is involved in the early event of TCR signalling (FIG. 4). The *PTPN22* risk allele contains a substitution of tryptophan for arginine at amino acid 620 (R620W) within a proline-rich SH3-binding site. This polymorphism has been shown to have two functional consequences: the binding between PTPN22 and CSK is disrupted[65,66] and the intrinsic enzymatic activity of PTPN22 is increased[70]. The effect of these biochemical changes seems to be an increase in the threshold level of stimulation that is required for TCR signalling, although this is based on a relatively small number of experiments on T cells from subjects with type 1 diabetes[70]. This result was unexpected and paradoxical, because autoimmunity is generally associated with overactivity of the immune system. Indeed, knockout mice for *Ptpn22* have a reduced threshold for TCR signalling, with expansion of memory T cells and increased production of immunoglobulins[71]. It is unclear whether overt autoimmunity would occur if these knockout mice were bred onto other strain backgrounds. As the *Ptpn22* 620W seems to be a gain-of-function risk allele, it will be of great interest to examine the phenotype of knock-in mice for this polymorphism.

The associations of the *PTPN22* 620W allele with type 1 diabetes, rheumatoid arthritis and autoimmune thyroid disease are remarkably robust, with numerous replications in Caucasian populations[72]. Odds ratios are in the range of 1.5–2.0 for a single copy of the risk allele, and are nearly twice as high for homozygotes. Therefore, there seems to be a dose effect on risk, consistent with a threshold effect on TCR signalling. There are two obvious ways of interpreting these findings in the context of mechanisms of T-cell tolerance. If carriers of the *PTPN22* 620W allele have higher thresholds for

Figure 4 | **Regulation of LCK tyrosine kinase activity by PTPN22 and CSK.** T-cell receptor (TCR) crosslinking leads to dephosphorylation of the tyrosine kinase LCK at Y505 by the membrane phosphatase CD45. Subsequent autophosphorylation of LCK at Y394 is an early event in T-cell activation, leading to phosphorylation (P) of downstream substrates by LCK. The tyrosine phosphatase PTPN22 and the intracellular kinase CSK both negatively regulate LCK by removal of the activating phosphate at Y394 (PTPN22) and rephosphorylation of Y505 (CSK). *PTPN22* polymorphisms (R620W) have been shown to disrupt the binding of PTPN22 to CSK, and to increase the intrinsic phosphatase activity of PTPN22 itself.

Natural killer cell
Large granular non-T, non-B-type lymphocytes. Natural killer cells are important for the early response to viruses. They produce cytokines, kill certain tumour cells and have appropriate receptors for antibody-dependent cell-mediated cytotoxicity.

Humoral
In mammals, humoral immunity describes B-cell-mediated immunity that fights bacteria and viruses in body fluids with antibodies that circulate in blood plasma and lymph.

Ancestry informative markers
A locus with several polymorphisms that exhibit substantially different frequencies between ancestral populations. For example, the Duffy null allele has a frequency of almost 100% of sub-Saharan Africans, but occurs rarely in other populations.

Complement system
A complex protein cascade that is involved in both innate and adaptive immunity. Three biochemical pathways activate the complement system: the classical complement pathway, the alternative complement pathway, and the mannose-binding lectin pathway.

TCR signalling (for example, require stronger TCR signals for activation), there might be a relative resistance to thymic deletion of potentially self-reactive thymocytes. Alternatively, reduced TCR signalling might lead to less effective T_{Reg} cell activity. Currently there is no direct evidence for either of these mechanisms. In addition, *PTPN22* is expressed in a variety of other haematopoietic cells[66], including B cells, macrophages and natural killer cells, where its function is not known. Therefore, it is still not entirely clear whether *PTPN22* alleles confer risk for autoimmunity solely by an effect on T-cell tolerance[73,74].

There are several other intriguing aspects to the genetics of *PTPN22*. First, the associations of variants in this gene with autoimmune disease primarily involve disorders with a prominent humoral component — for instance, neither multiple sclerosis nor inflammatory bowel disease exhibits an association with *PTPN22*, indicating different disease mechanisms in these disorders[75,76]. Second, the *PTPN22* 620W allele is almost absent in Asian populations; it remains to be seen whether other *PTPN22* alleles might be risk factors for autoimmunity in these ethnic groups. Comprehensive resequencing of *PTPN22* in Caucasians has provided hints that other polymorphisms might have a role in rheumatoid arthritis, but this has not been thoroughly studied[77]. Interestingly, a south-to-north gradient of increasing frequency of the 620W allele is observed in Europe[72]. This emphasizes the importance of properly matching cases and controls for European ancestry when performing case–control association studies of this allele, because hidden population stratification can be an important confounding variable in case–control studies[78]. Methods are now being developed to correct for this type of stratification using ancestry informative markers[79,80].

Genetic defects in B-cell tolerance mechanisms. Although the regulation of B cells is clearly important for preventing autoimmunity, most of the genetic evidence in humans implicates a role for T cells in maintaining tolerance by the adaptive immune system. However, it is likely that there are regulatory mechanisms specific to B cells that are also involved in human autoimmune disease. For example, one Fc receptor isotype, FCGR2B, has a central regulatory role in B-cell activation[81,82] and recent work provides some evidence that polymorphisms of *FCGR2B* might be associated with autoimmunity[83,84]. Other co-stimulatory molecules on B cells, such as CD40, might also contribute to disease susceptibility[85]. The human genetics of these B-cell-specific regulatory molecules is still in the early stages of investigation.

Innate and adaptive immune interactions

The innate immune response to microbial infection involves a large array of pattern-recognition receptors (PRRs) including some components of the complement system[86], Toll-like receptors (TLRs)[87], and a family of intracellular NOD-like receptors (NLRs)[88]. One member of the NLR family, NOD2, has been associated with Crohn disease[9,10], a form of inflammatory bowel disease which might have an autoimmune component. The genetics of Crohn disease was recently reviewed[89], and will not be discussed here. However, a significant development in the field of autoimmune disease research has been the realization that PRRs can also interact with molecules derived from self and facilitate a break in tolerance by the adaptive immune system.

A mechanism has been proposed for the cooperation of TLRs and autoantibodies to induce an inflammatory response to self, in which particulate antigens containing nucleic acids and proteins are released from dying cells and recognized by autoantibodies on the surface of B cells (FIG. 5). This alone will not normally lead to significant activation of B cells, but seminal studies have shown that co-stimulation of the B cell through TLR9 can lead to activation and differentiation into antibody-producing plasma cells[90,91]. The self ligands for TLR9 include both viral and self DNA, particularly in the form of CpG DNA fragments. Production of autoantibodies against nucleoprotein particulates can then lead to circulating immune complexes. These complexes in turn might trigger the activation of dendritic cells and of other APCs by virtue of co-stimulation through Fc receptors and TLR molecules that are also present on these cells.

Innate stimulation of interferon in SLE. Co-stimulation of dendritic cells through Fc receptors and TLR9 leads to the production of inflammatory cytokines, including type I interferon (FIG. 5). An important recent advance in understanding SLE pathogenesis has been the demonstration that interferon pathways are strongly upregulated in this disease[92,93]. These findings have formed the beginnings of a unifying hypothesis for SLE pathogenesis[93,94], in which combined stimulation of B cells and APCs by both the innate and adaptive immune system

leads to a pro-inflammatory state that is dominated by type 1 interferons. The adaptive immune system also has a role here, as the production of specific autoantibodies in SLE is under the control of HLA alleles that regulate the specificity of this humoral autoimmune response (TABLE 1). This new insight into the pathways that underlie SLE has led to the investigation of candidate genes that lie in the signalling pathways for TLRs and type 1 interferon production (FIG. 6) in an attempt to characterize the genetic source of the unifying pathogenic mechanism in this autoimmune disorder.

Two recent studies have shown that polymorphisms in the interferon regulatory factor 5 (*IRF5*) gene are strongly and convincingly associated with human SLE. An initial report demonstrated strong evidence for association in two Scandinavian SLE cohorts[95]. Subsequently, Graham *et al.* provided similarly strong evidence for association in four independent case–control cohorts from Sweden, Spain, Argentina and the United States[96]. Combined, in over 2,000 European Caucasian SLE cases and controls, the T allele of the SNP marker rs2004640, has an odds ratio of 1.47 at a high level of statistical significance. Interestingly, this allele is common in the Caucasian population as a whole, with a control allele frequency of 51%, and is enriched to 61% in SLE cases. Further study has revealed the existence of at least two distinct functional effects of the common *IRF5* risk haplotypes. First, a splice-site polymorphism encoded by rs2004640 results in the preferential use of an alternative first exon (exon 1B). Secondly, additional SNPs on this haplotype have a strong effect on the steady-state levels of *IRF5* mRNA. Interestingly, *IRF5* had been previously identified as having prominent *cis*-regulatory elements in a genome-wide screen for linkage and association with gene expression patterns in CEPH (Centre d'Etude du Polymorphisme Humain) cell lines[97]. Further work will be required to understand how both qualitative and quantitative differences in *IRF5* expression contribute to SLE.

Additional overlap between innate and adaptive pathways. Other intriguing cases of genetic variation acting at the overlap between innate and adaptive responses exist, although the extent to which such variants contribute to the overall burden of autoimmune disease is not yet clear. Complement components have roles in both adaptive and innate immunity, as PRRs (FIG. 1) and as mediators of immune effector mechanisms[98]. A variety of complement deficiencies have been associated with autoimmune diseases, particularly SLE[99]. SLE has also been associated with expression polymorphisms in the C-reactive protein response to inflammation, possibly through its role in the clearance of apoptotic cells[100] (FIG. 1). Polymorphisms in other PRRs such as the mannose-binding ligand have been associated with some autoimmune diseases[101,102], but convincing replication studies have not yet been published. Other members of the TLR family, such as TLR7, have recently been implicated in mouse models of autoimmunity[103,104], and offer intriguing candidate genes for follow-up studies in humans. Because the importance of innate immune mechanisms for human autoimmunity has only recently been appreciated, it seems likely that further genetic relationships with autoimmunity will be uncovered in the near future.

The future: multiple pathways, multistage models
It is becoming increasingly clear that multiple genes, which function in a variety of pathways, are involved in predisposition to autoimmune diseases, and several themes emerge from the current data. First, genetic variation in both the adaptive and innate immune systems is implicated, and these two systems interact. Second, there seem to be some genes, such as *PTPN22* and *CTLA4,* that are common underlying risk factors for multiple forms of autoimmune diseases. The fact that some autoimmune disorders do not share these associations indicates distinct underlying mechanisms in these diseases. The case of SLE illustrates a pattern in which a common underlying mechanism, in this case the dysregulation of interferon pathways, is modified by other genetic variants to result in a variety of phenotypic outcomes. As alluded to in the discussion on *PTPN22*, some alleles that are associated with autoimmune diseases are not present in certain ethnic groups, such as Asians. By contrast, there are examples of strong genetic associations in Asians[105] that are weak
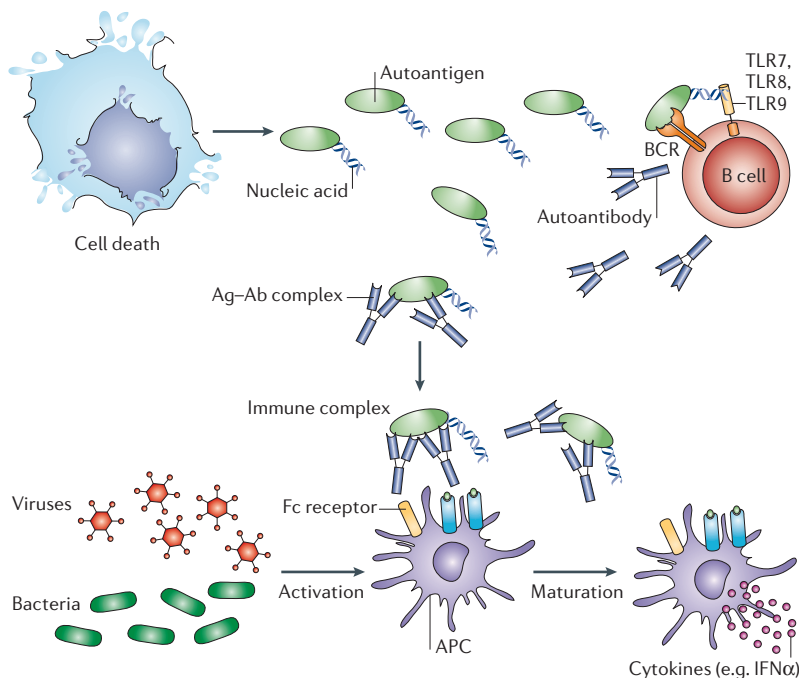

Figure 5 | **Model of innate immune responses in autoimmunity.** Innate immunity can be triggered by both endogenous and exogenous ligands, in cooperation with adaptive immune mechanisms (specific antibodies). Dying cells release particulate autoantigens containing protein and nucleic-acid fragments. Circulating B cells that have specificity for these antigens can be activated as a result of stimulation through the B-cell receptor (BCR) along with Toll-like receptors (TLR7, TLR8 and TLR9). This leads to further production of autoantibodies and immune complexes. These in turn might stimulate plasmacytoid dendritic cells, or other antigen-presenting cells (APCs) to produce type 1 interferons, such as interferon-α (IFNα). Similar triggering can occur directly through TLRs by microbial agents. These pathways are involved in the pathogenesis of human systemic lupus erythematosus.

in Caucasian populations[55]. It is currently unclear how much allelic and locus variability there is among different ethnic populations with regard to susceptibility genes for autoimmune disease. This might explain differences in disease manifestations and severity among these populations[106,107].

Most of the genetic variation that has been associated with autoimmune disease has been identified using candidate gene association studies. Such studies are plagued with problems of replication, mainly owing to inadequate sample sizes and a tendency to publish positive results without confirmation[108]. These are issues that will need to be tackled to achieve a wider and more reliable understanding of the genetic basis of autoimmunity. In addition, candidate gene studies are necessarily guided by assumptions about disease pathogenesis. We are now entering an era in which we can move beyond such assumptions and carry out more comprehensive searches for susceptibility genes, as genome-wide association studies are being extensively applied to the problem of complex disease, aided by resources such as the International HapMap Project[109] and rapidly evolving technologies for high-throughput SNP genotyping.

Importantly, although genome-wide association studies have had some notable recent successes[110], very large sample sizes and improved statistical methods will probably be required for a comprehensive identification of genes involved in autoimmunity. With the exception of the HLA alleles in the MHC, most genes are likely to have modest effects on disease risk. As discussed in this review, odds ratios for single risk alleles are generally less than two. It is also unclear to what extent rare alleles versus common alleles contribute to the overall genetic burden for autoimmunity in the population. So far, with the exception of the few Mendelian disorders with high penetrance, such as APS1 and IPEX, most risk alleles are common. If, however, large numbers of rare alleles contribute to 'sporadic' autoimmunity, it will be challenging to find them[111]. Another challenge for future studies is that nearly all of the genes involved in autoimmunity have been identified by using a univariate approach to association studies, without taking into account the gene–gene and gene–environment interactions that need to be considered in order to fully understand genetic risk. In terms of gene–gene interactions, this raises statistical problems that are related to multiple-hypothesis testing when carrying out gene discovery on a large scale[112]; a review of these issues has recently been published[113].

As highlighted at the beginning of this Review, autoimmune disorders have diverse phenotypes, even within a given disease category, and it is currently unclear to what degree genetics versus stochastic or environmental factors contribute to this diversity. With the exception of coeliac disease, in which exposure to dietary gluten is integral to pathogenesis, the environmental risk factors for autoimmunity are not well defined. The issue is further complicated by the fact that autoimmunity develops over time, and preclinical autoimmunity, generally detected in the form of circulating autoantibodies, can precede overt clinical disease by many years. This makes it particularly
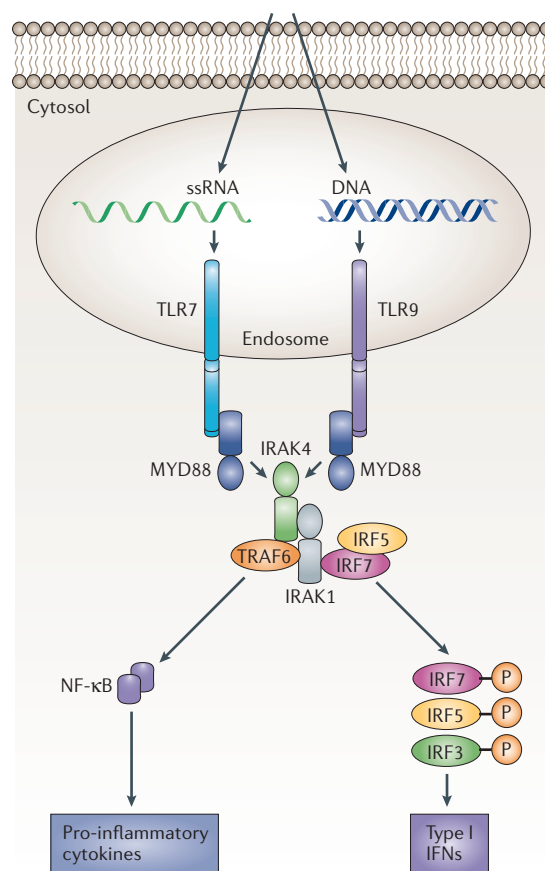
Penetrance
The proportion of individuals with a specific genotype who manifest the genotype at the phenotypic level. For example, if all individuals with a specific disease genotype show the disease phenotype, then the genotype is said to be 'completely penetrant'.

Multiple-hypothesis testing
Testing more than one hypothesis within an experiment. As a result, the probability of an unusual result from within the entire experiment occurring by chance is higher than the individual P-value associated with that result.

Figure 6 | **Activation of interferon production through TLR7 and TLR9 receptors.** A simplified schema shows activation pathways leading to the production of type I interferons (IFNs) in response to nucleic-acid fragments (exogenous or endogenous) through the Toll-like receptors TLR7 and TLR9 on plasmacytoid dendritic cells. Acting through the MYD88 adapter molecule, TLRs can activate the kinases IRAK4, IRAK1 and TRAF6, which in turn phosphorylate (P) transcription factors, especially the interferon regulatory factors IRF3, IRF5, and IRF7. The IRFs, together with additional transcription factors, upregulate transcription of type 1 interferons. Inflammatory pathways involving NF-κB can also be activated by TLRs. MYD88, myeloid differentiation factor 88; IRAK, interleukin-1 receptor-associated kinase; TRAF, tumour necrosis factor (TNF) receptor-associated factor.

difficult to identify relevant environmental exposures, because they might have occurred such a long time before disease onset. Nevertheless, environmental factors are beginning to be identified for some autoimmune disorders and integrated into models of pathogenesis. For example, smoking is an established risk factor for rheumatoid arthritis[114], and recent studies indicate that cigarette smoking might induce citrullination of proteins in pulmonary alveolar cells[115]. This is an intriguing observation because antibodies to citrullinated peptides are highly specific for rheumatoid arthritis[116], and the well-described HLA associations with this disease are closely related to the development of these autoantibodies[117]. Indeed, there is a strong

REVIEWS

interaction between these HLA alleles and smoking for the future development of rheumatoid arthritis[118], and these antibodies might be pathogenic[119]. Therefore, the emerging model indicates a need to place risk genes into several stages — risk genes that encode proteins involved in the enzymatic process of citrullination (one of these, which encodes the peptidyl arginine deiminase type IV (PADI4), seems to be a risk factor for rheumatoid arthritis[55,105]; risk genes involved in the development of these autoantibodies (HLA genes); and risk genes involved in the progression to clinical arthritis (possibly *PTPN22*). A similar multistage model is also applicable to type 1 diabetes, SLE and autoimmune thyroid diseases, in which autoantibodies predict individuals at risk for future disease[6–8], with progression to disease being dependent on unknown genetic, environmental, and presumably also stochastic factors. In view of the involvement of TLR pathways in lupus, viral infection is a likely environmental risk factor, and indeed there is now epidemiological support for a role for prior infection with Epstein–Barr virus[120].

Given these considerations, it will be important to put new genetic findings in the context of a longitudinal view that takes into account the various stages of disease development. This implies a need for the study of prospective cohorts of at-risk individuals, as well as subjects with disease, in order to fully understand how genetics affects the biology and pathogenesis of autoimmune diseases[121]. Indeed, it can be argued that longitudinal cohort studies will be the most important approach in the future, as they might lead to a practical multivariate way to carry out predictive testing and disease prevention. Currently, there are no good predictors for autoimmune disease. Even when autoantibodies are present in an asymptomatic individual, clinicians have no way of identifying who should receive preventive therapy. Similarly, there is a lack of biomarkers that can be used to predict outcome or response to therapy. Comprehensive studies of the genes involved in autoimmunity will at least partially contribute to addressing these pressing clinical problems.

1. Silverstein, A. M. Paul Ehrlich, archives and the history of immunology. *Nature Immunol.* **6**, 639 (2005).
2. Jacobson, D. L., Gange, S. J., Rose, N. R. & Graham, N. M. Epidemiology and estimated population burden of selected autoimmune diseases in the United States. *Clin. Immunol. Immunopathol.* **84**, 223–243 (1997).
3. Hallert, E., Husberg, M., Jonsson, D. & Skogh, T. Rheumatoid arthritis is already expensive during the first year of the disease (the Swedish TIRA project). *Rheumatology* **43**, 1374–1382 (2004).
4. Kobelt, G., Berg, J., Atherly, D. & Hadjimichael, O. Costs and quality of life in multiple sclerosis: a cross-sectional study in the United States. *Neurology* **66**, 1696–1702 (2006).
5. Ashwood, P., Wills, S. & Van de Water, J. The immune response in autism: a new frontier for autism research. *J. Leukoc. Biol.* **80**, 1–15 (2006).
6. Arbuckle, M. R. *et al.* Development of autoantibodies before the clinical onset of systemic lupus erythematosus. *N. Engl. J. Med.* **349**, 1526–1533 (2003).
7. Hoppu, S., Ronkainen, M. S., Kulmala, P., Akerblom, H. K. & Knip, M. GAD65 antibody isotypes and epitope recognition during the prediabetic process in siblings of children with type I diabetes. *Clin. Exp. Immunol.* **136**, 120–128 (2004).
8. Strieder, T. G., Prummel, M. F., Tijssen, J. G., Endert, E. & Wiersinga, W. M. Risk factors for and prevalence of thyroid disorders in a cross-sectional study among healthy female relatives of patients with autoimmune thyroid disease. *Clin. Endocrinol.* **59**, 396–401 (2003).
9. Hugot, J. P. *et al.* Association of NOD2 leucine-rich repeat variants with susceptibility to Crohn's disease. *Nature* **411**, 599–603 (2001).
10. Ogura, Y. *et al.* A frameshift mutation in *NOD2* associated with susceptibility to Crohn's disease. *Nature* **411**, 603–606. (2001).
11. Shiina, T., Inoko, H. & Kulski, J. K. An update of the HLA genomic region, locus information and disease associations: 2004. *Tissue Antigens* **64**, 631–649 (2004).
12. Criswell, L. A. *et al.* Analysis of families in the multiple autoimmune disease genetics consortium (MADGC) collection: the *PTPN22* 620W allele associates with multiple autoimmune phenotypes. *Am. J. Hum. Genet.* **76**, 561–571 (2005).
13. Michel, M. *et al.* Familial lupus erythematosus. Clinical and immunologic features of 125 multiplex families. *Medicine* **80**, 153–158 (2001).
14. Kyewski, B. & Klein, L. A central role for central tolerance. *Annu. Rev. Immunol.* **24**, 571–606 (2006).
15. Liston, A., Lesage, S., Gray, D. H., Boyd, R. L. & Goodnow, C. C. Genetic lesions in T-cell tolerance and thresholds for autoimmunity. *Immunol. Rev.* **204**, 87–101 (2005).

16. Sakaguchi, S., Setoguchi, R., Yagi, H. & Nomura, T. Naturally arising FOXP3-expressing CD25+CD4+ regulatory T cells in self-tolerance and autoimmune disease. *Curr. Top. Microbiol. Immunol.* **305**, 51–66 (2006).
References 14–16 provide a comprehensive discussion of the various mechanisms of T-cell tolerance. Sakaguchi, a pioneer in defining regulatory T cells, provides an up-to-date discussion of this fast-moving field.
17. Sakaguchi, S. & Sakaguchi, N. Regulatory T cells in immunologic self-tolerance and autoimmune disease. *Int. Rev. Immunol.* **24**, 211–226 (2005).
18. Apostolou, I. & von Boehmer, H. *In vivo* instruction of suppressor commitment in naive T cells. *J. Exp. Med.* **199**, 1401–1408 (2004).
19. Chen, W. *et al.* Conversion of peripheral CD4+CD25− naive T cells to CD4+CD25+ regulatory T cells by TGFβ induction of transcription factor FOXP3. *J. Exp. Med.* **198**, 1875–1886 (2003).
20. Hori, S., Nomura, T. & Sakaguchi, S. Control of regulatory T-cell development by the transcription factor FOXP3. *Science* **299**, 1057–1061 (2003).
21. Fontenot, J. D., Gavin, M. A. & Rudensky, A. Y. FOXP3 programs the development and function of CD4+CD25+ regulatory T cells. *Nature. Immunol.* **4**, 330–336 (2003).
22. Fontenot, J. D., Rasmussen, J. P., Gavin, M. A. & Rudensky, A. Y. A function for interleukin 2 in FOXP3-expressing regulatory T cells. *Nature. Immunol.* **6**, 1142–1151 (2005).
23. Bluestone, J. A. & Tang, Q. How do CD4+CD25+ regulatory T cells control autoimmunity? *Curr. Opin. Immunol.* **17**, 638–642 (2005).
24. Anderson, M. S. *et al.* Projection of an immunological self shadow within the thymus by the AIRE protein. *Science* **298**, 1395–1401 (2002).
This seminal paper demonstrates the importance of AIRE in regulating the thymic expression of self antigens and its role in autoimmunity.
25. Anderson, M. S. *et al.* The cellular mechanism of AIRE control of T-cell tolerance. *Immunity* **23**, 227–239 (2005).
26. Nagamine, K. *et al.* Positional cloning of the *APECED* gene. *Nature Genet.* **17**, 393–398 (1997).
27. Eisenbarth, G. S. & Gottlieb, P. A. Autoimmune polyendocrine syndromes. *N. Engl. J. Med.* **350**, 2068–2079 (2004).
28. Takase, H. *et al.* Thymic expression of peripheral tissue antigens in humans: a remarkable variability among individuals. *Int. Immunol.* **17**, 1131–1140 (2005).
29. Barratt, B. J. *et al.* Remapping the insulin gene/*IDDM2* locus in type 1 diabetes. *Diabetes* **53**, 1884–1889 (2004).
30. Pugliese, A. *et al.* The insulin gene is transcribed in the human thymus and transcription levels correlated with allelic variation at the *INS VNTR–IDDM2* susceptibility locus for type 1 diabetes. *Nature Genet.* **15**, 293–297 (1997).

31. 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).
32. Chentoufi, A. A. & Polychronakos, C. Insulin expression levels in the thymus modulate insulin-specific autoreactive T-cell tolerance: the mechanism by which the *IDDM2* locus may predispose to diabetes. *Diabetes* **51**, 1383–1390 (2002).
33. Kent, S. C. *et al.* Expanded T cells from pancreatic lymph nodes of type 1 diabetic subjects recognize an insulin epitope. *Nature* **435**, 224–228 (2005).
34. Giraud, M. *et al.* Transcriptional control of *CHRNA1* gene in thymus by a promoter SNP associated with onset of autoimmune myasthenia gravis. *Clinical Immunol.* **119**, 512 (2006).
35. Dechairo, B. M. *et al.* Association of the *TSHR* gene with Graves' disease: the first disease specific locus. *Eur. J. Hum. Genet.* **13**, 1223–1230 (2005).
36. Sakaguchi, N. *et al.* Altered thymic T-cell selection due to a mutation of the *ZAP70* gene causes autoimmune arthritis in mice. *Nature* **426**, 454–460 (2003).
37. Wildin, R. S., Smyk-Pearson, S. & Filipovich, A. H. Clinical and molecular features of the immunodysregulation, polyendocrinopathy, enteropathy, X linked (IPEX) syndrome. *J. Med. Genet.* **39**, 537–545 (2002).
38. Wildin, R. S. *et al.* X-linked neonatal diabetes mellitus, enteropathy and endocrinopathy syndrome is the human equivalent of mouse scurfy. *Nature Genet.* **27**, 18–20 (2001).
39. Brunkow, M. E. *et al.* Disruption of a new forkhead/winged-helix protein, scurfin, results in the fatal lymphoproliferative disorder of the scurfy mouse. *Nature Genet.* **27**, 68–73 (2001).
40. Khattri, R., Cox, T., Yasayko, S. A. & Ramsdell, F. An essential role for Scurfin in CD4+CD25+ T regulatory cells. *Nature Immunol.* **4**, 337–342 (2003).
41. Sakaguchi, S., Fukuma, K., Kuribayashi, K. & Masuda, T. Organ-specific autoimmune diseases induced in mice by elimination of T cell subset. I. Evidence for the active participation of T cells in natural self-tolerance; deficit of a T cell subset as a possible cause of autoimmune disease. *J. Exp. Med.* **161**, 72–87 (1985).
42. Sakaguchi, S. Naturally arising CD4+ regulatory T cells for immunologic self-tolerance and negative control of immune responses. *Annu. Rev. Immunol.* **22**, 531–562 (2004).
43. Setoguchi, R., Hori, S., Takahashi, T. & Sakaguchi, S. Homeostatic maintenance of natural FOXP3(+) CD25(+) CD4(+) regulatory T cells by interleukin (IL)-2 and induction of autoimmune disease by IL-2 neutralization. *J. Exp. Med.* **201**, 723–735 (2005).
44. Fehervari, Z., Yamaguchi, T. & Sakaguchi, S. The dichotomous role of IL-2: tolerance versus immunity. *Trends Immunol.* **27**, 109–111 (2006).

45. 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).
46. Keir, M. E. & Sharpe, A. H. The B7/CD28 costimulatory family in autoimmunity. *Immunol. Rev.* **204**, 128–143 (2005).
47. van Berkel, M. E. & Oosterwegel, M. A. CD28 and ICOS: similar or separate costimulators of T cells? *Immunol. Lett.* **105**, 115–122 (2006).
48. Chen, L. Co-inhibitory molecules of the B7-CD28 family in the control of T-cell immunity. *Nature Rev. Immunol.* **4**, 336–347 (2004).
49. Tivol, E. A. *et al.* Loss of CTLA-4 leads to massive lymphoproliferation and fatal multiorgan tissue destruction, revealing a critical negative regulatory role of CTLA-4. *Immunity* **3**, 541–547 (1995).
50. Waterhouse, P. *et al.* Lymphoproliferative disorders with early lethality in mice deficient in *CTLA-4*. *Science* **270**, 985–988 (1995).
51. Gough, S. C., Walker, L. S. & Sansom, D. M. *CTLA4* gene polymorphism and autoimmunity. *Immunol. Rev.* **204**, 102–115 (2005).
52. Yanagawa, T., Hidaka, Y., Guimaraes, V., Soliman, M. & DeGroot, L. J. *CTLA-4* gene polymorphism associated with Graves' disease in a Caucasian population. *J. Clin. Endocrinol. Metab.* **80**, 41–45 (1995).
53. Nistico, L. *et al.* The *CTLA-4* gene region of chromosome 2q33 is linked to, and associated with, type 1 diabetes. Belgian Diabetes Registry. *Hum. Mol. Genet.* **5**, 1075–1080 (1996).
54. Ueda, H. *et al.* Association of the T-cell regulatory gene *CTLA4* with susceptibility to autoimmune disease. *Nature* **423**, 506–511 (2003).
The first detailed, although still incomplete, analysis of the molecular basis of *CTLA* associations with autoimmune disease. See reference 51 for a more recent review.
55. Plenge, R. M. *et al.* Replication of putative candidate-gene associations with rheumatoid arthritis in > 4,000 samples from North America and Sweden: association of susceptibility with *PTPN22*, *CTLA4*, and *PADI4*. *Am. J. Hum. Genet.* **77**, 1044–1060 (2005).
56. Nishimura, H., Nose, M., Hiai, H., Minato, N. & Honjo, T. Development of lupus-like autoimmune diseases by disruption of the *PD-1* gene encoding an ITIM motif-carrying immunoreceptor. *Immunity* **11**, 141–151 (1999).
57. Nishimura, H. *et al.* Autoimmune dilated cardiomyopathy in PD-1 receptor-deficient mice. *Science* **291**, 319–322 (2001).
58. Prokunina, L. *et al.* A regulatory polymorphism in *PDCD1* is associated with susceptibility to systemic lupus erythematosus in humans. *Nature Genet.* **32**, 666–669 (2002).
59. Martin-Orozco, N. & Dong, C. New battlefields for costimulation. *J. Exp. Med.* **203**, 817–820 (2006).
60. Holets, L. M., Hunt, J. S. & Petroff, M. G. Trophoblast CD274 (B7-H1) is differentially expressed across gestation: influence of oxygen concentration. *Biol. Reprod.* **74**, 352–358 (2006).
61. Keir, M. E. *et al.* Tissue expression of PD-L1 mediates peripheral T cell tolerance. *J. Exp. Med.* **203**, 883–895 (2006).
62. Greenwald, R. J., Freeman, G. J. & Sharpe, A. H. The B7 family revisited. *Annu. Rev. Immunol.* **23**, 515–548 (2005).
63. Bottini, N., Vang, T., Cucca, F. & Mustelin, T. Role of PTPN22 in type 1 diabetes and other autoimmune diseases. *Semin. Immunol.* **18**, 207–213 (2006).
64. Wu, J. *et al.* Identification of substrates of human protein tyrosine phosphatase PTPN22. *J. Biol. Chem.* **281**, 11002–11010 (2006).
65. Bottini, N. *et al.* A functional variant of lymphoid tyrosine phosphatase is associated with type I diabetes. *Nature Genet.* **36**, 337–338 (2004).
66. Begovich, A. B. *et al.* A missense single-nucleotide polymorphism in a gene encoding a protein tyrosine phosphatase (*PTPN22*) is associated with rheumatoid arthritis. *Am. J. Hum. Genet.* **75**, 330–337 (2004).
67. Smyth, D. *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* **53**, 3020–3023 (2004).
68. Kyogoku, C. *et al.* Genetic association of the R620W polymorphism of protein tyrosine phosphatase *PTPN22* with human SLE. *Am. J. Hum. Genet.* **75**, 504–507 (2004).

69. Vandiedonck, C. *et al.* Association of the *PTPN22*R620W* polymorphism with autoimmune myasthenia gravis. *Ann. Neurol.* **59**, 404–407 (2006).
70. Vang, T. *et al.* Autoimmune-associated lymphoid tyrosine phosphatase is a gain-of-function variant. *Nature Genet.* **37**, 1317–1319 (2005).
71. Hasegawa, K. *et al.* PEST domain-enriched tyrosine phosphatase (PEP) regulation of effector/memory T cells. *Science* **303**, 685–689 (2004).
These two papers present contrasting evidence for the mechanisms by which *PTPN22* alleles predispose to autoimmunity. Vang *et al.* provide evidence for the gain of function of *PTPN22* risk alleles, whereas studies in knockout mice (Hasegawa *et al.*) indicate that *PTPN22* has predominantly negative regulatory effects in T cells.
72. Gregersen, P. K., Lee, H. S., Batliwalla, F. & Begovich, A. B. PTPN22: setting thresholds for autoimmunity. *Semin. Immunol.* **18**, 214–223 (2006).
73. Siminovitch, K. A. PTPN22 and autoimmune disease. *Nature Genet.* **36**, 1248–1249 (2004).
74. Gregersen, P. K. Gaining insight into PTPN22 and autoimmunity. *Nature Genet.* **37**, 1300–1302 (2005).
75. Begovich, A. B. *et al.* The R620W polymorphism of the protein tyrosine phosphatase *PTPN22* is not associated with multiple sclerosis. *Am. J. Hum. Genet.* **76**, 184–187 (2005).
76. van Oene, M. *et al.* Association of the lymphoid tyrosine phosphatase R620W variant with rheumatoid arthritis, but not Crohn's disease, in Canadian populations. *Arthritis Rheum.* **52**, 1993–1998 (2005).
77. Carlton, V. E. *et al.* PTPN22 genetic variation: evidence for multiple variants associated with rheumatoid arthritis. *Am. J. Hum. Genet.* **77**, 567–581 (2005).
78. Marchini, J., Cardon, L. R., Phillips, M. S. & Donnelly, P. The effects of human population structure on large genetic association studies. *Nature Genet.* **36**, 512–517 (2004).
79. Seldin, M. *et al.* European population substructure: clustering of northern and southern populations. *PLoS Genet.* **2**, 1339–1351 (2006).
80. Price, A. L. *et al.* Principal components analysis corrects for stratification in genome-wide association studies. *Nature Genet.* **38**, 904–909 (2006).
81. Nimmerjahn, F. & Ravetch, J. V. Fcγ receptors: old friends and new family members. *Immunity* **24**, 19–28 (2006).
82. Takai, T. Fc receptors and their role in immune regulation and autoimmunity. *J. Clin. Immunol.* **25**, 1–18 (2005).
83. Su, K. *et al.* A promoter haplotype of the immunoreceptor tyrosine-based inhibitory motif-bearing FcγRIIb alters receptor expression and associates with autoimmunity. II. Differential binding of GATA4 and Yin-Yang1 transcription factors and correlated receptor expression and function. *J. Immunol.* **172**, 7192–7199 (2004).
84. Blank, M. C. *et al.* Decreased transcription of the human *FCGR2B* gene mediated by the-343 G/C promoter polymorphism and association with systemic lupus erythematosus. *Hum. Genet.* **117**, 220–227 (2005).
85. Chadha, S. *et al.* Haplotype structure of *TNFRSF5–TNFSF5* (*CD40–CD40L*) and association analysis in systemic lupus erythematosus. *Eur. J. Hum. Genet.* **13**, 669–676 (2005).
86. Sontheimer, R. D., Racila, E. & Racila, D. M. C1Q: its functions within the innate and adaptive immune responses and its role in lupus autoimmunity. *J. Invest. Dermatol.* **125**, 14–23 (2005).
87. Akira, S., Uematsu, S. & Takeuchi, O. Pathogen recognition and innate immunity. *Cell* **124**, 783–801 (2006).
88. Meylan, E., Tschopp, J. & Karin, M. Intracellular pattern recognition receptors in the host response. *Nature* **442**, 39–44 (2006).
89. 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).
90. Leadbetter, E. A. *et al.* Chromatin-IgG complexes activate B cells by dual engagement of IgM and Toll-like receptors. *Nature* **416**, 603–607 (2002).
This groundbreaking study demonstrates that TLRs can co-stimumlate B-cell activation in the absence of T-cell help.
91. Rifkin, I. R., Leadbetter, E. A., Busconi, L., Viglianti, G. & Marshak-Rothstein, A. Toll-like receptors, endogenous ligands, and systemic autoimmune disease. *Immunol. Rev.* **204**, 27–42 (2005).

92. Baechler, E. C. *et al.* Interferon-inducible gene expression signature in peripheral blood cells of patients with severe lupus. *Proc. Natl Acad. Sci. USA* **100**, 2610–2615 (2003).
93. Ronnblom, L., Eloranta, M. L. & Alm, G. V. The type I interferon system in systemic lupus erythematosus. *Arthritis Rheum.* **54**, 408–420 (2006).
94. Baechler, E. C., Gregersen, P. K. & Behrens, T. W. The emerging role of interferon in human systemic lupus erythematosus. *Curr. Opin. Immunol.* **16**, 801–807 (2004).
95. Sigurdsson, S. *et al.* Polymorphisms in the tyrosine kinase 2 and interferon regulatory factor 5 genes are associated with systemic lupus erythematosus. *Am. J. Hum. Genet.* **76**, 528–537 (2005).
96. Graham, R. R. *et al.* A common haplotype of interferon regulatory factor 5 (*IRF5*) regulates splicing and expression and is associated with increased risk of systemic lupus erythematosus. *Nature Genet.* **38**, 550–555 (2006).
These two studies provide definitive evidence for the involvement of *IRF5* in genetic susceptibility to SLE, therefore linking TLR signalling with abnormalities of interferon regulation in this disease (also see reference 93 for a review).
97. Morley, M. *et al.* Genetic analysis of genome-wide variation in human gene expression. *Nature* **430**, 743–747 (2004).
98. Thurman, J. M. & Holers, V. M. The central role of the alternative complement pathway in human disease. *J. Immunol.* **176**, 1305–1310 (2006).
99. Pickering, M. C., Botto, M., Taylor, P. R., Lachmann, P. J. & Walport, M. J. Systemic lupus erythematosus, complement deficiency, and apoptosis. *Adv. Immunol.* **76**, 227–324 (2000).
100. Russell, A. I. *et al.* Polymorphism at the C-reactive protein locus influences gene expression and predisposes to systemic lupus erythematosus. *Hum. Mol. Genet.* **13**, 137–147 (2004).
101. Werth, V. P., Berlin, J. A., Callen, J. P., Mick, R. & Sullivan, K. E. Mannose binding lectin (*MBL*) polymorphisms associated with low MBL production in patients with dermatomyositis. *J. Invest. Dermatol.* **119**, 1394–1399 (2002).
102. Kravitz, M. S., Pitashny, M. & Shoenfeld, Y. Protective molecules — C-reactive protein (CRP), serum amyloid P (SAP), pentraxin3 (PTX3), mannose-binding lectin (MBL), and apolipoprotein A1 (Apo A1), and their autoantibodies: prevalence and clinical significance in autoimmunity. *J. Clin. Immunol.* **25**, 582–591 (2005).
103. Subramanian, S. *et al.* A *Tlr7* translocation accelerates systemic autoimmunity in murine lupus. *Proc. Natl Acad. Sci. USA* **103**, 9970–9975 (2006).
104. Pisitkun, P. *et al.* Autoreactive B cell responses to RNA-related antigens due to *TLR7* gene duplication. *Science* **312**, 1669–1672 (2006).
These two recent studies provide direct evidence for the involvement of TLRs (*Tlr7*) in mouse models of autoimmunity, setting the stage for a further investigation of the role of *TLR7* in humans, and further supporting the involvement of innate immune mechanisms in the development of autoimmunity.
105. Suzuki, A. *et al.* Functional haplotypes of *PADI4*, encoding citrullinating enzyme peptidylarginine deiminase 4, are associated with rheumatoid arthritis. *Nature Genet.* **34**, 395–402 (2003).
106. Alarcon, G. S. *et al.* Systemic lupus erythematosus in a multi-ethnic cohort (LUMINA) XXXII: [corrected] contributions of admixture and socioeconomic status to renal involvement. *Lupus* **15**, 26–31 (2006).
107. Krishnan, E. & Hubert, H. B. Ethnicity and mortality from systemic lupus erythematosus in the US. *Ann. Rheum. Dis.* **65**, 1500–1505 (2006).
108. Ioannidis, J. P., Ntzani, E. E., Trikalinos, T. A. & Contopoulos-Ioannidis, D. G. Replication validity of genetic association studies. *Nature Genet.* **29**, 306–309 (2001).
109. International HapMap Consortium. A haplotype map of the human genome. *Nature* **437**, 1299–320 (2005).
110. Klein, R. J. *et al.* Complement factor H polymorphism in age-related macular degeneration. *Science* **308**, 385–389 (2005).
111. Pritchard, J. K. & Cox, N. J. The allelic architecture of human disease genes: common disease-common variant. or not? *Hum. Mol. Genet.* **11**, 2417–2423 (2002).
112. Hirschhorn, J. N. & Daly, M. J. Genome-wide association studies for common diseases and complex traits. *Nature Rev. Genet.* **6**, 95–108 (2005).

REVIEWS

113. Wang, W. Y., Barratt, B. J., Clayton, D. G. & Todd, J. A. Genome-wide association studies: theoretical and practical concerns. *Nature Rev. Genet.* **6**, 109–118 (2005).
114. Costenbader, K. H., Feskanich, D., Mandl, L. A. & Karlson, E. W. Smoking intensity, duration, and cessation, and the risk of rheumatoid arthritis in women. *Am. J. Med.* **119**, 503.e1–503.e9 (2006).
115. Klareskog, L. *et al.* A new model for an etiology of rheumatoid arthritis: smoking may trigger HLA-DR (shared epitope)-restricted immune reactions to autoantigens modified by citrullination. *Arthritis Rheum.* **54**, 38–46 (2006).
    This report provides a compelling recent example of a gene–environment interaction in the development of autoimmunity.
116. Schellekens, G. A. *et al.* The diagnostic properties of rheumatoid arthritis antibodies recognizing a cyclic citrullinated peptide. *Arthritis Rheum.* **43**, 155–163 (2000).
117. Irigoyen, P. *et al.* Regulation of anti-cyclic citrullinated peptide antibodies in rheumatoid arthritis: contrasting effects of HLA–DR3 and the shared epitope alleles. *Arthritis Rheum.* **52**, 3813–3818 (2005).
118. Berglin, E. *et al.* A combination of autoantibodies to cyclic citrullinated peptide (CCP) and HLA–DRB1 locus antigens is strongly associated with future onset of rheumatoid arthritis. *Arthritis Res. Ther.* **6**, R303–R308 (2004).
119. Kuhn, K. A. *et al.* Antibodies against citrullinated proteins enhance tissue injury in experimental autoimmune arthritis. *J. Clin. Invest.* **116**, 961–973 (2006).
120. James, J. A. *et al.* Systemic lupus erythematosus in adults is associated with previous Epstein–Barr virus exposure. *Arthritis Rheum.* **44**, 1122–1126 (2001).
121. Manolio, T. A., Bailey-Wilson, J. E. & Collins, F. S. Genes, environment, and the value of prospective cohort studies. *Nature Rev. Genet.* **7**, 812–820 (2006).
122. Vyse, T. J. & Todd, J. A. Genetic analysis of autoimmune disease. *Cell* **85**, 311–318 (1996).
123. Redondo, M. J. & Eisenbarth, G. S. Genetic control of autoimmunity in type I diabetes and associated disorders. *Diabetologia* **45**, 605–622 (2002).
124. Ban, Y. & Tomer, Y. Susceptibility genes in thyroid autoimmunity. *Clin. Dev. Immunol.* **12**, 47–58 (2005).
125. Barcellos, L. F. *et al.* HLA–DR2 dose effect on susceptibility to multiple sclerosis and influence on disease course. *Am. J. Hum. Genet.* **72**, 710–716 (2003).
126. Graham, R. R. *et al.* Visualizing human leukocyte antigen class II risk haplotypes in human systemic lupus erythematosus. *Am. J. Hum. Genet.* **71**, 543–553 (2002).
127. Hall, F. C. *et al.* Influence of the HLA–DRB1 locus on susceptibility and severity in rheumatoid arthritis. *QJM* **89**, 821–829 (1996).
128. Prokunina, L. *et al.* Association of the PD-1.3A allele of the *PDCD1* gene in patients with rheumatoid arthritis negative for rheumatoid factor and the shared epitope. *Arthritis Rheum.* **50**, 1770–1773 (2004).

DATABASES
The following terms in this article are linked online to:
Entrez Gene: http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=gene
AIRE | CD28 | CSK | CTL4 | FCGR2B | FOXP3 | ICOS | IL2 | *IRF5* | LCK | NOD2 | PADI4 | PD1 | PDL1 | PDL2 | PTPN22 | TLR9 | *TSHR* | ZAP70
OMIM: http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=OMIM
Addison disease | APS1 | Crohn disease | Graves disease | IPEX | myasthenia gravis | SLE | type 1 diabetes | rheumatoid arthritis

FURTHER INFORMATION
International HapMap Project: http://www.hapmap.org
Access to this links box is available online.