Review

The genetics of SLE: an update in the light of genome-wide association studies

B. Rhodes¹ and T. J. Vyse¹

Understanding the pathogenesis of SLE remains a considerable challenge. Multiple abnormalities of both the innate and adaptive immune system have been described and, furthermore, immunological dysfunction precedes clinical presentation by many years. There is a strong genetic basis to SLE, which means that genetic studies can play a key role in furthering our understanding of this disease. Since susceptibility variants are present from birth and are unaffected by the course of the disease, or by its treatment, genetic analysis is, perhaps uniquely, capable of identifying fundamental, causative, disease mechanisms. Over the last 12 months, there has been a staggering increase in our understanding of SLE genetics. We have seen the identification of new and important SLE susceptibility genes through candidate gene studies, and we have seen the publication of two whole-genome association analyses. The ‘hypothesis free’ whole-genome studies have provided additional evidence in support of a number of existing susceptibility genes and have identified novel gene candidates. In this article, we review the current SLE genetics literature in the light of these recent advances and we discuss our current understanding of the functional role of the key susceptibility genes. By considering how these genes fall into clusters with shared function we can begin to understand how dysregulation at a number of key immunological steps may predispose to the development of SLE.

KEY WORDS: Systemic lupus erythematosus, Genetics, Whole-genome association.

Introduction

Systemic lupus erythematosus (SLE) is the archetypal systemic autoimmune disease. It is characterized by a diverse array of clinical symptoms, indicative of widespread immune-mediated damage. It is also a heterogeneous disease, presenting differently from patient to patient, and with no single clinical or immunological feature required to make a formal diagnosis [1]. Multiple components of both the innate and adaptive immune systems contribute to disease pathology and furthermore it is likely that immunological dysfunction precedes the onset of clinical disease by many years, making it a particularly challenging disease to study [2].

For at least 30 yrs we have known that there is a strong genetic component to SLE: a disease concordance of 2–5% in dizygotic compared with 24–57% in monozygotic twins and a sibling risk ratio (λs) of 20–29 [3–6]. Progress in identifying these genetic factors was initially slow. A scan of review articles from 2002 and 2003 shows that we had evidence of a role for the MHC and Fcγ receptor genes, but none of the other common gene variants currently recognized as being associated with greatest SLE risk had been identified [7, 8]. The publication of the human genome sequence in 2001 stimulated renewed interest in genetics and a general increase in the availability of funding for genetic studies across a range of diseases [9]. However, there were also critics of this ‘genohype’, arguing that the high costs involved would be better spent on other preventative or therapeutic measures, that it would be too difficult to dissect out the genes contributing to complex genetic disease, or, even if we were successful, that this knowledge would be of little practical benefit to patients [10]. Despite these voices of caution it has become clear that in a disease such as SLE, with a strong genetic basis and a complex immunological pathogenesis, genetic analysis is an immensely powerful tool. By the time a patient has overt disease it is hard to distinguish primary, causative, abnormalities from a multitude of secondary, downstream events. Disease-associated genetic variants are carried by patients from birth, and are therefore fundamental abnormalities unaffected by the disease process itself. It is analogous to investigating an engineering failure by reviewing thousands of wrecked parts from a collapsed structure. Ironically, the hypothesis-free nature of many genetic studies, criticized by some as ‘bad science’, may actually provide the key to unravelling SLE pathogenesis.

The last 5 yrs has seen a dramatic increase in our understanding of SLE genetics, and this has accelerated over the last 12 months with the publication of two high-density genome-wide association (GWA) analyses and the identification of several novel candidate genes through fine-mapping studies. In this review, we will discuss our current understanding of SLE as a complex genetic disorder. We will then review the current list of best substantiated SLE disease susceptibility genes, in particular considering how the known and potential function of these genes may allow us to formulate a ‘genetic hypothesis’ of SLE pathogenesis.

SLE genetic model

The inheritance of SLE does not follow simple Mendelian rules as we would expect for a single major gene effect, instead a polygenic model of susceptibility provides the best explanation for the familial clustering. In few cases, SLE is associated with rare but highly penetrant mutations, for example, homozygous deficiency of the complement components C1q, C2 or C4, complete FcγRIIIb deficiency or mutations in the DNA exonuclease TREX1 [11–14]. In the majority of cases, however, SLE genetic susceptibility is probably also determined by relatively common variants that are found throughout the population, each of which only contributes modestly to disease risk; this has been termed the ‘common disease, common variant’ hypothesis [15]. Data from human SLE linkage scans, and from the genetic dissection of mouse models of SLE, support this theory. These have been thoroughly reviewed elsewhere [16–18]. The four major groups to have performed

¹Section of Molecular Genetics and Rheumatology, Division of Medicine, Imperial College London, London, UK.

Submitted 29 April 2008; revised version accepted 6 June 2008.

Correspondence to: T. J. Vyse, Section of Molecular Genetics of Rheumatology, Faculty of Medicine, Imperial College London, Du Cane Road, London W12 0NN, UK. E-mail: t.vyse@imperial.ac.uk

© The Author 2008. Published by Oxford University Press on behalf of the British Society for Rheumatology. All rights reserved. For Permissions, please email: journals.permissions@oxfordjournals.org
human genome-wide linkage scans have identified at least nine independently replicated SLE linkage regions. Analyses conditioned on specific disease traits suggest that genetic effects arising from particular linkage regions may contribute to specific clinical or immunological features of SLE (i.e. the presence of haemolytic anaemia or the production of dsDNA antibodies). A similar picture has arisen from the study of mouse models. Studies also demonstrate that specific traits, such as autoantibody production or complement depletion, which are associated with SLE, but are not sufficient to make a diagnosis in isolation, may also be observed in the healthy relatives of patients of SLE, and indeed may be heritable within these families whether or not the family members actually have the disease [19, 20].

The model that emerges is therefore one in which SLE susceptibility is determined by multiple immunological abnormalities arising from genetic variation at multiple loci. Most of these variants are much more common than the prevalence of SLE in the population, so the immunological changes they cause, although sometimes measurable in isolation, are clearly insufficient to cause SLE per se. Presumably, a degree of inherited immunological dysregulation can be buffered by immune redundancy, regulatory pathways or even genetically determined protective effects. Although these genetic factors influence an individual’s lifetime risk of developing SLE, environmental factors are likely to provide the trigger for disease onset. The nature of these environmental triggers is largely unknown, although there is some evidence of an aetiological role for the EBV [21]. The environmental trigger is unlikely to be a rare and highly penetrant event; otherwise we would observe dramatic clustering of cases among individuals with the relevant exposure. Equally, the trigger is unlikely to be a single, common, environmental insult since SLE is not a highly prevalent disease. Perhaps the most likely scenario requires a number of triggers occurring together or sequentially over a limited period of time. The concept has therefore emerged of ‘threshold liability’ in which disease develops when a threshold of genetic and environmental susceptibility effects is reached [22]. Genetic liability is determined from birth by the complement of inherited susceptibility genes. Whether the disease threshold is then reached depends on the environmental influences; for those with many susceptibility genes only a minor environmental trigger may be required; for those with little genetic risk, disease may never develop despite strong or prolonged exposure to the relevant environmental triggers.

Key candidate genes

It is difficult to predict how many genes contribute to SLE susceptibility. The variants that are easiest to identify are those that are both reasonably common and also have a reasonably strong effect. Given the number of genes identified so far, and the density of genetic coverage provided by recent studies (which, although impressive, is not comprehensive enough to map all variation) we would estimate that there are roughly 20–50 of these. Many more variants may increase SLE susceptibility by variation (an indirect association). Until recently it has only been technologically and financially possible to genotype about 100 SNPs in a single project. Linkage disequilibrium usually operates over short physical distances, so this has restricted studies to a single gene or gene cluster, selected for their biological plausibility and/or position within a linkage region. Technological advances have now made it possible to simultaneously type hundreds of thousands of SNPs across the whole genome. This GWA analysis combines the power and resolution of a conventional association study with the hypothesis-free methodology of a linkage scan. Apart from the technological achievement, GWA studies have also been a product of unparalleled intellectual collaboration, with competitive investigators pooling patient cohorts to form multinational collaborations.

High-density whole-genome association scans have now been performed in a number of autoimmune diseases including RA and Type 1 diabetes mellitus (TID), and lower density scans in diseases such as AS and autoimmune thyroid disease [23–26]. These have been highly productive studies, resulting in the identification of novel gene candidates for each of these diseases.

In SLE, two high-density case–control GWA analyses have been published to date and a third, family-based study, is expected shortly. The first was the product of an European/US collaboration, The International Consortium for Systemic Lupus Erythematosus Genetics (SLEGEN) [27]. The initial study used a cohort of 720 SLE cases and 2337 controls; with replication of top hits in additional cohorts [28]. This study used the Illumina HumanHap 300 BeadChip, resulting in the successful genotyping of over 300 000 SNPs. The second study, published the same month, used a cohort of 1311 cases and 1783 controls, again with replication of top hits in additional cohorts [28]. This study used the Illumina HumanHap550 bead chip, resulting in over 500 000 SNPs successfully genotyped. Although impressive, it should be noted that the SNP densities achieved by these chips are not dense enough to achieve definitive genome-wide coverage. It should also be noted that the cohorts employed in these studies were of white, northern European ancestry, and thus represent only a small proportion of the human population. Population-specific data, such as that collected by the HapMap project reveal marked differences in SNP frequency and genetic architecture. In particular, the more ancient African genome appears to be genetically more diverse, and show evidence of more recombination than populations of European or East Asian ancestry [29]. The study of cohorts of non-European ancestry may therefore reveal additional novel and important SLE susceptibility variants.

The key gene discoveries are summarized in Table 1. These are categorized according to the current level of evidence: we have defined overwhelming evidence as genome-wide significance in both GWA studies; strong evidence as a high level of significance in one of the GWA studies supported by evidence from a meta-analysis of fine-mapping studies; good evidence as a high level of significance in one of the GWA studies or high level significance in a fine-mapping study in at least two independent cohorts. Genes requiring further work have either lower levels of significance attached to them, or the literature is inconsistent (inconsistent effect or inconsistencies over the key polymorphisms). We now consider the most important genes in detail. We will not discuss...
The genetics of SLE

The SLEGEN GWA analysis has now provided evidence of independent signals arising from both Class II and Class III regions, with markers from both regions being retained in a model choice analysis [27]. Recent data from a mapping study in 314 white European SLE families has refined these associations [32]. The first signal arose from a small 180 kb Class II region containing HLA-DRB1, HLA-DQA1 and HLA-DQB1 only. A clear second signal was observed for a marker in the Class III gene SKIV2L [supperkiller viriladic activity 2-like (Saccharomyces cerevisiae)]. By reference to the linkage disequilibrium around this marker it is likely that the true functional variant associated with this second signal is either in SKIV2L itself, or in a 40 kb interval also containing the genes CFB (complement factor B), RDBP (RD RNA binding protein), DOM3Z (dom-3 homolog Z (Caenorhabditis elegans)) and STK19 (serine–threonine kinase 19). It is notable that this study gives evidence against a functional effect from TNF-308A.

Further work will be required to determine the effect arising from C4A. C4 is encoded by two genes, C4A and C4B, which have minor sequence differences. The resulting proteins have different functional characteristics, with C4A better able to bind immune complexes [33]. The C4 genes are inherited in a discrete ‘RCCX module’, which contains one C4 gene (either C4A or C4B) along with three neighbouring genes. The Class III MHC carries between one and four copies of this module; hence each diploid genome has between two and eight C4 genes, which may be either C4A or C4B [34]. Carrying less than two copies of C4A has been identified as a risk factor for SLE [34]. MHC sequencing, however, reveals that the common AH81 (DR3), SLE associated, haplotype contains a monomodular RCCX unit with a single C4B gene and no C4A [35]. It is therefore possible that the observed C4A association is present simply due to linkage disequilibrium between the gene copy number and functional variants in other MHC genes. We will only be able to answer this question when we have a comprehensive analysis incorporating both C4A copy number analysis and a dense MHC SNP map.

**ITGAM**

The identification of ITGAM as a major susceptibility gene was perhaps the greatest surprise of the GWA analyses; not because it lacks biological plausibility, but because it has been subject to expression studies in the past with little convincing evidence for a role in SLE [36]. ITGAM encodes the integrin αM protein that dimerizes with integrin β2 to form the cell surface receptor known as complement receptor 3 (CR3) or Mac-1. CR3 is expressed primarily on neutrophils, macrophages and dendritic cells and can bind a variety of ligands including intercellular adhesion molecule 1 (ICAM-1) and the C3bi fragment of activated complement C3. An initial fine-mapping study of ITGAM in populations of European and African ancestry was published alongside the two GWA studies [37]. The proposed candidate functional mutation results in an arginine/histidine amino acid switch at position 77 within the β-propeller domain. This amino acid does not lie within any known ligand binding site, but may alter the confirmation of the I/A domain to which many ligands do bind. This variant could therefore influence leucocyte trafficking mediated via ICAM-1, or equally it could influence the CR3-mediated uptake of apoptotic cells or immune complexes. Functional data is awaited with interest.

**IRF5**

Even before the GWA studies, there was growing evidence to implicate IRF5 as an important SLE susceptibility gene. Interest in type 1 (IFN-α and -β) IFN pathways was stimulated by the discovery that there is a general up-regulation IFN-inducible genes.
in SLE [38, 39]. IRF5 itself encodes an IFN-induced transcription factor, responsible for mediating expression of many of these genes. The genetics of IRF5 expression is complex and has been reviewed elsewhere [40]. The best current genetic model proposes an SLE risk haplotype carrying multiple functional mutations. In vitro functional evidence exists for at least two of these. Polymorphism at SNP rs2004640 creates a novel splice site in exon 1B allowing the expression of a novel IRF5 isoform [41, 42]. A second mutation rs10954213 located in the 3′ UTR creates a functional polyadenylation site and hence a shorter and more stable gene transcript [43]. The genetic effect arising from IRF5 may therefore have both quantitative and qualitative components. Adding to the complexity at this locus are reports proposing functional insertion–deletion polymorphisms in exon 6 and upstream of exon 1A, and evidence for differing SLE susceptibility effects across ethnic populations [40, 44–47]. As fine-mapping proceeds with other major susceptibility genes it will be interesting to see whether the presence of multiple functional mutations at each locus proves to be the exception or the rule.

To understand how IRF5 variants may predispose to SLE we need to understand the physiological role of the Type 1 IFN pathways. The majority of cells produce Type 1 IFNs as part of their early response to viral infection. Particularly large amounts are produced by plasmacytoid dendritic cells, perhaps stimulated by the recognition of viral RNA and DNA through TLR7 and TLR9. In SLE, it is possible that this is also triggered in response to inadequately cleared nucleic acid antigens released from apoptotic cells [48]. Type 1 IFNs exert a multitude of downstream effects on the immune system. Perhaps critically, they stimulate Th1 pathways and sustain activated T cells, while also lowering the threshold for B-cell activation through the B-cell receptor (BCR) and promoting B-cell survival and differentiation [48–51]. It can therefore be seen that genetic variants that prolong or alter the actions of IRF5 could result in a prolonged pro-inflammatory response, and potentially break immunological tolerance. Interestingly, IRF5 signalling has also been shown to play a role in the regulation of cell cycle and apoptosis raising the possibility that susceptibility variants of IRF5 exert their effects at multiple levels [52].

**STAT4**

The association between signal transducer and activator of transcription 4 (STAT4) variation and SLE risk was initially reported in 2007 from a case–control association study [53]. This was subsequently confirmed in both GWA studies. No functional candidate polymorphism has yet been clearly established, so further study is required. STAT4 encodes a transcription factor that mediates the expression of genes in a number of key immunological pathways; the overall effects are likely to be complex. STAT4 is known to mediate signals induced by immunologically relevant cytokines including, like IRF5, the Type 1 IFNs, but also IL-12 and -23 [54, 55]. In response to these cytokines, STAT4 activation plays an important role in directing a Th1 T-cell response, and mediates the production of Th1-type cytokines such as IFN-γ [56–58]. STAT4 may also play a role in the differentiation of the potentially pathogenic Th17 T-cell subset [59]. In addition, STAT4 signalling also mediates type 1 IFN signalling in antigen-presenting cells, and may be necessary for the production of IFN-γ by these cells [60, 61].

**BLK, BANK1 and LYN**

The identification that three key genes involved in BCR signalling are all strongly associated with SLE attests to the importance of this pathway in disease pathogenesis. B-lymphoid tyrosine kinase (BLK) was one of the top hits in both GWA analyses, while LYN was associated with high significance in the SLEGEN study only. B-cell scaffold protein with ankyrin repeats (BANK1) was identified as a potential candidate following an early low-density genome-wide study and this was confirmed by higher density mapping in additional case-control cohorts [62]. No functional data exist for LYN. For BLK, preliminary data suggest that the risk allele is associated with reduced BLK transcription, although further fine-mapping is required [28]. For BANK1 a number of potentially functional SNPs were identified, but best functional evidence was found for rs17266594, which altered a branch point upstream of exon 2, resulting in the generation of a novel short isoform, but little quantitative difference in BANK1 expression overall [62].

All three of these genes play a critical role in controlling the activation of B cells following signalling through the BCR. Following ligand binding and BCR aggregation, an early intracellular event is the recruitment and activation of Src-family protein tyrosine kinases, including BLK and LYN, which mediate further intracellular signalling [63]. The exact role of these kinases in determining cellular events has yet to be determined with certainty. The best characterized functionally is LYN. Activated LYN mediates B-cell activation by phosphorylating the ITAM domain of the BCR-associated Igα/β signalling molecules, in turn recruiting and activating the tyrosine kinase SYK, which initiates multiple activating signals. LYN also mediates inhibitory signals by phosphorylating inhibitory receptors such as CD22 and FcγRIIB and may therefore have a critical role as a modulator of B-cell activation thresholds [64]. Lyn knockout mice develop a lupus-like autoimmune disease [65]. The role of BLK is less well understood and Blk knockout mice appear to have a relatively normal B-cell phenotype [66]. BANK1 is a scaffold protein that acts downstream of LYN by mediating an interaction between LYN and the IP(3)R calcium channel, thus facilitating release of intracellular calcium as an important event in B-cell activation [67].

**PTPN22**

Protein tyrosine phosphatase non-receptor 22 (PTPN22) has received a great deal of attention as an important susceptibility gene for RA and a number of other autoimmune diseases. However, it is also an important candidate gene in SLE, having been identified with high levels of significance in the SLEGEN GWAS analysis. Prior to this a number of candidate gene studies found an association between SNP rs2476601 and SLE, and this has been confirmed in a meta-analysis [68]. The lymphoid tyrosine phosphatase protein (LYP), which is encoded by PTPN22, is involved in the down-regulation of T-cell activation through its interaction with Csk; this interaction is prevented by the arginine to tryptophan amino acid substitution consequent upon the associated mutation [69, 70]. Intuitively one would expect this to result in increased T-cell signalling and activation; however, experimental evidence suggests the opposite with TCR signalling actually reduced in cells carrying the tryptophan variant protein [71]. A number of explanations have been proposed for this paradox including an effect of the mutation on the tyrosine phosphatase activity of LYP, or an effect on the binding of other ligands or the conformation of LYP in response to these ligands [72]. At a cellular level the mechanism by which reduced T-cell activation may actually increase the potential for autoimmunity remains a matter for speculation, although the suppression of regulatory T-cells is a possibility [72].

**FCGR2A and FCGR3B**

SNPs at each of the low-affinity Fcγ-receptor genes have been associated with SLE, although early reports are from smaller studies and results are often inconsistent [73]. Best genetic evidence is for a missense mutation at FCGR2A. FcγIIa is expressed on neutrophils, monocytes, macrophages, dendritic cells and platelets. The most widely described polymorphism, rs1801274, results in a histidine(H) to arginine(R) switch at position 131 within the ligand binding domain of FcγRIIa. The 131R variant has a lower binding affinity for its primary
lupus nephritis (OR 2.43) and SLE without renal involvement (OR 2.21) [14].

The Fcγ receptor region on chromosome 1q23-24 is the high level of sequence similarity between each of the Fcγ-receptor genes. This presents a problem in designing genotyping assays specific to the gene of interest and may lead to SNPs and associations being incorrectly assigned to the wrong gene. Another complicating factor at this locus is the presence of copy number variation. We are increasingly realizing that large-scale duplication or deletion of DNA segments is a major source of human genetic variation, with up to 12% of the genome exhibiting CNV [78]. Sequence similarity between genes suggests that the whole Fcγ-receptor gene cluster arose from the duplication of a single ancestral gene. Within the modern human genome there is variation in the number of copies of a large segment of this gene cluster incorporating FCGR3B and FCGR2C, with the diploid genome having between zero and six copies. An initial study demonstrated the association between low copy number and immune-mediated nephritis, both in humans and rats [79]. Subsequently, in a larger follow-up study, it was shown that carrying less than two copies is a major risk factor both for lupus nephritis (OR 2.43) and SLE without renal involvement (OR 2.21) [14]. FCGR3B is expressed solely on neutrophils. It lacks an ITAM domain, so the transmission of intracellular signalling through CR3 (integrin αMβ2) [80–82]. The realization that two major SLE susceptibility genes cooperate in mediating phagocytosis and activation by neutrophils suggests that this cell may be more important in SLE than has previously been appreciated. Recent work demonstrating that neutrophils are a source of T-cell regulating cytokines such as IL-18 suggests a potential role [83].

**TNFSF4**

Also known as OX40 ligand, TNF super-family 4 (TNFSF4) is a membrane-bound protein expressed on the surface of antigen-presenting cells. It provides a strong activating signal when bound to its ligand OX40 on the surface of CD4+ T cells. Both protective and risk haplotypes at TNFSF4 were identified in a study of two SLE case-control cohorts and two independent family-based cohorts, with high levels of significance observed across the combined data sets [84]. In vitro studies demonstrated that the risk haplotype was associated with increased TNFSF4 transcript and cell surface expression of OX40L. Signalling through OX40L/OX40 mediates proliferation and expansion of T cells during the primary immune response and the generation of T-cell memory [85]. It has not been fully established how OX40/OX40L interactions influence T-cell subset profiles. Most evidence suggests a bias towards a Th2 pattern of cytokine release, although there is also evidence for a down-regulation of regulatory T-cell subsets [86, 87]. B cells express OX40L, along with other antigen-presenting cells. There is also good evidence that signalling through OX40L can induce B-cell activation and differentiation [88, 89].

**Genes with poorly defined function**

So far we have discussed key candidate genes for which we have a reasonable understanding, at least in outline, of how they may be working. We also need to consider a number of genes whose functional role is poorly understood. There may be false positives among this group, but the levels of significance attributed to their association with SLE are high, and most are likely to be genuine disease genes. Understanding how they work will provide new and exciting insights into the pathogenesis of SLE.

**Methylation and ubiquitination**

Some of the SLE candidate genes appear to play a role in the regulation of gene expression, or be involved in the post-translational expression of proteins. MECP2 encodes a protein involved in DNA methylation. This is a potentially important mechanism in the regulation of gene expression and abnormal T cell DNA methylation has been implicated in the pathogenesis of SLE [90]. MECP2 is on the X-chromosome, which is also of interest given that SLE is predominantly a disease of women.

Ubiquitination is another incompletely understood process by which gene and protein expression may be controlled. The primary role for ubiquitination is the tagging of proteins for proteolytic degradation. Through a network of ubiquitin-carrier enzymes and ubiquitin ligases this process may be quite selective, say with specific cell surface receptors targeted for rapid degradation. Ubiquitination also has a role in regulating transcription factor and intracellular kinase activity, thus influencing intracellular signalling pathways [91]. The ubiquitin-carrier enzyme gene UBE2L3 has been identified as an SLE candidate gene. There is little work on UBE2L3 specifically, but it appears to be widely expressed and has a role in the maturation of transcription factors such as the nuclear factor-xB precursor p105 [92, 93].

**Control of apoptosis**

A failure to adequately clear apoptotic cell debris has been postulated as a key pathogenic mechanism in SLE [94]. Less understood is whether apoptosis itself is abnormal [95]. A number of mouse models with apoptotic defects develop autoimmune disease, presumably due to a failure to delete auto-reactive lymphocytes. Studies of human lymphocytes and neutrophils from SLE patients, in contrast, generally show increased rates of apoptosis [96, 97]. Since apoptotic cells are proposed as a source of auto-antigen there is also biological plausibility for accelerated apoptosis as a primary disease mechanism. It is therefore notable that many of the key SLE candidates play a role in apoptosis and cell-cycling decisions. We have already discussed IRF5. Signalling through CR3 (integrin αMβ2) has also been shown to mediate neutrophil apoptosis in the resolution of inflammation [98]. Autophagy is a regulatory process by which long-lived intracellular proteins and organelles are degraded, but excessive autophagy activity can trigger apoptosis. This process is controlled by the autophagy genes that act in an ubiquitin-like conjugation system. The autophagy gene ATG5 has been identified as an SLE candidate, and encodes a protein with ubiquitin-carrier enzyme-like activity. Intriguingly ATG5 may act as the switch that determines whether autophagy progresses to apoptosis, and over-expression may result in accelerated apoptosis [99].

**Other genes**

The association with gene KIAA1542 seems likely to be an indirect association. This gene lies adjacent to IRF7 that encodes IFN regulatory factor-7. IRF7 has functions that overlap with IRF5, and again this emphasizes the importance of the Type 1 IFN pathways in SLE. SCUBE1 (signal peptide, CUB domain, EGF-like 1) encodes a glycoprotein that is expressed both on platelets and endothelial cells [100, 101]. It plays a role in cell adhesion and...
may have potential as a regulator of inflammation. For the remaining associated genes the function is either completely unknown or appears related to ubiquitous cellular housekeeping functions. It is therefore premature to hypothesize a functional role in SLE.

**SLE pathogenesis—a genetic hypothesis**

With the advance in our understanding of SLE genetics it becomes possible to postulate a ‘genetic hypothesis’ of SLE, based not on the numerous immunological abnormalities that have been described, but on the key candidate genes (Fig. 1). As discussed earlier this genetic approach has the advantage of highlighting fundamental disease mechanisms rather than secondary events.

The association of the **FCGR2A** and **FCGR3B** genes and **ITGAM** (along with the much rarer C1q deficiency) suggests that the early response of the innate immune system to the recognition of autoantigens is a critical step in SLE pathogenesis. **FCGR3B** is expressed on neutrophils only, but **FCGR2A** and **ITGAM** are expressed much more widely on antigen-presenting cells including the macrophages and dendritic cells. It is possible that susceptibility variants simply fail to bind auto-antigenic debris efficiently, thus impairing clearance. Alternatively, the susceptibility effect may arise from a more subtle alteration in intracellular signalling, and subsequent whole cell response (for example, the production of Type 1 IFN) to ligand binding through these receptors. These mechanisms may be important, both at the onset of autoimmunity, but also as an amplification event as auto-antigen containing immune complexes are continually recognized and reacted to.

The second key step appears to dysregulation of the response of both the innate and adaptive immune system to initial cytokine signals. In particular, critical variants of **IRF5**, **STAT4** and possibly **IRF7** may result in an overactive, pro-inflammatory response to early signals released following the encounter of auto-antigen.

The final immunological event in SLE is the generation of overactive, auto-reactive B cells, producing high-affinity auto-reactive B cells and autoantibody.
auto-antibody. B cells may over-react in response to direct binding of antigen to the BCR, and the importance of this B-cell pathway is emphasized by the association of three genes involved in early BCR signalling, namely BLK, LYN and BANK1 along with FCGR2B. Additionally, B-cells may receive inappropriately strong and pro-inflammatory CD4+ T-cell help. The importance of this pathway is suggested by an additional cluster of associated genes involved in antigen-presenting cell/T-cell cross-talk and T-cell activation, including Class II MHC, STAT4 and PTPN22. The CTLA4/ICOS locus, not discussed in depth in this review would also fall in this category. TNFSF4 (OX40L) probably plays a complex role in both B-cell and T-cell activation through a bi-directional signalling mechanism.

Interestingly, the genes involved in T-cell activation are often shared between autoimmune diseases. For example, PTPN22 has also been associated with RA, juvenile idiopathic arthritis, Graves disease and T1D; STAT4 has been associated with RA; CTLA4 has been associated with Graves disease, T1D, autoimmune hypothyroidism and possibly RA; while the Class II MHC has been associated with multiple autoimmune disorders [23, 53, 69, 70, 102–105]. This suggests that a cluster of genes increasing the activation potential of CD4+ T cells may increase susceptibility to autoimmune disease generally, while other genes, associated with antigen recognition and B-cell responses may determine that autoimmunity is of a systemic SLE type. It is perhaps interesting to compare the pattern of genetic association seen in SLE with that seen in the most intensively studied autoimmune disease, T1D. Aside from variation of the insulin gene, genetic susceptibility in T1D appears largely dependent on variation of key T-cell regulating genes [especially Class II MHC, PTPN22 and IL2RA (CD25)] [26]. The risk of disease associated with these variants is high, particularly for the Class II MHC, but the risk associated with all the other T1D susceptibility genes is much lower (OR < 1.25). In contrast, in SLE, the risk associated with shared T-cell regulating genes is modest, but overall there are many more genes associated with a moderate disease risk (OR > 1.25), suggesting that T-cell regulation is not the key step in SLE pathogenesis, but simply one of a number of pathways that can influence disease susceptibility. This observation may influence how we target immunoregulatory pathways for therapeutic intervention in the future.

The association of a number of genes with less well-defined function shows how we have really only started to touch on the physiological control of immune responses, and how they may go wrong in SLE. Control of gene expression by ubiquitination or methylation may be much more important and specific than previously thought. Similarly control of apoptosis both as a mechanism of immune down-regulation or as a source of auto-antigen may be a key mechanism. Considerable progress has been made but there is still clearly a great deal we do not yet understand about this disease.

Rheumatology key messages

- Major advances in the identification of disease susceptibility genes in SLE
- The genetic risk is conferred by many genes.
- In comparison with some autoimmune diseases, the genetic risk is more evenly distributed among the disease genes.

Acknowledgements

Funding: B.R. is funded by an Arthritis Research Campaign Clinical Research Fellowship and T.J.V. is supported by a Wellcome Trust Senior Fellowship.

Disclosure statement: The authors have declared no conflicts of interest.

References

23 The Wellcome Trust Case-Control Consortium. Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. Nature 2007;447:661–78.
erythromas (SLE): low copy number is a risk factor for and high copy number is a protective factor against SLE susceptibility in European Americans. Am J Hum Genet 2005;76:528–37.
40 Kozyrev SV, Alarcon-Riquelme ME. The genetics and biology of Irf5-mediated type 1 interferons in the interferon-gamma response to viral infection. Science 2008;321:187–94.


