Review

Genetics of rheumatic disease
Alex Clarke and Timothy J Vyse

Imperial College London, Faculty of Medicine, Section of Molecular Genetics and Rheumatology, Fifth Floor, Commonwealth Building, Hammersmith Hospital, London W12 0NN, UK

Corresponding author: Timothy J Vyse, t.vyse@imperial.ac.uk

Published: 14 October 2009
This article is online at http://arthritis-research.com/content/11/5/248
© 2009 BioMed Central Ltd

Abstract

Many of the chronic inflammatory and degenerative disorders that present to clinical rheumatologists have a complex genetic aetiology. Over the past decade a dramatic improvement in technology and methodology has accelerated the pace of gene discovery in complex disorders in an exponential fashion. In this review, we focus on rheumatoid arthritis, systemic lupus erythematosus and ankylosing spondylitis and describe some of the recently described genes that underlie these conditions and the extent to which they overlap. The next decade will witness a full account of the main disease susceptibility genes in these diseases and progress in establishing the molecular basis by which genetic variation contributes to pathogenesis.

Genetics of rheumatic disease

The spectrum of rheumatic disease is wide and includes conditions with diverse pathology, although most have in common a heritable risk with a complex genetic basis. There has therefore been intense effort to understand the contribution of genotype to the expression of disease in terms of both basic pathogenesis and clinical characteristics. Recent technical advances in genotyping and statistical analysis and international collaborations assembling large cohorts of patients have led to a wealth of new data. In this review we describe insights gained into the pathogenesis of autoimmune rheumatic disease by the techniques of modern genetics, in particular evidence from genome-wide association (GWA) studies, which provide support for the existence of a common genetic risk basis to several diseases. To reflect the new data from GWA studies, our discussion will be confined to rheumatoid arthritis (RA), systemic lupus erythematosus (SLE), and ankylosing spondylitis (AS), which in some cases share a common autoimmune pathogenesis. Osteoarthritis and osteoporosis are also complex genetic traits but limitations of space are such that these two conditions will not be considered in this review.

The concept of a systematic, GWA study became practical with the cataloguing of libraries of common polymorphisms. Currently, over 20 million single nucleotide polymorphisms (SNPs) have been identified [1] and platforms are available to type up to 1 million of these in a single reaction. Although not all SNPs are currently genotyped, as the human genome is arranged into haplotype blocks in linkage disequilibrium, it is only necessary to type so-called tag SNPs, which identify these areas of limited variability [2], to achieve good representation of the total amount of genetic variation. Most typed SNPs are relatively common (minor allele frequency of >5%) and if associated with disease are likely, therefore, to have only modest pathogenic effects (odds ratios (ORs) usually between 1.2 and 2), as otherwise they would become depleted in a population due to natural selection. It is necessary, therefore, to invoke the ‘common-disease common-variant’ (CD-CV) model [3], which assumes an accumulation of risk caused by the carriage of multiple deleterious alleles, to explain current experimental findings.

One of the revolutionary advantages of the GWA study is the freedom from a required gene-centric hypothesis, which provides an unprecedentedly effective technique for risk gene discovery. Many disease-associated genes identified by GWA studies were completely unsuspected to be relevant - for example, the autophagy system in Crohn’s disease [4]. However, because in essence up to 1 million independent hypotheses are being tested in each genotyping reaction, sample sizes powered to detect even the stronger associations must be large, and criteria for significance stringent. The general consensus is that significance can be defined as a P-value smaller than 5 × 10^-7, which in a cohort such as the Wellcome Trust Case Control Consortium (WTCCC) of 2,000 cases, for example, approximates to a power of 43% rising to 80% to detect alleles with ORs of 1.3

AS = ankylosing spondylitis; BCR = B cell receptor; CNV = copy number variation; GWA = genome-wide association; HLA = human leukocyte antigen; IFN = interferon; IL = interleukin; IRF = interferon regulatory factor; MHC = major histocompatibility complex; NF = nuclear factor; OR = odds ratio; RA = rheumatoid arthritis; SLE = systemic lupus erythematosus; SNP = single nucleotide polymorphism; STAT = signal transducer and activation of transcription factor; TLR = toll-like receptor; TNF = tumour necrosis factor; TNFAIP = TNFα-induced protein-3; TRAF = TNF receptor associated factor; WTCCC = Wellcome Trust Case Control Consortium.
and 1.5, respectively [5]. However, the genome is subject to variation at more than the SNP level, and individuals also differ in the copy number of sections of DNA of greater than several kilobases in size, so called copy number variation (CNV), which in fact accounts for more total nucleotide difference between individuals than SNPs [6,7]. CNV can affect gene expression levels [8] and has been linked to autoimmune disease [9,10], including SLE [11]. Whilst the latest genotyping platforms include assessment of CNV, earlier products actively excluded SNPs within regions of the most variation as they were more likely to fail quality control steps. Association studies based on CNV are, therefore, in their relative infancy. Finally, the genome is subject to modification without a change in DNA sequence; epigenetic mechanisms can have profound effects on gene expression. These include DNA methylation and changes in chromatin structure [12].

It has become apparent that SLE, RA, and AS, which have divergent clinical features, may share a common genetic risk framework, and we aim in our review to illustrate this.

The MHC region and antigen processing

The major histocompatibility complex (MHC) region on chromosome 6 contributes to the risk of almost all autoimmune diseases, and its role in immunity in mice was recognized over 60 years ago. In humans, the MHC locus is also known as the HLA (human leukocyte antigen) region, reflecting the initial identification of MHC gene products on the surface of white blood cells. The classical MHC extends over around 4 megabases, and comprises three clusters: class I, II, and III. Class I and II regions include genes that encode the α- and β-chains of the MHC I and II complexes, and flank the class III region, which contains an assortment of immunologically relevant genes. Despite extensive study, the mechanisms that link the MHC to disease are largely unknown, although it is supposed that variation in the MHC peptide binding cleft facilitates presentation of self-antigen to autoreactive lymphocytes.

These difficulties in understanding the MHC are not without reason; it contains some of the most polymorphic loci described in the genome, and has a highly complicated genetic architecture, with some regions exhibiting extended linkage disequilibrium [13].

In RA, the MHC accounts for around a third of the genetic liability [14]. Alleles at HLA-DRB1 contribute much of this risk - for example, DRB1*0401 carries an OR of 3. GWA studies confirm the strong association with MHC variants; risk alleles confer an OR of around 2 to 3 in homozygotes [15], with very high statistical significance (P<10^{-100}). Additional loci contributing to the risk of RA identified by high-density genotyping include HLA-DP in patients with anti-cyclic citrullinated peptide antibodies [16]. SLE not only has strongly associated alleles in the class II region, HLA-DR2 (DRB1*1501) and DR3 (DRB1*0301) [14], with ORs of 2 [17], but also risk variants in the class III cluster, which encodes genes such as TNF and the complement components C2, C4A and C4B. C4 is crucial in the classical and mannose-binding lectin pathways of complement activation, and complete deficiency of C4 or indeed other components of the classical pathway are rare, but strong, risk factors for SLE [18]. The C4 gene is subject to CNV and is of two isoatypes, C4A and C4B. It is an attractive hypothesis that CNV at C4 affects expression and contributes to SLE risk. However, it remains to be established whether haplotypes carrying partial C4 deficiency exert their risk via an influence on complement or through other genetic variants that are in linkage disequilibrium. Other loci in the class III region have been implicated in SLE, including the SKIV2L gene, SNPs in which carry an OR of 2 in a family-based analysis [19]. SKIV2L encodes superkiller viralicidic activity 2-like, the human homologue of which is a DEAD box protein that may have nucleic acid processing activity. The second MHC III signal for SLE we will consider was identified in the International Consortium on the Genetics of Systemic Lupus Erythematosus (SLEGEN) GWA study [17,20]. The SNP rs3131379 in mutS homologue 5 (MSH5) has an OR of 1.82. There is evidence that MSH5 has a role in immunoglobulin class switch variation [21]. Again, further work is required to definitively implicate this gene rather than variants in linkage disequilibrium, which include HLA-DRB1*0301 and C4A deletions.

Clearly, HLA-B27 is the overwhelming association in AS, with an OR of 200 to 300. In the MHC, other genetic risk variants have been identified, including HLA-B60 (OR 3.6) [22] and various HLA-DR genes with relatively minor contributions [23]. The pathogenic mechanism for these risk alleles is unknown. Outside of the MHC, two significant genes have so far been identified in AS: ARTS1 and IL-23R [24], the latter of which will be discussed below and has been associated with several different autoimmune diseases. ARTS1 has two identified functions. Its first is in the processing of peptide for presentation via MHC I. It is localised in the endoplasmic reticulum, and is upregulated by IFNγ. It acts as an amino-terminal aminopeptidase and in mice is essential for the display of the normal peptide repertoire. In its absence, many unstable and highly immunogenic MHC-peptide complexes are presented [25]. A hypothetical connection with HLA-B27 can thus be drawn. Its other function is to downregulate signalling by IL-1, IL-6, and TNFα through surface receptor cleavage [26-28]. The most associated SNP rs30187 risk allele has an OR of 1.4, and is of unknown functional significance.

Innate-adaptive interface

Interferon signalling: IRF5

It is clear that type 1 interferons (IFNα and IFNβ) are of great importance in the pathogenesis of SLE. Patients with active disease have high levels of IFNα, which has multiple immuno-
modulatory actions [29], including the induction of dendritic cell differentiation, the upregulation of innate immune receptors such as toll-like receptors (TLRs), the polarization of T cells towards a Th1 phenotype, and the activation of B cells. Type I interferons are produced by all cells in response to viral infection, but particularly by plasmacytoid dendritic cells in response to unmethylated CpG oligonucleotides binding to TLR-9, or RNA to TLR-7. Using a candidate gene approach targeting the IFN signalling pathway, the SNP rs2004640 in IRF5 (interferon regulatory factor 5) was found to be significantly associated with SLE (OR 1.6) [30], a risk gene confirmed in several other studies [17,31-35]. The functional consequences for IRF5 of the identified mutations are variable, but include the creation of a 5’ donor splice site in an alternative exon 1, allowing the expression of several isoforms [35], a 30 base-pair in-frame insertion/deletion variant of exon 6, a change in the 3’ untranslated region, and a CGGGG insertion-deletion (indel) polymorphism, the latter two affecting mRNA stability [32,36]. Interestingly, these mutations may occur together in a haplotype, with varying degrees of associated risk. The exact role of IRF5 in IFN signalling has not been fully elucidated, but it is also critical for the gene induction programme activated by TLRs [37], providing further biological plausibility for its importance in the pathogenesis of SLE. Haplotypes of IRF5 are also implicated in RA, and may confer either protection (OR 0.76) or predisposition (OR 1.8) [38]. The same CGGGG indel allele described above also carries risk for multiple sclerosis and inflammatory bowel disease [36].

**TNF-associated signalling pathway: TNFAIP3 and TRAF1-C5**

TNF-associated signalling pathway genes play a prominent role in the risk for both SLE and RA, and associations with variants in TNFAIP3, and the TRAF1-C5 locus have been identified [39,40]. TNFα-induced protein-3 (TNFAIP3; also known as A20) is a ubiquitin editing enzyme that acts as a negative regulator of NFκB. A20 can disassemble Lys63-linked polyubiquitin chains from targets such as TRAF6 and RIP1. A second region of A20 catalyses Lys48-linked ubiquitination that targets the molecule for degradation by the proteasome [41]. A20 modifies key mediators in the downstream signalling of TLRs that use MyD88, TNF receptors, the IL-1 receptor family, and nucleotide-oligomerization domain protein 2 (NOD2) [42]. Tnfaip3 knockout mice develop severe multi-organ inflammatory disease, and the phenotype is lethal [43]. The SNP rs10499194 in TNFAIP3 carries an OR of 1.33 for RA, and rs5029939 an OR of 2.29 for SLE [44], the latter also conferring an increased risk of haematologic or renal complications [45].

On chromosome 9, the region containing TRAF1 (TNF receptor associated factor 1) and C5 (complement component 5) genes is associated with significant risk for RA (risk SNP OR of approximately 1.3) in most [15,40,46-48], but not all [5], studies. Due to linkage disequilibrium, the functional variant remains elusive. TRAF1 is principally expressed in lymphocytes, and inhibits NFκB signalling by TNF. This pathway is blocked in TRAF1 overexpression [49] whilst, conversely, Traf1-/- mice are sensitized to TNF and have exaggerated TNF-induced skin necrosis [50].

The complement system has long been known to be involved in the pathogenesis of RA. In the collagen-induced arthritis model of RA, C5 deficiency prevents disease de novo and ameliorates existing symptoms and signs [51,52]. Interestingly, GG homoygotes at the TRAF1-C5 SNP rs3761847 with RA have a significantly increased risk of death (hazard ratio 3.96, 95% confidence interval 1.24 to 12.6, \( P = 0.02 \)) from malignancy or sepsis, potentially allowing identification of patients for appropriate screening [53].

**Immunomodulatory adhesion molecule: ITGAM**

Integrin-α-M (ITGAM), variants of which are strongly associated with SLE, forms a heterodimer with integrin-β2 to produce αMβ2-integrin (also known as CD11b, Mac-1, or complement receptor-3), which mediates the adhesion of myeloid cells to the endothelium via ICAM-1 (Intercellular adhesion molecule-1) and recognizes the complement component iC3b. It not only has a role in cell trafficking and phagocytosis [54], but also has other immunomodulatory functions. Antigen-presenting cells produce tolerogenic IL-10 and transforming growth factor-β on iC3b binding to CD11b [55], and mice deficient in this receptor upregulate expression of IL-6, favouring a pro-inflammatory Th17 response [56]. Despite its implication in defective immune complex clearance in SLE, experimental evidence for a role was lacking. GWA studies, however, demonstrate a strong and significant association [17,33,44], with an OR of 1.83 (\( P = 7 \times 10^{-50} \)) in meta-analysis [57]. The implicated SNP rs1143679 is non-synonymous, causing the substitution of histidine for arginine at amino acid 77, although this change does not affect the iC3b binding site [58]. Furthermore, although this SNP is disease associated in European and Hispanic patients, it is monomorphic in Japanese and Korean populations [59]; an explanation of its effect is therefore outstanding. It has been mentioned that CNV is important in the binding and clearance of immune complexes, thereby marking itself as a potential SLE risk gene. There is a significant association between low FCGR3B copy number and SLE. Patients with two or fewer copies of FCGR3B have an OR of 2.43 for SLE with nephritis, and 2.21 for SLE without nephritis [61].

**Lymphocyte differentiation**

**T cell receptor signalling: PTPN22**

Outside the HLA region, the first reproducible genetic association for RA came with the implication of PTPN22 from a candidate gene approach [62] based on linkage analysis.
identification of a susceptibility locus at 1p13 [63]. It has remained the strongest and most consistent association mapped by GWA studies in RA. A role in SLE has also been identified [17]. The OR for the risk allele is around 1.75 in RA, and 1.5 in SLE. However, it should be noted that this allele (encoding the R620W mutation) is monomorphic or not disease associated in Korean or Japanese patients [64,65].

PTPN22 encodes lymphotoxin tyrosine phosphatase (LYP), a protein tyrosine phosphatase that inhibits T cell receptor signalling, decreasing IL-2 production. The disease associated SNP is responsible for a change from arginine to tryptophan at position 620, which inhibits binding to the SH3 domain of carboxy-terminal Src kinase. This in turn appears to enhance dephosphorylation of tyrosine residues in the Src family kinases Lck, FynT, and ZAP-70 [66,67]. The overall effect of the mutation is a reduction in T cell receptor signalling. The pathogenic effect of this is unclear, but may relate to impaired negative selection in the thymus, or lead to signalling. The pathogenic effect of this is unclear, but may relate to impaired negative selection in the thymus, or lead to a reduction in regulatory T cells [68]. Conversely, the R623Q variant of PTPN22, which is a loss-of-function mutation affecting the phosphatase activity of LYP, is protective against SLE [69]. PTPN22 does not appear to be a risk gene for AS [70].

Polarization towards Th1 and Th17 phenotypes: STAT4 and IL23R

STAT4 encodes signal transducer and activation of transcription factor-4, responsible for signalling by IL-12, IL-23, and type 1 IFNs [71]. STAT4 polarizes T cells towards Th1 and Th17 phenotypes, which has the potential to promote autoimmunity [72]. In RA the OR for the risk allele of SNP rs7574865 is 1.32 in one case-control study [73], with a less strong disease association at rs11893432 in a meta-analysis of GWA studies (OR 1.14) [15]. There is convincing evidence that STAT4 is a risk locus for SLE in multiple racial groups [33,74], and it may be theorized that interference in type I IFN signalling may be the underlying pathogenic mechanism in this case. Distinctive disease pathways could, therefore, emerge from mutations in a single gene. The WTCCC AS study identified IL23R as a risk gene in AS [24]. IL-23 is instrumental in the development of T cells with the pro-inflammatory Th17 phenotype [75], and IL23R has been linked to psoriasis, ulcerative colitis, and Crohn’s disease in GWA studies [5,76,77]. An interesting connection between these conditions, all of which may share common clinical features, is thus made. In AS the risk SNP rs11209032 confers an OR of 1.3.

B cell activation

B cells are a population long suspected to be important in autoimmune rheumatic disease, and the benefits of their depletion in RA and SLE has resurrected interest in their pathogenic role. The risk genes identified so far are involved in signalling from the B cell receptor (BCR). BLK encodes a Src family tyrosine kinase restricted to the B cell lineage and is poorly understood. Risk alleles in the region upstream of the transcription initiation site are associated with SLE (OR 1.39, \( P=1 \times 10^{-10} \)) and reduce levels of BLK mRNA [33]. BANK1 (B cell scaffold protein with ankyrin repeats-1) undergoes tyrosine phosphorylation upon B cell activation by the BCR, leading to an increase in intracellular calcium through the inositol trisphosphate mechanism [78]. The non-synonymous SNP rs10516487 in BANK1, which substitutes histidine for arginine at amino acid 61, also has disease association (OR 1.38) [79]. The functional consequence of this may be higher affinity for the inositol trisphosphate receptor, as the substitution is located in the binding site.

Lyn, another Src tyrosine kinase, is important in determining signalling thresholds for myeloid and B cells. On BCR ligation, it phosphorylates tyrosine residues of Syk, an activating tyrosine kinase, CD19, and the immunoreceptor tyrosine-based activation motifs (ITAMs) of the Igγ/µβ subunits of the BCR. However, it also has a critical regulatory role, mediated by phosphorylation of the inhibitory motifs of CD22 and FCγRIIB, which in turn activate SH2-domain containing phosphatases, leading to dephosphorylation and deactivation of a number of signalling intermediaries [80]. Lyn/- mice develop severe autoimmunity associated with glomerulonephritis [81]. An association between SNPs in LYN and SLE, identified initially in the SLEGEN GWA study [17], has been recently confirmed in a case-control study [82]. The most associated SNP, rs6983130, is near the primary transcription initiation site.

OX40L, a member of the TNF super-family encoded by TNFSF4 (TNF superfamily 4), is associated with SLE. The cross-talk between B lymphocytes and dendritic cells expressing OX40L, and T cells that express its receptor, OX40, serves to enhance the adaptive immune response [83]. An upstream TNFSF4 haplotype, associated with SLE, enhances gene expression in vitro [84,85], although the mechanism responsible for the deleterious effects observed remains to be established.

Despite the importance of B cells in the pathogenesis of RA, none of the gene effects described above have been identified in the current generation of GWA studies. However, variants at CD40 in European patients do carry risk [15]. CD40 expressed on B cells, via interaction with its ligand CD154 on CD4+ T cells, promotes immunoglobulin class switching, and germinal centre formation. B cells, however, also have a regulatory role, likely to be mediated by IL-10, and disruption of this function may be another route to autoimmune disease [86].

Post-translational modification: PADI4

Peptidyl arginine deiminase-4 (PADI4) is a member of the enzyme family responsible for the post-translational citrullination of arginine residues in RA synovium, subsequently recognized by anti-cyclic citrullinated protein antibodies. In Japanese [87] and Korean patients [88], case-control asso-
cation studies have identified functional haplotypes of PADI4 conferring risk of RA. However, in Caucasian populations this association is inconsistent [89-91].

**Conclusion**

Even with the proliferation of new genetic associations discovered in the past few years by GWA studies, only around 10 to 15% of the inherited risk for SLE and RA can be currently explained. This may be accounted for, in part, by a number of factors, some related to limitations of recent study design. As mentioned above, even the largest current GWA cohorts have limited power to detect associations with ORs <1.3, potentially losing multiple risk genes. By definition, most genotyped SNPs are common, and so rare but causal variants have a tendency to be missed. These rarer SNPs may be either those with a low minor allele frequency (<5%), or occur de novo, of which 200 to 500 non-synonymous SNPs are expected per individual [92]. In many cases, it is far from certain if the associated SNP is functional, or in linkage disequilibrium with the true cause. Finally, the great majority of GWA studies have been conducted on European populations, thereby excluding carriers of many potential risk variants from analysis. However, it is unfortunately the case that current genotyping platforms often have poor coverage of tagging SNPs within populations that exhibit low levels of genomic linkage disequilibrium, such as those of African ancestry [93]. For example, the latest high-density genotyping chips from Affymetrix (6.0) and Illumina (1M) may capture fewer than half the SNPs identified through re-sequencing in Yoruban Nigerians [94]. Given that clear differences exist in the risk of autoimmune disease according to ethnicity, and that not all disease risk alleles are in common, it is imperative that full account of this variation is made. Structural genetic differences have only recently begun to be assessed by modern genotyping platforms, and the contribution of, for example, CNV to inherited disease risk is largely unquantified. Even more difficult to appreciate is the influence of heritable epigenetic factors, and the exact relationship between genotype and phenotype. Nevertheless, although it will probably not be possible to explain all the observed genetic risk in the near future, we are rapidly moving towards the ability to quickly and cheaply fully sequence individual genomes [95], with all the advantages that brings [96]. In the meantime, understanding the functional basis of the disease risk variants so far identified presents an outstanding challenge. Integration of genotypic with RNA and protein expression data in a systems biologic approach represents one potentially valuable methodology [97]. Exploring and therapeutically utilizing the genetic differences between individuals is axiomatic to personalized medicine, and will undoubtedly lead to better outcomes in the management of autoimmune disease.

**Competing interests**

The authors declare that they have no competing interests.

**References**

1. Entrez SNP [http://www.ncbi.nlm.nih.gov/sites/entrez?db=snp]
2. International HapMap Consortium, Frazer KA, Ballinger DG, Cox DR, Hinds DA, Stuve LL, Gibbs RA, Belmont JW, Boudreau A, Hardenbol P, Leal SM, Møstervik S, Wheeler DA, Willis TD, Yu F, Yang H, Zeng C, Gao Y, Hu H, Hu W, Li C, Lin W, Liu S, Pan H, Tang X, Wang J, Wang W, Wu J, Zhang B, Zhang Q, et al.: A second generation human haplotype map of over 3.1 million SNPs. *Nature* 2007, 449:851-861.
3. Risch N, Merikangas K: The future of genetic studies of complex human diseases. *Science* 1996, 273:1516-1517.
4. Rioux JD, Xavier RJ, Taylor KD, Silverberg MS, Goyette P, Huett A, Green T, Kuballa P, Barnard MM, Datta LW, Shugart YY, Griffiths AM, Targan SR, Ippoliti AF, Bernard JJ, Mei L, Nicolae DL, Regueiro M, Schumm LP, Steinhardt AH, Rotter JI, Duer R, Cho JH, Daly MJ, Brant SR: Genome-wide association study identifies new susceptibility loci for Crohn disease and implicates autophagy in disease pathogenesis. *Nat Genet* 2007, 39:596-604.
5. Wellcome Trust Case Control Consortium: Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. *Nature* 2007, 447:661-678.
6. Sebat J, Lakshmi B, Troge J, Alexander J, Young J, Lundin P, Mánér S, Massa H, Walker M, Chi M, Navin N, Lucito R, Healy J, Hicks J, Ye K, Reiner A, Gilliam TC, Trask B, Patterson N, Zetterberg A, Wigler M: Large-scale copy number polymorphism in the human genome. *Science* 2004, 305:529-528.
7. Redon R, Ishikawa S, Fitch KR, Feuk L, Perry GH, Andrews TD, Fiegler H, Shapero MH, Caron AR, Chen W, Cho EK, Dallaire S, Freeman JL, Gonzalez JR, Gratacos M, Huang J, Kalaizipolpos M, Komura D, MacDonald JR, Marshall CR, Mei R, Montgomery L, Nishimura K, Okamura K, Shen F, Somerville MJ, Tchinda J, Valsesia A, Woodward C, Yang F, et al.: Global variation in copy number in the human genome. *Nature* 2006, 444:444-454.
8. Stranger BE, Forrest MS, Dunning M, Ingle CE, Beazley C, Thorne N, Redon R, Bird CP, de Graaf A, Lee C, Tyler-Smith C, Carter N, Scherer SW, Tavares S, Deloukas P, Hurles ME, Dermitzakis ET: Relative impact of nucleotide and copy number variation on gene expression phenotypes. *Science* 2007, 315:848-853.
9. Altman TJ, Dong R, Vyse TJ, Norsworthy PJ, Johnson MD, Smith J, Mangion J, Robertson-Lowe C, Marshall AJ, Petretto E, Hodges MD, Bhangal G, Patel SG, Sheehan-Rooney K, Duda M, Cook PR, Evans DJ, Domín J, Flint J, Boyle JJ, Pusey CD, Cook HT: Copy number polymorphism in Fcgr3 predisposes to glomerulonephritis in rats and humans. *Nature* 2006, 439:851-855.
10. Hollox EJ, Huffmeier U, Zeeuwen PL, Palla R, Lascorz J, Rodijk-Olthuis D, van der Kerkhof PC, Traupe H, de Jongh G, den Heijer PR, Evans DJ, Domín J, Flint J, Boyle JJ, Pusey CD, Cook HT: Copy number polymorphism in Fcgr3 predisposes to glomerulonephritis in rats and humans. *Nature* 2006, 439:851-855.
11. Yang Y, Chung EK, Wu YL, Savelli SL, Nagaraja HN, Zhou B, Hebert M, Jones KN, Shu Y, Kitzmiller K, Blanchong CA, McBride KL, Higgins GC, Rennebohm RM, Rice RR, Hackshaw KV, Roubey RA, Grossman JM, Tsao BP, Birmingham DJ, Rovin BH, Hebert LA, Yu CY: Gene copy-number variation and associ-
present many unstable and highly immunogenic peptides. Nat Immunol 2007, 8:101-108.

26. Cui X, Rouhani FN, Hawari F, Levine SJ: Identification of the type II IL-1 decoy receptor as a novel TNFR1-binding protein that promotes TNFR1 ectodomain shedding. J Clin Invest 2002, 110:515-526.

27. Behrens TW, Alarcón-Riquelme ME, Alarcón S, Brown MM, et al.: The type I interferon system in systemic lupus erythematosus. Arthritis Rheum 2006, 54: 408-420.

28. Sigurdsson S, Nordmark G, Göring HH, Lindroos K, Timms K, et al.: The type I interferon system in systemic lupus erythematosus. Arthritis Rheum 2007, 50:358-365.

29. Behrens TW, Alarcón-Riquelme ME, Alarcón S, Brown MM, et al.: The type I interferon system in systemic lupus erythematosus. Arthritis Rheum 2006, 54: 408-420.

30. Sigurdsson S, Nordmark G, Göring HH, Lindroos K, Timms K, et al.: The type I interferon system in systemic lupus erythematosus. Arthritis Rheum 2007, 50:358-365.
39. Plenge RM, Cotopas C, Davies L, Price AL, de Bakker PI, Maler J, Pe’er I, Burtt NP, Blumenstiel B, DelFedele M, Parkin M, Barry R, Winslow W, Healy C, Graham RE, Neale BM, Izmailova E, Rotenbouf R, Parker AN, Glasner R, Karlan EW, Maher N, Haifler DA, Lee DM, Seldin MF, Remmers EF, Lee AT, Padyukov L, Alfredsson L, Coblyn J, et al.: Two independent alleles at 6q23 associated with risk of rheumatoid arthritis. Nat Genet 2007, 39:1477-1482.

40. Plenge RM,Siegelstedt M, Padyukov L, Lee AT, Remmers EF, Ding B, Liew A, Khalili H, Chandrasekaran A, Davies LR, Li W, Tan AK, Bonnard C, Ong RT, Thalamuthu A, Pettersson S, Liu C, Chan VV, Carulli JP, Beckman EM, Altschuler D, Alfredsson L, Criswell LA, Amos CI, Seldin MF, Kastner DL, Klæreskog L. Study of rheumatoid arthritis - a genomewide study. N Engl J Med 2007, 357:1199-1209.

41. Komander D, Barford D: Structure of the A20 OTU domain and mechanistic insights into deubiquitination. Biochem J 2008, 409:77-85.

42. SC: Deubiquitilation and regulation of the immune response. Nat Rev Immunol 2008, 8:501-511.

43. Lee EG, Boone DL, Chai S, Libby SL, Chien M, Lodolce JP, Ma A: Failure to regulate TNF-induced NF-kappaB and cell death responses in A20-deficient mice. Science 2000, 289:2350-2353.

44. Graham RR, Cotopas C, Davies L, Hackett R, Lessard CJ, Leon JM, Burtt NP, Guiducci C, Parkin M, Gates C, Plenge RM, Behrens TW, Wither JE, Roux JD, Fortin PR, Graham DC, Wong TC, Vye SJ, Bryant AR, Altshuler D, Gaffney PM. Genetic variants near TNFAIP3 on 6q23 are associated with systemic lupus erythematosus. Nat Genet 2008, 40:1059-1061.

45. Bates JS, Lessard CJ, Leon JM, Nguyen T, Battiest LJ, Rodgers J, Kaufman KM, James JA, Gilkeson GS, Kelly JA, Humphrey MB, Harleb J, Gray-McGuire C, Moser KL, Gaffney PM: Meta-analysis and imputation identifies a 108 kb risk haplotype spanning TNFAIP3 associated with lupus nephritis and hematologic manifestations. Genes Immun 2009, 10:470-477.

46. Kurreeman FA, Padyukov L, Marques RB, Schrödter J, Sédighzadeh M, Stoecken-Risbergens GA, van der Helm-van Mil AH, Allaart CF, Verduyn W, Houwing-Duistermaat J, Alfredsson L, Begovich AB, Klareskog L, Huizinga TW, Vye SJ: A candidate gene approach identifies the TRAF1/C5 region as a susceptibility locus for rheumatoid arthritis. PLoS Med 2007, 4:e267.

47. Kurreeman FA, Padyukov L, Marques RB, Schroeder HW Jr, Schrödter J, Sedighzadeh M, Stoeken-Risbergen G, van der Helm-van Mil AH, Allaart CF, Verduny W, Houwing-Duistermaat J, Alfredsson L, Begovich AB, Klæreskog L, Huizinga TW, Torres RE: A candidate gene approach identifies the TRAF1/C5 region as a risk factor for rheumatoid arthritis. PLoS Med 2007, 4:e267.

48. Barton A, Thomson W, Ke X, Eyer S, Hinks A, Bowes J, Matis LA, Smith AM, Goolden L, Morgan AW, Nash G, Smith KG: Copy number of FCGR3B, which is associated with systemic lupus erythematosus (SLE), correlates with protein expression and immune complex uptake. J Exp Med 2008, 205:1573-1582.

49. Fanciulli M, Norris JW, Pettretto E, Dong R, Harper L, Jolly EC, Watts R, Lau YL, Morgan AW, Nash G, Smith KG: A role from malignancy or sepsis, in patients with rheumatoid arthri...
66. Cloutier JV, Veillette A: Cooperative inhibition of T-cell antigen receptor signaling by a complex between a kinase and a phosphatase. J Exp Med 1999, 189:111-121.

67. Gjörloff-Wingren A, Saxena M, Williams S, Hammi D, Mustelin T: Characterization of TCR-induced receptor-proximal signaling events negatively regulated by the protein tyrosine phosphatase PTPN22. Eur J Immunol 1999, 29:3845-3854.

68. Vang T, Miletic AV, Anmura Y, Tautz L, Rickert RC, Mustelin T: Receptor tyrosine phosphatases in autoimmunity. Annu Rev Immunol 2008, 26:29-55.

69. Örru V, Tsai SJ, Rueda B, Fiorillo E, Stanford SM, Dickson GN, Tsai SJ, Rueda B, Fiorillo E, Stanford SM, Dasgupta J, Ferrari A, Young ID, Aboushi K, Chen XW, Martin J, Bottini N: A loss-of-function variant of PTPN22 is associated with reduced risk of systemic lupus erythematosus. Hum Mol Genet 2009, 18:569-579.

70. Orozco G, García-Porrúa C, Lépez-Nevot MA, Raya E, González-Gay MA, Martin J: Genomic structure and functional analysis of the PTPN22 superfamily of spondylopathies and a functional polymorphism of PTPN22 proposed as a general susceptibility marker for autoimmunity. Arthritis Res Dis 2006, 8:687-688.

71. Watford WT, Hiscott KM, Stanford SM, Desgupta J, Arntzen JU, Desgupta J, Stanford SM, Desgupta J, Arntzen JU, Desgupta J, Stanford SM, Desgupta J, Stanford SM: Multiple defects in the immune system of mice deficient in the co-stimulator OX40L with systemic lupus erythematosus. J Autoimmun 2009, 32:216-220.

72. Mathur AN, Chang HC, Zisoulis DG, Stritesky GL, Yu Q, O'Malley JJ, Kapur R, Levy DE, Kansas GS, Kaplan MH: Signaling by IL-12 and IL-23 and the immunoregulatory role of STAT4. Immunol Rev 2006, 202:139-156.

73. Remmers EF, Plenge RM, Lee AT, Graham RR, Hom G, Behrens MJ, Abraham C, Bayless TM, Bossa F, Griffiths AM, Ippoliti AF, Matarin M, Reveille JD, Vyse TJ, Gilkeson GS, Kamen DL, Mehrian-Shai R, Silverman E, Klein-Gitelman M, McCurdy D, N, Ciobanu V, Kaufman KM, Ojwang JO, Ziegler J, Quismorio FP, N Engl J Med 2007, 357:977-986.

74. Namjou B, Sestak AL, Armstrong DL, Zidovetzki R, Kelly JA, Jacob JT, Kapur R, Levy DE, Kansas GS, Kaplan MH: High-density genotyping reveals 38 independent genome-wide significant associations with systemic lupus erythematosus in different racial groups. Arthritis Rheum 2009, 60:1085-1095.

75. Kastelein RA, Hunter CA, Cua DJ: Discovery and biology of IL-23 and the IL-23 receptor in the immune system. Annu Rev Immunol 2007, 25:687-688.

76. Nair RP, Duffin KC, Helms C, Ding J, Stuart PE, Goldgar D, Gudat F, Targan SR, Schumm LP, Kistner EO, Lee AT, Gregersen PK, Targan SR, Schumm LP, Kistner EO, Lee AT, Gregersen PK, Targan SR, Schumm LP, Kistner EO, Lee AT, Gregersen PK: STAT4 and the risk of rheumatoid arthritis and systemic lupus erythematosus. N Engl J Med 2007, 357:977-986.

77. Namjou B, Sestak AL, Armstrong DL, Zidovetzki R, Kelly JA, Jacob JT, Kapur R, Levy DE, Kansas GS, Kaplan MH: High-density genotyping reveals 38 independent genome-wide significant associations with systemic lupus erythematosus in different racial groups. Arthritis Rheum 2009, 60:1085-1095.

78. Kozyrev SV, Abelson AK, Wojcik J, Zaghlool A, Linga Reddy MV, Bhandari S, Kontopidou F, Smith GP, Milton J, Brown CG, Hall KP, Evers DJ, Barnes CL, Bignell HR, Boutel JM, Bryant JT, Carter RJ, Keira Ceetham R, Cox AJ, Ellis DJ, Flat-bush MR, Gormley NA, Humphray SJ, Irving LJ, Karbelashvili MS, Kirk SM, Li H, Liu X, Massinger KS, Murray LJ, Obradovic B, Ots T, Parkinson BL, Pratt MR, et al: Accurate whole human genome sequencing using reversible terminator chemistry. Nature 2008, 456:53-59.

79. Ykrukov GV, Shpunt A, Stamatoyanopoulos JA, Sunyaev SR: Power of deep, all-exon resequencing for discovery of human trait genes. Proc Natl Acad Sci USA 2009, 106:3871-3876.
97. Chen Y, Zhu J, Lum PY, Yang X, Pinto S, MacNeil DJ, Zhang C, Lamb J, Edwards S, Sieberts SK, Leonardson A, Castellini LW, Wang S, Champy MF, Zhang B, Emilsson V, Doss S, Ghazalpour A, Horvath S, Drake TA, Lusis AJ, Schadt EE: Variations in DNA elucidate molecular networks that cause disease. Nature 2008, 452:429-435.