

Genetics of SLE: mechanistic insights from monogenic disease and disease-associated variants

Carola G. Vinuesa^{1,2,3}✉, Nan Shen^{4,5,6} & Thuvāraka Ware^{1,2}

Abstract

The past few years have provided important insights into the genetic architecture of systemic autoimmunity through aggregation of findings from genome-wide association studies (GWAS) and whole-exome or whole-genome sequencing studies. In the prototypic systemic autoimmune disease systemic lupus erythematosus (SLE), monogenic disease accounts for a small fraction of cases but has been instrumental in the elucidation of disease mechanisms. Defects in the clearance or digestion of extracellular or intracellular DNA or RNA lead to increased sensing of nucleic acids, which can break B cell tolerance and induce the production of type I interferons leading to tissue damage. Current data suggest that multiple GWAS SLE risk alleles act in concert with rare functional variants to promote SLE development. Moreover, introduction of orthologous variant alleles into mice has revealed that pathogenic X-linked dominant and recessive SLE can be caused by novel variants in *TLR7* and *SAT1*, respectively. Such bespoke models of disease help to unravel pathogenic pathways and can be used to test targeted therapies. Cell type-specific expression data revealed that most GWAS SLE risk genes are highly expressed in age-associated B cells (ABCs), which supports the view that ABCs produce lupus autoantibodies and contribute to end-organ damage by persisting in inflamed tissues, including the kidneys. ABCs have thus emerged as key targets of promising precision therapeutics.

Sections

Introduction

Monogenic SLE defects

Common SLE variants

Rare variants

Integrating genetic information

Conclusion

¹The Francis Crick Institute, London, UK. ²University College London, London, UK. ³China Australia Centre for Personalized Immunology (CACPI), Renji Hospital, Shanghai Jiao Tong University School of Medicine (SJTUSM), Shanghai, China. ⁴Shanghai Institute of Rheumatology, Renji Hospital, Shanghai Jiao Tong University School of Medicine (SJTUSM), Shanghai, China. ⁵Center for Autoimmune Genomics and Aetiology, Division of Human Genetics, Cincinnati Children's Hospital Medical Center, Cincinnati, OH, USA. ⁶Department of Paediatrics, University of Cincinnati, Cincinnati, OH, USA. ✉e-mail: carola.vinuesa@crick.ac.uk

Key points

- Monogenic systemic lupus erythematosus (SLE) commonly results from defects in degradation or increased sensing of nucleic acids and have led to the development of in vivo models of disease and mechanistic exploration of pathogenesis.
- The study of high frequency and variably penetrant common variants must be integrated with the study of low frequency, highly penetrant rare, ultra-rare and novel variants to clarify the genetic contribution to disease.
- Common variants identified through genome-wide association studies (GWAS) highlight a role for complement, immune complex handling, endosomal Toll-like receptors, type I interferon signalling, B cell tolerance and B cell receptor signalling pathways in SLE.
- The use of online transcriptomic databases to integrate GWAS data highlighted an increased expression of SLE-associated genes in age-associated B cells (ABCs).
- ABCs are expanded peripherally in patients with SLE, especially in those with lupus nephritis, and in the kidney during episodes of lupus nephritis.
- Therapeutic regimens that target CD19⁺ B cells are highly effective at depletion of ABCs, which express high levels of CD19, and had a promising effect in a small study of patients with refractory SLE.

Introduction

Systemic autoimmune diseases such as systemic lupus erythematosus (SLE) are caused by the inability of the immune system to distinguish self from non-self arising from loss of immune tolerance, which leads to aberrant and unregulated inflammation. These are lifelong, often devastating, illnesses with heterogeneous clinical phenotypes associated with multiple morbidity and disease courses that wax and wane with time¹. Many autoimmune diseases have female preponderance and represent a high burden in certain communities (for example, the prevalence of SLE is high in Afro-Caribbean populations).

Most autoimmune diseases are polygenic in nature. Over the past few decades, genome-wide association studies (GWAS) have identified hundreds of risk variants that are thought to act additively to cause disease². The relevant genes can exhibit epistasis³ or pleiotropy⁴, be susceptible to epigenetic changes^{5,6} and non-coding variation⁷, and display incomplete penetrance or variable expressivity⁸. However, GWAS only identify a minor portion of the inherited genetic risk associated with disease phenotypes, leading to 'missing heritability'.

By contrast, the rise of high-throughput technologies that enable whole-genome sequencing (WGS) and whole-exome sequencing (WES) has allowed the discovery of novel and rare or ultra-rare variants with strong effects, especially in SLE^{9–16}. These variants are being catalogued in databases of physiological and pathogenic human variation, increasingly allowing large-scale population research. These data also underscore the importance of genetic contribution to disease and the natural selective pressures that have led to these variations¹⁷. Functional validation of variant genes and their effects on downstream transcriptomic and proteomic pathways clarifies disease pathogenesis¹⁸. Now, the

pressing challenge is to integrate insights from monogenic causes of autoimmune disease with the contributions from common, rare or novel variants, in both coding and non-coding regions, their possible epistasis, and the epigenetic regulation of SLE risk loci.

SLE is the archetypal systemic autoimmune disease and is caused by a breach of tolerance to nucleic acids that leads to formation of autoantibodies to nucleic components such as double-stranded DNA (dsDNA) and RNA, thereby triggering inflammation in multiple organs¹⁹. Whether different molecular pathways to disease are responsible for the observed clinical heterogeneity and response to available therapies remains unclear. Transcriptomic analyses have identified key immune cells central to pathogenesis, such as B cells²⁰, and have defined patient subgroups according to molecular signatures²¹. These signatures seem to distinguish quiescent from active disease states²², although not consistently²³, making it difficult to dissect their functional relevance.

In this Review, we explore the genetic variants (monogenic, rare, ultra-rare and common) that cause, contribute to and are associated with SLE. We also discuss how common variants can be integrated with gene expression datasets to evaluate key immune cells involved in disease. Age-associated B cells (ABCs) have emerged from these analyses as key cells, and we examine them further, including their clinical relevance and their potential as therapeutic targets.

Monogenic SLE defects

Severe forms of SLE can be caused by alterations in single genes (thus termed monogenic SLE) and tend to present in childhood. Broadly, these alterations affect the complement system, extracellular and intracellular nucleic acid sensing and processing, RAS signalling and metabolic pathways (Table 1). Monogenic SLE is rare but increasingly recognized – damaging variants in genes known to cause monogenic SLE are identified in ~3–10% of patients with early-onset SLE^{9,10}. Nevertheless, the list of genes causing monogenic SLE is rapidly increasing with the use of next-generation sequencing technologies and monogenic SLE might be more prevalent than currently thought.

Complement defects

Homozygosity for loss-of-function (LoF) variants in the early proteins of the classic pathway of complement activation have long been known to cause SLE – complement defects explain <1% of SLE cases²⁴. Deficiency of C1 complex proteins and total C4 deficiency are responsible for most complement-related SLE. Of note, over 90% and up to 75% of patients with these immunodeficiencies develop SLE, respectively²⁵. Complement deficiency has been linked to SLE through multiple mechanisms. One of the best understood functions of complement is its ability to process and clear immune complexes, and apoptotic or dead cells. This function prevents the presentation of autoantigens to the immune system in the context of inflammatory signals. For example, research from the 1990s showed that the surface of apoptotic cells contains SLE autoantigens²⁶ and clearance of apoptotic cells is defective in macrophages from C1q-deficient mice and humans²⁷. Moreover, in the absence of C1q, immune complexes bind to plasmacytoid dendritic cells (pDC) and readily stimulate production of type I IFN²⁸. Notably, although C1q can bind DNA and cooperate with DNaseI to clear chromatin from dying cells²⁹, it can also exacerbate SLE by preventing the degradation of neutrophil extracellular traps (NETs) that promote production of type I IFN³⁰.

Inherited complement defects are very rare but complement proteins are very often reduced in patients with SLE, particularly in those with active disease owing to complement activation and

Table 1 | Monogenic causes of SLE

Gene (protein) ^a	Polarity	Location	Ligand or interacting protein ^b	Associated disease	OMIM ^c
Complement defects					
<i>C1QA, C1QB, C1QC</i>	LoF	Extra	IgG Fc, IgM Fc, LAIR1, C1S, C1R	C1Q deficiency	613652
<i>C1S</i>	LoF	Extra	C1R, C1Q	C1S deficiency	613783
<i>C1R</i>	LoF	Extra	C1S, C1Q	Ehlers–Danlos syndrome type 1	130080
<i>C2</i>	LoF	Extra	C4B	C2 deficiency	217000
<i>C4A, C4B</i>	LoF	Extra	Immune complexes	C4a and C4b deficiencies	614380, 614379
<i>C3</i>	GoF	Extra	CD21/CD35	C3 deficiency	613779
<i>FCN3</i>	LoF	Intra	GalNAc, C1S, C1R, C2, C4a, C4bB	Immunodeficiency owing to ficolin 3 deficiency	604973
Extracellular nucleic acid sensing					
<i>TLR7</i>	GoF	Endo	Guanosine, ssRNA, MyD88	SLE 17	301080
<i>DNASE1</i>	LoF	Extra	NET DNA	SLE	152700
<i>DNASE1L3</i>	LoF	Extra	dsDNA (apoptotic bodies)	SLE 16 and HUV	614420
<i>DNASE2</i>	LoF	Lys	Exogenous DNA	Autoinflammatory–pancytopenia syndrome	619858
<i>ADA2</i>	LoF	Lys/extra	Adenosine, 2'-deoxyadenosine	DADA2	28830446
Intracellular nucleic acid sensing					
<i>TREX1</i>	LoF	Intra	DNA (retroviral/transposon)	AGS 1, FCL, SSd, SS	225750, 610448, 152700
<i>TMEM173 (STING)</i>	LoF	Intra	cGAMP and excess DNA	SAVI, FCL	615934
<i>RNASEH2B</i>	LoF	Intra	RNA	AGS 2	610181
<i>RNASEH2C</i>	LoF	Intra	RNA	AGS 3	610329
<i>RNASEH2A</i>	LoF	Intra	RNA	AGS 4	610333
<i>SAMHD1</i>	LoF	Intra	dNTPs	AGS 5, FCL 2, CLL (somatic)	612952, 614415
<i>ADAR1</i>	LoF	Intra	dsRNA	AGS 6	615010
<i>IFIH1 (MDA5)</i>	GoF	Intra	dsRNA	AGS 7, FCL, Singleton–Merten syndrome type 1	615846, 182250
<i>ACP5 (TRAP)</i>	LoF	Intra	OPN	SPEND	607944
MAPK/PI3K					
<i>PTPN11 (SHP2)</i>	LoF	Intra	MAPK, PI3K	Noonan syndrome type 1	163950
<i>KRAS</i>	LoF	Intra	MAPK, PI3K	Noonan syndrome type 3	609942
<i>NRAS</i>	LoF	Intra	MAPK, PI3K	Noonan syndrome type 6	613224
<i>SHOC2</i>	LoF	Intra	MAPK, PI3K	Noonan syndrome-like	607721
<i>PIK3CD (PI3K p110δ)</i>	GoF	Intra	AKT	Immunodeficiency 14	615513
<i>P2RY8</i>	LoF	Memb	ERK, AKT	NA	NA
ALPS					
<i>PRKCD</i>	LoF	Intra	STAT1, NCF1	ALPS, type III	615559
<i>FAS</i>	LoF	Memb	FASL, CASP8, CASP10, RIPK1	ALPS, type IA	601859
<i>FASLG</i>	LoF	Memb	FAS	ALPS, type IB	601859
Metabolic					
<i>SLC7A7</i>	LoF	Intra	L-Arginine, leucine, glutamine	Lysinuric protein intolerance	222700
<i>PEPD</i>	LoF	Intra	Xaa-Pro amino acids	Prolidase deficiency	170100
<i>MAN2B1</i>	LoF	Lys	N-linked carbohydrates	α-Mannosidosis, types I and II	248500
<i>SAT1</i>	LoF	Intra	Polyamines	SLE	NA
<i>CYBB</i>	LoF	Intra	NADPH	Chronic granulomatous disease	306400

Table 1 (continued) | Monogenic causes of SLE

Gene (protein) ^a	Polarity	Location	Ligand or interacting protein ^b	Associated disease	OMIM ^c
Other					
<i>RAG1, RAG2</i>	LoF	Intra	V(D)J genes	SCID, ID with granulomas	233650
<i>IKZF1</i> (Ikaros)	LoF	Intra	Target genes	CVID 13, ITP	616873
<i>TNFAIP3</i> (A20)	LoF	Intra	TRAF6, TAX1BP1, IKKKG	Autoinflammatory syndrome, Behçet-like	616744

AGS, Aicardi-Goutières syndrome; ALPS, autoimmune lymphoproliferative syndrome; CLL, chronic lymphocytic leukaemia; CVID, common variable immunodeficiency; DADA2, adenosine deaminase 2 deficiency; dNTP, deoxyribonucleoside triphosphate; dsDNA, double-stranded DNA; dsRNA, double-stranded RNA; Endo, endosome; Extra, extracellular; FCL, familial chilblain lupus; GoF, gain-of-function; HUV, hypocomplementaemic urticarial vasculitis; ID, immune dysregulation; Intra, intracellular; ITP, immune thrombocytopenic purpura; LoF, loss-of-function; Lys, lysosome; MAPK, mitogen-activated protein kinase; Memb, plasma membrane; NA, not available; NET, neutrophil extracellular trap; OMIM, Online Mendelian Inheritance in Man; PI3K, phosphoinositide 3-kinase; SAVI, STING-associated vasculopathy with onset in infancy; SCID, severe combined immunodeficiency; SLE, systemic lupus erythematosus; SPEND, spondyloenchondrodysplasia with immune dysregulation; SS, systemic sclerosis; SSd, Sjögren syndrome; ssRNA, single-stranded RNA. ^aProtein name specified in instances where it differs from the gene name. ^bCorresponds to ligands or interacting protein identified in mice or humans according to the OMIM number. ^cOMIM catalogue gene or locus number.

consumption³¹. Such reduction in complement factors has been linked to defective complement-mediated opsonization of bacteria in patients with SLE³² and probably also impairs the removal of apoptotic cells, creating an excess of ligands for extracellular DNases and RNases.

Degradation of extracellular nucleic acids

Defects in a growing number of enzymes with important roles in the degradation of extracellular DNA – including those due to variants in *DNASE1* (ref. 33) and *DNASE1L3* (ref. 34) – can cause severe, yet typical SLE. DNase1 degrades dsDNA within NETs³⁵, whose physiological function is to capture extracellular pathogens and limit the spread of infections, although NET formation can also be triggered by sterile inflammation³⁶. Intriguingly, ~40% of patients with SLE fail to degrade NETs, although in most cases this defect is not due to *DNASE1* variants, but can instead be explained by the presence of antibodies that either bind to DNase1 and inhibit its activity³⁷, or bind to DNA in NETs and prevent DNase1 access³⁵. DNase1L3 degrades extracellular dsDNA that can be found on exposed nucleosomes within microparticles or apoptotic bodies. As with DNase1, 30% of all patients with SLE (and >50% of those with lupus nephritis (LN)) have antibodies to DNase1L3 with neutralizing activity against DNase1L3, thus phenocopying genetic DNase1L3 deficiency³⁸. Abundance of undigested extracellular nucleic acids, which can engage endosomal Toll-like receptors (TLRs), is therefore likely to be a very important trigger in SLE.

Endosomal nucleic acid sensing: TLR7 and TLR9

Increasing evidence suggests that nucleic acid signalling in endosomes is important in SLE pathogenesis³⁹. Accumulation of nucleic acids owing to inefficient degradation can activate the endosomal nucleic acid sensors TLR7 and TLR9. In mice lacking DNase1L3, splenomegaly is dependent on MyD88 (ref. 40), which is a key TLR signalling adaptor, with TLR7 and TLR9 driving anti-DNA responses and autoimmunity redundantly⁴¹. In addition to NETs and nucleosomal dsDNA exposed in apoptotic bodies or microparticles, mitochondrial DNA extruded from neutrophils, eosinophils and platelets represents an important source of endosomal TLR ligands in SLE, particularly for TLR9 (refs. 19,42).

Abundant evidence in mice and humans supports a role for increased TLR7 expression or activation in the development of SLE⁴³. Increased TLR7 expression causes lupus-like disease in mice by promoting the survival of immature self-reactive B cells⁴⁴ and most patients with SLE have signs of TLR7 hyperactivity within circulating B cells⁴⁵. In 2022, we reported a TLR7 gain-of-function (GoF) variation as a cause of human monogenic lupus¹⁸. A *TLR7* Y264H de novo

variant was identified in a child with SLE, including central nervous system (CNS) and kidney manifestations. Mice carrying the Y264H variant develop lupus-like disease consistent with an X-linked dominant pattern of inheritance. Intriguingly, the lupus-causing variant augmented the affinity for guanosine and its derivative 2',3'-cGMP, eliminating the need to bind uridine for TLR7 activation. This finding indicates that self-nucleic acids (including guanosine derivatives), rather than RNA from exogenous microbes, can trigger the disease.

TLR7 is highly expressed in pDCs (and, to a lesser extent, in DCs and macrophages) and is an important trigger of type I IFN secretion. Whether type I IFN is an initiating or driving factor, or a perpetuating or exacerbating factor in SLE has been a longstanding question. Notably, in experiments in *Dnase1l3*^{-/-} *Ifnar1*^{-/-} mice, lack of type I IFN signalling did not affect the initial emergence of anti-dsDNA antibodies at 1–3 months but the animals had reduced autoantibody titres at 9–12 months compared with *Dnase1l3*^{-/-} animals⁴¹. Similarly, deficiency in the receptor for type I IFN reduced but did not abolish autoantibody formation in TLR7-dependent lupus mouse models^{38,46}. Type I IFN can increase the expression of TLR7 and MHC, and therefore serves as a positive feedback loop that enhances nucleic acid sensing and antigen-presenting capacity, as well as plasma cell differentiation and thus autoantibody production¹⁹.

The role of TLR9 is complex and can elicit effects that oppose those of TLR7, depending on the cell type. For example, stimulation of B cells with immune complexes containing TLR7 ligands induces B cell-activating factor (BAFF)-independent proliferation and survival, whereas stimulation with immune complexes containing TLR9 ligands induces apoptosis of proliferating cells, which can be rescued by BAFF⁴⁷. In line with this protective effect, TLR9 stimulation can limit nephritis in the MRL/*lpr* model⁴⁸. By contrast, provision of NET-derived TLR9 ligands to human pDCs in vitro leads to the production of type I IFN⁴⁹.

Cytoplasmic nucleic acid sensing

Defects in extracellular DNases or increased endosomal sensing of nucleic acids have been associated with classic SLE phenotypes, whereas defects in genes that encode intracellular nucleic acid sensors or enzymes that detect or degrade cytoplasmic nucleic acids (*TREX1*, *STING1*, *RNASEH2B*, *RNASEH2C*, *RNASEH2A*, *SAMHD1* and *IFIH1* (also known as *MDA5*)) tend to cause severe interferonopathies⁵⁰. These defects are often associated with autoinflammatory diseases including Aicardi-Goutières syndrome (AGS)⁵⁰, STING-associated vasculopathy with onset in infancy (SAVI)⁵¹, MDA5 associated Singleton-Merten syndrome⁵² and familial chilblain lupus⁵³. Nevertheless, 0.5–2% of adult

patients with SLE have damaging *TREX1* rare variants^{50,54}. Of note, *TREX1* has been reported to digest DNA from endogenous retroviruses⁵⁵ that are normally kept in check via TLR7 signalling⁵⁶. Importantly, extracellular and intracellular sensing of nucleic acids is not strictly compartmentalized and crosstalk between these two pathways has been described. Under certain circumstances, nucleic acids can escape the endosome or lysosome, and activate cytoplasmic sensors⁵⁷. Moreover, type I IFN signalling induced by cytoplasmic nucleic acid sensing can increase expression of endosomal TLRs⁵⁸, thus lowering their signalling threshold.

RASopathies and elevated ERK and AKT signalling

RASopathies due to mutations in *PTPN11*, *KRAS*, *NRAS* and *SHOC2* that activate mitogen-activated protein kinase (MAPK) and phosphoinositide 3-kinase (PI3K) typically cause Noonan syndrome⁵⁹, which is characterized by the presence of atypical facial features, short stature, heart defects and developmental delays. Occasionally, patients with variants in these genes present with SLE⁶⁰, which is thought to result from increased activation of extracellular-signal-regulated kinase (ERK) and/or PI3K in lymphoid cells. Elevated RAC serine/threonine protein kinase (AKT) and ERK signalling also occurs in patients with SLE who have damaging hypomorphic variants in *P2RY8* (ref. 61). In addition to confining B cells to germinal centres, *P2RY8* exerts other roles in human B cell tolerance, including limiting plasma cell formation and enforcing negative selection of self-reactive B cells⁶¹. Increased AKT signalling probably also underlies autoimmunity seen in patients with GoF variants in *PIK3CD*. Although these patients typically develop type 1 activated PI3K δ syndrome⁶², an association with paediatric SLE has also been reported⁶³.

Defects in lymphocyte apoptosis

Autoimmune lymphoproliferative syndrome (ALPS), caused by germline or somatic variants in *FAS* or *FASLG*, is usually distinct from SLE. Patients present with enlarged lymph nodes, spleen and liver, and do not typically develop antinuclear antibodies, although two-thirds of patients present with autoimmune haemolytic anaemia and thrombocytopenia. However, SLE-like presentations have been noted in a few case reports⁶⁴. Similarly, LoF missense variants in *PRKCD* were initially reported in a family with a history of juvenile SLE characterized by increased B cell survival⁶⁵. Subsequent studies found that *PRKCD* LoF variants typically cause an ALPS type III, which is characterized by pronounced B cell proliferation and survival, particularly of immature B cells⁶⁶.

Metabolic syndromes

Defects in metabolic pathways can also give rise to SLE-like phenotypes. Examples include variants in *SLC7A7* (lysine/arginine transporter)⁶⁷, *PEPD* (proline dehydrogenase)⁶⁸, *MAN2B1* (α -mannosidosis)⁶⁹, *CYBB* (defective phagocyte NADPH oxidase complex, although patients rarely have autoantibodies)⁷⁰ and, the most recent addition, *SAT1* (encodes a rate-limiting enzyme in polyamine catabolism)⁷¹. In two unrelated families, an unaffected mother transmitted the X-linked *SAT1* variant to two affected male children. One of the variants, when introduced in mice, caused a lupus-like disease with kidney damage, neutrophil defects and reduced regulatory CD4 T cell subsets⁷¹. It has been proposed that excess polyamine due to *SAT1* deficiency triggers SLE owing to its ability to bind and modify DNA⁷². Polyamines are found largely bound to RNA and DNA in mammalian cells and, given the high expression of *SAT1* in neutrophils, they might stabilize NETs. Polyamines also

promote a change in DNA structure to left-handed (Z-DNA), which is a target of dsRNA-specific adenosine deaminase (encoded by *ADAR*) and has been linked to SLE and AGS⁷³.

Common SLE variants

The most common gene variants are single nucleotide polymorphisms (SNPs), followed by insertion/deletion (indel) and structural variants⁷⁴. SNPs are classified as common if the minor allele frequency (MAF) is >0.05 (5%). If SNPs occur within the coding exon without a resultant change in the amino acid, the protein encoded will remain the same and the variants are designated synonymous. By contrast, a missense variant arises from a SNP that leads to an amino acid alteration with subsequent change in protein structure and/or function. Nonsense variants refer to those that introduce a premature stop codon, often leading to nonsense-mediated RNA decay or formation of a truncated (that is, smaller) protein. However, most SNPs occur in non-coding introns that are found near to, or in linkage disequilibrium with, coding pathogenic variants⁷. Importantly, these SNPs can also occur in promoter or enhancer regions, and exert roles in the regulation of gene transcription⁷⁵.

SLE GWAS

GWAS attempt to identify common risk variants that are present in a disease cohort at a significantly higher frequency than in healthy individuals. Additive effects of many risk alleles underpin the polygenic nature and non-Mendelian inheritance of disease. However, the information provided by GWAS is a double-edged sword. Because the risk variants are only associated with the clinical phenotype, clarifying the biological functional consequence and establishing causality is difficult. Firstly, mapping relevant genes to SNPs based on proximity alone misses genes with long-range effects⁷⁶ and can incorrectly identify genes that are co-segregated in linkage disequilibrium⁷⁷. Post-GWAS analyses such as statistical fine mapping can mitigate these limitations⁷⁸.

Glossary

Burden analyses

Using a gene-based approach, these analyses evaluate the aggregate burden of variants in a gene, comparing disease and control cohorts.

Epistasis

The interaction between independently inherited genes leading to a change in the phenotypic expression.

Expressivity

The degree to which a particular genotype is expressed phenotypically, reflecting individual variability owing to the environment, modifier genes and epigenetics.

Neutrophil extracellular traps

(NETs). Net-like structures primarily composed of chromatin and

antimicrobial proteins released by dying neutrophils.

Pleiotropy

Denotes that a genetic variation can affect multiple phenotypic traits.

Statistical fine mapping

A process of refinement after a genome-wide association study to identify variants that are likely to be causal; includes rigorous quality control, high confidence in genotyped single nucleotide polymorphisms and analysis of a population large enough to identify genes in linkage disequilibrium.

Trios

Refer to an affected proband and their parents.

Use of expression quantitative trait loci (eQTL), which are genomic loci that can explain the variance of expression phenotypes in different cell types, can improve the validity of the mapping⁷⁹. However, owing to inconsistencies between the various models integrating these factors, GWAS genomic annotation is often still based on proximity to the SNP.

Multiple ancestry SLE GWAS that include hundreds of thousands of cases and controls are now available. Summary data on SLE (accession ID 0002690) can be obtained from the [GWAS Catalogue](#), which is a curated online GWAS dataset delivered and developed by the European Molecular Biology Laboratory–European Bioinformatics Institute and the National Human Genome Research Institute (Fig. 1). In this database, the genomic annotations for SNPs are provided by an Ensembl mapping pipeline that assigns SNPs to genes based on proximity.

Common SLE genetic variants are consistently associated with the highly polymorphic and gene-dense region encoding the HLA⁸⁰. The haplotype and linkage disequilibrium inheritance patterns of HLA complicate interpretation of SNPs in this region, but, given its centrality to antigen presentation, the risk association is likely to be more substantial than currently appreciated. Other notable SLE-associated immune genes include *STAT4*, *IRF5* and *BLK*, which are all involved in the regulation of type I IFN responses (reviewed in refs. 15,81–83). Broadly, SLE GWAS suggest roles for complement activation, handling of immune complexes, endosomal TLRs and type I IFN signalling, as well as B cell tolerance and B cell receptor (BCR) signalling pathways.

Over the past five years, genetic association and linkage studies have been adapted to the study of whole-exome sequences from SLE pedigrees and burden analyses. One study searched for gene-disrupting effects in two large SLE pedigrees from Iceland¹². Although no single variant segregated with disease in all affected individuals, a small set of coding variants were shared by these families, with the most robust effect found at the *FAM71E1/EMC10* locus (*FAM71E1*p.L7F, rs185418641). This is a low frequency variant, with a MAF in non-Finnish Europeans of 0.0096 (~1 in 50 individuals, gnomAD genomes v3.1.2). A subsequent study of 24 SLE pedigrees identified *CDC27* as a gene that passed the burden test based on variants with a MAF of 0.1 (present in 1 in 20 individuals)⁸⁴. Although no individual variant underwent functional validation to establish disease causation, *CDC27* expression was low in patients with SLE and correlated with disease activity indicators, such as elevated C-reactive protein and erythrocyte sedimentation rate, and low complement C3 and C4.

Rare variants

For decades, the prevailing view was that non-Mendelian SLE would predominantly result from the additive effect of many common variants but in the past 5 years it has become clear that rare variants (MAF <0.005 (0.5%)), which are present in <1% individuals, are also likely to contribute to disease, and can often have large effects⁸⁵. WGS and WES have enabled the interrogation of the entire non-coding and coding genome in health and disease. The combination of segregation studies with functional analyses can often establish a causal relationship between genotype and phenotype, even for previously undescribed variants, referred to as ‘novel’, or that are very rare in the population and thus unlikely to be revealed by GWAS.

Nearly a decade has passed since the first report of the use of WES to uncover an SLE gene defect, which revealed a homozygous rare *TREX1* missense variant in a child from a consanguineous family that had severe SLE with CNS manifestations¹³. This report was the first demonstration of childhood-onset SLE caused by a *TREX1* variant. An increasing number of studies are now using WES, WGS or sequencing

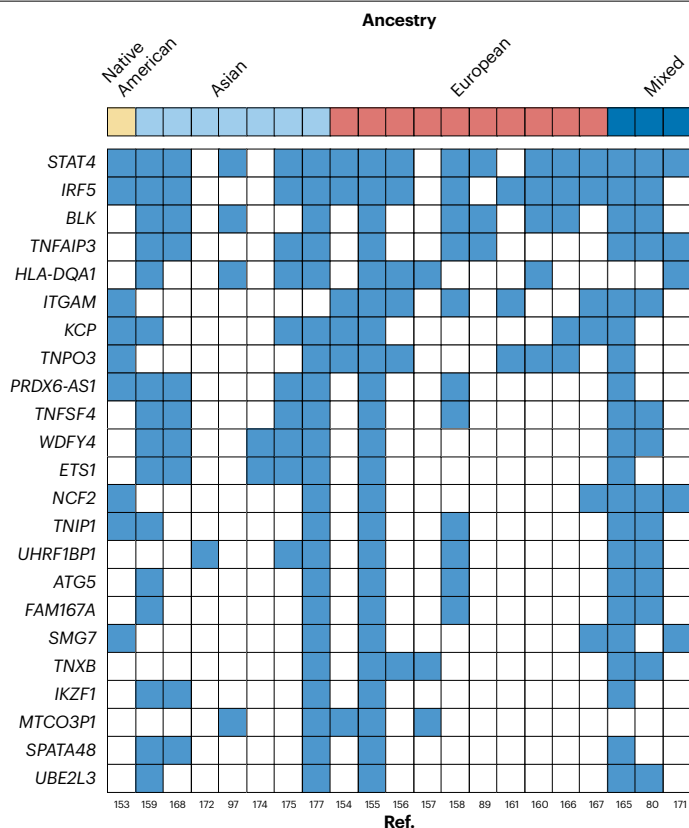


Fig. 1 | Mapped common variants identified by SLE GWAS. The figure illustrates selected common variants that have been identified by systemic lupus erythematosus (SLE) genome-wide association studies (GWAS). Genes associated with significant single nucleotide polymorphisms (that is, with a genome-wide significance $<5 \times 10^{-8}$)¹⁵² in the summary data from the GWAS Catalogue are listed by ancestry^{89,96,97,153–177}. This figure includes only the most common genes identified by five or more studies to improve validity of findings across multiple ancestries – a comprehensive list of all SLE-associated genes found in two or more SLE GWAS can be found in Supplementary Fig. 1. Studies exploring shared susceptibility or response to treatment have been excluded. A filled-in box denotes that the gene has been mapped by that GWAS.

gene panels to screen for novel or rare coding and non-coding variants in genes that are either associated with SLE based on GWAS or known to cause monogenic SLE^{9–11,14,62}.

In our 2019 study, we found that >80% of patients with SLE carry variants in genes associated with SLE, including functional *BLK* variants, which were detected in up to 12% of patients¹⁴. Importantly, although the rare variant burden was not significantly different between patients and healthy individuals after correction for multiple testing and inferred ethnicity, the functional consequences differed. Variants that were found only in patients failed to repress IFN β activity in a luciferase assay, in contrast to rare *BLK* variants that were found only in healthy individuals. Notably, a 2021 study of the same cohorts found that the rare variant burden was higher in patients when the test was restricted to variants that affected proteins within the endosomal and/or TLR pathway¹⁵. Our 2019 study¹⁴ highlighted the importance of functional validation for strengthening the link to causation. This work also uncovered novel mechanisms by which B cell scaffold protein with

ankyrin repeats (BANK1) and tyrosine protein kinase BLK prevent autoimmunity – BANK1 silences TNF receptor-associated factor 6 (TRAF6) in sequestosomes and BLK limits the nuclear translocation of IRF5.

In a Swedish study, sequencing of 71 trios revealed a significant enrichment in ultra-rare (MAF <0.001 (0.1%)) nonsense and missense

variants in 22 genes known to cause monogenic SLE, with 21 variants found in 18 patients, and 8 of them thought to have high potential to contribute to SLE⁹. The UK Juvenile-Onset SLE Cohort Study¹⁰ sequenced a panel of 23 genes associated with Mendelian SLE, as well as 40 genomic regions and 59 genes associated with SLE by GWAS. *C3*, *SAMHD1*, *TREX1* and the *RNASEH2* family harboured the most variants. This study also noted that patients with genetic risk variants developed neuropsychiatric symptoms over time; this effect might have been linked to an observed enrichment in interferonopathy-causing gene variants.

A small WES study in seven children with SLE and lymphoproliferation, revealed somatic *NRAS* activating variants in four of the seven patients, and nonsense and missense mutations in *TNFAIP3* and *PIK3CD* in two of the remaining three patients¹⁶. These findings highlight that lymphoproliferation is probably part of the ALPS spectrum and less likely to be driven by interferonopathy-causing genes. Rare variants affecting the expression of SLE gene-associated loci have also been found in non-coding regulatory regions of the genome, including functional SNPs regulating the *ITGAM1*, *IRF5*, *ETS1* and *TNIP1* loci⁸⁶.

Trio studies can reveal gene variants that have high potential to be pathogenic such as de novo and compound heterozygous (also known as biallelic) variants, as highlighted above with the discovery of monogenic SLE caused by TLR7 GoF or SAT1 LoF. Two-thirds of severe genetic diseases that occur in the neonatal period and infancy are caused by de novo variants⁸⁷. A de novo frameshift variant in *TNFAIP3*, which encodes TNFAIP3 (also known as zinc-finger protein A20), was found in a proband with neuropsychiatric SLE⁸⁸. This gene has been linked to multiple autoimmune diseases by GWAS, including SLE⁸⁹. The variant impaired K63-linked ubiquitylation of TRAF6 in vitro, which suggests a contribution to disease, although a causal relationship cannot be inferred.

Using trio sequencing, we identified a de novo *P2RY8* LoF variant in childhood-onset SLE and several *P2RY8* rare variants in adult SLE and primary antiphospholipid syndrome⁶¹. The variants were damaging to protein function in vitro. Of note, although *P2RY8* is not present in mouse B cells, these could be retrovirally transduced to express the wild-type or mutant protein. These assays showed impaired germinal centre confinement of B cells expressing variant *P2RY8* and uncovered a role for *P2RY8* in enforcing negative selection of B cells at the transition from immature to follicular recirculating stage and limiting plasma cell formation.

In another case of childhood-onset SLE, we identified a missense de novo variant in *PACSIN1* (ref. 90). This *PACSIN1* variant (termed Q59K) interferes with the ability of TRAF4 to repress TRAF6, leading to increased type I IFN production. Of note, although the biochemical assays in human cells and cell lines showed that the human Q59K variant has profound effects, such as augmented IFN β production in response to TLR7 ligands, the orthologous variant only causes mild phenotypes in mouse B cells. Thus, although the results support a contributing effect of the variant to human SLE, whether it is sufficient to cause disease cannot be determined. Differences in mouse versus human phenotypes are frequently seen in autoimmune disease and might be explained by redundancy between gene family members in mice, or the requirement for environmental triggers or specific genetic backgrounds, including HLA haplotype for expressivity.

The challenge of variant interpretation

The growing use of genomic sequencing in the clinic is rapidly producing an ever-increasing list of putative disease-causing variants. Their clinical utility relies on our ability to classify those variants according

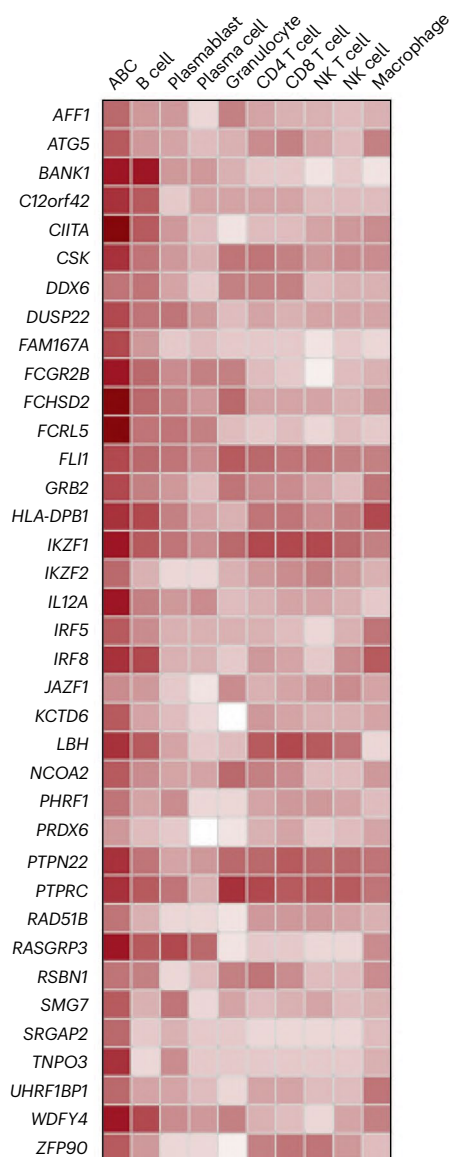


Fig. 2 | Cell-specific expression of SLE GWAS genes. The heatmap illustrates the log2-transformed expression of genes associated with significant single nucleotide polymorphisms in two or more systemic lupus erythematosus (SLE) genome-wide association studies (GWAS) in ten key types of immune cell^{89,96,97,153–177} (Supplementary Fig. 1). These immune cell subsets include age-associated B cells (ABCs), which have been defined as CD19⁺CD27⁺CD21⁺CD10⁺ B cells. Genes that are most highly expressed in ABCs relative to any other cell type are shown – a more comprehensive figure for the cell-specific expression of all significantly associated SLE GWAS genes can be found in Supplementary Fig. 2. The heatmap was created using the online transcriptome database GENEVESTIGATOR (Affymetrix Human Genome U133 Plus 2.0 array) and a compendium-wide analysis using conditional search of cell types. Hierarchical clustering performed by cell type. NK cell, natural killer cell.

to the likelihood that they are disease-causing. Variant interpretation has important ramifications for clinicians, patients and their families, as it can provide a genomic diagnosis (thus ending the diagnostic odyssey and offering reproductive options in families with inherited severe disease) and guide the use of precision therapies. In a clinical setting, the American College of Medical Genetics and Genomics established a framework for variant interpretation that is useful to establish variant pathogenicity, particularly for variants occurring in genes that are already known to be disease-causing⁹¹. Within this framework, causal relationships between a gene variant and disease can often be established in the absence of functional validation. For example, if a de novo variant is predicted to cause LoF in a gene whose loss is a known mechanism of disease, then the variant can already be classified as pathogenic. As the number of sequencing studies identifying rare variants in patients with SLE and the sizes of cohorts increase, the accuracy of variant classification and ability to predict pathogenicity will also increase.

An National Institutes of Health initiative known as the Clinical Genome Resource (ClinGen) consortium was launched in 2013 with the aim of creating an authoritative public resource to further refine and standardize the criteria used to evaluate sequence variant pathogenicity for application to clinical diagnosis and precision therapies. This resource was tasked with aggregating genomic and health data, and curating disease genes and gene variants according to gene disease validity, variant pathogenicity, gene dosage sensitivity and clinical actionability⁹². In 2022, a Rheumatologic Autoimmune Clinical Domain Working Group was created, and plans are underway to incorporate gene and variant curation expert panels focused on SLE. This multi-institutional team will classify the clinical validity of SLE genes and determine variant pathogenicity.

Integrating genetic information

The genetic risk associated with common, rare, ultra-rare and novel variants, and their respective weighted effect sizes, can be integrated to create polygenic risk scores (PRS) to help predict the likelihood of a particular clinical phenotype⁹³. Risk variants are assumed to be inherited independently, so the effect sizes associated will be additive, which is how the PRS is currently calculated⁹⁴. Although this approach does not take into account inter-gene or gene–environment interactions, twin studies suggest that an additive paradigm probably reflects true risk heritability⁹⁵. In SLE, ancestry-matched predictors (that is, PRS applied to a patient of the same ancestry as that from which the PRS was created) are significantly better at predicting disease risk than ancestry-mismatched predictors⁹⁶. Some translatability exists between similar ancestry groups – for example, a GWAS dataset from Chinese patients has been used to create a valid PRS model for Thai patients that can distinguish individuals with SLE from healthy individuals⁹⁷.

Analysis of cell-type-specific gene expression

Another approach to integrating and distilling information gleaned by GWAS is to compare gene expression data across multiple cell types. Several transcriptome datasets are available, including GENEVESTIGATOR, an online database incorporating information from multiple public and private gene expression studies, including the Gene Expression Omnibus managed by the National Center for Biotechnology Information, and The Cancer Genome Atlas supervised by the National Human Genome Research Institute. Integrating gene expression profiles with GWAS of autoimmune diseases highlighted an important role for B cells in SLE⁹⁸. Similarly, analysis of epigenetic data using a marker of active enhancers

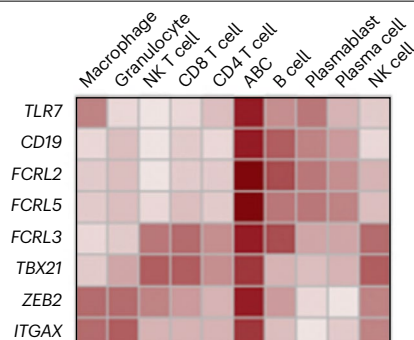


Fig. 3 | Cell-specific expression of genes characteristic of ABCs. The heatmap illustrates the log2-transformed expression of age-associated B cell (ABC) markers in ten key types of immune cells. The heatmap was created using the online transcriptome database GENEVESTIGATOR (Affymetrix Human Genome U133 Plus 2.0 array) and a compendium-wide analysis using conditional search of cell types. ABCs were defined as CD19⁺CD27⁺CD21⁺CD10⁺ B cells. NK cell, natural killer cell.

(acetylated histone 3 lysine 27 (H3K27Ac)) across different cell types revealed that B cells are likely to be the most important cell type in SLE⁵.

Across ten human immune cell subsets, a large fraction of SLE-associated genes are most highly expressed in CD19⁺CD21⁺CD27⁺CD10⁺ B cells, followed by neutrophils (Fig. 2). Although these cells were termed ‘anergic B cells’ in the study that generated the transcriptomic data⁹⁹, the CD19⁺CD21⁺CD27⁺CD10⁺ B cell subset is now known to correspond to ABCs, which are atypical B cells that are also known as CD11c⁺ or double negative 2 (DN2) B cells^{45,100,101}. Specifically, these cells express high levels of *TBX21* (encoding T-bet), *FCRL2*, *FCRL5*, *FCRL3*, *ITGAX* (encoding CD11c), which are typical ABC markers, and other genes known to be highly expressed in ABCs, including *ITGAM* (encoding CD11b), *TLR7* and *ZEB2* (Fig. 3).

Further evidence that CD19⁺CD21⁺CD27⁺ B cells are ABCs comes from their characterization in a 2021 study, in which nearly 95% of CD19⁺CD21⁺CD27⁺ human circulating B cells were IgM⁺ and CD11c⁺ (ref. 100). These CD19⁺CD21⁺CD27⁺ cells also express the phenotype (Fig. 4) and transcriptional profile that characterize ABCs in patients with SLE¹⁰⁰. Further, the study showed that ABCs with this same transcriptomic profile are found in patients with other autoimmune diseases such as rheumatoid arthritis¹⁰², in patients with common variable immunodeficiency who also present with autoimmune manifestations, and in individuals with HIV infection or malaria¹⁰⁰.

ABC biology. ABCs were initially described in 2011 by two independent groups of researchers and were found in aged or autoimmune-prone mice^{103,104}. Subsequent studies in humans detected these atypical, activated B cell subsets in individuals with autoimmune or infectious diseases, and noted that they have phenotypic markers that overlap with those of mouse ABCs and activated naive B cells^{105,106}. ABCs are typically class-switched antigen-experienced B cells that are characterized by the expression of CD11c (traditionally associated with DCs) and Fc receptor-like protein 5 (FcRL5); unlike recirculating naive follicular B cells, ABCs express very low amounts of CD21, CD23 and CXC-chemokine receptor 5 (CXCR5)^{45,104,107–109}. Remarkably, ABCs found in the context of autoimmunity are phenotypically similar to those identified in the context of infection and ageing^{100,110}, although they might be functionally different¹¹¹.

CD11c⁺ B cells are heterogeneous and whether they represent a distinct B cell subset remains controversial. Nonetheless, transcriptomic signatures obtained from both bulk and single-cell data, can distinguish ABCs from other B cell subsets fairly convincingly, particularly in terms of unique migration and activation requirements^{6,20,45,100,108,112}. The ABC signature includes high expression of specific surface integrins required for cell adhesion and migration (α X integrin (also known as CD11c and encoded by *ITGAX*) and β 2 integrin (encoded by *ITGB2*)), unique transcription factors (T-bet, zinc-finger E-box-binding homeobox 2 (ZEB2) and zinc-finger and BTB domain-containing protein 32 (ZBTB32)), cytokine receptors (IL-10 receptor (IL-10R) and IL-21R) and chemokines (CXCL9 and CXCL10), and low expression of markers that characterize other B cell subsets (CD27, CXCR5, complement receptor type 2 (CR2) and low affinity immunoglobulin- ϵ Fc receptor (FCER2)) (Fig. 5). Although both human and mouse ABCs can be identified by high expression of CD11c and FcRL5, as well as low expression of CD21 and CXCR5, some inter-species differences might exist. For example, mouse splenic ABCs cells comprise both IgM⁺ and IgM⁻ cells, whereas

human circulating ABCs are largely isotype-switched^{100,108,113}. Nevertheless, some evidence suggests that spleen and circulating ABC pools are not in equilibrium, possibly because some ABCs can become spleen-resident¹⁰⁴. ABC formation is driven by chronic inflammation and sustained exposure to stimuli such as excessive autoantigens, TLR ligands, cytokines and T cell help¹¹⁴. TLR7 signals are crucial for ABC formation, both in vitro and in vivo, which can be influenced by regulators of the TLR7 pathway^{18,45,104,108,115–117}. The need for CD40 signals for ABC formation and survival highlights a dependence on T cell help^{45,115,118}. Cytokine signals, particularly IFN γ and IL-21, are also essential for ABC differentiation^{45,113,115,119}. Mice lacking the RHO GTPase-regulatory proteins DEF6 and SWAP70, which bind to and sequester IRF5 in the nucleus of ABCs, spontaneously develop an SLE-like disease with accumulation of ABCs in an IRF5-dependent¹²⁰ and IRF8-dependent manner¹¹⁶. In mice, ABCs can be derived from follicular B cells and show a developmental trajectory that differs from that of germinal centre

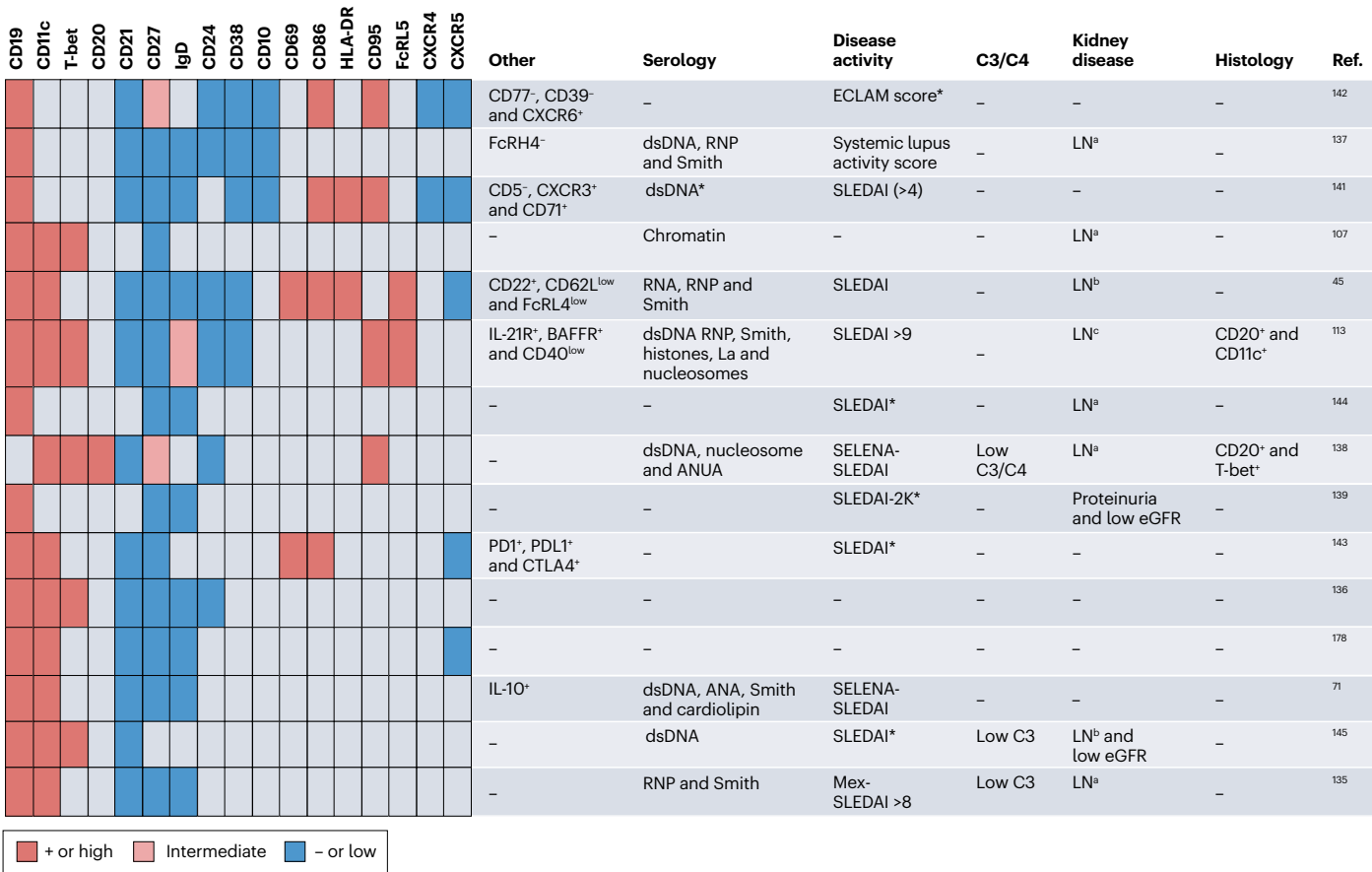


Fig. 4 | Human ABC phenotypes and their clinical correlates in SLE. Several reports suggest that age-associated B cell (ABC) phenotypes are expanded in patients with systemic lupus erythematosus (SLE) compared with controls. Some studies found that disease activity either correlated (rows where no cut-off is given for disease activity scores) or was associated (where discrete scores are defined) significantly with ABC frequencies. Disease activity scores and serology marked with an asterisk represent instances where no significant correlation nor association with ABCs was identified. Of note, complement C3 and C4 correlate negatively with ABC frequencies, and lupus nephritis

(LN) is significantly associated with an expanded ABC population. Grey boxes represent untested or unreported results. ANUA, anti-nucleosome antibody; dsDNA, double-stranded DNA; ECLAM, European Consensus Lupus Activity Measurement; eGFR, estimated glomerular filtration rate; Mex-SLEDAI, Mexican SLEDAI; RNP, ribonucleoprotein particle; SELENA-SLEDAI, Safety of Estrogens in Lupus Erythematosus National Assessment – SLEDAI; SLEDAI, Systemic Lupus Erythematosus Disease Activity Index. ^aLN classified according to American College of Rheumatology (ACR) guidelines. ^bLN diagnosed through kidney biopsy. ^cLN samples obtained from a biobank.

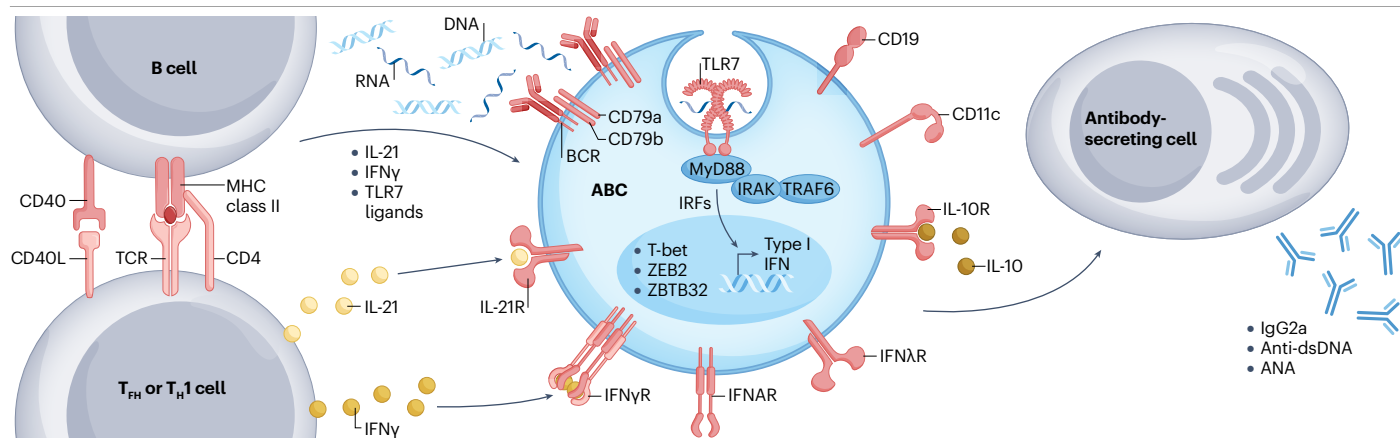


Fig. 5 | Formation and key molecular phenotype of APCs and their interactions. B cells present MHC-restricted antigens to T follicular helper (T_{fh}) cells or effector T helper 1 (T_H1) cells, with co-stimulation via CD40–CD40L. T_H1 cells can also secrete IL-21, which, in combination with Toll-like receptor 7 (TLR7) ligands, drive the B cell to adopt an age-associated B cell (ABC) phenotype. APCs express high levels of CD19 and CD11c, receptors for IL-21, IFN γ , type I interferon, IFN α and IL-10, and the transcription factors T-bet, ZEB2 and ZBTB32;

these cells also secrete IFN γ . Detection of nucleic antigens via the endosomal receptor TLR7 can also induce APCs to become antibody-secreting cells. ANA, antinuclear antibodies; BCR, B cell receptor; dsDNA, double-stranded DNA; CD40L, CD40 ligand; IFNAR, interferon- α/β receptor; IL-21R, IL-21 receptor; IRAK, IL-1 receptor-associated kinase; IRF, interferon regulatory factor; TCR, T cell receptor; TRAF6, TNF receptor-associated factor 6; ZBTB32, zinc-finger and BTB domain-containing protein 32; ZEB2, zinc-finger E-box-binding homeobox 2.

B cells¹⁰⁴. However, whether other B cell subsets can give rise to APCs and contribute to their accumulation in vivo is unclear. APCs can rapidly differentiate into antibody-secreting cells (ASCs), also known as plasma cells, upon encounter with TLR stimuli^{107,108}. Although APCs produce fewer antibodies than CD138⁺ plasma cells, CD138⁺ APCs seem capable of antibody secretion¹²¹. In lupus-prone mice, APCs produce anti-chromatin autoantibodies and ex vivo cultures of mouse APCs showed that they can secrete autoantibodies of multiple specificities^{103,107,120,122}. APCs also have a hyperactivated state with upregulated CD80, CD86 and MHCII expression, which can promote immune synapse formation and interaction with T cells^{103,104}. This hyperactivated state of APCs is consistent with the hypermetabolic feature ascribed to them – transcriptome analysis suggests high glucose uptake and mitochondrial respiration^{123,124}, and real-time metabolic measures, such as oxygen consumption rate, are increased^{123,125}. APCs can also migrate and infiltrate affected organs in multiple diseases^{126–129}.

Elucidating the mechanisms of transcriptional regulation and their effects on the formation of APCs is a crucial area of study in APC biology. T-bet has been identified as a regulator of APC formation in autoimmune conditions and has been associated with the increased presence of APCs expressing IgG2A or IgG2C^{109,122,130}. Furthermore, patients with inherited T-bet deficiency have low frequency of CD11c⁺CD21^{low} B cells in the circulation¹³¹. T-bet has been implicated in regulating APC formation through metabolic control¹²³. However, although T-bet promotes class switching to IgG2A or IgG2C, it does not seem to have a specific role in the differentiation of APCs^{116,132}. By contrast, we and others have found that ZEB2 has a crucial role in driving APC formation in both mice and humans (findings reported in preprint studies^{133,134}). B cell-intrinsic ZEB2 deficiency impairs APC differentiation in both autoimmune and infectious diseases, suggesting that it might be the master regulator that determines APC identity and pathogenicity in SLE.

APCs in SLE. Considerable evidence shows an expanded circulating APC population in SLE compared with healthy individuals (Fig. 4), in an age-independent way^{45,113}, with higher frequencies in patients of

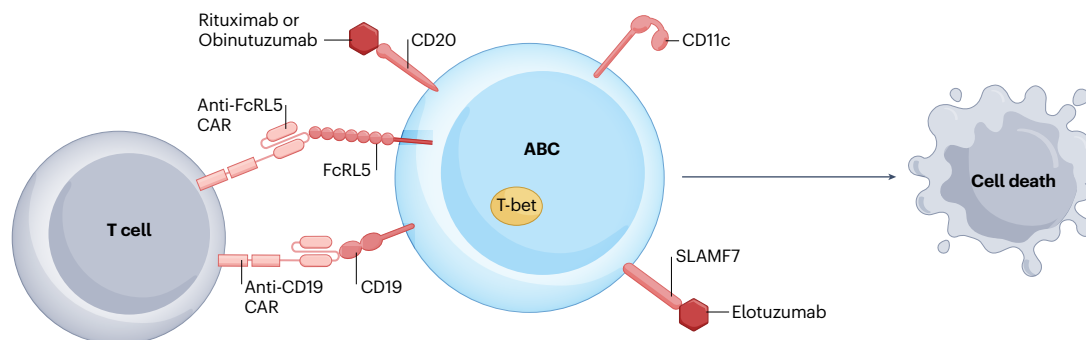
African ancestry compared with those of European ancestry^{45,113,135}. Beyond the pronounced type I IFN signature seen in SLE peripherally^{21,23} and at sites of disease²⁰, type II IFN is required for APC formation and differentiation^{45,119} and type III IFN λ is high in SLE and positively correlates significantly with frequency of APCs¹³⁶. Interestingly, *IFNLRI* is also upregulated in CD19⁺CD11c⁺IgD⁺CD27⁺CD21⁺CXCR5⁺ APCs⁴⁵ and in activated B cells from mice carrying an orthologue of a human SLE-causing *TLR7* variant¹⁸.

The frequency of APCs in human SLE associates significantly and correlates positively with autoantibodies to nuclear antigens, including dsDNA, ribonucleoprotein particle, RNA, Smith, chromatin, nucleosomes, La and histones^{45,107,113,135,137,138}, and is in line with the frequency of plasma cells^{45,139}. Human APCs, which are similar to activated naive cells¹⁴⁰, seem to have the potential to become ASCs. In vitro, APCs sorted from patients with SLE produce IgG after stimulation with IL-21, TLR7 and IL-10 (ref. 45), or co-culture with activated CD4⁺ T cells¹¹³. These cells can also acquire a CD27⁺CD38⁺ plasma cell-like phenotype following co-culture with activated CD4⁺ T cells¹¹³; the culture supernatant contained autoantibodies to a plethora of self-antigens, including histones and Smith protein. This plasma cell precursor phenotype is consistent with the transcriptomic signature of bulk sequenced APCs in which there is upregulation of *IRF4* (and its target gene *SLAMF7*) and *PRDM1*, and downregulation of *IRF8* and *BACH2*. This transcriptomic signature is consistent with a plasma cell fate^{45,113}.

APCs associate significantly and correlate positively with disease activity scores defined by measures such as the Systemic Lupus Erythematosus Disease Activity Index (SLEDAI), and low complement levels^{45,135,137,138,141}. These associations were not found in some studies, but in these studies, the APC population was either incorrectly defined (for example, it included CD27⁺ B cells¹⁴² or non-isotype-switched cells¹⁴³) or only partially characterized using limited surface markers^{139,144}.

Most strikingly, the frequency of APCs is consistently and significantly higher in patients with LN than in healthy individuals or in patients with non-LN SLE^{45,107,113,135,137,138,144,145}. The frequency of APCs

a Depletion



b Suppression

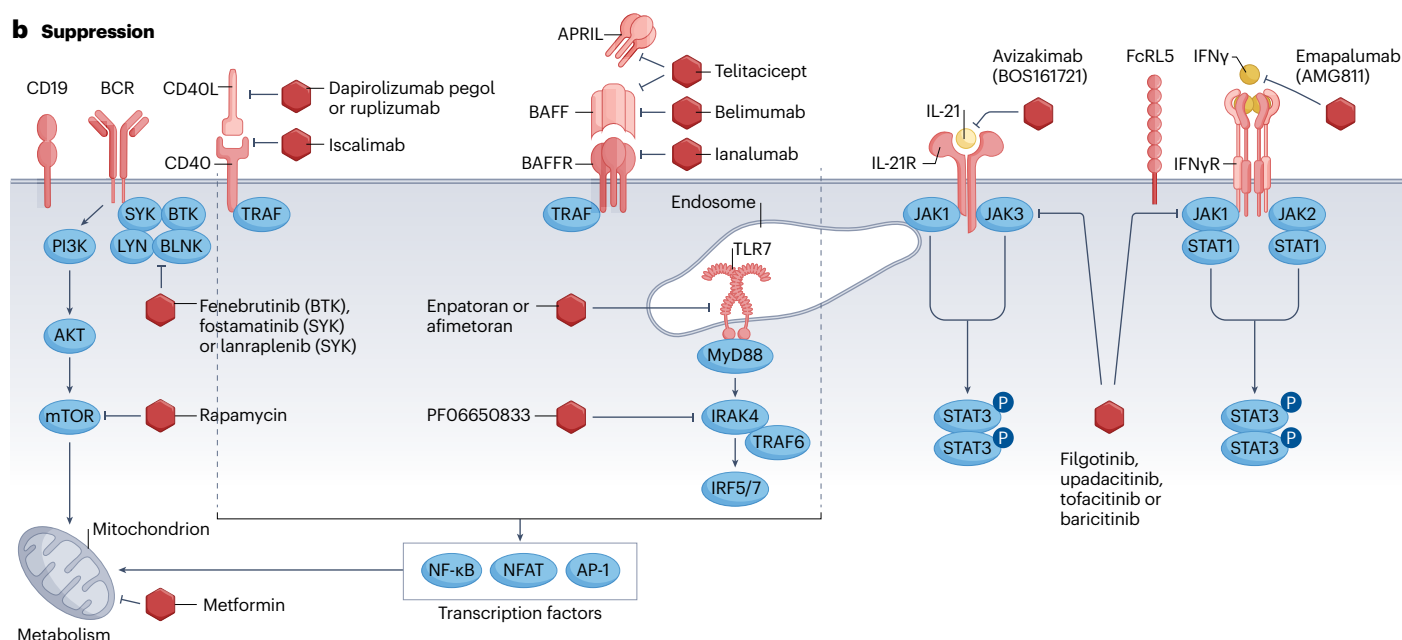


Fig. 6 | Therapeutic targets of ABCs. Several approaches might enable the therapeutic targeting of age-associated B cells (ABCs). **a**, Depletion strategies include CD19⁺-targeted chimeric antigen receptor (CAR) T cells and anti-CD20⁺ monoclonal antibodies. **b**, ABCs might also be targetable through extracellular suppression strategies that target CD40–CD40 ligand (CD40L) co-stimulation, B cell activation via B cell-activating factor (BAFF; also known as TNFSF13B), belonging to the TNF family, and stimulation via IL-21 and IFN γ . Moreover, intracellular suppression of ABCs could be achieved through inhibition of Toll-like

receptor 7 (TLR7) binding, or of signalling downstream of the B cell receptor (BCR), TLR7, IL-21 receptor (IL-21R) and IFN γ receptor (IFN γ R). AP-1, activator protein 1; APRIL, a proliferation-inducing ligand; BAFFR, BAFF receptor; BLNK, B cell linker protein; BTK, Bruton's tyrosine kinase; FcRL5, Fc receptor-like protein 5; IRAK, IL-1 receptor-associated kinase; IRF, interferon regulatory factor; JAK, Janus kinase; NF- κ B, nuclear factor- κ B; NFAT, nuclear factor of activated T cells; PI3K, phosphoinositide 3-kinase; STAT, signal transducer and activator of transcription protein; TRAF, TNF receptor-associated factor.

in these patients correlates negatively with estimated glomerular filtration rate¹⁴⁵. Circulating ABC frequencies are highest in patients with LN classes III, IV and V (according to histopathological classification of kidney biopsy samples)^{137,138}. Patients in disease remission after treatment have significantly lower frequencies of circulating ABCs than patients with active disease¹³⁹. These findings support a highly correlated relationship between ABCs and LN.

Single-cell transcriptomic analysis of kidney biopsy samples during flares of LN has shown that tissue from patients with disease (but not healthy tissue from living donor kidney biopsies) contains a large and highly heterogeneous B cell cluster with some characteristics consistent with an ABC signature²⁰. An infiltration of CD20⁺CD11c⁺ (ref. 113) and CD20⁺T-bet⁺ B cells¹³⁸ was also detected on histology of

kidney biopsies from patients with LN. The substantial frequencies of ABCs in the inflamed kidney might be recruited from the circulation, as suggested by the significant positive correlation between ABC frequency in the circulation and ABC frequency in kidney tissue¹³⁸. The similar transcriptomic profile of naive B cells and ABCs, together with increased expression of activation markers in ABCs that form part of the ABC transcriptomic signature²⁰ suggests a continuum of activation and in situ differentiation from naive cells to the distinct ABC subset. This further supports the assertion that human ABCs are also likely to develop in a germinal centre-independent manner and represent an effector B cell response tailored to peripheral tissues⁴⁵. These data show that ABCs can be present in the kidneys of patients with LN, and the significant correlations between circulating ABCs

and disease activity, low complement, and high autoantibody titres in LN, suggest that these cells are involved in the pathogenesis of kidney damage in SLE.

Implications for treatment. Therapeutic approaches that disturb ABC accumulation have beneficial effects in patients with SLE, which suggests that they could be a promising therapeutic target (Fig. 6). Monoclonal antibodies such as belimumab (targets BAFF) and rituximab (anti-CD20 antibody) had moderate effects on ABC reduction in small studies^{146,147}. Of note, these effects were only observed early and transiently in the case of rituximab. By contrast, anti-CD19 chimeric antigen receptor (CAR) T cell therapy, which uses autologous T cells to target CD19⁺ cells only, seemed to eliminate CD21^{low}CD11c⁺ ABCs efficiently, as well as other memory B cell subsets in all of five patients with refractory SLE¹⁴⁸. The high CD19 expression levels on ABCs, combined with the higher penetration of CAR T cells into tissues compared with antibodies, might explain why CD19-targeted CAR T cells are more effective than belimumab or rituximab. All patients had LN classes III, IV and V, which remarkably responded to CAR T cells as reflected by a fall in proteinuria at 3 months and a reduction in autoantibody titres; overall disease activity also improved¹⁴⁸. In addition to high expression of CD19, ABCs also have high surface levels of FcRL5 and SLAMF7. Of note, FcRL5 is not expressed on other B cell subsets and SLAMF7 is only expressed in plasma cells in addition to ABCs; therefore, targeting FcRL5 and SLAMF7 have potential as relatively selective targets for ABC depletion^{45,149}.

Signals and regulators that are crucial for ABC development also provide therapeutic targets. For example, preprint findings showed the JAK–STAT pathway involved in cytokine signal transduction can be targeted to suppress ABCs¹³⁴. Similarly, antagonists of molecules that support ABC differentiation (for example, TLR7, IL-21 and CD40L), and inhibitors of elements of the BCR pathway might also be treatment options. Although mouse ABCs seem to rely less on BAFF signals for survival than follicular or marginal zone B cells¹⁰⁴, BAFF promotes human ABC induction⁴⁵, highlighting inter-species differences and the potential for including BAFF in combination therapies to increase efficacy in humans. Targeting immunometabolism is another interesting therapeutic option as ABCs seem to acquire hypermetabolic features. Although therapies precisely targeting ABCs are still lacking, many studies in lupus mouse models are underway to evaluate this proof-of-concept for SLE therapy and will provide insights to help define the best targets^{150,151}.

Conclusion

Autoimmune diseases such as SLE have complex heritability patterns, shaped by environmental factors and infection history, and manifest as heterogeneous disease phenotypes. The study of common genetic variants associated with disease is possible through the unbiased screening of large populations of patients and healthy individuals via GWAS. Use of rigorous post-GWAS analyses, such as statistical fine mapping and eQTLs, can improve the validity of the associations identified. However, this approach describes lowly penetrant variants and explains only a minor portion of individual risk heritability – few functional inferences can be made. By contrast, monogenic causes of SLE, and rare, ultra-rare and novel variants identified by sequencing of the exome or the whole genome, lead to highly penetrant and severe clinical phenotypes. Their study enabled the development of in vivo models of disease, the mechanistic exploration of molecular pathways and, ultimately, functional validation, establishment of causality and testing

of therapeutic targets. The learning from common, rare, ultra-rare and novel variants and monogenic causes of disease needs to be integrated to gain a more comprehensive understanding of the genetic architecture of SLE.

Published online: 12 July 2023

References

- Kaul, A. et al. Systemic lupus erythematosus. *Nat. Rev. Dis. Primers* **2**, 16039 (2016).
- Taylor, K. E. et al. Risk alleles for systemic lupus erythematosus in a large case-control collection and associations with clinical subphenotypes. *PLoS Genet.* **7**, e1001311 (2011).
- Rose, A. M. & Bell, L. C. Epistasis and immunity: the role of genetic interactions in autoimmune diseases. *Immunology* **137**, 131–138 (2012).
- Ellinghaus, D. et al. Analysis of five chronic inflammatory diseases identifies 27 new associations and highlights disease-specific patterns at shared loci. *Nat. Genet.* **48**, 510–518 (2016).
- Farh, K. K. et al. Genetic and epigenetic fine mapping of causal autoimmune disease variants. *Nature* **518**, 337–343 (2015).
- Scharer, C. D. et al. Epigenetic programming underpins B cell dysfunction in human SLE. *Nat. Immunol.* **20**, 1071–1082 (2019).
- Aune, T. M. et al. Expression of long non-coding RNAs in autoimmunity and linkage to enhancer function and autoimmune disease risk genetic variants. *J. Autoimmun.* **81**, 99–109 (2017).
- Groopman, E. E., Povysil, G., Goldstein, D. B. & Gharavi, A. G. Rare genetic causes of complex kidney and urological diseases. *Nat. Rev. Nephrol.* **16**, 641–656 (2020).
- Almlof, J. C. et al. Whole-genome sequencing identifies complex contributions to genetic risk by variants in genes causing monogenic systemic lupus erythematosus. *Hum. Genet.* **138**, 141–150 (2019).
- Charras, A. et al. Panel sequencing links rare, likely damaging gene variants with distinct clinical phenotypes and outcomes in juvenile-onset SLE. *Rheumatology* **62**, SI210–SI225 (2023).
- Dasdemir, S. et al. Genetic screening of early-onset patients with systemic lupus erythematosus by a targeted next-generation sequencing gene panel. *Lupus* **31**, 330–337 (2022).
- Delgado-Vega, A. M. et al. Whole exome sequencing of patients from multicase families with systemic lupus erythematosus identifies multiple rare variants. *Sci. Rep.* **8**, 8775 (2018).
- Ellyard, J. I. et al. Identification of a pathogenic variant in *TREX1* in early-onset cerebral systemic lupus erythematosus by whole-exome sequencing. *Arthritis Rheumatol.* **66**, 3382–3386 (2014).
- Jiang, S. H. et al. Functional rare and low frequency variants in *BLK* and *BANK1* contribute to human lupus. *Nat. Commun.* **10**, 2201 (2019).
- Lea-Henry, T. N. et al. Increased burden of rare variants in genes of the endosomal Toll-like receptor pathway in patients with systemic lupus erythematosus. *Lupus* **30**, 1756–1763 (2021).
- Li, G. et al. Genetic heterogeneity of pediatric systemic lupus erythematosus with lymphoproliferation. *Medicine* **99**, e20232 (2020).
- Ramos, P. S., Shedlock, A. M. & Langefeld, C. D. Genetics of autoimmune diseases: insights from population genetics. *J. Hum. Genet.* **60**, 657–664 (2015).
- Brown, G. J. et al. TLR7 gain-of-function genetic variation causes human lupus. *Nature* **605**, 349–356 (2022).
- Caielli, S., Wan, Z. & Pascual, V. Systemic lupus erythematosus pathogenesis: interferon and beyond. *Annu. Rev. Immunol.* **41**, 533–560 (2023).
- Arazi, A. et al. The immune cell landscape in kidneys of patients with lupus nephritis. *Nat. Immunol.* **20**, 902–914 (2019).
- Banchereau, R. et al. Personalized immunomonitoring uncovers molecular networks that stratify lupus patients. *Cell* **165**, 551–565 (2016).
- Nakano, M. et al. Distinct transcriptome architectures underlying lupus establishment and exacerbation. *Cell* **185**, 3375–3389.e21 (2022).
- Perez, R. K. et al. Single-cell RNA-seq reveals cell type-specific molecular and genetic associations to lupus. *Science* **376**, eabf1970 (2022).
- 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).
- Arason, G. J., Jorgensen, G. H. & Ludviksson, B. R. Primary immunodeficiency and autoimmunity: lessons from human diseases. *Scand. J. Immunol.* **71**, 317–328 (2010).
- Casciola-Rosen, L. A., Anhalt, G. & Rosen, A. Autoantigens targeted in systemic lupus erythematosus are clustered in two populations of surface structures on apoptotic keratinocytes. *J. Exp. Med.* **179**, 1317–1330 (1994).
- Taylor, P. R. et al. A hierarchical role for classical pathway complement proteins in the clearance of apoptotic cells in vivo. *J. Exp. Med.* **192**, 359–366 (2000).
- Santer, D. M., Wiedeman, A. E., Teal, T. H., Ghosh, P. & Elkon, K. B. Plasmacytoid dendritic cells and C1q differentially regulate inflammatory gene induction by lupus immune complexes. *J. Immunol.* **188**, 902–915 (2012).
- Gaipl, U. S. et al. Cooperation between C1q and DNase I in the clearance of necrotic cell-derived chromatin. *Arthritis Rheum.* **50**, 640–649 (2004).

30. Leffler, J. et al. Neutrophil extracellular traps that are not degraded in systemic lupus erythematosus activate complement exacerbating the disease. *J. Immunol.* **188**, 3522–3531 (2012).
31. Sturfelt, G. & Truedsson, L. Complement in the immunopathogenesis of rheumatic disease. *Nat. Rev. Rheumatol.* **8**, 458–468 (2012).
32. Mitander, A. et al. Complement consumption in systemic lupus erythematosus leads to decreased opsonophagocytosis in vitro. *J. Rheumatol.* **45**, 1557–1564 (2018).
33. Yasutomo, K. et al. Mutation of *DNASE1* in people with systemic lupus erythematosus. *Nat. Genet.* **28**, 313–314 (2001).
34. Al-Mayouf, S. M. et al. Loss-of-function variant in *DNASE1L3* causes a familial form of systemic lupus erythematosus. *Nat. Genet.* **43**, 1186–1188 (2011).
35. Hakkim, A. et al. Impairment of neutrophil extracellular trap degradation is associated with lupus nephritis. *Proc. Natl Acad. Sci. USA* **107**, 9813–9818 (2010).
36. Rosazza, T., Warner, J. & Sollberger, G. NET formation – mechanisms and how they relate to other cell death pathways. *FEBS J.* **288**, 3334–3350 (2021).
37. Yeh, T. M., Chang, H. C., Liang, C. C., Wu, J. J. & Liu, M. F. Deoxyribonuclease-inhibitory antibodies in systemic lupus erythematosus. *J. Biomed. Sci.* **10**, 544–551 (2003).
38. Hartl, J. et al. Autoantibody-mediated impairment of *DNASE1L3* activity in sporadic systemic lupus erythematosus. *J. Exp. Med.* **218**, e20201138 (2021).
39. Lind, N. A., Rael, V. E., Pestal, K., Liu, B. & Barton, G. M. Regulation of the nucleic acid-sensing Toll-like receptors. *Nat. Rev. Immunol.* **22**, 224–235 (2022).
40. Sisirak, V. et al. Digestion of chromatin in apoptotic cell microparticles prevents autoimmunity. *Cell* **166**, 88–101 (2016).
41. Soni, C. et al. Plasmacytoid dendritic cells and type I interferon promote extrafollicular B cell responses to extracellular self-DNA. *Immunity* **52**, 1022–1038.e7 (2020).
42. Lood, C. et al. Neutrophil extracellular traps enriched in oxidized mitochondrial DNA are interferogenic and contribute to lupus-like disease. *Nat. Med.* **22**, 146–153 (2016).
43. Fillatreau, S., Manfroi, B. & Dorner, T. Toll-like receptor signalling in B cells during systemic lupus erythematosus. *Nat. Rev. Rheumatol.* **17**, 98–108 (2021).
44. Giltiay, N. V. et al. Overexpression of TLR7 promotes cell-intrinsic expansion and autoantibody production by transitional T1 B cells. *J. Exp. Med.* **210**, 2773–2789 (2013).
45. Jenks, S. A. et al. Distinct effector B cells induced by unregulated Toll-like receptor 7 contribute to pathogenic responses in systemic lupus erythematosus. *Immunity* **49**, 725–739.e6 (2018).
46. Jackson, S. W. et al. B cell IFN- γ receptor signaling promotes autoimmune germinal centers via cell-intrinsic induction of BCL-6. *J. Exp. Med.* **213**, 733–750 (2016).
47. Nundel, K. et al. Cell-intrinsic expression of TLR9 in autoreactive B cells constrains BCR/TLR7-dependent responses. *J. Immunol.* **194**, 2504–2512 (2015).
48. Leibler, C. et al. Genetic dissection of TLR9 reveals complex regulatory and cryptic proinflammatory roles in mouse lupus. *Nat. Immunol.* **23**, 1457–1469 (2022).
49. Garcia-Romo, G. S. et al. Netting neutrophils are major inducers of type I IFN production in pediatric systemic lupus erythematosus. *Sci. Transl. Med.* **3**, 73ra20 (2011).
50. Crow, Y. J. et al. Characterization of human disease phenotypes associated with mutations in *TREX1*, *RNASEH2A*, *RNASEH2B*, *RNASEH2C*, *SAMHD1*, *ADAR*, and *IFIH1*. *Am. J. Med. Genet. A* **167A**, 296–312 (2015).
51. Fremont, M. L. et al. Overview of STING-associated vasculopathy with onset in infancy (SAVI) among 21 patients. *J. Allergy Clin. Immunol. Pract.* **9**, 803–818.e11 (2021).
52. Rutsch, F. et al. A specific *IFIH1* gain-of-function mutation causes Singleton-Merten syndrome. *Am. J. Hum. Genet.* **96**, 275–282 (2015).
53. Lee-Kirsch, M. A. et al. Familial chilblain lupus, a monogenic form of cutaneous lupus erythematosus, maps to chromosome 3p. *Am. J. Hum. Genet.* **79**, 731–737 (2006).
54. Lee-Kirsch, M. A. et al. Mutations in the 3'-5' DNA exonuclease *TREX1* are associated with systemic lupus erythematosus. *Nat. Genet.* **39**, 1065–1067 (2007).
55. Stetson, D. B., Ko, J. S., Heidmann, T. & Medzhitov, R. *TREX1* prevents cell-intrinsic initiation of autoimmunity. *Cell* **134**, 587–598 (2008).
56. Young, G. R. et al. Resurrection of endogenous retroviruses in antibody-deficient mice. *Nature* **491**, 774–778 (2012).
57. Okude, H., Ori, D. & Kawai, T. Signaling through nucleic acid sensors and their roles in inflammatory diseases. *Front. Immunol.* **11**, 625833 (2020).
58. Green, N. M. et al. Murine B cell response to TLR7 ligands depends on an IFN- β feedback loop. *J. Immunol.* **183**, 1569–1576 (2009).
59. Tartaglia, M. et al. Mutations in *PTPN11*, encoding the protein tyrosine phosphatase SHP-2, cause Noonan syndrome. *Nat. Genet.* **29**, 465–468 (2001).
60. Bader-Meunier, B. et al. Are RASopathies new monogenic predisposing conditions to the development of systemic lupus erythematosus? Case report and systematic review of the literature. *Semin. Arthritis Rheum.* **43**, 217–219 (2013).
61. He, Y. et al. P2RY8 variants in lupus patients uncover a role for the receptor in immunological tolerance. *J. Exp. Med.* **219**, e20211004 (2022).
62. Thouenon, R., Moreno-Corona, N., Poggi, L., Durandy, A. & Kracker, S. Activated PI3Kinase Delta syndrome – a multifaceted disease. *Front. Pediatr.* **9**, 652405 (2021).
63. Li, G. M. et al. A mutation in *PIK3CD* gene causing pediatric systemic lupus erythematosus: a case report. *Medicine* **98**, e15329 (2019).
64. Wu, J. et al. Fas ligand mutation in a patient with systemic lupus erythematosus and lymphoproliferative disease. *J. Clin. Invest.* **98**, 1107–1113 (1996).
65. Belot, A. et al. Protein kinase c δ deficiency causes Mendelian systemic lupus erythematosus with B cell-defective apoptosis and hyperproliferation. *Arthritis Rheum.* **65**, 2161–2171 (2013).
66. Oliveira, J. B. The expanding spectrum of the autoimmune lymphoproliferative syndromes. *Curr. Opin. Pediatr.* **25**, 722–729 (2013).
67. Noguchi, A. & Takahashi, T. Overview of symptoms and treatment for lysinuric protein intolerance. *J. Hum. Genet.* **64**, 849–858 (2019).
68. Butbul Aviel, Y. et al. Prolidase deficiency associated with systemic lupus erythematosus (SLE): single site experience and literature review. *Pediatr. Rheumatol. Online J.* **10**, 18 (2012).
69. Lipinski, P. et al. Long-term outcome of patients with alpha-mannosidosis – a single center study. *Mol. Genet. Metab. Rep.* **30**, 100826 (2022).
70. Winkelstein, J. A. et al. Chronic granulomatous disease. Report on a national registry of 368 patients. *Medicine* **79**, 155–169 (2000).
71. Xu, L. et al. Loss-of-function variants in *SAT1* cause X-linked childhood-onset systemic lupus erythematosus. *Ann. Rheum. Dis.* **81**, 1712–1721 (2022).
72. Mizui, M. Implications of polyamine metabolism for the pathogenesis of SLE: from cases of monogenic lupus with *SAT1* gene loss-of-function. *21st Century Pathol.* **2**(5), 130 (2022).
73. Jiao, H. et al. Z-nucleic-acid sensing triggers ZBP1-dependent necroptosis and inflammation. *Nature* **580**, 391–395 (2020).
74. Sachidanandam, R. et al. A map of human genome sequence variation containing 1.42 million single nucleotide polymorphisms. *Nature* **409**, 928–933 (2001).
75. Hon, C. C. et al. An atlas of human long non-coding RNAs with accurate 5' ends. *Nature* **543**, 199–204 (2017).
76. Park, L. Population-specific long-range linkage disequilibrium in the human genome and its influence on identifying common disease variants. *Sci. Rep.* **9**, 11380 (2019).
77. Slatkin, M. Linkage disequilibrium – understanding the evolutionary past and mapping the medical future. *Nat. Rev. Genet.* **9**, 477–485 (2008).
78. Schaid, D. J., Chen, W. & Larson, N. B. From genome-wide associations to candidate causal variants by statistical fine-mapping. *Nat. Rev. Genet.* **19**, 491–504 (2018).
79. Vosa, U. et al. Large-scale cis- and trans-eQTL analyses identify thousands of genetic loci and polygenic scores that regulate blood gene expression. *Nat. Genet.* **53**, 1300–1310 (2021).
80. Morris, D. L. et al. Unraveling multiple MHC gene associations with systemic lupus erythematosus: model choice indicates a role for HLA alleles and non-HLA genes in Europeans. *Am. J. Hum. Genet.* **91**, 778–793 (2012).
81. Ghodke-Puranik, Y. & Niewold, T. B. Immunogenetics of systemic lupus erythematosus: a comprehensive review. *J. Autoimmun.* **64**, 125–136 (2015).
82. Sandling, J. K. et al. Molecular pathways in patients with systemic lupus erythematosus revealed by gene-centred DNA sequencing. *Ann. Rheum. Dis.* **80**, 109–117 (2021).
83. Teruel, M. & Alarcon-Riquelme, M. E. The genetic basis of systemic lupus erythematosus: what are the risk factors and what have we learned. *J. Autoimmun.* **74**, 161–175 (2016).
84. Shang, S. et al. A novel gene *CDC27* causes SLE and is associated with the disease activity. *Front. Immunol.* **13**, 876963 (2022).
85. Jiang, S. H., Stanley, M. & Vinuesa, C. G. Rare genetic variants in systemic autoimmunity. *Immunol. Cell Biol.* **98**, 490–499 (2020).
86. Jones, S. A. et al. Rare variants in non-coding regulatory regions of the genome that affect gene expression in systemic lupus erythematosus. *Sci. Rep.* **9**, 15433 (2019).
87. French, C. E. et al. Whole genome sequencing reveals that genetic conditions are frequent in intensively ill children. *Intensive Care Med.* **45**, 627–636 (2019).
88. Duan, R. et al. Correction to: a de novo frameshift mutation in *TNFAIP3* impairs A20 deubiquitination function to cause neuropsychiatric systemic lupus erythematosus. *J. Clin. Immunol.* **40**, 1062–1063 (2020).
89. Graham, R. R. et al. Genetic variants near *TNFAIP3* on 6q23 are associated with systemic lupus erythematosus. *Nat. Genet.* **40**, 1059–1061 (2008).
90. Xie, C. et al. A de novo *PACSIN1* gene variant found in childhood lupus reveals a role for *PACSIN1*/TRAF4 complex in Toll-like receptor 7 activation. *Arthritis Rheumatol.* **75**, 1058–1071 (2023).
91. Richards, S. et al. Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genet. Med.* **17**, 405–424 (2015).
92. Reh, M. H. L. et al. ClinGen – the clinical genome resource. *N. Engl. J. Med.* **372**, 2235–2242 (2015).
93. Choi, S. W., Mak, T. S. & O'Reilly, P. F. Tutorial: a guide to performing polygenic risk score analyses. *Nat. Protoc.* **15**, 2759–2772 (2020).
94. Lewandowski, L. et al. Determining a polygenic risk score in pediatric systemic lupus erythematosus [abstract]. *Arthritis Rheumatol.* **71** (Suppl. 10), 1965 (2019).
95. Polderman, T. J. et al. Meta-analysis of the heritability of human traits based on fifty years of twin studies. *Nat. Genet.* **47**, 702–709 (2015).
96. Wang, Y. F. et al. Identification of 38 novel loci for systemic lupus erythematosus and genetic heterogeneity between ancestral groups. *Nat. Commun.* **12**, 772 (2021).
97. Tangtanatukul, P. et al. Meta-analysis of genome-wide association study identifies *FBN2* as a novel locus associated with systemic lupus erythematosus in Thai population. *Arthritis Res. Ther.* **22**, 185 (2020).
98. Hu, X. et al. Integrating autoimmune risk loci with gene-expression data identifies specific pathogenic immune cell subsets. *Am. J. Hum. Genet.* **89**, 496–506 (2011).
99. Isnardi, I. et al. Complement receptor 2/CD21^{low} human naive B cells contain mostly autoreactive unresponsive clones. *Blood* **115**, 5026–5036 (2010).
100. Holla, P. et al. Shared transcriptional profiles of atypical B cells suggest common drivers of expansion and function in malaria, HIV, and autoimmunity. *Sci. Adv.* **7**, eabg8384 (2021).
101. Thorarindottir, K. et al. CD21^{low} B cells associate with joint damage in rheumatoid arthritis patients. *Scand. J. Immunol.* **90**, e12792 (2019).

102. Vidal-Pedrola, G. et al. Characterization of age-associated B cells in early drug-naïve rheumatoid arthritis patients. *Immunology* **168**, 640–653 (2023).
103. Rubtsov, A. V. et al. Toll-like receptor 7 (TLR7)-driven accumulation of a novel CD11c⁺ B-cell population is important for the development of autoimmunity. *Blood* **118**, 1305–1315 (2011).
104. Hao, Y., O'Neill, P., Naradikian, M. S., Scholz, J. L. & Cancro, M. P. A B-cell subset uniquely responsive to innate stimuli accumulates in aged mice. *Blood* **118**, 1294–1304 (2011).
105. Karnell, J. L. et al. Role of CD11c⁺ T-bet⁺ B cells in human health and disease. *Cell Immunol.* **321**, 40–45 (2017).
106. Mouat, I. C. & Horwitz, M. S. Age-associated B cells in viral infection. *PLoS Pathog.* **18**, e1010297 (2022).
107. Liu, Y. et al. T-bet⁺CD11c⁺ B cells are critical for antichromatin immunoglobulin G production in the development of lupus. *Arthritis Res. Ther.* **19**, 225 (2017).
108. Rubtsov, A. V., Rubtsova, K., Kappler, J. W. & Marrack, P. TLR7 drives accumulation of ABCs and autoantibody production in autoimmune-prone mice. *Immunol. Res.* **55**, 210–216 (2013).
109. Rubtsova, K., Rubtsov, A. V., van Dyk, L. F., Kappler, J. W. & Marrack, P. T-box transcription factor T-bet, a key player in a unique type of B-cell activation essential for effective viral clearance. *Proc. Natl Acad. Sci. USA* **110**, E3216–E3224 (2013).
110. Ambegaonkar, A. A., Holla, P., Dizon, B. L., Sohn, H. & Pierce, S. K. Atypical B cells in chronic infectious diseases and systemic autoimmunity: puzzles with many missing pieces. *Curr. Opin. Immunol.* **77**, 102227 (2022).
111. Mouat, I. C. et al. Gammaherpesvirus infection drives age-associated B cells toward pathogenicity in EAE and MS. *Sci. Adv.* **8**, eade6844 (2022).
112. Nehar-Belaid, D. et al. Mapping systemic lupus erythematosus heterogeneity at the single-cell level. *Nat. Immunol.* **21**, 1094–1106 (2020).
113. Wang, S. et al. IL-21 drives expansion and plasma cell differentiation of autoreactive CD11c⁺T-bet⁺ B cells in SLE. *Nat. Commun.* **9**, 1758 (2018).
114. Cancro, M. P. Age-associated B cells. *Annu. Rev. Immunol.* **38**, 315–340 (2020).
115. Naradikian, M. S. et al. Cutting edge: IL-4, IL-21, and IFN- γ interact to govern T-bet and CD11c expression in TLR-activated B cells. *J. Immunol.* **197**, 1023–1028 (2016).
116. Ricker, E. et al. Altered function and differentiation of age-associated B cells contribute to the female bias in lupus mice. *Nat. Commun.* **12**, 4813 (2021).
117. Yu, B. et al. B cell-specific XIST complex enforces X-inactivation and restrains atypical B cells. *Cell* **184**, 1790–1803.e17 (2021).
118. Keller, B. et al. The expansion of human T-bet^{high}CD21^{low} B cells is T cell dependent. *Sci. Immunol.* **6**, eabh0891 (2021).
119. Zumaquero, E. et al. IFN γ induces epigenetic programming of human T-bet^{hi} B cells and promotes TLR7/8 and IL-21 induced differentiation. *eLife* **8**, e41641 (2019).
120. Manni, M. et al. Regulation of age-associated B cells by IRF5 in systemic autoimmunity. *Nat. Immunol.* **19**, 407–419 (2018).
121. Zhou, S. et al. A novel humanized cutaneous lupus erythematosus mouse model mediated by IL-21-induced age-associated B cells. *J. Autoimmun.* **123**, 102686 (2021).
122. Rubtsova, K. et al. B cells expressing the transcription factor T-bet drive lupus-like autoimmunity. *J. Clin. Invest.* **127**, 1392–1404 (2017).
123. Han, X. et al. Modulating T-bet-promoted energy metabolism in pathogenic age/autoimmune-associated B cells ameliorates lupus autoimmunity. *Arthritis Rheumatol.* **75**, 1203–1215 (2023).
124. Maul, R. W. et al. Transcriptome and IgH repertoire analyses show that CD11c^{hi} B cells are a distinct population with similarity to B cells arising in autoimmunity and infection. *Front. Immunol.* **12**, 649458 (2021).
125. Frasca, D., Romero, M., Garcia, D., Diaz, A. & Blomberg, B. B. Hyper-metabolic B cells in the spleens of old mice make antibodies with autoimmune specificities. *Immun. Ageing* **18**, 9 (2021).
126. Ramesh, A. et al. A pathogenic and clonally expanded B cell transcriptome in active multiple sclerosis. *Proc. Natl Acad. Sci. USA* **117**, 22932–22943 (2020).
127. Trivedi, N. et al. Liver is a generative site for the B cell response to *Ehrlichia muris*. *Immunity* **51**, 1088–1101.e1085 (2019).
128. Zhang, F. et al. Defining inflammatory cell states in rheumatoid arthritis joint synovial tissues by integrating single-cell transcriptomics and mass cytometry. *Nat. Immunol.* **20**, 928–942 (2019).
129. Zhang, C. et al. Novel loss-of-function mutations in *TNFAIP3* gene in patients with lupus nephritis. *Clin. Kidney J.* **15**, 2027–2038 (2022).
130. Peng, S. L., Szabo, S. J. & Glimcher, L. H. T-bet regulates IgG class switching and pathogenic autoantibody production. *Proc. Natl Acad. Sci. USA* **99**, 5545–5550 (2002).
131. Yang, R. et al. Human T-bet governs the generation of a distinct subset of CD11c^{high}CD21^{low} B cells. *Sci. Immunol.* **7**, eabq3277 (2022).
132. Levack, R. C., Newell, K. L., Popescu, M., Cabrera-Martinez, B. & Winslow, G. M. CD11c⁺ T-bet⁺ B cells require IL-21 and IFN- γ from type 1 T follicular helper cells and intrinsic Bcl-6 expression but develop normally in the absence of T-bet. *J. Immunol.* **205**, 1050–1058 (2020).
133. Gao, X. et al. ZEB2 regulates the development of CD11c⁺ atypical B cells. Preprint at *bioRxiv* <https://doi.org/10.1101/2022.09.01.506173> (2022).
134. Gu, S. et al. The transcription factor Zeb2 drives differentiation of age-associated B cells. Preprint at *bioRxiv* <https://doi.org/10.1101/2021.07.24.453633> (2021).
135. Hurtado, C. et al. Altered B cell phenotype and CD27⁺ memory B cells are associated with clinical features and environmental exposure in Colombian systemic lupus erythematosus patients. *Front. Med.* **9**, 950452 (2022).
136. Barnas, J. L. et al. B cell activation and plasma cell differentiation are promoted by IFN- λ in systemic lupus erythematosus. *J. Immunol.* **207**, 2660–2672 (2021).
137. Wei, C. et al. A new population of cells lacking expression of CD27 represents a notable component of the B cell memory compartment in systemic lupus erythematosus. *J. Immunol.* **178**, 6624–6633 (2007).
138. Wu, C. et al. Lupus-associated atypical memory B cells are mTORC1-hyperactivated and functionally dysregulated. *Ann. Rheum. Dis.* **78**, 1090–1100 (2019).
139. You, X. et al. Double negative B cell is associated with renal impairment in systemic lupus erythematosus and acts as a marker for nephritis remission. *Front. Med.* **7**, 85 (2020).
140. Tipton, C. M. et al. Diversity, cellular origin and autoreactivity of antibody-secreting cell population expansions in acute systemic lupus erythematosus. *Nat. Immunol.* **16**, 755–765 (2015).
141. Jacobi, A. M. et al. Activated memory B cell subsets correlate with disease activity in systemic lupus erythematosus: delineation by expression of CD27, IgD, and CD95. *Arthritis Rheum.* **58**, 1762–1773 (2008).
142. Wehr, C. et al. A new CD21^{low} B cell population in the peripheral blood of patients with SLE. *Clin. Immunol.* **113**, 161–171 (2004).
143. Rincon-Arevalo, H. et al. Deep phenotyping of CD11c⁺ B cells in systemic autoimmunity and controls. *Front. Immunol.* **12**, 635615 (2021).
144. Zhu, L. et al. Altered frequencies of memory B cells in new-onset systemic lupus erythematosus patients. *Clin. Rheumatol.* **37**, 205–212 (2018).
145. Sosa-Hernandez, V. A. et al. CD11c⁺ T-bet⁺ CD21^{hi} B cells are negatively associated with renal impairment in systemic lupus erythematosus and act as a marker for nephritis remission. *Front. Immunol.* **13**, 892241 (2022).
146. Ramskold, D. et al. B cell alterations during BAFF inhibition with belimumab in SLE. *EBioMedicine* **40**, 517–527 (2019).
147. Faustini, F. et al. Rituximab in systemic lupus erythematosus: transient effects on autoimmunity associated lymphocyte phenotypes and implications for immunogenicity. *Front. Immunol.* **13**, 826152 (2022).
148. Mackensen, A. et al. Anti-CD19 CAR T cell therapy for refractory systemic lupus erythematosus. *Nat. Med.* **28**, 2124–2132 (2022).
149. Nellore, A. et al. A transcriptionally distinct subset of influenza-specific effector memory B cells predicts long-lived antibody responses to vaccination in humans. *Immunity* **56**, 847–863.e8 (2023).
150. Hocevar, S. et al. PEGylated gold nanoparticles target age-associated B cells in vivo. *ACS Nano* **16**, 18119–18132 (2022).
151. Levack, R. C. et al. Adenosine receptor 2a agonists target mouse CD11c⁺T-bet⁺ B cells in infection and autoimmunity. *Nat. Commun.* **13**, 452 (2022).
152. Fadista, J., Manning, A. K., Florez, J. C. & Groop, L. The (in)famous GWAS P-value threshold revisited and updated for low-frequency variants. *Eur. J. Hum. Genet.* **24**, 1202–1205 (2016).
153. Alarcon-Riquelme, M. E. et al. Genome-wide association study in an Amerindian ancestry population reveals novel systemic lupus erythematosus risk loci and the role of European admixture. *Arthritis Rheumatol.* **68**, 932–943 (2016).
154. Armstrong, D. L. et al. GWAS identifies novel SLE susceptibility genes and explains the association of the HLA region. *Genes Immun.* **15**, 347–354 (2014).
155. Benthani, J. et al. Genetic association analyses implicate aberrant regulation of innate and adaptive immunity genes in the pathogenesis of systemic lupus erythematosus. *Nat. Genet.* **47**, 1457–1464 (2015).
156. Chung, S. A. et al. Differential genetic associations for systemic lupus erythematosus based on anti-dsDNA autoantibody production. *PLoS Genet.* **7**, e1001323 (2011).
157. Demirci, F. Y. et al. Identification of a new susceptibility locus for systemic lupus erythematosus on chromosome 12 in individuals of European ancestry. *Arthritis Rheumatol.* **68**, 174–183 (2016).
158. Gateva, V. et al. A large-scale replication study identifies *TNIP1*, *PRDM1*, *JAZF1*, *UHRF1BP1* and *IL10* as risk loci for systemic lupus erythematosus. *Nat. Genet.* **41**, 1228–1233 (2009).
159. Han, J. W. et al. Genome-wide association study in a Chinese Han population identifies nine new susceptibility loci for systemic lupus erythematosus. *Nat. Genet.* **41**, 1234–1237 (2009).
160. Hom, G. et al. Association of systemic lupus erythematosus with *C8orf13-BLK* and *ITGAM-ITGAX*. *N. Engl. J. Med.* **358**, 900–909 (2008).
161. International Consortium for Systemic Lupus Erythematosus Genetics, et al. Genome-wide association scan in women with systemic lupus erythematosus identifies susceptibility variants in *ITGAM*, *PXK*, *KIAA1542* and other loci. *Nat. Genet.* **40**, 204–210 (2008).
162. Joo, Y. B. et al. Publisher Correction: Genetic variants in systemic lupus erythematosus susceptibility loci, *XKR6* and *GLT1D1* are associated with childhood-onset SLE in a Korean cohort. *Sci. Rep.* **8**, 11713 (2018).
163. Julià, A. et al. Genome-wide association study meta-analysis identifies five new loci for systemic lupus erythematosus. *Arthritis Res. Ther.* **20**, 100 (2018).
164. Kozzyrev, S. V. et al. Functional variants in the B-cell gene *BANK1* are associated with systemic lupus erythematosus. *Nat. Genet.* **40**, 211–216 (2008).
165. Langefeld, C. D. et al. Transancestral mapping and genetic load in systemic lupus erythematosus. *Nat. Commun.* **8**, 16021 (2017).
166. Lee, Y. H., Bae, S. C., Choi, S. J., Ji, J. D. & Song, G. G. Genome-wide pathway analysis of genome-wide association studies on systemic lupus erythematosus and rheumatoid arthritis. *Mol. Biol. Rep.* **39**, 10627–10635 (2012).

167. Leffers, H. C. B. et al. Established risk loci for systemic lupus erythematosus at *NCF2*, *STAT4*, *TNPO3*, *IRF5* and *ITGAM* associate with distinct clinical manifestations: a Danish genome-wide association study. *Jt. Bone Spine* **89**, 105357 (2022).
168. Lessard, C. J. et al. Identification of a systemic lupus erythematosus risk locus spanning *ATG16L2*, *FCHSD2*, and *P2RY2* in Koreans. *Arthritis Rheumatol.* **68**, 1197–1209 (2016).
169. Morris, D. L. et al. Genome-wide association meta-analysis in Chinese and European individuals identifies ten new loci associated with systemic lupus erythematosus. *Nat. Genet.* **48**, 940–946 (2016).
170. Okada, Y. et al. A genome-wide association study identified *AFF1* as a susceptibility locus for systemic lupus erythematosus in Japanese. *PLoS Genet.* **8**, e1002455 (2012).
171. Sakaue, S. et al. A cross-population atlas of genetic associations for 220 human phenotypes. *Nat. Genet.* **53**, 1415–1424 (2021).
172. Song, Q. et al. Genome-wide association study on Northern Chinese identifies *KLF2*, *DOT1L* and *STAB2* associated with systemic lupus erythematosus. *Rheumatology* **60**, 4407–4417 (2021).
173. Yang, J. et al. *ELF1* is associated with systemic lupus erythematosus in Asian populations. *Hum. Mol. Genet.* **20**, 601–607 (2011).
174. Yang, W. et al. Genome-wide association study in Asian populations identifies variants in *ETS1* and *WDFY4* associated with systemic lupus erythematosus. *PLoS Genet.* **6**, e1000841 (2010).
175. Yang, W. et al. Meta-analysis followed by replication identifies loci in or near *CDKN1B*, *TET3*, *CD80*, *DRAM1*, and *ARID5B* as associated with systemic lupus erythematosus in Asians. *Am. J. Hum. Genet.* **92**, 41–51 (2013).
176. Zhang, Y. M. et al. Shared genetic study gives insights into the shared and distinct pathogenic immunity components of IgA nephropathy and SLE. *Mol. Genet. Genomics* **296**, 1017–1026 (2021).
177. Yin, X. et al. Meta-analysis of 208370 East Asians identifies 113 susceptibility loci for systemic lupus erythematosus. *Ann. Rheum. Dis.* **80**, 632–640 (2021).
178. Wangriatisak, K. et al. The expansion of activated naive DNA autoreactive B cells and its association with disease activity in systemic lupus erythematosus patients. *Arthritis Res. Ther.* **23**, 179 (2021).

Author contributions

All authors researched data for the article, made substantial contributions to discussions of the content and wrote the manuscript. C.G.V. and T.W. reviewed or edited the manuscript before submission.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41581-023-00732-x>.

Peer review information *Nature Reviews Nephrology* thanks F. Hiepe and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.

Related links

ClinGen: <https://clinicalgenome.org>
GENEVESTIGATOR: <https://genevestigator.com/>
GWAS Catalog: <https://www.ebi.ac.uk/gwas/>

© Springer Nature Limited 2023