Ann Rheum Dis 000 (2026) 1–16

Contents lists available at ScienceDirect

Annals of the Rheumatic Diseases

journal homepage: https://www.sciencedirect.com/journal/annals-of-the-rheumatic-diseases

Systemic lupus erythematosus

Progression from at-risk state to clinical and severe systemic lupus erythematosus involves molecular dysregulations potentially reversible by biologics: implications for early diagnosis and treatment

Sofia Papanikolaou[1,2], Evgenia Emmanouilidou[1], Christina Adamichou[1,3], Eleni Kalogiannaki[1], Dionysios Nikolopoulos[4,5,6], Micaela Fredi[7,8], Lucy Marie Carter[9,10], Chiara Tani[11], Noemin Kapsala[3], Argiro Repa[1], Nikos Malissovas[5], Panagiotis Garantziotis[12], Nestor Avgoustidis[1], Dimitra Nikoleri[1,13], Aggelos Banos[4,5], Giannis Vatsellas[14], Prodromos Sidiropoulos[1,13], Dimitrios Konstantopoulos[2], Dimitrios Boumpas[4,5], Edward M Vital[10], Laura Andreoli[7,8], Marta Mosca[11], Luís Inês[15,16], Christoforos Nikolaou[2], George Bertsias[1,13,*]

[1] *Rheumatology and Clinical Immunology, University of Crete Medical School, Iraklio, Greece*
[2] *Institute for Bio-Innovation, Biomedical Sciences Research Center "Alexander Fleming", Vari, Greece*
[3] *4th Department of Internal Medicine, Hippokration University Hospital, Aristotle University of Thessaloniki, Thessaloniki, Greece*
[4] *4th Department of Internal Medicine, Attikon University Hospital, National and Kapodistrian University of Athens Medical School, Athens, Greece*
[5] *Laboratory of Autoimmunity and Inflammation, Center of Clinical, Experimental Surgery and Translational Research, Biomedical Research Foundation Academy of Athens, Athens, Greece*
[6] *Division of Rheumatology, Department of Medicine Solna, Karolinska Institutet, Karolinska University Hospital, and Center for Molecular Medicine (CMM), Stockholm, Sweden*
[7] *Rheumatology and Clinical Immunology Unit—ERN ReCONNET, ASST Spedali Civili, Brescia, Italy*
[8] *Department of Clinical and Experimental Sciences, University of Brescia, Brescia, Italy*
[9] *Newcastle-upon-Tyne Hospitals NHS Foundation Trust, Freeman Hospital, Newcastle, UK*
[10] *Leeds Institute of Rheumatic and Musculoskeletal Medicine, University of Leeds, Leeds, UK*
[11] *Department of Clinical and Experimental Medicine, University of Pisa, Pisa, Italy*
[12] *Department of Internal Medicine 3-Rheumatology and Immunology, Friedrich Alexander University Erlangen-Nuremberg and Universitätsklinikum Erlangen, Erlangen, Germany*
[13] *Division of Immunity, Institute of Molecular Biology and Biotechnology, Foundation for Research and Technology-Hellas (FORTH), Heraklion, Greece*
[14] *Greek Genome Center, Biomedical Research Foundation, Academy of Athens, Athens, Greece*
[15] *Rheumatology Department, Unidade Local de Saúde de Coimbra, Coimbra, Portugal*
[16] *Faculty of Medicine, University of Coimbra, Coimbra, Portugal*

*Correspondence to Dr. George Bertsias, Rheumatology, Clinical Immunology, University of Crete Medical School and University Hospital of Heraklion, Voutes, Heraklion, Greece.

E-mail address: gbertsias@uoc.gr (G. Bertsias).

Handling editor Josef Smolen.

https://doi.org/10.1016/j.ard.2026.02.003

ARTICLE INFO

Article history:
Received 11 July 2025
Received in revised form 10 December 2025
Accepted 4 February 2026

ABSTRACT

Objectives: Understanding the molecular events underlying systemic lupus erythematosus (SLE) onset and progression can facilitate early intervention strategies. We explored transcriptomic changes during progression from preclinical to clinical and advanced SLE, and assessed their potential reversibility by targeted agents.

Methods: This includes multicentre prospective study of individuals with nondiagnostic features suggestive of SLE. Baseline blood RNA-sequencing was performed in groups who progressed to classifiable SLE or not (n = 36 each), and compared with healthy (n = 42) and early SLE (n = 43). Transcriptome analysis included supervised methods, gene module-based clustering, and drug reversibility/repurposing pipelines with publicly available data. An independent at-risk cohort (n = 15 progressors, n = 20 nonprogressors) was used to validate molecular classifiers.

Results: At-risk individuals exhibited deregulated metabolic, cytokine signalling, haematologic, and stress-response pathways. Progressors vs nonprogressors showed heightened interferon (IFN)-alpha (α)/gamma (γ) and inflammatory cytokine activity, corroborated by Gene Set Enrichment Analysis and coexpression analysis, and intensified during SLE classification. Unsupervised modelling of gene module-eigengene patterns revealed molecular heterogeneity among progressors, differing in p53/insulin activity and Toll-like receptor (TLR) signalling. Importantly, we characterised a 17-gene susceptibility signature that predicted transition with an area under the curve 0.80 of the receiver operating characteris (95% CI: 0.65-0.94) in the external cohort. Analysis across the continuum—from healthy and at-risk states to early SLE and lupus nephritis—revealed stepwise upregulation of IFN-α/γ, haem metabolism, and oxidative phosphorylation pathways, with additional IFN-related genes activated in established disease. SLE susceptibility and progression signatures demonstrated in silico reversibility by anifrolumab and belimumab.

Conclusions: IFN-α/γ and inflammatory cytokine signatures characterise evolving/early SLE, whereas amplification of IFN signalling and oxidative phosphorylation denotes severity. Blood molecular profiling may aid risk stratification and rationalise early biologic approaches.

WHAT IS ALREADY KNOWN ON THIS TOPIC

- Systemic lupus erythematosus (SLE) is preceded by a preclinical phase characterised by nonspecific autoantibodies and/or elevated inflammatory mediators, yet prediction to full-blown disease remains suboptimal.
- The molecular programmes driving transition from lupus at-risk state to early SLE, and its progression to severe forms, remain poorly defined—hindering targeted prevention strategies.

WHAT THIS STUDY ADDS

- Irrespective of clinical outcome, individuals at risk for SLE exhibit substantial blood molecular deregulation.
- The baseline transcriptome of those transitioning to classifiable SLE is marked by activated interferon-alpha, interferon-gamma, and inflammatory cytokine signalling pathways.
- The at-risk state and its evolution to clinical SLE demonstrate molecular heterogeneity, differing in the activation of Toll-like receptor (TLR) and insulin/p53 signalling pathways.
- A compact gene set—enriched in interferon and inflammatory cytokine pathways—predicted transition to SLE with an area under the curve of 0.80 (95% CI: 0.65-0.94) in an external validation cohort.
- Progressive amplification of interferon-alpha/-gamma, haem metabolism, and oxidative phosphorylation signatures underlie the continuum from healthy to at risk, early SLE, and severe disease (lupus nephritis).
- Belimumab and anifrolumab may counteract molecular perturbations underlying evolution from lupus at risk to clinical disease, and its progression to full-blown/severe forms.

HOW THIS STUDY MIGHT AFFECT RESEARCH, PRACTICE OR POLICY

- Blood transcriptome profiling holds potential in identifying individuals at high risk for SLE and its severe forms.
- Type I/II interferon and oxidative phosphorylation pathways may represent actionable targets by existing and novel therapies towards the prevention or delay of SLE disease progression.

INTRODUCTION

Systemic lupus erythematosus (SLE) develops in a multistep fashion, starting with asymptomatic self-reactivity and inflammation and advancing to full-blown disease [1]. This is evidenced by antinuclear antibodies (ANAs) being detectable on average 2.3 to 3.3 years before clinical onset, with ANA subtypes reactive to specific nuclear components accruing closer to diagnosis [2,3]. Circulating inflammatory hallmarks, including interferon (IFN)-associated mediators, gradually accumulate in individuals moving towards SLE classification/diagnosis [4,5]. However, early signs of autoimmunity such as ANAs may be present in otherwise healthy subjects (14%-27%), some of whom exhibit other lupus-related immune/serologic abnormalities [6]. Consistently, ANA-positive individuals exhibit significant gene expression deregulations in blood [7]. These data suggest that progression from at-risk state to manifest SLE is dynamically regulated by factors that accelerate or slow disease development.

Previously, we mapped the transcriptional landscape of patients with established SLE and defined signatures linked to the disease and its activity [8]. Dysregulated genes involved in

neutrophil activation and humoral responses denoted kidney involvement [8]. We also characterised coexpression gene modules that correlated with clinical traits, including the cumulative number of classification criteria, presence of anti-DNA antibodies, and overall or organ-specific activity [9]. These findings suggest quantifiable molecular variability among individuals in different clinical states. Nonetheless, how blood immune deregulations evolve across the disease spectrum—from healthy to at-risk, early full-blown, and severe lupus autoimmunity—remains ill-defined.

Here, we leveraged a prospective cohort of individuals at risk for SLE, of whom approximately 20% progressed to classifiable disease during follow-up [10]. To uncover molecular features of subclinical autoimmunity and early SLE, we performed blood RNA-sequencing (RNA-seq) on a nested case-control sample of progressors and nonprogressors, followed by differential expression and modular transcriptomic analyses by clinical trajectory. Unsupervised clustering identified molecular subgroups within the at-risk population, and predictive modelling using concise gene sets derived in this study was externally validated for their association with transition. We extended the analysis across the full spectrum from health to lupus nephritis, revealing stage-specific transcriptional programmes and immune cell shifts. Finally, we integrated transcriptomic data with pharmacological databases to nominate candidate targets for early interception.

METHODS

At-risk cohort and subjects

This is a nested analysis within a multicentre prospective cohort of adults presenting with nondiagnostic signs and symptoms suggestive of SLE. Between 10/2016 and 12/2021, participating centres consecutively enrolled individuals based on criteria reflecting a spectrum of at-risk states (Supplementary Table S1). Eligibility was determined by ANA positivity in combination with additional features (Supplementary Table S2), or a family history of SLE accompanied by suggestive serologic or clinical findings. Individuals were monitored regularly for ≥12 months (median 18 months) using a prespecified protocol that included reassessment of the classification criteria. *Progressors* were defined as those who accrued sufficient features to meet the 1997 American College of Rheumatology (ACR) [11] or 2019 European Alliance of Associations for Rheumatology (EULAR)/ACR [12] criteria. Subjects who did not evolve to classifiable SLE were defined as *nonprogressors*, though they could still have been diagnosed with undifferentiated autoimmune disease. This analysis was conducted in a balanced subcohort of progressors and nonprogressors (n = 36 each) on the basis of sample availability. Control groups included healthy individuals (n = 42) and patients with early SLE (n = 43), defined as physician-diagnosed cases meeting classification criteria, with a duration of ≤3 years since diagnosis/classification and active disease (Supplementary Table S3).

RNA isolation, library preparation, and sequencing

Blood was stored in PaxGene tubes, and total RNA was extracted using the RNeasy kit (Qiagen). Following quality control, libraries were prepared using the NEBNext Ultra II directional RNA library prep kit. Sequencing was performed on Illumina NovaSeq6000 platforms. Preprocessing and alignment pipelines are provided in Supplementary Methods.

Differential expression analysis

DESeq2 [13] was used for pairwise transcriptome comparisons. Genes with an average read count of ≤5 across all samples were filtered out. Batch effects were corrected using *limma* [14]. P value <.05 and an absolute log2-fold change (FC) > 0.58 defined differentially expressed genes (DEGs). In addition to the Wald test, the likelihood ratio test was applied to identify gene sets whose expression varied across the healthy, at-risk, and early SLE groups. Overrepresentation analysis (ORA) of DEGs and gene set enrichment analysis (GSEA) of the ranked gene lists were performed with *clusterProfiler*.

Coexpression analysis

CEMiTool [15] was used to identify reproducible gene coexpression modules and generate gene-gene interaction networks within each module [16]. Module activity across sample groups was assessed using sample annotation metadata. Genes in each module were subjected to ORA using *clusterProfiler* and *Enrichr*.

Cell-type estimation

Immune cell composition was inferred using CIBERSORTx [17] (Supplementary Methods).

Prioritisation of IFN-related genes, predictive modelling, and external validation

To identify informative IFN-related genes, we applied elastic-net regression with 5-fold cross-validation across a grid of α and λ values. Model stability was assessed by refitting the elastic-net in 100 bootstrap resamples, and genes selected in ≥60% of iterations were retained. To extend the analysis to pathway-level representations, we used pathMED [18] to summarise gene expression profiles into Z-scored activity for various gene sets derived from our differential expression analyses. These activity scores were used as features in machine-learning classifiers to evaluate their relative predictive contribution. For gene-level modelling of the core-susceptibility signature, elastic-net regression was applied to variance-stabilised and batch-corrected expression data, with 5-fold cross-validation used to select optimal penalisation parameters. The final classifier was trained on our data and tested in an independent RNA-seq dataset from the Leeds at-risk cohort [7]. Details are described in the Supplementary Methods.

Gene signature reversibility analysis

We assessed the reversibility of gene signatures by interrogating publicly available blood RNA-seq data from patients with SLE treated with belimumab or anifrolumab. For belimumab, we analysed transcriptomic changes in clinical responders (meeting SLE Responder Index-4 and Lupus Low Disease Activity State) from an observational cohort, by comparing RNA profiles at baseline and after 6 months of treatment [19]. A second dataset, obtained from the TULIP-1/2 (Treatment of Uncontrolled Lupus via the Interferon Pathway) trials, compared transcriptomes of anifrolumab- vs standard therapy-treated patients at 52 weeks [20]. Drug-responsive genes were defined using $|\log2FC| > 0.58$ and P <.05. To evaluate whether these genes were enriched in the transcriptional profiles of progressors vs nonprogressors, we performed GSEA using *fgsea* [21]. The susceptibility-associated gene list was ranked by log2FC, and

enrichment was tested using the parameters: *minSize* = 10, *maxSize* = 500, and *nperm* = 1000.

Drug-repurposing analysis

We explored potential drug candidates using 2 complementary resources based on the Library of Integrated Network-Based Cellular Signatures L1000 dataset: the L1000CDS^2 search engine [22] and the Connectivity Map (CMap) [23]. Both frameworks compare multigene differential expression signatures against large-scale perturbagen profiles, enabling gene set-level inference of compounds predicted to mimic or reverse the observed transcriptional patterns, as detailed in the Supplementary Methods.

Statistical analysis

Group comparisons were performed using the Mann-Whitney and Fisher's exact test. Generalised linear models linked module M1-M5 eigengene expression to demographic, clinical, and serologic features. To identify latent clusters based on standardised module eigengenes, we applied finite mixture modelling using a Gaussian Mixture Model framework (Supplementary Methods). All analyses were performed using the statistical software STATA (version 19.0)..

Patient and public involvement

Patients and/or the public were not involved in the design, conduct, reporting, or dissemination plans of this research.

RESULTS

Individuals at risk for SLE exhibit extensive molecular deregulations irrespective of their clinical outcome

The characteristics of individuals at risk who progressed (n = 36) or did not (n = 36) to SLE are summarised in Table. At inclusion, nonprogressors had a lower frequency of ANA positivity, but other clinical features were comparable to those of progressors. Progressors met the 1997 ACR (n = 23) and/or the 2019 EULAR/ACR (n = 30) criteria, and primarily accumulated mucocutaneous, musculoskeletal, haematological, and immunological manifestations. We first analysed the baseline blood transcriptome in the entire cohort of at-risk individuals and compared it to that of healthy individuals. A total of 1256 genes showed deregulation (≥1.5-fold gene expression difference: $|\log2FC| > 0.58$, adjusted $P < .05$), with 902 upregulated and 359 downregulated genes in the former vs the latter group (Fig 1A, Supplementary Table S4A). Genes with differential expression in at-risk subjects were associated with haem metabolism, oxidative phosphorylation (OxPhos), reactive oxygen species pathway, IFN-alpha/-gamma response, DNA repair and apoptosis (Fig 1B). Comparing progressors and healthy subjects, we identified 1622 DEGs (Fig 1C, Supplementary Table S4B). Genes with higher expression in progressors were enriched in heme metabolism, coagulation, IFN-alpha response, KRAS (Kirsten Rat Sarcoma Viral Oncogene Homolog) signalling, OxPhos, the p53 pathway, and myogenesis (Fig 1D). By contrast, genes with lower expression in progressors were associated with androgen response, tumour necrosis factor alpha (TNFα) signalling via NF-κB (Nuclear Factor kappa-light-chain-enhancer of activated B-cells), mitotic spindle (cell division), protein secretion, and UV response.

A similar comparison between nonprogressors and healthy subjects revealed 1020 DEGs (Fig 1E, Supplementary Table S4C). Activated genes in non-progressors were linked to haem metabolism, coagulation, myogenesis, epithelial-to-mesenchymal transition, and OxPhos (Fig 1F), whereas genes with lower expression were associated with early 2 factor targets (cell cycle regulation), TNFα signalling, and protein secretion. Collectively, broad molecular perturbations occur in the at-risk population—regardless of clinical evolution—spanning metabolic, cytokine signalling, haematologic, cell cycle, and stress-response pathways. Among the shared processes, progressors showed both a higher number of associated genes and stronger overexpression patterns compared to nonprogressors (Supplementary Table S4, Supplementary Fig S1A–G).

IFN and inflammatory cytokine gene signatures are present in the at-risk state evolving to SLE

To quantify these differences and delineate signatures specifically associated with imminent SLE, we examined the overlap of DEGs in progressors and nonprogressors using the healthy transcriptome as a reference. Consistent with the common clinical and immunological characteristics of the 2 at-risk groups, we found many shared DEGs (n = 790; Fig 2A). DEGs exclusively upregulated in progressors (n = 368) were over-represented in the IFN-alpha pathway, whereas downregulated DEGs (n = 464) were involved in protein secretion, UV response, and androgen response (Fig 2B).

We directly compared progressors and nonprogressors, to examine transcriptional differences within the at-risk state. Nominal *P* values ($P < .05$) and an absolute log2FC > 0.58 were used to retain sensitivity to subtle signals suggested by the clinical similarities between the 2 groups (Table, Fig 2A). Using these criteria, we identified 91 *susceptibility* DEGs (Fig 2C, Supplementary Table S4D). Among the most upregulated ($\log2FC \geq 1$) were genes involved in IFN response (*FI27, IFI44, IFI44L, IFIT1, ISG15, RSAD2, SIGLEC1, USP18, CMPK2, OAS3*, and *HERC5*), complement (*C1QC*), epithelial cell structure/keratinisation (*KRT8, KRT79*), and extracellular matrix remodelling (*FAM83A, S100A16*). The top downregulated genes ($\log2FC \leq -1$) participated in extracellular matrix organisation and fibrosis (*MMP2, DCN, FN1*, and *COL3A1*). GSEA supported these patterns (Fig 2D). Notably, a subset of 14 IFN-related genes (Fig S2) demonstrated predictive capacity for progression, with an odds ratio per 1-standard deviation increase in the composite IFN gene score of 4.96 (95% CI: 2.00-12.27) (Fig 2E).

To assess the reproducibility of our findings, we analysed an external RNA-seq dataset from peripheral blood mononuclear cell of at-risk individuals (Leeds cohort) [7]. Using an identical workflow, GSEA reproduced the top enrichment of the IFN-related pathways, with 14 of 21 enriched Hallmark pathways showing concordant trends across cohorts (Fig S3A,B), indicating that these represent consistent transition-associated signatures. Analysis of log2FC for genes expressed in both datasets demonstrated a positive correlation among DEGs (Pearson $r = 0.547$; $R^2 = 0.299$), with 17 shared DEGs representing a highly significant overlap ($P = 1.2 \times 10^{-22}$) (Fig S3C). Nonetheless, given the within-state comparison context, our gene-level findings were interpreted as exploratory and used primarily for downstream analyses. When applying an adjusted *P* value <.05 threshold, 4 DEGs were detected (Supplementary Table S4D).

Twenty-eight DEGs (24 upregulated) were common to both progressors vs healthy and progressors vs nonprogressors but absent in nonprogressors vs healthy comparisons. This core-

Table

Demographic and clinical characteristics of individuals at risk for SLE who evolved or not into classifiable SLE

	Nonprogressors (n = 36)		Progressors (n = 36)		*P* value[a]
Age (y)	34.8 (10.2) [2]		36.9 (11.1)		.448
Sex (female)	32 (88.9%)		36 (100.0%)		.115
Body mass index (kg/m^2)	23.8 (3.5)		25.1 (4.2)		.147
Race (White)	36 (100.0%)		36 (100.0%)		–
FDR(s) with SLE	5 (13.9%)		5 (13.9%)		1.000
FDR(s) with other autoimmune disease	12 (33.4%)		11 (30.5%)		.800
Nulliparous	11 (34.4%)		15 (44.1%)		.459
Smoking					
Never	16 (44.4%)		16 (44.4%)		
Past	8 (22.2%)		6 (16.7%)		
Current	12 (33.3%)		14 (38.9%)		.803
	Baseline	Cumulative	Baseline	Cumulative	
1997 ACR criteria					
Malar rash	5 (13.9%)	5 (13.9%)	8 (22.2%)	15 (41.7%)	.541
Discoid rash	0	0	1 (2.8%)	1 (2.8%)	1.000
Photosensitivity	8 (22.2%)	8 (22.2%)	12 (33.3%)	13 (36.1%)	.430
Mucosal ulcers	3 (8.3%)	4 (11.1%)	5 (13.9%)	8 (22.2%)	.710
Synovitis	11 (30.6%)	16 (44.4%)	12 (33.3%)	33 (91.7)	1.000
Serositis	0	0	2 (5.6%)	2 (5.6%)	.493
Renal	0	0	0	0	–
Neurological	0	1 (2.8%)	1 (2.8%)	2 (5.6%)	1.000
Haematological	9 (25.0%)	10 (27.8%)	7 (19.4%)	13 (36.1%)	.778
Immunological	9 (25.0%)	9 (25.0%)	7 (19.4%)	14 (38.9%)	.778
ANA	25 (69.4%)	25 (69.4%)	34 (94.4%)	35 (97.2%)	.012
Total number	1.9 (0.6)	2.2 (0.5)	2.5 (0.7)	3.8 (0.9)	<.001
2019 EULAR/ACR criteria					
Fever score	0.1 (0.3)	0.1 (0.3)	0.0 (0.0)	0.0 (0.0)	.321
Cutaneous score	1.5 (2.4)	1.6 (2.3)	2.0 (2.6)	3.4 (2.7)	.393
Synovitis score	1.8 (2.8)	2.7 (3.0)	2.0 (2.9)	5.5 (1.7)	.804
Neurological score	0.0 (0.0)	0.1 (0.5)	0.1 (0.8)	0.2 (1.0)	.321
Serosal score	0.0 (0.0)	0.0 (0.0)	0.3 (1.3)	0.3 (1.3)	.158
Haematological score	0.8 (1.5)	0.9 (1.5)	0.7 (1.4)	1.0 (1.6)	.745
Renal score	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	1.000
Antiphospholipid score	0.2 (0.6)	0.2 (0.6)	0.2 (0.6)	0.2 (0.6)	.838
Complement score	1.3 (1.7)	1.6 (1.8)	1.3 (1.7)	2.1 (1.8)	1.000
Autoantibodies score	0.3 (1.4)	0.3 (1.4)	0.3 (1.4)	0.9 (2.1)	.238
Total score	6.0 (3.4)	7.3 (3.3)	6.9 (2.7)	13.7 (3.6)	.238
Immunological features					
All subjects					
Anti-RNP	0	1 (2.8%)	0	0	–
Anti-Ro/SSA	3 (8.3%)	1 (2.8%)	5 (13.9%)	4 (11.1%)	.710
Anti-La/SSB	0	0	0	1 (2.8%)	–
Low C3/C4	14 (38.9%)	3 (8.3%)	13 (36.1%)	8 (22.2%)	.808
Anti-dsDNA	3 (8.3%)	0	3 (8.3%)	5 (13.9%)	1.000
Anti-Sm	0	0	0	0	–
Antiphospholipid Abs	7 (19.4%)	0	8 (22.2%)	1 (2.8%)	.772
Direct Coombs	0	0	1 (2.8%)	2 (5.6%)	1.000
ANA-negative subjects					
Anti-RNP	0	0	0	0	–
Anti-Ro/SSA	0	0	0	0	–
Anti-La/SSB	0	0	0	0	–
Low C3/C4	8 (72.7%)	0	2 (100%)	0	1.000
Anti-dsDNA	1 (9.1%)	0	0	0	1.000
Anti-Sm	0	0	0	0	–
Antiphospholipid Abs	2 (18.2%)	0	1 (50.0%)	0	.432
Direct Coombs	0	0	0	0	

ANA, antinuclear antibody; ACR, American College of Rheumatology; dsDNA, double-stranded DNA; EULAR, European Alliance of Associations for Rheumatology; FDR, first-degree relative; La/SSB, La/Sjögren's-syndrome-related antigen B autoantibodies; SLE, systemic lupus erythematosus; Sm, Smith antigen; RNP, ribonucleoprotein; Ro/SSA, Ro/Sjögren's-syndrome-related antigen A autoantibodies.

[a] Comparison of baseline characteristics between progressors and nonprogressors. The Mann-Whitney test was used for numerical variables and Fisher's exact test for categorical variables.

[b] Data are presented as median (IQR) values or n (%).

susceptibility signature showed enrichment in IFN-alpha, IFN-gamma, and inflammatory responses (Fig 2F), highlighting their central role in SLE evolution. Among the upregulated DEGs, 2 were specific to IFN-alpha, 1 to IFN-gamma, and 6 were common to both pathways.

To corroborate these findings, we took advantage of available RNA-sequencing data from a subset of progressors (n = 4) sampled both at baseline and at the time of SLE classification. A total of 589 genes were differentially expressed (|log2FC| > 0.58, *P* <.05; Supplementary Table S4E), the majority (n = 424)

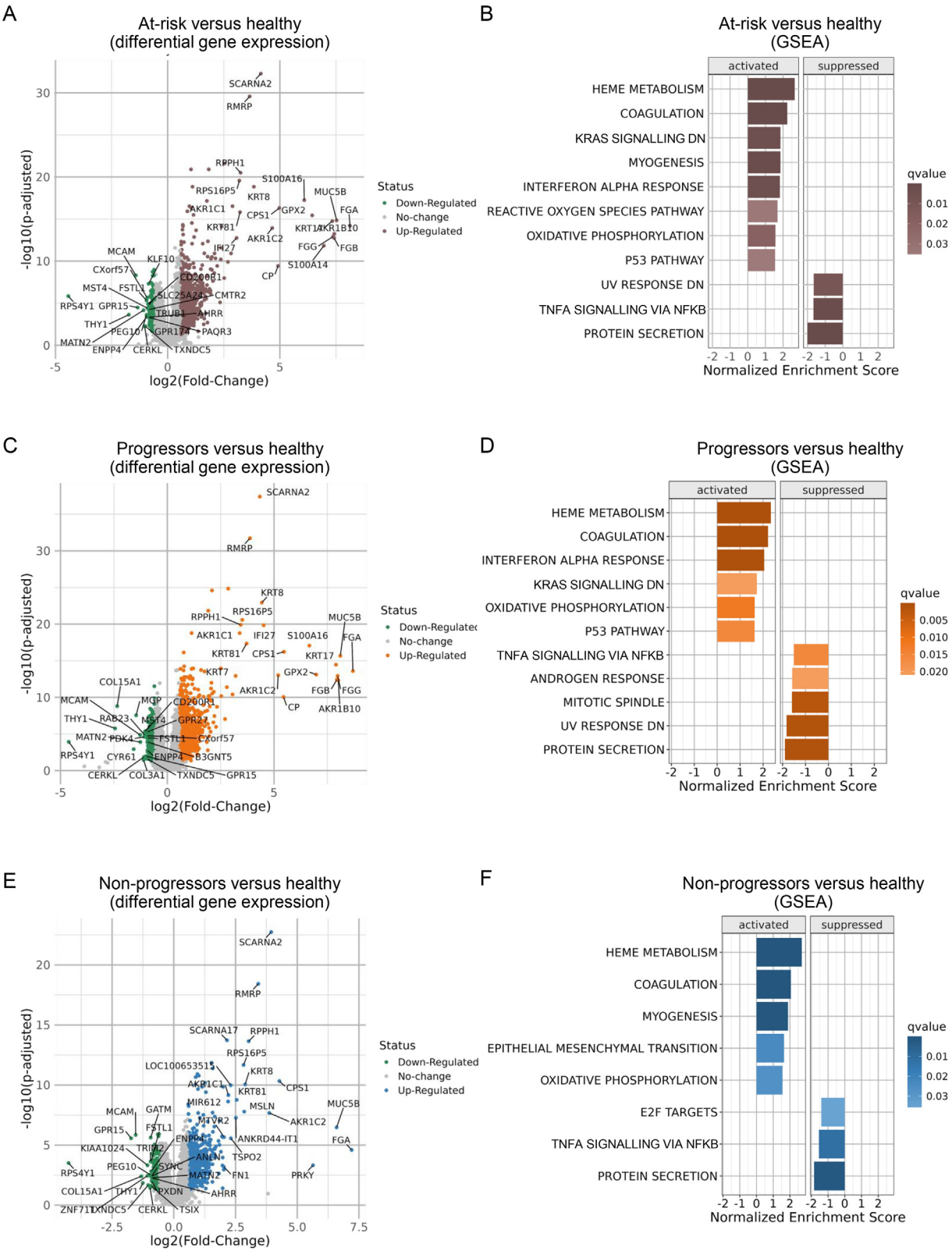


Figure 1. Perturbations in blood transcriptome and biological processes in individuals at risk for SLE as compared to healthy subjects. (A) Volcano plot showing differential gene expression between the entire cohort of at-risk individuals (n = 72) and healthy counterparts (n = 42). A total of 1256 genes were significantly deregulated (|log2-fold change [FC]| >0.58; adjusted *P* value <.05), including 902 upregulated and 359 downregulated genes. Purple and green dots denote significantly upregulated and downregulated genes, respectively. The x-axis shows log2FC; the y-axis shows $-\log_{10}$ of adjusted *P* values. (B) Gene set enrichment analysis (GSEA) of all genes ranked by log2FC in at-risk individuals vs healthy controls, using Hallmark C50 gene sets (MSigDB). Positively and negatively enriched pathways are ranked by normalised enrichment score (NES) and colour-coded by false discovery rate (FDR) q-value. Colour intensity reflects statistical significance. (C) Volcano plot of differentially expressed genes (DEGs) between individuals at risk who transitioned to classifiable SLE (progressors; n = 36) and healthy subjects. A total of 1622 DEGs reached statistical significance (adjusted *P* value <.05), with 993 upregulated (log2FC >0.58; red) and 629 downregulated (log2FC < −0.58; green). (D) GSEA of all genes ranked by log2FC in progressors vs healthy controls, using Hallmark C50 gene sets. Positively and negatively enriched pathways are ranked by NES and colour-coded by FDR q-value. Colour intensity reflects statistical significance. (E) Volcano plot showing DEGs between individuals at risk who did not transition to SLE (nonprogressors; n = 36) and healthy subjects. A total of 1020 DEGs reached statistical significance (adjusted *P* value

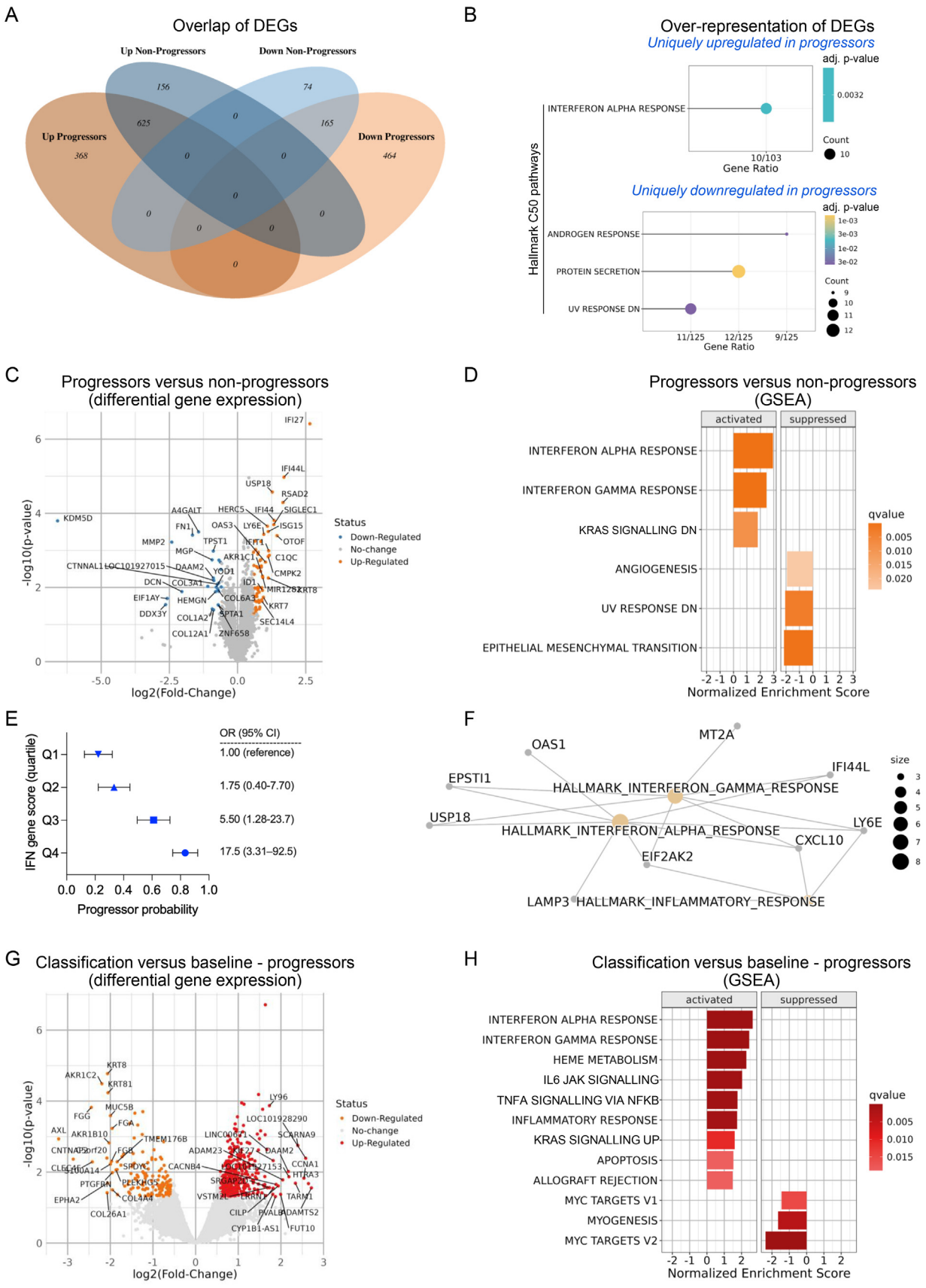


Figure 2. Type I/II interferon and inflammatory cytokine signatures underlie the transition from the at-risk state to SLE. (A) Venn diagram showing the overlap of differentially expressed genes (DEGs: |log2FC| >0.58; adjusted *P* value <.05) derived from the comparisons of progressors vs healthy and nonprogressors vs healthy counterparts. A total of 625 genes were commonly upregulated and 165 downregulated. The diagram also displays

<.05), with 781 upregulated (log2FC >0.58; red) and 239 downregulated (log2FC < −0.58; green). (F) GSEA of all genes ranked by log2FC in non-progressors vs healthy controls, using Hallmark C50 gene sets. Pathways are ranked by NES and coloured by FDR q-value, with intensity reflecting statistical significance. KRAS, kirsten rat sarcoma viral oncogene homolog; NFKB, nuclear factor kappa-light-chain-enhancer of activated B cells; SLE, systemic lupus erythematosus; TNFA, tumor necrosis factor alpha; UV, ultraviolet.

being upregulated (Fig 2G). These genes were enriched in IFN-alpha, IFN-gamma, and inflammatory response pathways (Fig 2H), suggesting their intensification as SLE becomes imminent. interleukin-6/Janus kinase/Signal Transducer and Activator of Transcription 3 and TNFα signalling via NF-κB pathways were also activated.

Unsupervised gene modular analysis reveals biological processes accelerating or protecting against transition from at risk to lupus classification

To explore the broader biological context of the at-risk state and its potential progression to SLE, we applied CEMiTool [15,16] to the full dataset of healthy and at-risk RNA profiles. This tool enables the integrated discovery and functional analysis of coexpressed gene modules and their networks. Five modules were identified, consisting of 591, 165, 75, 43, and 34 genes, respectively (Fig 3A). Of particular interest was *Module M4*, positively enriched in progressors (normalised enrichment score [NES] = 2.76) and negatively enriched in nonprogressors (NES = −2.75). M4 genes were involved in IFN-alpha/gamma and broader cytokine signalling pathways, with interaction network analysis highlighting the RNA-sensor *DDX58/RIG-I* (Fig 3B).

Module M1 was positively enriched in healthy individuals but negatively enriched in nonprogressors and progressors. M1 genes were involved in the TLR signalling (Fig 3A,B). *Modules M2, M3*, and *M5* were enriched in progressors and suppressed in healthy individuals, with nonprogressors showing intermediate activation of M3/M5 and mild suppression of M2. Functionally, M2 was linked to haematopoiesis and MHC class I antigen presentation; M3 to mRNA translation, splicing, and regulation; M5 to cAMP (cyclic adenosine monophosphate)/protein kinase A signalling, particularly in the regulation of insulin secretion, and the p53 pathway (Fig S4A–F). Using the module eigengenes to represent the overall expression pattern of each module, we observed modest relationships with demographics or classification criteria items (Fig S5). Altogether, analysis of coexpressed genes reveals distinct biological processes underlying the stable or evolving at-risk state and reinforces the role of IFN gene signatures in progression to SLE.

Modular transcriptomic clustering reveals heterogeneous at-risk states associated with progression

Although enrichment scores differed between progressors and nonprogressors (Fig 3A), module eigengenes were not individually associated with progression (Supplementary Table S5). Moreover, their intercorrelations (Supplementary Table S6) suggested partially overlapping transcriptional programmes. We therefore applied Gaussian Mixture Modelling to identify transcriptional subgroups within the at-risk population that might better reflect clinical trajectories.

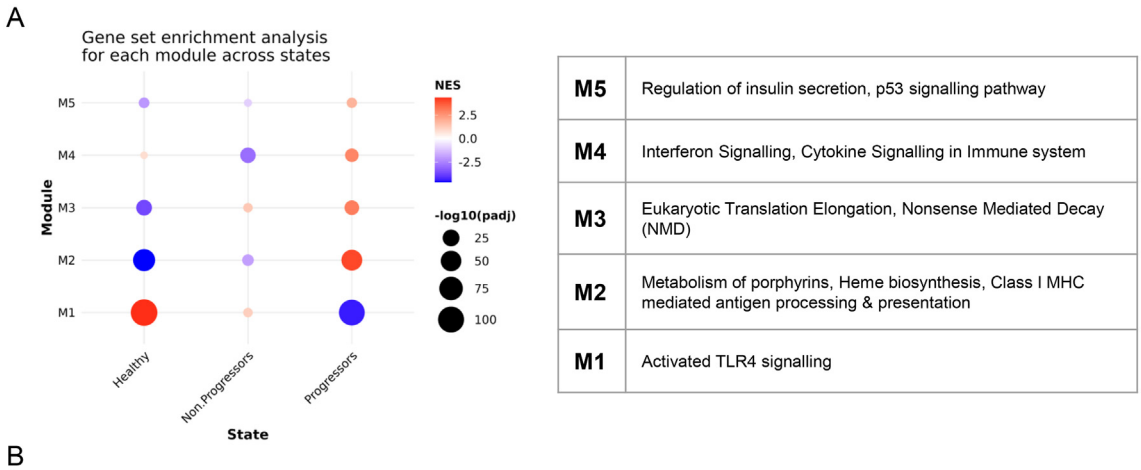
Four clusters were recognised as optimal, with cluster A being the largest (n = 46) (Fig 3C,D). Compared with this cluster, clusters B (n = 8) and C (n = 7) exhibited overall lower average activity in modules M2-M4. Module M1 activity was elevated in cluster C but suppressed in clusters B and D (n = 11), whereas M5 activity was increased in both clusters C and D. Notably, progressors were significantly over-represented in clusters A/D compared with clusters B/C (Fig 3D). These findings suggest that the at-risk condition evolving to SLE—enriched in clusters A and D—may represent a molecularly heterogeneous state unified by elevated IFN/cytokine signalling (M4) and relatively preserved activity in haematopoiesis/antigen processing (M2) and mRNA translation/splicing (M3), while differing in other immuno-transcriptional features such as dampened TLR signalling (M1) or elevated insulin secretion/p53 pathway activity (M5).

Transcriptional dynamics across healthy, at-risk, and early SLE reveal progressively amplified immune-metabolic signatures linking susceptibility to severity

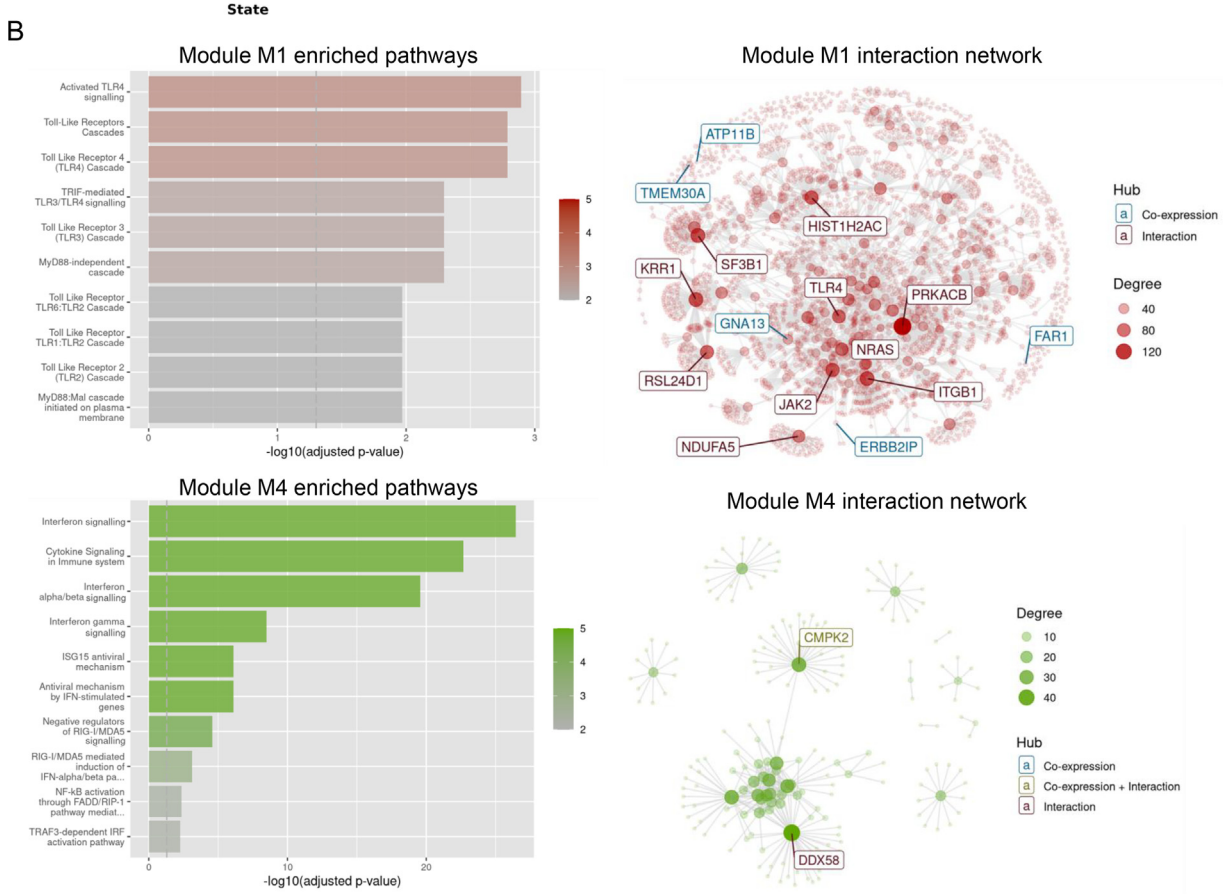
Progressors and nonprogressors shared gene signatures, with a gradient of enrichment linked to clinical outcome (Fig 3A). Moreover, certain signatures unique to progressors were intensified upon disease classification (Fig 2G,H). These observations prompted us to examine how the molecular landscape evolves from health to the preclinical phase and, ultimately, to established disease. We expanded our blood RNA-sequencing to include 43 patients with early, active SLE (Supplementary Table S3), including 9 with kidney involvement. Applying a likelihood ratio test, we identified 6 non-overlapping gene clusters with differential expression patterns across the entire spectrum: healthy individuals, nonprogressors, progressors, early SLE, and lupus nephritis (Fig 4A, Fig S6A, Supplementary Table S7).

Cluster 1 genes (n = 525) demonstrated stepwise activation across the spectrum. These transcripts were enriched in IFN-alpha and IFN-gamma, heme metabolism, and OxPhos pathways, implicating them in both early and advanced stages of lupus (Fig 4B). Two genes were unique to the IFN-alpha pathway, 12 to IFN-gamma, and 21 were shared by

uniquely upregulated and downregulated genes in each at-risk group (progressors, nonprogressors). (B) Overrepresentation analysis (ORA) of genes uniquely upregulated or downregulated in progressors vs healthy individuals was performed using Hallmark gene sets. Pathways are ranked by gene ratio and statistical significance (adjusted *P* value). (C) Volcano plot showing differential gene expression between progressors and nonprogressors. A total of 91 genes were significantly deregulated (|log2FC| >0.58; *P* value <.05), including 68 upregulated (red) and 23 downregulated (blue). (D) Gene set enrichment analysis (GSEA) of genes ranked by log2FC in progressors vs nonprogressors using Hallmark gene sets (MSigDB). Positively and negatively enriched pathways are ranked by NES and colour-coded by FDR q-value. (E) An interferon-related gene score (IFN-score) was derived from 14 IFN-related genes selected through elastic-net regression and stability analysis. A logistic regression model was applied using IFN-score quartiles as the predictor, allowing estimation of both predicted probabilities (with 95% CIs) of being a progressor and the corresponding odds ratios (95% CI) across quartiles. The figure illustrates the graded increase in progression risk across IFN-score quartiles. (F) ORA was performed in a core set of 28 genes displaying differential expression (same thresholds as above) in both progressors vs healthy and progressors vs nonprogressors, but not in nonprogressors vs healthy subjects. The *cnetplot* function was used to illustrate enriched Hallmark pathways (adjusted *P* value <.05) and their associated genes, highlighting also shared gene-pathway relationships. (G) Volcano plot showing blood transcriptome changes in 4 progressors, assessed at baseline and at the time of SLE classification. A total of 589 genes were significantly deregulated (|log2FC| >0.58; *P* value <.05), including 424 upregulated (dark orange) and 165 downregulated (mint). (H) GSEA of genes ranked by log2FC between the SLE classification and baseline timepoints using Hallmark gene sets. Pathways are ranked by NES and coloured by FDR q-value, with intensity reflecting statistical significance. DEGs, differentially expressed genes; log2FC, log2-fold change; NES, normalised enrichment score; SLE, systemic lupus erythematosus.



M5	Regulation of insulin secretion, p53 signalling pathway
M4	Interferon Signalling, Cytokine Signalling in Immune system
M3	Eukaryotic Translation Elongation, Nonsense Mediated Decay (NMD)
M2	Metabolism of porphyrins, Heme biosynthesis, Class I MHC mediated antigen processing & presentation
M1	Activated TLR4 signalling



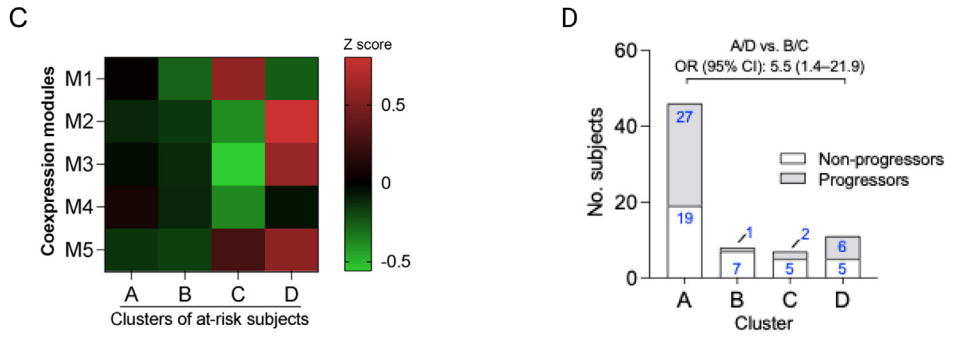


Figure 3. Unsupervised detection and functional annotation of coexpressed gene modules reveal biological processes that potentially accelerate or protect against the at-risk-to-SLE transition. (A) Gene set enrichment analysis (GSEA) of each study group (healthy, nonprogressors, progressors) across modules of coexpressed genes identified by the CEMiTool package as detailed in *Methods*. Modules M1−M5 contain 591, 165, 75, 43, and 34 genes, respectively. Functional annotations were derived via overrepresentation analysis (ORA), identifying key biological pathways associated with each module (outlined in the adjacent table). (B) For Modules M1 and M4, left panels present the top enriched pathways from the ORA, ranked by $-\log_{10}$(adjusted *P* value). Right panels display gene coexpression networks, highlighting the top 10 hub genes. One network graph is shown per module. M1 genes were involved in TLR signalling, with interaction network analysis highlighting *TLR4*. M4 genes were implicated in interferon (both interferon-alpha and -gamma) signalling with interaction network analysis revealing a key role for *DDX58* (encoding the viral RNA-sensor retinoic acid-inducible gene I [RIG-I]) and the mitochondrial gene *CMPK2* (encoding for cytidine/uridine monophosphate kinase 2). (C) Gaussian Mixture Modelling (GMM) of modules M1−M5 eigengenes in the full at-risk cohort identified 4 clusters (A−D) as outlined in the *Methods*. The heatmap

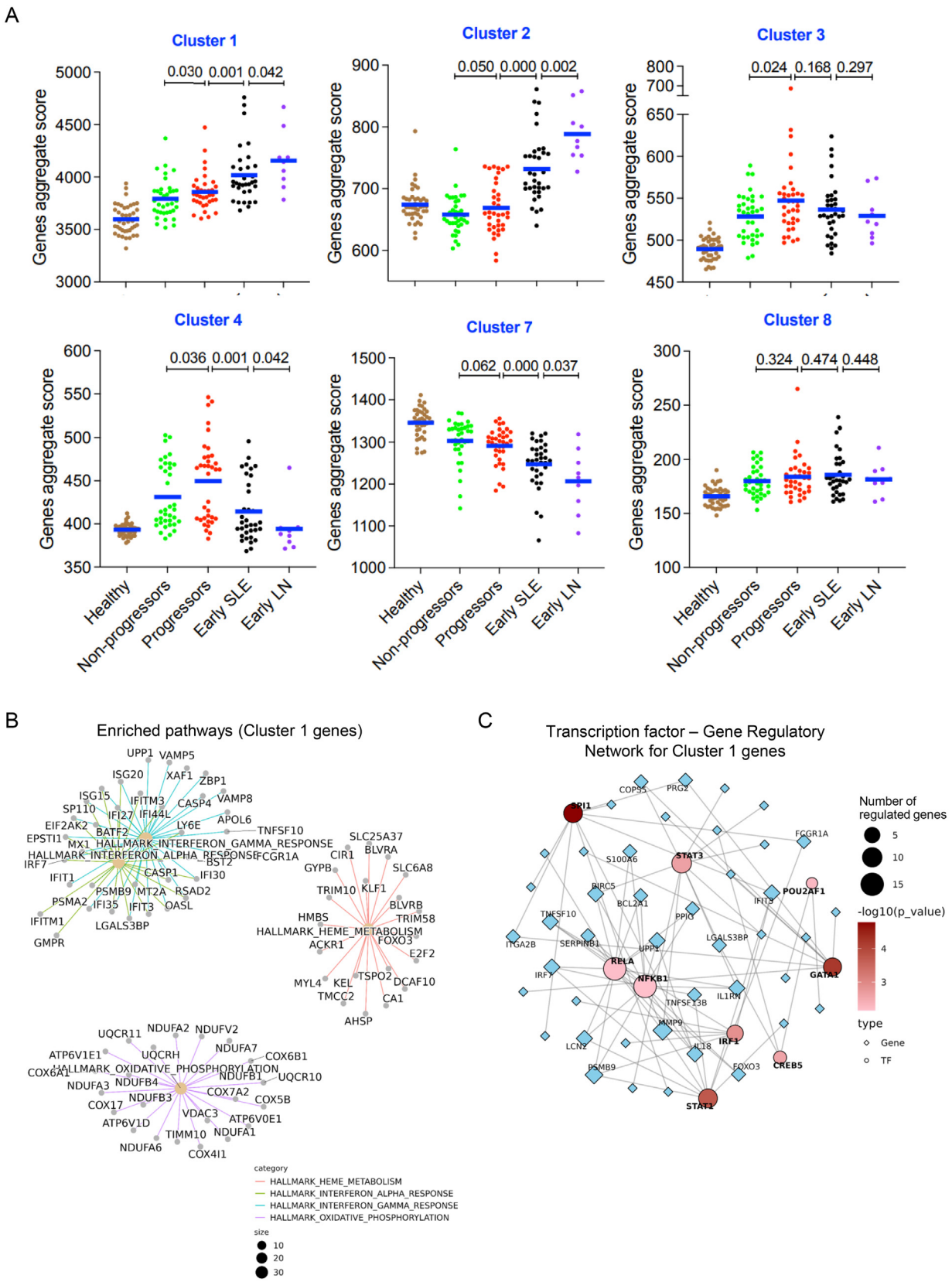


Figure 4. Transcriptional dynamics across healthy individuals, at-risk individuals, and individuals with early SLE reveal layered molecular programmes underlying susceptibility and progression to severe disease. (A) Gene aggregation scores were calculated for each individual as the sum of normalised expression values for all genes within each of the 6 clusters identified by likelihood ratio testing. Dot plots show group-specific expression

displays the mean standardised expression of M1−M5 eigengenes across these clusters. (D) Frequency of progressors (grey) and nonprogressors (white) within each GMM-derived cluster (A−D). Clusters A and D contained significantly more progressors than clusters B and C (odds ratio [OR] 5.5; 95% CI: 1.4-21.9). DEGs, differentially expressed genes; FADD, Fas-associated death domain; JAK-STAT3, Janus kinase-signal transducer and activator of transcription 3; KRAS, Kirsten rat sarcoma viral oncogene homolog; IL6, Interleukin-6; MYC, MYC proto-oncogene; MHC, major histocompatibility complex; MDA5, melanoma differentiation-associated protein 5; NF-κB, Nuclear Factor kappa-light-chain-enhancer of activated B cells; NES, normalized enrichment score; SLE, systemic lupus erythematosus; TLR, Toll-like receptor; TRAF3, Tumor necrosis factor receptor-associated factor 3.

both. A transcription factor-regulatory network analysis predicted upstream regulators of this cluster (Fig 4C). The heatmap in Fig 4D illustrates the relative expression of pathway-enriched genes across disease stages. A similar stepwise trend across all groups, albeit with decreasing expression, was observed in *Cluster 7* genes (n = 175), which showed nominally significant overrepresentation in IL2/STAT5 (P = .0398) (Supplementary Table S7) implicated in regulatory T-cell function [24].

Cluster 2 comprised 109 genes specifically activated in overt disease, enriched in IFN-alpha and IFN-gamma (adjusted $P \leq 4.32 \times 10^{-10}$), complement activation (adjusted P = .008), and inflammatory response pathways (P = .0482) (Fig S6B). Among the IFN-responsive transcripts, 2 were unique to the IFN-alpha pathway, 5 to IFN-gamma, and 11 were shared by both. Notably, these differed from the IFN-driven genes in *cluster 1*, suggesting that beyond the gradual upregulation of IFN signalling, a qualitative expansion of IFN signature is associated with SLE establishment. *Cluster 4* (n = 70 genes), and to a lesser extent *cluster 3* (n = 77 genes), comprised genes induced in progressors but downregulated in full-blown disease. These were enriched in the p53 pathway and glycolysis-related processes, respectively (nominal P <.05).

The above-described molecular patterns were mirrored by shifts in CIBERSORT (Cell-type Identification By Estimating Relative Subsets Of RNA Transcripts)-inferred immune cell populations across disease stages (Fig S7). Collectively, these data support the concept of SLE as a multistep disease process, marked by the progressive emergence and amplification of immune cell abnormalities and transcriptomic changes, predominantly involving IFN-signalling, inflammation, and metabolic pathways.

A compact gene signature distinguishes progressors from nonprogressors with external validity

To assess whether transcriptional patterns could aid in predicting transition from the at-risk state to SLE, we applied pathMED [18], which summarises gene-level expression into module activity scores (Fig 5A). Across 7 gene sets identified above, machine-learning classifiers prioritised the core-susceptibility signature (Fig 2F) as the most informative (Fig S8). To examine the performance of this signature at the gene level, we trained an elastic-net regression model on our dataset using 5-fold cross-validation to select optimal penalisation parameters. For external validation, we used the Leeds cohort of ANA-positive individuals at risk for SLE (15 progressors, 20 nonprogressors) [7]. Of the 24 core-susceptibility DEGs, 7 were expressed in both datasets and were used as model features. The resulting classifier achieved an area under the curve of 0.797 (95% CI: 0.649-0.944) in the external cohort (Fig 5B). Thus, a compact gene set captures transcriptional tendencies associated with imminent SLE and maintains good discriminatory capacity in an independent dataset.

Potential modulation of the SLE susceptibility and progression gene signatures by existing biologics and candidate agents

Having defined signatures underlying the evolution from at-risk state to SLE and to severe forms, we assessed whether therapeutic compounds could induce transcriptional changes opposite to these signatures. We retrieved blood RNA-sequencing data on transcripts downregulated in patients with active SLE following treatment with belimumab (n = 242 genes [19]) or anifrolumab (n = 219 genes [20]). Testing their enrichment within a ranked list of genes from the progressors vs nonprogressors comparison, we found significant enrichment for both anifrolumab (NES = 3.28, P <.0001) and belimumab (NES = 2.28, P <.0001) (Fig 6A, Fig S9), suggesting that these agents may counteract key molecular programmes underlying evolution from at-risk-to-SLE state. We next examined whether the *progression*-associated signatures may be potentially modulated by these biologics. For this, we tested for enrichment of *cluster 1* genes in drug-response gene rankings, where genes were ordered by expression changes (log2FC) in anifrolumab- or belimumab-treated patients. Anifrolumab showed significant negative enrichment of *cluster 1* (NES = −2.38, P <.001) and positive enrichment of *cluster 7* (NES = 2.13, P <.001), indicating transcriptional responses opposite to both *progression*-associated signatures (Fig 6B). Belimumab showed a similar, nonsignificant trend for *cluster 1* (NES = −1.01, P = .46) and reversal for *cluster 7* (NES = 1.82, P < 0001) (Fig 6C).

Drug-repurposing analysis using L1000CDS^2 identified fedratinib (TG-101348) and thiostrepton (T8902) as candidate small molecules predicted to reverse molecular perturbations associated with at-risk-to-SLE transition (Supplementary Table S8). Using CMap, 105 compounds were predicted to induce expression changes opposite to the *susceptibility* DEGs (Supplementary Table S9). The top 20 were visualised and grouped by mechanism of action (Fig 6D), including MG-132 (proteasome inhibitor), PFI-1 (BET inhibitor), and everolimus (mTOR inhibitor). Enrichment analysis of drug classes and pathways associated with these perturbagens highlighted bromodomain inhibitors and agents modulating NF-κB activation in B-cells (Fig 6E).

DISCUSSION

We characterised molecular alterations in individuals at risk for SLE, stratified by clinical trajectory, and identified targetable signatures underlying disease susceptibility and progression. Compared to healthy controls, the at-risk state

patterns across healthy, nonprogressors, progressors, early SLE, and early lupus nephritis (LN). Each dot represents an individual, and each panel corresponds to 1 gene cluster. Horizontal lines indicate median values. Statistical significance was assessed using 1-way ANOVA with multiple testing correction; false discovery rate (FDR) q-values are shown. (B) *Cnetplot* showing functional enrichment of cluster 1 genes using the Hallmark C50 gene sets. Edges link genes to their enriched pathways, based on overrepresentation analysis. (C) Transcription factor (TF)−gene regulatory network analysis was performed based on TRRUST-inferred regulators of cluster 1 genes. TFs were selected by enrichment significance (Q < 0.05). The bipartite network shows TF-target gene interactions, with TF nodes coloured by $-\log_{10}$ (P value). Target genes regulated by 2 or more TFs are labelled to indicate potential coregulation. Key upstream regulators included *GATA1, SPI1, IRF1, STAT1, STAT3, NF-κB, CREB5*, and *POU2AF1*. (D) Heatmap of Cluster 1 genes involved in the *interferon-alpha response* (23 of 97 genes), *interferon-gamma response* (33 of 200 genes), *oxidative phosphorylation* (23 of 200 genes) and *haem metabolism* (20 of 200 genes), showing expression across groups. Expression values are z-score scaled and clustered to highlight group-specific patterns. Eight of the 24 upregulated core-susceptibility signature genes (*IFI44L, EPSTI1, MT2A, MT1L, CEACAM6, OTOF, EIF2AK2, LY6E*) belonged also to *Cluster 1*. ANOVA, analysis of variance; IFN, interferon; SLE, systemic lupus erythematosus; TRRUST, transcriptional regulatory relationships unraveled by sentence-based text mining.

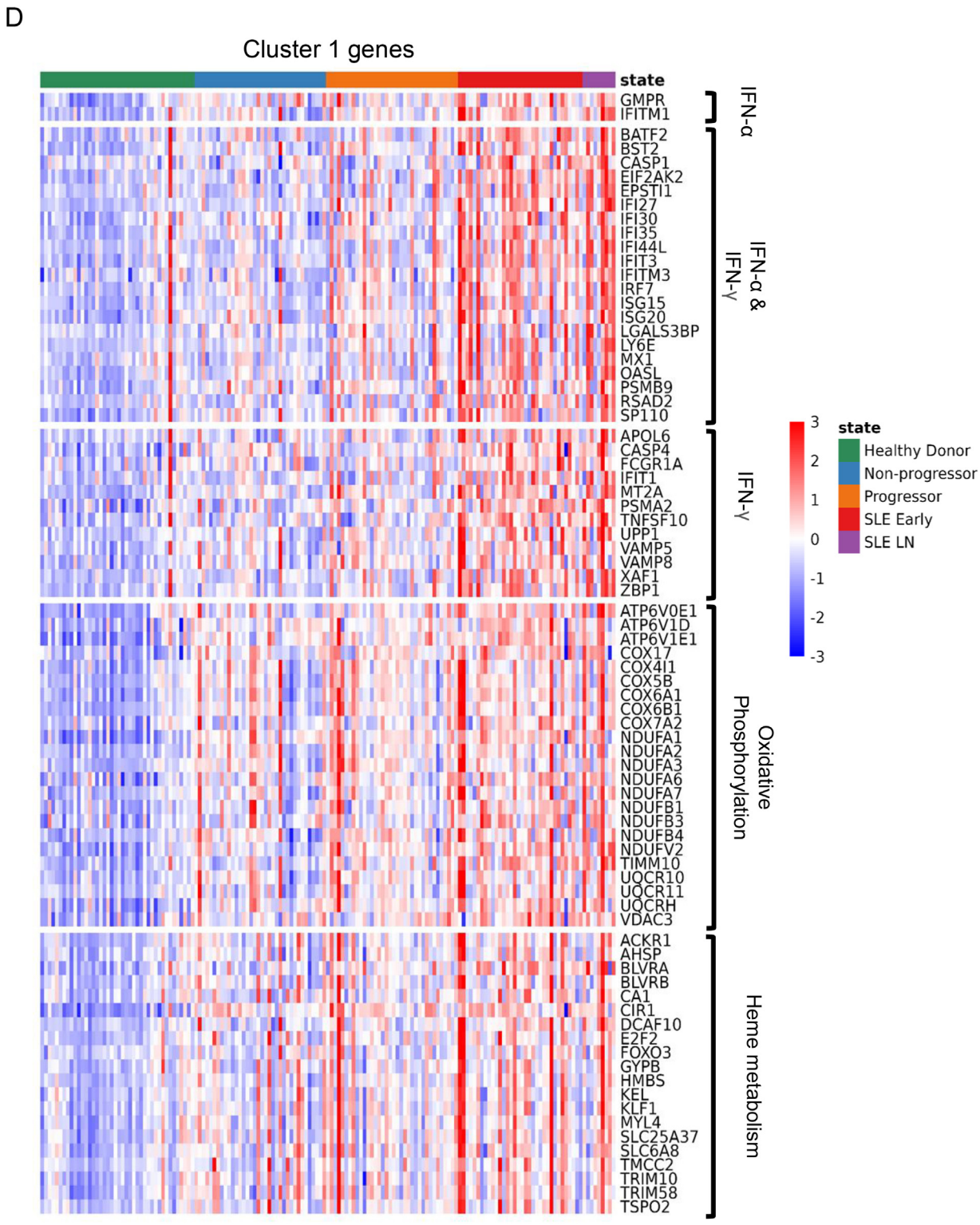


Figure 4. Continued.

showed deregulation of metabolic, cytokine signalling, haematologic, cell cycle, and stress-response pathways, whereas imminent SLE was marked by activation of type I/II IFN and inflammatory gene sets. Gene-module analysis clustered at-risk individuals into molecularly distinct subsets, and predictive modelling approaches indicated that blood transcriptomic features may aid risk stratification. Notably, perturbations evident in the at-risk stage—such as IFN-α/γ signalling and OxPhos—also characterise active established and severe SLE and may be reversed by biologics, supporting their potential for early intervention.

SLE is preceded by a phase in which clinical and/or immunological features are present but insufficient for definitive diagnosis or classification [1,25]. We found that the blood transcriptome in these individuals diverges substantially from that of healthy controls, with deregulation of multiple immune

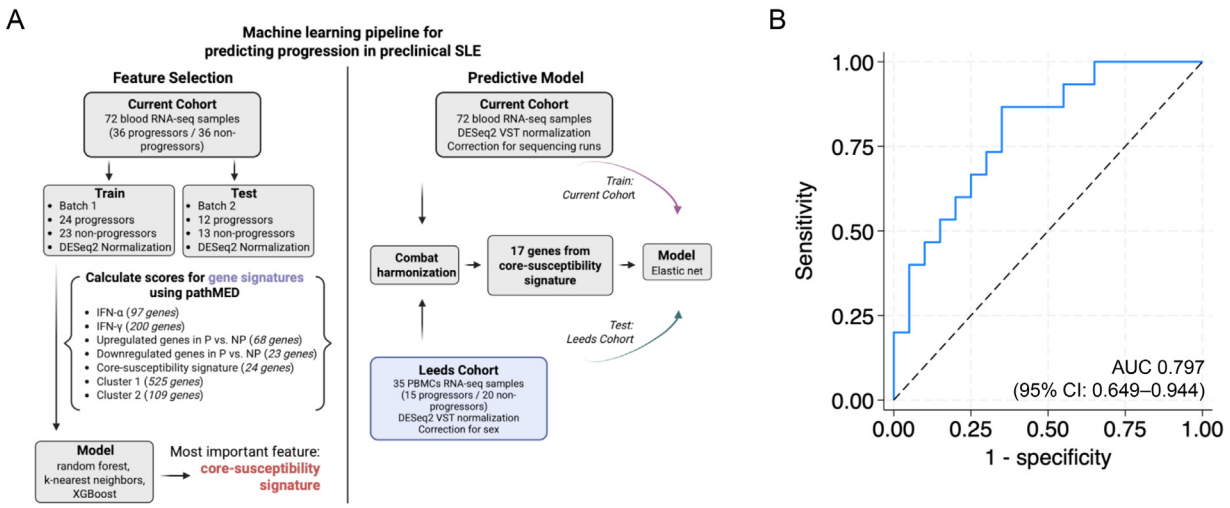


Figure 5. External validation of the core-susceptibility gene signature in distinguishing at-risk individuals who progressed or did not progress to classifiable SLE. (A) Overview of the feature selection workflow and external validation strategy used for the prediction model. Illustration was created in Biorender. (B) For gene-level modelling of the susceptibility signature, elastic-net regression was applied to variance-stabilised and batch-corrected expression data, with 5-fold cross-validation used to select optimal penalisation parameters. The final classifier was trained on the present study dataset and evaluated in an independent RNA-seq cohort [7]. Details of preprocessing, model fitting, and validation procedures are provided in the Supplementary Methods. AUC, area under the curve; DESeq2, differential expression analysis based on the negative binomial distribution; IFN, interferon; NP, non-progressors; P, progressors; PBMC, peripheral blood mononuclear cells, RNA-seq, RNA-sequencing; SLE, systemic lupus erythematosus; VST, variance stabilizing transformation.

and nonimmune pathways. This aligns with previous studies showing that ANA-positive or incomplete lupus subjects exhibit diverse cellular and molecular alterations [6,7,26–28]. These changes are consistent with the clinical burden often reported in this population [25,29].

Using supervised and unsupervised approaches, we implicate both type I and type II IFNs in the at-risk state progressing to SLE. This is consistent with prior transcriptomic [7,30] and proteomic [5,31,32] studies assessing type I IFN bioactivity, supporting its pathogenic role in SLE [33,34]. Our findings align with those of Munroe et al [5] who reported elevated type II IFN mediators and autoantibodies preceding a surge in IFN-α activity closer to SLE classification. These data are also in line with the shift towards $CD4^+$ Th1 cells with IFN-γ production, previously observed in patients with undifferentiated connective tissue disease, who progressed to definitive diseases including SLE [28].

By analysing coexpressed gene modules, we observed heterogeneity within the at-risk state transitioning to SLE. A subset of these individuals exhibited dampened TLR4 signalling, possibly reflecting skewing towards type-2 immunity or representing a compensatory mechanism [35]. This group was also characterised by increased p53 signalling, a finding previously reported in patients with SLE [8], which might relate to the multifaceted immune effects of p53 [36]. Our observation of stepwise induction of specific gene signatures from health to active SLE and lupus nephritis merits discussion. Chiche et al [37] showed sequential activation of IFN-related modules in longitudinal SLE samples, where a basal module remained stable, whereas a high-threshold module (induced by type I and II IFNs) tracked with renal flares. These data align with our implication of both IFN types in evolution from health to mild and severe autoimmunity. High baseline expression of IFN-stimulated genes has been linked to more severe or active disease trajectories [38,39], supporting a role in disease progression beyond acute flares [40–42].

Genes involved in OxPhos showed progressive perturbation across disease stages. This aligns with prior RNA-sequencing data from our group implicating this pathway in active/severe SLE, including nephritis [8]. Similarly, Nakano et al [43] identified OxPhos as a core SLE signature across multiple cell types. Mitochondrial metabolism has been linked to functional maladaptation of T-cells [44], B-cells [45], and monocytes [46] in SLE. The specific cellular and functional impact of upregulated OxPhos gene expression requires further investigation.

Given the clinical similarity of the 2 at-risk groups, relaxed statistical thresholds were applied to preserve sensitivity to subtle transcriptional differences. While this increases the likelihood of uncertain or false–positive signals, the consistency of our findings was supported by convergent evidence from GSEA, coexpression analysis, and validation in an independent dataset.

We previously showed that belimumab treatment in active SLE suppresses transcriptional perturbations related to B-cell activation, type I/II IFN signalling, interleukin-6/Signal Transducer and Activator of Transcription 3, and neutrophil pathways [19]. Similarly, anifrolumab modulates multiple lupus-related processes, including apoptotic, innate, and adaptive mechanisms [20]. We demonstrate that both agents are associated with transcriptional shifts that oppose the core molecular programmes linked to progression from the at-risk state to classifiable and severe SLE. These findings suggest potential relevance for modifying early disease trajectories but require formal validation before therapeutic conclusions can be drawn.

Recent studies of preclinical/at-risk SLE converge on pathway-level alterations, including type I/II IFN signalling, TNF/mitogen-activated protein kinase/NF-κB pathways, and immunometabolic rewiring (including autophagy/OxPhos), observed

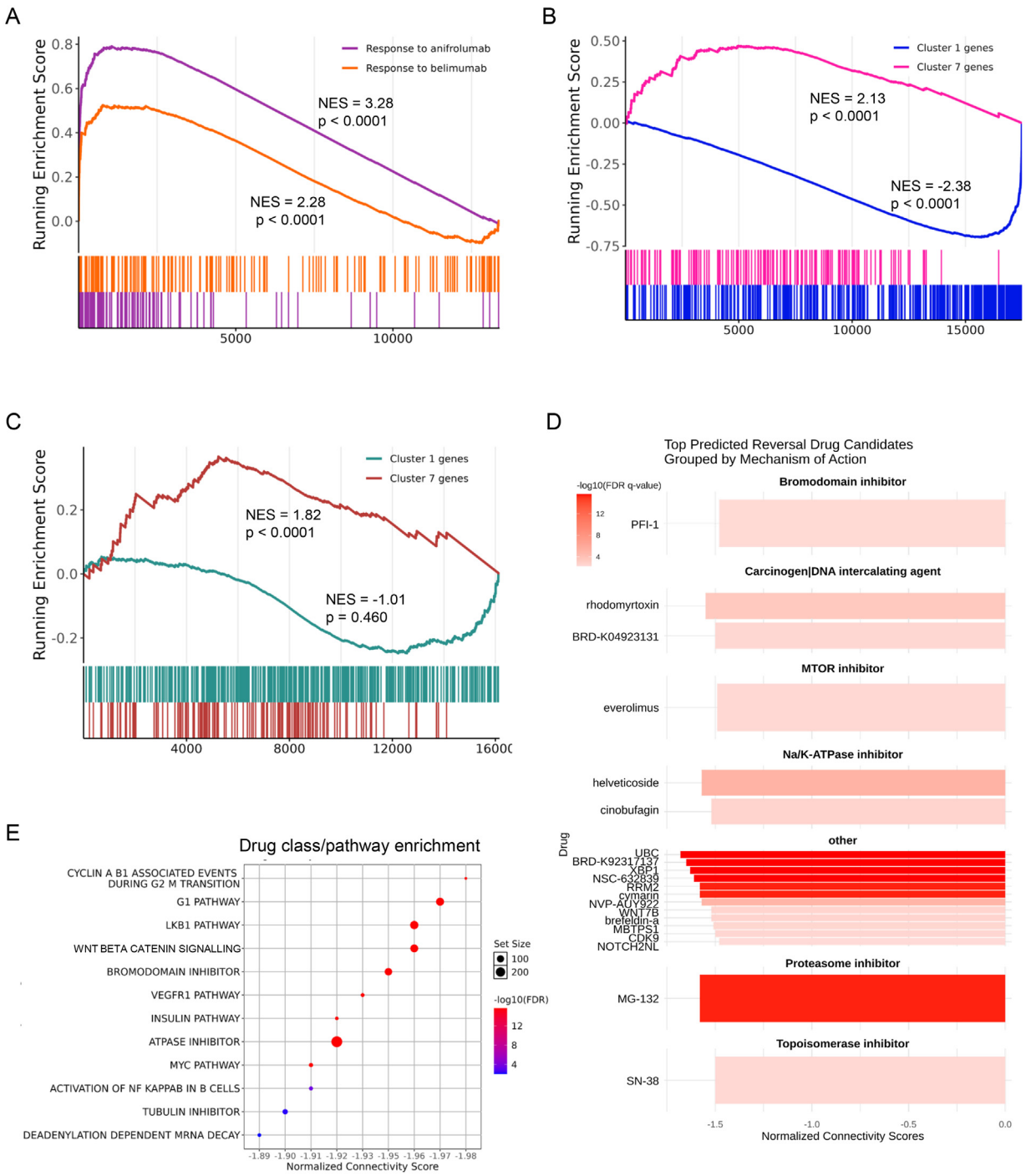


Figure 6. Predicted reversibility analysis of SLE susceptibility and progression signatures by approved biologics and candidate compounds identified through drug-repurposing pipelines. (A) Gene set enrichment analysis (GSEA) of 2 gene signatures: (i) genes significantly downregulated (n = 219, log2FC < −0.58; *P* value <.05) in anifrolumab-treated patients vs placebo (at 52 weeks) and (ii) genes significantly downregulated (n = 242, log2FC < −0.58; *P* value <.05) in belimumab responders after 6 months of treatment vs baseline. Both signatures were tested against a ranked gene list (by log2FC) from the progressor vs nonprogressor comparison. The anifrolumab response signature (purple) and the belimumab response signature (orange) showed significant enrichment, with normalised enrichment scores (NES) of 3.28 (*P* <.0001) and 2.28 (*P* <.0001), respectively. (B) GSEA of *Cluster 1* (blue) and *Cluster 7* (pink) genes against the ranked gene list (by log2FC) from the anifrolumab vs placebo comparison showed strong negative enrichment for cluster 1 (NES = −2.23, *P* value <.0001) and strong positive enrichment for cluster 7 (NES = 2.13, *P* value < 0.0001). (C) GSEA of *Cluster 1* (green) and *Cluster 7* (red) genes against the ranked gene list (by log2FC) from the belimumab (6-month vs baseline) comparison showed strong positive enrichment for cluster 7 (NES = 1.82, *P* value <.0001) and negative—yet nonsignificant—enrichment for *Cluster 1* (NES = −1.01, *P* value =.46). (D) Bar plot showing the top 20 compounds predicted to reverse the susceptibility signature (progressor vs nonprogressor comparison), as identified by CMap analysis. The input included 91 DEGs (|log2FC| >0.58; *P* value <.05). Compounds are grouped vertically by annotated mechanisms of action. The x-axis indicates the normalised connectivity score (norm_cs); bar colour reflects prediction confidence ($-\log_{10}$ FDR). (E) Pathway enrichment analysis of drug signatures identified by the CMap query. Enriched pathways were derived from compounds with significant negative connectivity (FDR <0.05; norm_cs < −1), indicating the potential of each corresponding drug class to reverse the susceptibility signature. CMap, Connectivity Map; DEGs, differentially expressed genes; FDR, false discovery rate; log2FC, log2-fold change; NES, normalised enrichment score; SLE, systemic lupus erythematosus; NES, normalized enrichment score; MTOR, mechanistic target of rapamycin; MYC, MYC proto-oncogene; NF, nuclear factor; VEGFR1, vascular endothelial growth factor receptor 1.

across proteomic and single-cell modalities [47–50]. Our findings concord to these reports and extend them by demonstrating reproducible, ranked-list pathway signals in an independent cohort and by linking these signals to drug-reversal analyses.

In conclusion, our study defines molecular events driving the transition from health to early and severe SLE, anchored in type I/II IFN signalling and OxPhos that intensify with disease progression. These signatures, detectable before clinical classification, are both prognostic and potentially actionable. Through system-level analysis and pharmacologic mapping, we highlight therapeutic targets—including belimumab, anifrolumab, and repurposed agents—that could be explored for early interception in high-risk individuals.

Competing interests

GB reports grants from GSK, AstraZeneca, MSD (outside the context of the current work), and honoraria and/or consulting fees from Lilly, Novartis, AstraZeneca, GSK, Abbvie, Pfizer, and Otsuka. LI reports honoraria and/or consulting fees from GSK, AstraZeneca, and Otsuka. LA reports honoraria and/or consulting fees from GSK, Pfizer, and UCB. LMC has received consultancy fees from UCB and Alumis. EMV has received honoraria from Novartis, AstraZeneca, Otsuka, Roche, UCB, Aurinia, Lilly, Alumis, BMS, GSK, and Pfizer and research grants from AstraZeneca and Sandoz. The remaining authors declare no conflicts of interest. One of the coauthors (DB) is a member of the Journal Editorial Board.

Acknowledgements

We thank the individuals who participated in the study as well as the physicians and nurses of the participating Rheumatology Departments (Crete, Athens, Coimbra, Brescia, Pisa) for their dedicated patient care. The graphical abstract and Figure 5A were created with Biorender.

Contributors

SP performed data curation, and bioinformatics analysis, prepared figures, and drafted parts of the manuscript. EE performed reviewed medical charts and data entry. CA assisted in the assembly of the at-risk cohort. EK assisted in the assembly of the at-risk cohort and in blood samples collection. DN and NK assisted in the assembly of the at-risk cohort and collected data on patients with SLE. MF, CT, AR, NA, LA, MM, and LI recruited individuals at risk and performed monitoring and data entry. LMC and EMV shared RNA-seq data, provided guidance on data preprocessing/analysis and drafted parts of the manuscript. NM and DN assisted with the processing of blood samples and RNA isolation. PG contributed to the drug-repurposing analysis. AB and PS assisted with collection of specimens and clinical data from patients with SLE. GV performed RNA-sequencing. DK performed RNA mapping. DB contributed to data interpretation and drafting of the manuscript. CN supervised the bioinformatics analyses and drafted parts of the manuscript. GB conceived and supervised the study, recruited at-risk individuals and patients with SLE, performed statistical analysis, and drafted the manuscript.

All authors had unrestricted access to all study data and assumed final responsibility for the decision to publish.

Funding

The study received funding from the FOREUM (Foundation for Research in Rheumatology; project id: 016) and the Pancretan Health Association.

Provenance and peer review

This was an unsolicited, externally peer reviewed submission.

Patient consent for publication

Patients' informed consent was obtained for data collection, biospecimens collection, analysis, and publication of results.

Ethics approval

This study involves human participants, and data collection and analysis were approved by the ethics committees of the participating centres (protocol no. 7/2018/16-05-2018; NP3212/21-09-2018; CHUC-109-17). All patients gave informed consent upon inclusion in the respective registries.

Data availability statement

The RNA-sequencing data generated for this study have been deposited in GEO under accession number GSE311055.

Supplementary materials

Supplementary material associated with this article can be found in the online version at doi:10.1016/j.ard.2026.02.003.

Orcid

George Bertsias: http://orcid.org/0000-0001-5299-1406

REFERENCES

[1] Choi MY, Costenbader KH. Understanding the concept of pre-clinical autoimmunity: prediction and prevention of systemic lupus erythematosus: identifying risk factors and developing strategies against disease development. Front Immunol 2022;13:890522.
[2] Arbuckle MR, McClain MT, Rubertone MV, Scofield RH, Dennis GJ, James JA, et al. Development of autoantibodies before the clinical onset of systemic lupus erythematosus. N Engl J Med 2003;349(16):1526–33 Oct 16.
[3] Eriksson C, Kokkonen H, Johansson M, Hallmans G, Wadell G, Rantapaa-Dahlqvist S. Autoantibodies predate the onset of systemic lupus erythematosus in northern Sweden. Arthritis Res Ther 2011;13(1):R30. Feb 22.
[4] Lu R, Munroe ME, Guthridge JM, Bean KM, Fife DA, Chen H, et al. Dysregulation of innate and adaptive serum mediators precedes systemic lupus erythematosus classification and improves prognostic accuracy of autoantibodies. J Autoimmun 2016;74:182–93 Nov.
[5] Munroe ME, Lu R, Zhao YD, Fife DA, Robertson JM, Guthridge JM, et al. Altered type II interferon precedes autoantibody accrual and elevated type I interferon activity prior to systemic lupus erythematosus classification. Ann Rheum Dis 2016;75(11):2014–21 Nov.
[6] Slight-Webb S, Lu R, Ritterhouse LL, Munroe ME, Maecker HT, Fathman CG, et al. Autoantibody-positive healthy individuals display unique immune profiles that may regulate autoimmunity. Arthritis Rheumatol 2016;68(10):2492–502 Oct.
[7] Carter LM, Md Yusof MY, Wigston Z, Plant D, Wenlock S, Alase A, et al. Blood RNA-sequencing across the continuum of ANA-positive autoimmunity reveals insights into initiating immunopathology. Ann Rheum Dis 2024;83(10):1322–34 Sep 30.
[8] Panousis NI, Bertsias GK, Ongen H, Gergianaki I, Tektonidou MG, Trachana M, et al. Combined genetic and transcriptome analysis of patients with SLE:

distinct, targetable signatures for susceptibility and severity. Ann Rheum Dis 2019;78(8):1079–89 Aug.
[9] Ntasis VF, Panousis NI, Tektonidou MG, Dermitzakis ET, Boumpas DT, Bertsias GK, et al. Extensive fragmentation and re-organization of transcription in Systemic Lupus Erythematosus. Sci Rep 2020;10(1):16648 Oct 6.
[10] Emmanouilidou E, Adamichou C, Nikoloudaki M, Kalogiannaki E, Repa A, Avgustidis N, et al. POS0768 presence of anti-Ro/SSA autoantibodies, hypocomplementemia and photosensitivity indicate individuals with connective tissue disease features who are at increased risk for transition to systemic lupus erythematosus. Ann Rheum Dis 2022;81(Suppl 1):671.
[11] Hochberg MC. Updating the American College of Rheumatology revised criteria for the classification of systemic lupus erythematosus. Arthritis Rheum 1997;40(9):1725. Sep.
[12] Aringer M, Costenbader K, Daikh D, Brinks R, Mosca M, Ramsey-Goldman R, et al. 2019 European League Against Rheumatism/American College of Rheumatology classification criteria for systemic lupus erythematosus. Ann Rheum Dis 2019;78(9):1151–9 Sep.
[13] Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol 2014;15(12):550.
[14] Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, et al. limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Res 2015;43(7):e47. Apr 20.
[15] Russo PST, Ferreira GR, Cardozo LE, Burger MC, Arias-Carrasco R, Maruyama SR, et al. CEMiTool: a bioconductor package for performing comprehensive modular co-expression analyses. BMC Bioinformatics 2018;19(1):56. Feb 20.
[16] Cheng CW, Beech DJ, Wheatcroft SB. Advantages of CEMiTool for gene co-expression analysis of RNA-seq data. Comput Biol Med 2020;125:103975 Oct.
[17] Newman AM, Steen CB, Liu CL, Gentles AJ, Chaudhuri AA, Scherer F, et al. Determining cell type abundance and expression from bulk tissues with digital cytometry. Nat Biotechnol 2019;37(7):773–82 Jul.
[18] Toro-Dominguez D, Martorell-Marugan J, Martinez-Bueno M, Lopez-Dominguez R, Carnero-Montoro E, Barturen G, et al. Scoring personalized molecular portraits identify systemic lupus erythematosus subtypes and predict individualized drug responses, symptomatology and disease progression. Brief Bioinform 2022;23(5):bbac332. Sep 20.
[19] Moysidou GS, Garantziotis P, Sentis G, Nikoleri D, Malissovas N, Nikoloudaki M, et al. Molecular basis for the disease-modifying effects of belimumab in systemic lupus erythematosus and molecular predictors of early response: blood transcriptome analysis implicates the innate immunity and DNA damage response pathways. Ann Rheum Dis 2025;84(2):262–73 Feb.
[20] Baker T, Sharifian H, Newcombe PJ, Gavin PG, Lazarus MN, Ramaswamy M, et al. Type I interferon blockade with anifrolumab in patients with systemic lupus erythematosus modulates key immunopathological pathways in a gene expression and proteomic analysis of two phase 3 trials. Ann Rheum Dis 2024;83(8):1018–27 Jul 15.
[21] Korotkevich G, Sukhov V, Budin N, Shpak B, Artyomov MN, Sergushichev A. An algorithm for fast preranked gene set enrichment analysis using cumulative statistic calculation. bioRxiv. 2021:060012.
[22] Duan Q, Reid SP, Clark NR, Wang Z, Fernandez NF, Rouillard AD, et al. L1000CDS(2): LINCS L1000 characteristic direction signatures search engine. NPJ Syst Biol Appl. 2016;2:16015.
[23] Subramanian A, Narayan R, Corsello SM, Peck DD, Natoli TE, Lu X, et al. A next generation connectivity map: L1000 platform and the first 1,000,000 profiles. Cell 2017;171(6) Nov 301437–52.e1417.
[24] He J, Zhang X, Wei Y, Sun X, Chen Y, Deng J, et al. Low-dose interleukin-2 treatment selectively modulates CD4(+) T cell subsets in patients with systemic lupus erythematosus. Nat Med 2016;22(9):991–3 Sep.
[25] Gatto M, Saccon F, Zen M, Iaccarino L, Doria A. Preclinical and early systemic lupus erythematosus. Best Pract Res Clin Rheumatol 2019;33(4):101422 Aug.
[26] Aberle T, Bourn RL, Munroe ME, Chen H, Roberts VC, Guthridge JM, et al. Clinical and serologic features in patients with incomplete lupus classification versus systemic lupus erythematosus patients and controls. Arthritis Care Res (Hoboken) 2017;69(12):1780–8 Dec.
[27] Henning S, Lambers WM, Doornbos-van der Meer B, Abdulahad WH, Kroese FGM, Bootsma H, et al. Proportions of B-cell subsets are altered in incomplete systemic lupus erythematosus and correlate with interferon score and IgG levels. Rheumatology (Oxford) 2020;59(9):2616–24 Sep 1.
[28] Szodoray P, Nakken B, Barath S, Gaal J, Aleksza M, Zeher M, et al. Progressive divergent shifts in natural and induced T-regulatory cells signify the transition from undifferentiated to definitive connective tissue disease. Int Immunol 2008;20(8):971–9 Aug.
[29] Tani C, Trentin F, Parma A, Zucchi D, Cardelli C, Stagnaro C, et al. Disease evolution and organ damage accrual in patients with stable UCTD: a long-term monocentric inception cohort. RMD Open 2024;10(2):e003967 Apr 25.
[30] Md Yusof MY, Psarras A, El-Sherbiny YM, Hensor EMA, Dutton K, Ul-Hassan S, et al. Prediction of autoimmune connective tissue disease in an at-risk cohort: prognostic value of a novel two-score system for interferon status. Ann Rheum Dis 2018;77(10):1432–9 Oct.
[31] Lambers WM, Westra J, Arends S, Doornbos-van der Meer B, Horvath B, Bootsma H, et al. Persistent low complement levels and increased interferon gene expression are predictive for disease progression in patients with incomplete systemic lupus erythematosus. Joint Bone Spine 2022;89(5):105381 Oct.
[32] Kim ST, Munoz-Grajales C, Dunn SE, Schneider R, Johnson SR, Touma Z, et al. Interferon and interferon-induced cytokines as markers of impending clinical progression in ANA(+) individuals without a systemic autoimmune rheumatic disease diagnosis. Arthritis Res Ther 2023;25(1):21. Feb 10.
[33] Ronnblom L, Leonard D. Interferon pathway in SLE: one key to unlocking the mystery of the disease. Lupus Sci Med 2019;6(1):e000270.
[34] Psarras A, Wittmann M, Vital EM. Emerging concepts of type I interferons in SLE pathogenesis and therapy. Nat Rev Rheumatol 2022;18(10):575–90 Oct.
[35] Kawai T, Ikegawa M, Ori D, Akira S. Decoding Toll-like receptors: recent insights and perspectives in innate immunity. Immunity 2024;57(4):649–73 Apr 9.
[36] Wu HH, Leng S, Eisenstat DD, Sergi C, Leng R. Targeting p53 for immune modulation: Exploring its functions in tumor immunity and inflammation. Cancer Lett 2025;617:217614 May 1.
[37] Chiche L, Jourde-Chiche N, Whalen E, Presnell S, Gersuk V, Dang K, et al. Modular transcriptional repertoire analyses of adults with systemic lupus erythematosus reveal distinct type I and type II interferon signatures. Arthritis Rheumatol 2014;66(6):1583–95 Jun.
[38] Mai L, Asaduzzaman A, Noamani B, Fortin PR, Gladman DD, Touma Z, et al. The baseline interferon signature predicts disease severity over the subsequent 5 years in systemic lupus erythematosus. Arthritis Res Ther 2021;23(1):29. Jan 16.
[39] Northcott M, Jones S, Koelmeyer R, Bonin J, Vincent F, Kandane-Rathnayake R, et al. Type 1 interferon status in systemic lupus erythematosus: a longitudinal analysis. Lupus Sci Med 2022;9(1):e000625 Feb.
[40] Feng X, Wu H, Grossman JM, Hanvivadhanakul P, FitzGerald JD, Park GS, et al. Association of increased interferon-inducible gene expression with disease activity and lupus nephritis in patients with systemic lupus erythematosus. Arthritis Rheum 2006;54(9):2951–62 Sep.
[41] Gomez-Banuelos E, Goldman DW, Andrade V, Darrah E, Petri M, Andrade F. Uncoupling interferons and the interferon signature explains clinical and transcriptional subsets in SLE. Cell Rep Med 2024;5(5):101569 May 21.
[42] Chasset F, Mathian A, Dorgham K, Ribi C, Trendelenburg M, Huynh-Do U, et al. Serum interferon-alpha levels and IFN type I-stimulated genes score perform equally to assess systemic lupus erythematosus disease activity. Ann Rheum Dis 2022;81(6):901–3 Jun.
[43] Nakano M, Ota M, Takeshima Y, Iwasaki Y, Hatano H, Nagafuchi Y, et al. Distinct transcriptome architectures underlying lupus establishment and exacerbation. Cell 2022;185(18) Sep 13375–89.e3321.
[44] Buang N, Tapeng L, Gray V, Sardini A, Whilding C, Lightstone L, et al. Type I interferons affect the metabolic fitness of CD8(+) T cells from patients with systemic lupus erythematosus. Nat Commun 2021;12(1):1980. Mar 31.
[45] Takeshima Y, Iwasaki Y, Nakano M, Narushima Y, Ota M, Nagafuchi Y, et al. Immune cell multiomics analysis reveals contribution of oxidative phosphorylation to B-cell functions and organ damage of lupus. Ann Rheum Dis 2022;81(6):845–53 Jun.
[46] Gkirtzimanaki K, Kabrani E, Nikoleri D, Polyzos A, Blanas A, Sidiropoulos P, et al. IFNalpha impairs autophagic degradation of mtDNA promoting autoreactivity of SLE monocytes in a STING-dependent fashion. Cell Rep 2018;25(4) Oct 23921–33.e925.
[47] Bylinska A, Smith M, Lu R, Jones B, Guthridge C, Marlin M, et al. Identifying predictive serum soluble mediators signatures specific to ANA+ at risk of SLE individuals with next generation proteomics. Arthritis Rheumatol 2024;76:1900–2.
[48] Bylinska A, Smith M, Slight-Webb S, Guthridge C, Marlin C, Thomas K, et al. Single-cell multi-omic evaluation of differences in immune cell populations in progression toward systemic lupus erythematosus. Arthritis Rheumatol 2023;75:1834–7.
[49] Munroe ME, Young K, Lu R, Guthridge JM, Kamen DL, Gilkeson GS, et al. Dysregulated soluble immune mediators and lupus-associated autoantibody specificities inform the development of immune indexes that characterise classified SLE transition and SLE disease activity. Lupus Sci Med 2025;12(2):e001753 Nov 28.
[50] Psarras A, Arocha S, Tamon L, Ahern D, Yusof MM, Vital E, et al. Characterisation of immunometabolic reprogramming at the single cell level in patients with systemic lupus erythematosus and preclinical autoimmunity. Arthritis Rheumatol 2025;77:5028–9.