




Targeted multiomics in childhood-onset SLE reveal distinct biological phenotypes associated with disease activity: results from an explorative study

Mohamed Javad Wahadat ^{1,2}, Sander J van Tilburg,¹ Yvonne M Mueller,¹ Harm de Wit,¹ Cornelia G Van Helden-Meeuwsen,¹ Anton W Langerak,¹ Marike J Gruijters,² Amani Mubarak,² Marleen Verkaaik,² Peter D Katsikis,¹ Marjan A Versnel ¹, Sylvia Kamphuis ²

To cite: Wahadat MJ, van Tilburg SJ, Mueller YM, *et al.* Targeted multiomics in childhood-onset SLE reveal distinct biological phenotypes associated with disease activity: results from an explorative study. *Lupus Science & Medicine* 2023;**10**:e000799. doi:10.1136/lupus-2022-000799

► Additional supplemental material is published online only. To view, please visit the journal online (<http://dx.doi.org/10.1136/lupus-2022-000799>).

MAV and SK are joint senior authors.

Received 8 August 2022
Accepted 10 February 2023



© Author(s) (or their employer(s)) 2023. Re-use permitted under CC BY-NC. No commercial re-use. See rights and permissions. Published by BMJ.

¹Department of Immunology, Erasmus MC, Rotterdam, The Netherlands

²Department of Paediatric Rheumatology, Erasmus MC Sophia Children's Hospital, Rotterdam, The Netherlands

Correspondence to

Dr Sylvia Kamphuis; s.kamphuis@erasmusmc.nl

ABSTRACT

Objective To combine targeted transcriptomic and proteomic data in an unsupervised hierarchical clustering method to stratify patients with childhood-onset SLE (cSLE) into similar biological phenotypes, and study the immunological cellular landscape that characterises the clusters.

Methods Targeted whole blood gene expression and serum cytokines were determined in patients with cSLE, preselected on disease activity state (at diagnosis, Low Lupus Disease Activity State (LLDAS), flare). Unsupervised hierarchical clustering, agnostic to disease characteristics, was used to identify clusters with distinct biological phenotypes. Disease activity was scored by clinical SELENA-SLEDAI (Safety of Estrogens in Systemic Lupus Erythematosus National Assessment-Systemic Lupus Erythematosus Disease Activity Index). High-dimensional 40-colour flow cytometry was used to identify immune cell subsets.

Results Three unique clusters were identified, each characterised by a set of differentially expressed genes and cytokines, and by disease activity state: cluster 1 contained primarily patients in LLDAS, cluster 2 contained mainly treatment-naïve patients at diagnosis and cluster 3 contained a mixed group of patients, namely in LLDAS, at diagnosis and disease flare. The biological phenotypes did not reflect previous organ system involvement and over time, patients could move from one cluster to another. Healthy controls clustered together in cluster 1. Specific immune cell subsets, including CD11c+ B cells, conventional dendritic cells, plasmablasts and early effector CD4+ T cells, differed between the clusters.

Conclusion Using a targeted multiomic approach, we clustered patients into distinct biological phenotypes that are related to disease activity state but not to organ system involvement. This supports a new concept where choice of treatment and tapering strategies are not solely based on clinical phenotype but includes measuring novel biological parameters.

WHAT IS ALREADY KNOWN ON THIS TOPIC

⇒ The quest for personalised therapies, which are more efficacious while avoiding unnecessary side-effects, has fuelled omic studies investigating biological pathways in primarily adults with SLE to find novel biomarkers for disease activity states. Children could benefit the most from better disease understanding as they have a more severe disease than adults with SLE.

WHAT THIS STUDY ADDS

⇒ Targeted multiomics, combined with unsupervised hierarchical clustering analysis performed in children with SLE, led to the identification of clusters of patients with distinct biological phenotypes that were associated with disease activity states.

HOW THIS STUDY MIGHT AFFECT RESEARCH, PRACTICE OR POLICY

⇒ Clinical decision-making in SLE and patient outcome will be improved by better understanding of the biological pathways underlying active disease and disease remission. Multiomic studies have shown their value in adult-onset SLE where the first steps are made towards stratification of patients to guide treatment choices. With the results of our study, we made a start to achieve the same for children with SLE.

INTRODUCTION

Childhood-onset SLE (cSLE) is a devastating relapsing remitting autoimmune disease characterised by significant heterogeneity between patients.¹ Before the age of 30 years, 5%–10% of the patients with cSLE will have died as a result of severe disease.^{2–4} At 30 years of age, the majority of the patients will have developed irreversible damage, due to the disease or its treatment. Prednisone accounts for

around 50% of the damage, which is a reason to limit or withdraw prednisone completely.^{2,3} Response to therapy and when to taper or withdraw medication cannot be clearly predicted by clinical disease phenotypes. Therefore, personalised treatment strategies are needed, especially those that not only use clinical characteristics but also novel biomarkers reflecting underlying immunological pathway activation.

The quest for personalised therapies that are more efficacious while avoiding unnecessary side-effects has fuelled studies investigating biological pathways in (c) SLE to stratify patients in distinct groups. These studies mostly used single-layer biological techniques (primarily transcriptomics, next to genetics and flow cytometry), which could differentiate patients from healthy controls (HCs) and were moderately successful in stratifying patients in clinically distinct groups.^{5–7} However, the overlap between patient groups was considerable and the clinical usefulness of the identified biological parameters remains to be determined.

The clinical complexity of cSLE may be impossible to dissect by using single-layer biological pathway analysis. Combining data from multiple biological layers to stratify patients with (c)SLE in groups with similar biological pathway activation has been done and is a promising new option to improve patient stratification.^{8–11} Type I interferons (IFNS) play a pivotal role in the pathogenesis of SLE and blocking their activity results in improvement of disease activity.^{12,13} Additionally, other pro-inflammatory and anti-inflammatory cytokines, like interleukin (IL)-6, IL-10, IL-8, IL-17 and chemokine (C-C motif) ligand 2 (CCL2), have been associated with SLE disease activity and proposed as potential biomarkers.^{14–16} Moreover, biological agents targeting cytokines and their receptors have shown their value for the treatment of SLE, thus underlining the contribution of these proteins to disease pathogenesis.^{13,17} Hence, combining transcriptomic with proteomic (eg, cytokine) data would be a promising option. Unsupervised computational models integrate high-dimensional data from multiple ‘-omes’ (eg, transcriptome and proteome) and are a method to find combinations of biological markers that stratify patients into homogeneous subgroups.

Here we combined targeted transcriptomic and proteomic data in an unsupervised hierarchical clustering method, agnostic to disease characteristics, to stratify patients into similar biological phenotypes. Moreover, we identified unique immune cell subsets that characterise these clusters. Lastly, we studied the relation of these biological phenotypes to the patient’s clinical characteristics.

METHODS

Study population

In this exploratory study, two cohorts of patients with cSLE were studied. These patients were preselected in three equally sized groups based on disease activity state

(treatment naïve at diagnosis, Low Lupus Disease Activity State (LLDAS), disease flare) (table 1). The discovery cohort consisted of 17 patients with cSLE. Blood samples from 12 of 17 patients with cSLE were collected at a single time point. For 5 of 17 patients, two to three blood samples per patient were used; these were taken at different time points when the patient was in a different disease activity state. The replication cohort consisted of 12 other patients with cSLE with blood samples taken at a single time point (online supplemental table 1). These patients were either in LLDAS or had active disease (clinical SELENA-SLEDAI (Safety of Estrogens in Systemic Lupus Erythematosus National Assessment-Systemic Lupus Erythematosus Disease Activity Index) ≥ 4). All patients fulfilled the Systemic Lupus International Collaborating Clinics classification criteria.¹ Additionally, eight age-matched HCs, without symptoms of underlying viral infections or the use of any medications, were included (online supplemental table 1).

Gene signatures

A selection of genes from four previously described gene modules was determined by reverse transcription-PCR.⁵ These gene modules have been associated with disease features and recently been translated into applicable gene signatures represented by the selected genes.¹⁸ The genes are listed in table 2.

Cytokine measurement

Fifteen cytokines were measured using an ELLA Simple Plex system (Protein simple, San Jose, California, USA) and four subsequent cytokines were measured using a conventional ELISA according to the manufacturer’s protocol (table 2). For further analysis, we used cytokines with a value above the lower limit of quantification in at least 20% of the patients to prevent skewing of the data. Twelve of the 19 cytokines that were measured fulfilled this requirement.

High-dimensional flow cytometry

Peripheral blood mononuclear cells from patients and HCs were stained with a 40-colour antibody panel as previously described¹⁹ and analysed using a five-laser Aurora spectral flow cytometer (Cytek Biosciences, California, USA). The unsupervised and statistical inference portions of the flow cytometry analysis were performed using the OMIQ data analysis software (www.omiq.ai). The unsupervised analysis method based on surface markers without any two-dimensional gating was used.²⁰ The workflow included running flowAI to check for changes in channels over acquisition time, uniform manifold approximation and projection for dimensionality reduction, flowSOM for clustering and edgeR for statistical inference. For the statistical comparisons of abundance, the Flow Cytometry Standard files were subsampled to ensure the same number of events was included per group (either per biological phenotype or disease activity state).

Table 1 Patient characteristics in the discovery cohort

ID	Gender	Age (years)	Ethnicity	Disease duration (years)	Clinical SELENA-SLEDAI	BILAG-2004	Anti-dsDNA (IU/mL)*	Treatment at visit				Disease activity	Biological phenotype		
								Autoantibodies	Treatment naive	HCQ	Prednisone (mg/kg/day)			NSAIDs	DMARD/biologicals state†
Pt1	Female	15.7	White	1.2	12	21	380	Anti-RNP, anti-Sm	No	Yes	0.12	No	Mycophenolate mofetil	Flare	HBP
Pt1	Female	17.6	White	3.2	0	1	53	Anti-RNP, anti-Sm	No	Yes	–	No	Mycophenolate mofetil/belimumab	LLDAS	MBP
Pt1	Female	14.5	White	0	10	19	548	Anti-RNP, anti-Sm	Yes	No	–	No	No	At diagnosis	LBP
Pt2	Female	17.9	White	3	12	12	24	Anti-SSA, anti-Sm	No	Yes	0.66	No	Mycophenolate mofetil	Flare	MBP
Pt2	Female	19	White	4.1	0	0	33	Anti-SSA, anti-Sm	No	Yes	–	No	Mycophenolate mofetil	LLDAS	LBP
Pt3	Female	9.7	Mixed	1.3	14	22	479	–	No	Yes	0.26	No	Mycophenolate mofetil	Flare	MBP
Pt3	Female	12.5	Mixed	4.1	0	0	46	–	No	Yes	0.05	No	Mycophenolate mofetil	LLDAS	LBP
Pt4	Female	18.5	White	2.8	2	12	0.4	Anti-SSA/B	No	Yes	0.15	Yes	Mycophenolate mofetil	Flare	MBP
Pt4	Female	17.7	White	1.9	2	2	1.5	Anti-SSA/B	No	Yes	0.08	Yes	Mycophenolate mofetil/belimumab	LLDAS	MBP
Pt5	Female	17.7	White	3.4	3	20	7.1	–	No	Yes	–	No	Mycophenolate mofetil	Flare	MBP
Pt5	Female	18	White	3.7	0	0	3.8	–	No	Yes	0.06	No	Mycophenolate mofetil	LLDAS	MBP
Pt6	Female	15.4	White	3.2	0	0	11	Anti-RNP	No	Yes	–	No	Mycophenolate mofetil	LLDAS	LBP
Pt7	Female	15	African-American	6	6	28	113	Anti-RNP	No	Yes	–	No	Mycophenolate mofetil	Flare	HBP
Pt8	Female	15.9	Mixed	0	6	22	32	Anti-RNP	No	No	0.33	No	No	At diagnosis	MBP
Pt9	Male	18.5	South-Asian	3.1	0	0	1.1	–	No	Yes	–	No	Mycophenolate mofetil	LLDAS	LBP
Pt10	Female	5.5	White	0.8	0	0	0.5	Anti-SSA	No	Yes	–	No	No	LLDAS	LBP
Pt11	Female	9.8	African-American	0	10	29	354	Anti-SSA	Yes	No	–	No	No	At diagnosis	HBP
Pt12	Female	14.5	Mixed	0	6	13	11	Anti-SSA/B, anti-RNP, anti-Sm	Yes	No	–	No	No	At diagnosis	LBP
Pt13	Female	11.8	Hispanic	0	4	21	211	Anti-RNP, anti-Sm	Yes	No	–	No	No	At diagnosis	HBP
Pt14	Female	14.2	White	0	2	8	1.1	Anti-RNP	Yes	No	–	No	No	At diagnosis	LBP
Pt15	Female	16.6	Asian	0	14	64	33	Anti-SSA/B, anti-RNP, anti-Sm	No	No	0.99	No	No	At diagnosis	MBP
Pt16	Female	17.1	White	0	5	13	2.9	Anti-RNP	Yes	No	–	No	No	At diagnosis	MBP

Continued

Table 1 Continued

ID	Gender	Age (years)	Ethnicity	Disease duration (years)	Clinical SLEDAI	BILAG-2004	Anti-dsDNA (IU/mL)*	Autoantibodies	Treatment naive	Treatment at visit			Disease activity state†	Biological phenotype
										HCQ	Prednisone (mg/kg/day)	NSAIDs		
Pt17	Female	10.3	African-American	0	15	36	35	Anti-SSA/B	Yes	No	-	No	At diagnosis	HBP

Mixed ethnicity: pt3 (Hispanic-White), pt8 (African-American-Hispanic), pt12 (African-American-White).

*Measured with a fluorometric enzyme immunoassay, cut-off is 10 IU/mL.

†At diagnosis refers to patients sampled within 2 weeks from diagnosis.

BILAG, British Isles Lupus Assessment Group; DMARD, disease-modifying antirheumatic drug; HBP, high biological phenotype; LBP, low biological phenotype; NSAIDs, non-steroidal anti-inflammatory drugs; SLEDAI, Safety of Estrogens in Systemic Lupus Erythematosus National Assessment-Systemic Lupus Erythematosus Disease Activity Index.

Statistical analysis

For the unsupervised hierarchical clustering and principal component analysis (PCA), we loaded the patient’s clinical, gene expression and cytokine data into R (V.4.0.4). Only cytokines which were positive in >20% of the patients were used in the analysis. If a cytokine was included in the analysis and a sample had a value below detection limit, this value was set as the lower limit of quantification. Standardised scores (z-scores) were calculated based on the log10-transformed gene expression and cytokine levels. Unsupervised hierarchical clustering was performed using Ward’s Hierarchical Agglomerative Clustering Method (ward.d2) on the Euclidean distances of the z-scores. The optimal number of clusters for both cohorts was assigned with the NbClust (V.1.0.12) package in R. Subsequently, heatmaps were plotted using the R package Complexheatmap (V.2.10.0).²¹ PCA was performed and visualised based on the z-scores using the R function prcomp. The grouping of patients in the different clusters was agnostic to parameters reflecting clinical disease phenotype.

The Kruskal-Wallis test (three groups) was used to analyse comparisons between medians. Levene’s test was used for homogeneity of variance across the groups. Dunn’s multiple comparisons post hoc test was used to correct for multiple testing as indicated in the legend of each figure. Spearman’s correlation coefficient was calculated to assess correlation. Values of p<0.05 were considered statistically significant. Non-significant results are not shown in figures. The false discovery rate (FDR) was used to compare immune cell subsets between the biological phenotypes. A cut-off value of FDR <0.1 was used to indicate significant differences between the populations. Graphpad Prism V.8.0 (Graphpad Software, La Jolla, California, USA) and R statistical software were used for graph design and statistical analysis.

RESULTS

Unsupervised hierarchical clustering identifies three unique biological phenotypes associated with disease activity state

Gene expression and cytokine levels were measured in 23 samples from 17 patients with cSLE. These data were combined and subjected to an unsupervised hierarchical clustering approach agnostic to any clinical or laboratory disease characteristics. This approach identified three clusters, further termed as biological phenotypes (figure 1A). Patients within each cluster had a unique and remarkably similar combination of differentially expressed genes and cytokine profiles. Patients in cluster 1 had a low gene expression and cytokine profile, and were termed low biological phenotype (LBP). Patients in cluster 2 had a high gene expression and high cytokine profile, and were termed high biological phenotype (HBP). Lastly, patients in cluster 3 had a high gene expression and a low cytokine profile and were termed mixed biological phenotype (MBP) (figure 1A and online supplemental figure 1).

Table 2 Overview of genes and cytokines

Modules	Selected genes*	Cytokines measured on ELLA Simple Plex	Cytokines measured with ELISA
<i>M1.2 IFN</i>	IFI44	IL-6	CXCL13†
	IFIT1	TNF α	IFN λ ‡
	IFIT3	IL-8†	IL-17‡
	Ly6e	CCL2†	IL-23‡
	MxA	IL-10	
<i>M5.12 IFN</i>	PSMB9	IFN γ	
	ISG20	IFN α (pan)	
	NCOA7	IFN β †	
	SP140	CXCL10†	
	TAP1	IFN α 2†	
<i>Neutrophil</i>	DEFA4	TGF- β 1†	
	BPI	IL-1 β ‡	
	DEFA3/DEFA1	IL-12‡	
	CEACAM8	IL-2‡	
	CEACAM6	IL-4‡	
<i>Plasma cell</i>	IGJ		
	TXNDC5		

*Details can be found in Wahadat *et al.*¹⁸

†Cytokines which were measured in the replication cohort.

‡Cytokines that were positive in less than 20% of the samples patients and thus were excluded from further analysis.

CCL2, chemokine ligand 2; IFN, interferon; IL, interleukin; TGF- β 1, transforming growth factor beta 1; TNF, tumour necrosis factor.

To confirm that our clustering approach was not skewed by the patients from whom we had multiple samples collected at different time points, we repeated the unsupervised hierarchical clustering with only data from the sample collected at the first time point of each patient (online supplemental figure 2). The clusters did not change, confirming that the biological phenotypes were linked to disease activity.

Next, to assess whether these biological phenotypes were patient or disease activity state dependent, we studied the five patients from whom multiple time points in different disease activity states were available. In three out of five patients, a change in the biological phenotype was directly associated with a change in disease activity state. The biological phenotypes from the other two patients did not change, but stayed in the MBP group, while the disease activity state changed from flare to LLDAS or vice versa. Notably, these two patients were the only patients in LLDAS in the MBP group, and of all patients in LLDAS, they were the only two to have a clinical SELENA-SLEDAI above 0 (figure 1A).

Cluster stability in the study cohort was assessed by resampling the dataset and computing the Jaccard index to the original cluster.²² A stable cluster should yield a Jaccard index of >0.75.²² After 1000 iterations, the LBP, HBP and MBP group showed a Jaccard index of 0.86, 0.84 and 0.91, respectively (figure 1B), indicating the stability of the three clusters. Next, a PCA showed a clear difference between the identified clusters (figure 1C). PC1

and PC2 combined indicated a proportion of variation of 62.64%. These findings underscored the robustness of the identified clusters.

Subsequently, we studied the association between the gene expression and cytokine data to explore the driving mechanisms underlying the identified biological phenotypes. While a strong correlation was observed between the tested genes as well as between the tested cytokines, a much lower correlation was present between these two datasets (figure 1D). Additionally, while the gene expression and cytokine data are discordant in the MBP group, gene expression and cytokine data in the other two groups are concordant (online supplemental figure 1). This supports the additive value of combined analysis of transcriptional and cytokine data in a single analysis and indicates that genes and cytokines have independent roles in the network that underlies the biological phenotypes.

Validation of biological phenotypes in an independent cSLE cohort

In order to validate the biological phenotypes, we applied our clustering approach on a second cohort of patients with cSLE (n=12), who were either in LLDAS or had active disease (clinical SLEDAI >4). From these patients, we had transcriptional data of all targeted genes and serum analysis of 7 of the 12 cytokines studied in the discovery cohort. Therefore, we adjusted the clustering algorithm in our discovery cohort regarding the number of cytokines. All patients had the same biological phenotype as they were

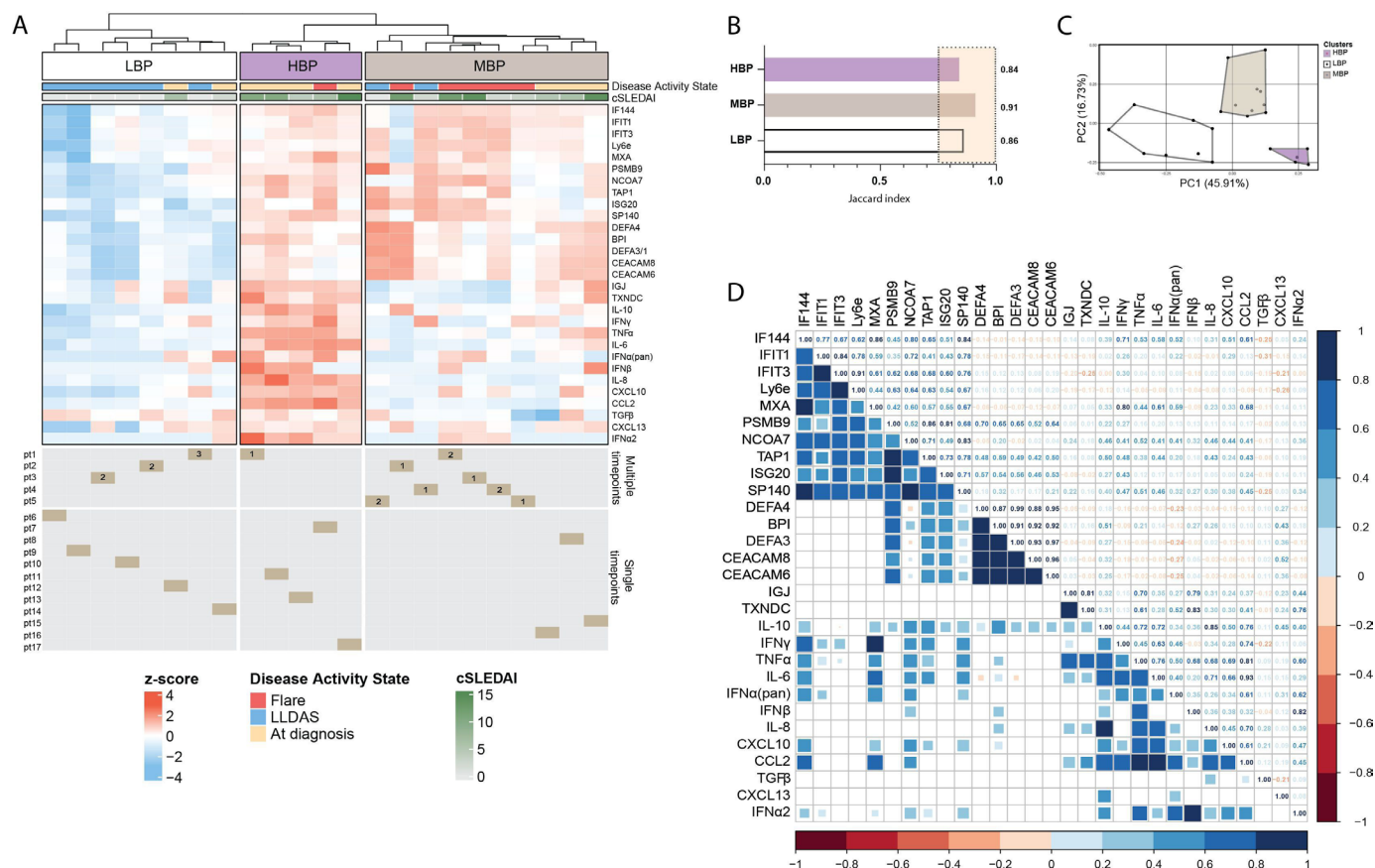


Figure 1 Unsupervised hierarchical clustering identifies three unique biological phenotypes associated with disease activity state. (A) Unsupervised hierarchical clustering using Ward's agglomerative method and passing the Euclidean distance between samples and using row-based log-transformed z-scores identified three clusters. Visit numbers per patient are depicted at the bottom. (B) Jaccard index of the three identified clusters after k-means bootstrapping (measures how many times cells consistently clustered together). Orange dashed bar indicates the best range for cluster stability. (C) Principal component analysis showing that the three clusters lie apart from each other. The first two components with their percentage of variance are shown in parentheses. (D) Correlation plot depicting the inter/intracorrelation between genes and cytokines (N=23). Blue indicates a positive and red indicates a negative correlation. Numbers indicate Spearman's r, squares indicated a significant correlation ($p < 0.05$). Red-blue colour indicates the z-scores. HBP, high biological phenotype; LBP, low biological phenotype; LLDAS, Low Lupus Disease Activity State; MBP, mixed biological phenotype; SELENA-SLEDAI, Safety of Estrogens in Systemic Lupus Erythematosus National Assessment-Systemic Lupus Erythematosus Disease Activity Index.

initially assigned to (online supplemental figure 3A). To find out whether a specific cytokine may be dominant in the cluster analysis, we reanalysed the data by removing one cytokine at a time to see the effect on the clusters. In this analysis, any cytokine could be removed except for IL-8 and CCL2, which seem to be the main drivers of these clusters (data not shown). Applying our clustering method to the replication cohort showed a clear distinction between the patients in LLDAS and with active disease (online supplemental figure 3B). PCA, representing the gene expression and cytokine profiles, indicated that measurements in the replication cohort showed good agreement with those in the discovery cohort (online supplemental figure 3C). Together, these data confirmed the validity and robustness of the study outcome, showing identical clustering despite the relative low patient numbers in each cohort.

Unsupervised hierarchical clustering groups HCs within the low biological phenotype cluster of patients with cSLE

When data from HCs were added to our clustering approach, patients with cSLE in the discovery cohort remained in the same cluster (online supplemental figure 4A). Remarkably, all HCs grouped together with patients with cSLE in the LBP cluster and PCA showed a clear overlap between these two groups (online supplemental figure 4A,B). PCA of the combination of HCs and patients in the LBP group from the discovery and replication cohort also showed overlap (online supplemental figure 4C). Patients with cSLE, who clustered together with HCs, were primarily patients in LLDAS who were on medication, with the majority treated with both hydroxychloroquine (HCQ) and mycophenolate mofetil (MMF). Only one of these patients with cSLE was on low-dose prednisone (0.05 mg/kg/day) (table 1). Interestingly, all patients in LBP with LLDAS without prednisone remained in LLDAS for at least 1 year after stopping prednisone.

These data illustrate that the biological phenotype of HCs with the used biological markers was not different from patients with cSLE with low disease activity in this explorative study.

Biological phenotypes reflect disease activity states

Most patients in the LBP group were in LLDAS (six of eight). The other two patients were included at diagnosis, with only involvement of the mucocutaneous domain next to serological activity (SELENA-SLEDAI 2 and 6). The median clinical SELENA-SLEDAI in this cluster was 0 (0–6) (figure 2A). The HBP group contained patients at diagnosis and one patient with a flare. The median clinical SELENA-SLEDAI in this cluster was 10 (4–15; $p=0.011$ (Kruskal-Wallis test with Dunn's multiple comparisons post hoc)) (figure 2A). Lastly, for patients in the MBP group, which concerns a mix of patients in LLDAS, disease flare or at diagnosis, the median clinical SELENA-SLEDAI was 5.5 (0–14; $p=0.026$ (Kruskal-Wallis test with Dunn's multiple comparisons post hoc)) (figure 2A). When disease duration was compared between the three clusters, patients in the HBP group had median disease duration of 0 (0–6.0) years, while patients in the MBP and LBP cluster had median disease duration of 1.6 (0–3.7) and 3.2 (0–4.1) years, respectively (figure 2B). This could suggest that disease duration could be the driving factor for the biological phenotypes. However, this is contradicted by the finding that two patients who were at diagnosis were also found in the LBP group and patients with a disease flare were in the HBP group. When the days in LLDAS were compared between LLDAS patients in the LBP and MBP group, we found that LLDAS patients in the LBP group were already in LLDAS for a much longer period of time (median 715 days) when compared with the two LLDAS patients in the MBP group (median time in LLDAS 198 days) (figure 2B). Together, these data underline that not only disease duration but particularly (biological) disease activity might be the driving factor for the biological disease phenotypes.

Medication use does not explain the biological phenotypes identified by hierarchical clustering

When comparing medication use of the patients in the different clusters, HCQ and MMF usage were similar in the LBP and MBP groups, yet patients clustered separately (figure 2B). Notably, four out of five patients in the HBP group were treatment naïve at blood sampling (figure 2B). Interestingly, almost all (8 of 10) patients in the MBP group were treated with prednisone, where one out of eight and zero out of five were treated with prednisone in the LBP and HBP groups, respectively (figure 2B). In the MBP group, no correlation was found between prednisone dosage and gene expressions or cytokine levels (online supplemental figure 5 and table 1) suggesting that medication use does not explain the biological phenotypes identified by our clustering approach.

Biological phenotypes are not driven by organ system involvement or currently used laboratory parameters for SLE disease activity

To assess whether biological phenotypes reflect specific organ system involvement, we studied the organ systems that were ever affected in each patient. No difference was observed between the number of organ systems ever involved between the three clusters (figure 2C). Studying the autoantibody profiles of the patients in the clusters, no difference between patients with cSLE in the different clusters was found (table 1). Lastly, we studied whether routine laboratory parameters to measure SLE disease activity were different between the three biological phenotypes (figure 2D). Not surprisingly, anti-dsDNA, C reactive protein, erythrocyte sedimentation rate and IgG levels were highest in the HBP group, the group with newly diagnosed patients (figure 2D). However, C3, C4, thrombocyte, lymphocyte and neutrophil counts did not differ between any of the groups and importantly could not differentiate between patients in LBP and MBP. These data indicate that the identified biological phenotypes are not primarily driven by organ system involvement, and currently used laboratory parameters to measure cSLE disease activity are of limited influence on biological disease phenotype (LBP, MBP) during follow-up of the patient.

High-dimensional flow cytometry identifies immune cell subsets that differ between biological phenotypes

To investigate the cellular landscape that characterised the three identified biological phenotypes, we applied 40-colour spectral flow cytometry (figure 3A). FlowSOM analysis identified 32 distinct immune cell subsets (online supplemental figure 6). In line with previous findings,⁶ there was a clear difference in the immune cell subsets from HCs compared with the patients. In total, 14 of 32 populations differed significantly between these two groups, with central memory CD8⁺ T cells being the cell population that showed the highest fold difference (figure 3B,C and online supplemental table 2).

When we compared the difference between the biological phenotypes in patients with cSLE, the HBP group clearly differed from the other two biological phenotypes: 16 of 32 of the identified immune cell subsets differed between the HBP and LBP groups (figure 3D and online supplemental table 3), and 10 of 32 populations differed between the HBP and MBP groups (figure 3E,F and online supplemental table 4). When comparing the LBP and MBP groups, only 2 of 32 populations differed (figure 3F and online supplemental table 5). Of interest, CD11c⁺ CD16[−] dendritic cells (DCs) were increased in the LBP group, while a unique CD11c⁺ B cell population was increased in the MBP and HBP group (figure 3F). To assess associations with the biological phenotypes, we studied the correlation between the 32 immune cell subsets and the clinical SELENA-SLEDAI. Nine immune cell populations were significantly associated with the clinical SELENA-SLEDAI (online supplemental figure 7). Of these, IgG+

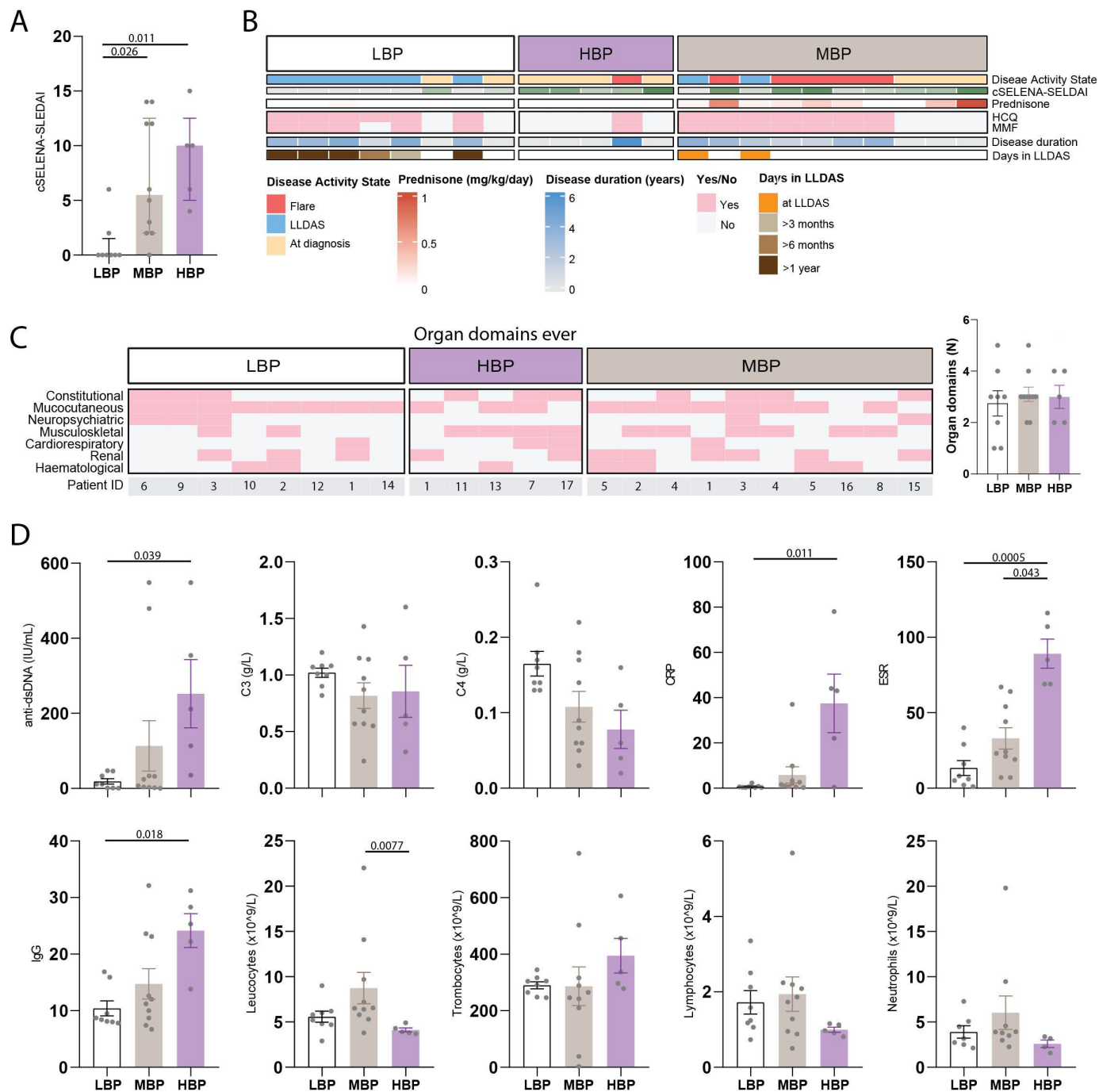


Figure 2 Biological phenotypes are associated with disease states and not influenced by the use of medication. (A) Clinical SELENA-SLEDAI score of patients with different biological phenotypes. Each dot represents individual patients. (B) Heatmap indicating the disease activity state, clinical SELENA-SLEDAI, prednisone usage in mg/kg/day, disease duration in years, days in LLDAS, hydroxychloroquine (HCQ) and mycophenolate mofetil (MMF) use. (C) Heatmap indicating organ domain involvement ever. Bar plot indicates the number of organ domains involved per patient in each cluster. (D) Bar plots indicating 10 laboratory parameters in patients with different biological immune phenotypes. Kruskal-Wallis test with Dunn's multiple comparisons post hoc test was used for relation between the clusters. CRP, C reactive protein; ESR, erythrocyte sedimentation rate; HBP, high biological phenotype; LBP, low biological phenotype; LLDAS, Low Lupus Disease Activity State; MBP, mixed biological phenotype; SELENA-SLEDAI, Safety of Estrogens in Systemic Lupus Erythematosus National Assessment-Systemic Lupus Erythematosus Disease Activity Index.

memory B cells and the central memory CD8⁺ T cells did not show any significant difference between the biological phenotypes (online supplemental figure 7). The seven other populations significantly differed between

the identified biological phenotypes (figure 3G). Here, CD11c⁺ B cells, plasmablasts and early effector CD4⁺ T cells were increased in the HBP group compared with the LBP group. On the other hand, conventional DCs,

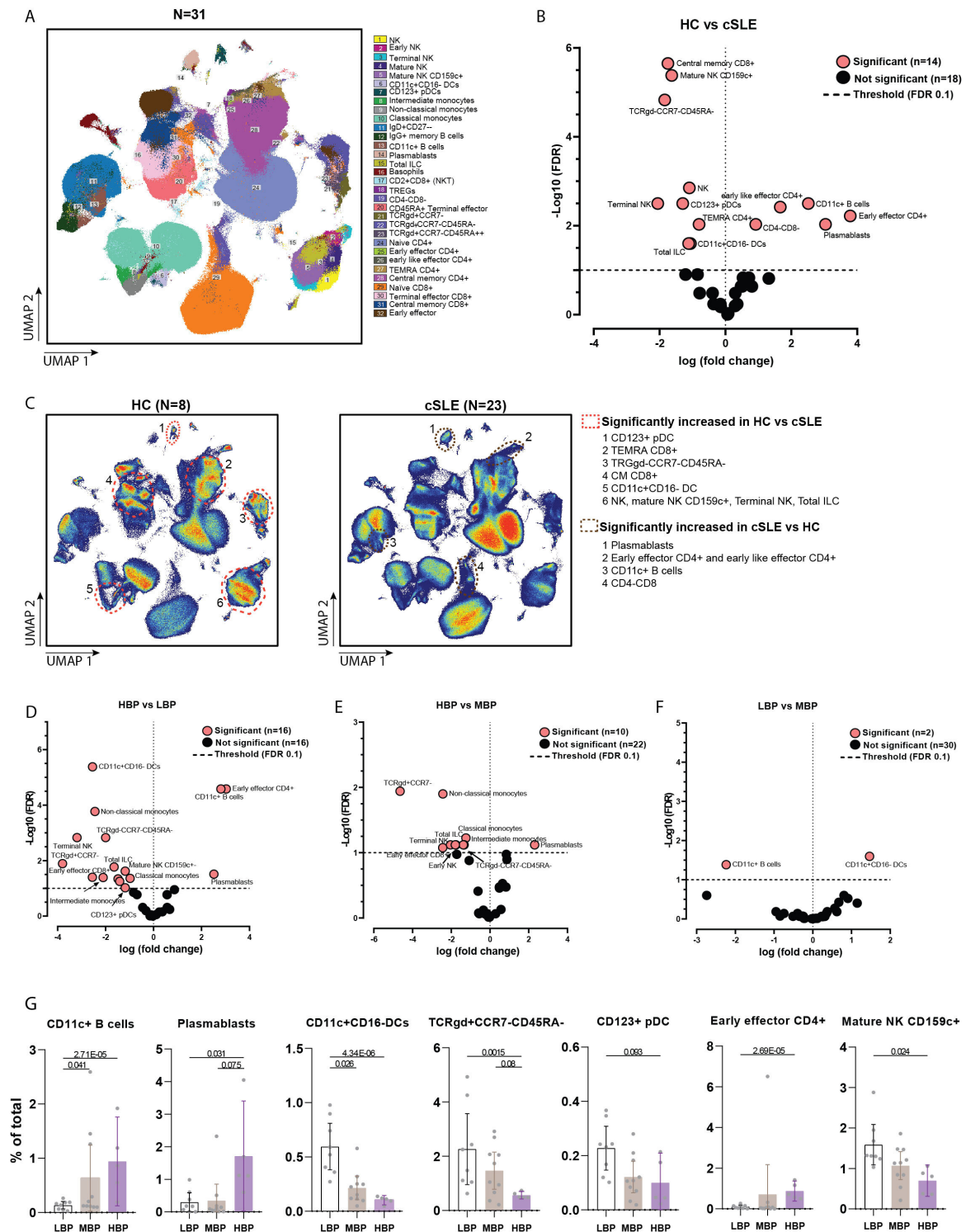


Figure 3 High-dimensional flow cytometry identifies immune subsets that differ between biological phenotypes. (A) Population cluster identification in high-dimensional 40-colour flow cytometry data using uniform manifold approximation and projection (UMAP) dimensionality reduction. Visualisation is performed on combined patients with childhood-onset (cSLE) and healthy control (HCs) data (n=31). (B) Volcano plot comparing patients/HCs. Positive fold changes indicate cellular populations increased in patients, while negative fold changes indicate decreases in patients. (C) UMAP plots of HCs (n=8) and all patients with cSLE (n=23). Populations that are increased in each group are indicated by a dashed line. (D) Volcano plot comparing HBP (N=5) versus LBP (N=8). Positive fold changes indicate cellular populations increased in HBP compared with LBP. (E) Volcano plot comparing HBP (N=5) versus MBP (N=10). Positive fold changes indicate cellular populations increased in HBP compared with MBP. (F) Volcano plot comparing LBP (N=8) versus MBP (N=10). Positive fold changes indicate cellular populations increased in MBP compared with LBP. (G) Bar plots depict the percentage of cells in total population. The FDR is depicted as statistical analysis. DCs, dendritic cells; FDR, false discovery rate; HBP, high biological phenotype; LBP, low biological phenotype; MBP, mixed biological phenotype; NK, natural killer; pDCs, plasmacytoid DCs.

TCR $\gamma\delta$ + CCR7- CD45RA-, plasmacytoid DCs and mature natural killer CD159+ cells were decreased in the HBP group compared with the LBP group (figure 3G).

Taken together, multiple immune cell subsets differed significantly between the biological phenotypes and were associated with disease activity. This cellular analysis underscores the connection between immune cell subsets and targeted transcriptomic/proteomic data, indicating the additional relevance of these cells in defining biological phenotypes.

DISCUSSION

Our study illustrates that combining targeted transcriptomic and proteomic data is a reliable method to cluster patients in similar clinical and biological phenotypes that represent a specific disease activity state rather than organ system involvement. Moreover, high-dimensional flow cytometry revealed immune cell subsets that were characteristic for these phenotypes. Thus, these biological phenotypes can facilitate patient stratification for optimal treatment choices and provide information to improve interventional clinical trial design.

Transcriptomic studies in SLE have been able to discriminate between patients with low and high disease activity. Yet, gene signatures are often expressed in binary values and remain relatively stable over time, and therefore are not sensitive to changes that occur during clinical follow-up of patients.²³ Multiomics have led to groundbreaking findings in the field of oncology indicating their usefulness for other fields like autoimmunity.²⁴ One study, in patients with adult-onset SLE, combined different gene modules and various plasma soluble mediators and identified seven unique molecular profiles.¹¹ However, these profiles did not distinguish between patients with different disease activity scores. Furthermore, this study included multiple gene modules from which the majority had no relation with SLE disease activity.⁵ In contrast, we chose a targeted approach and selected genes and cytokines that previously were associated with disease activity^{5 18} expecting a higher clinical relevance from these genes and cytokines. Indeed, this approach enabled a more precise stratification of patients in groups with similar biological phenotypes that are related to disease activity state.

For the first time, we were able to link almost all patients with cSLE in LLDAS to a specific biological disease phenotype, namely the LBP group. Even more, HCs clustered with these LLDAS patients in the LBP group. This shows that patients on medication can achieve comparable combinations of gene expression and cytokine profiles as HCs and additionally increases our understanding of LLDAS being associated with reduced risk of flares, less damage development and corticosteroid sparing.^{25 26} Hence, we speculate that for more accurate characterising of LLDAS, analysis of these novel biological parameters could be used. This concept is in line with findings in patients with ulcerative colitis in which histological

remission at the gut level, in addition to clinical remission, is associated with corticoid sparing and less hospitalisation for disease flares.²⁷

Medication usage can be an important confounding factor in analysing biological profiles in patients. For example, high-dose corticosteroids have been reported to have an inhibitory effect on multiple proinflammatory cytokines and the type I IFN signature.^{28 29} In our study, patients who used prednisone (primarily in the MBP group) still had high levels of type I IFN-induced genes and prednisone dose did not correlate with gene expression or protein levels. Moreover, between the different biological phenotypes, no major differences in HCQ and MMF usage were observed. Remarkably, almost all patients in the LBP group, who have low gene and cytokine profiles, were in LLDAS and used HCQ and disease-modifying antirheumatic drugs but not prednisone. Interestingly, these patients remained in LLDAS for at least 1 year after stopping prednisone. This underscores that continued prednisone use is not needed to remain in LLDAS.

As expected, pro-inflammatory cytokines are upregulated in the HBP group compared with the LBP group. An interesting finding is the upregulation of transforming growth factor beta (TGF- β) in HCs, patients in the LBP and also in patients in the HBP group. This latter finding is somewhat contradictory to literature showing that low concentrations of TGF- β are associated with high disease activity.³⁰ However, we also know that TGF- β can have dual roles with the activation of resting monocytes but also the inhibition of activated macrophages.^{31 32} This is in line with the theory that often the effects of TGF- β are context dependent and more research is needed to understand the biological role of TGF- β .³³ Additionally, when analysing whether one specific cytokine was dominant in driving the clusters, we found that IL-8 and CCL2 were the main drivers of the clusters in our explorative study. Although these findings are suggesting that the analyte number could be reduced, additional studies with larger datasets should first confirm this.

In SLE, gene signatures and cytokines have been associated with involvement of certain organ systems.^{23 34 35} Currently, organ system involvement is seen as an important trait that guides treatment strategies. Here we show that biological phenotypes are able to change and are more a reflection of the patients' disease activity state than that of the specific organ system(s) involved. This supports a new concept where choice of treatment and tapering strategies are not solely based on clinical phenotype but includes measuring these novel biological parameters. This concept is supported by our recent work showing that patients with cSLE, irrespective of organ system involvement, reached LLDAS in a median time of 6 months. This was achieved by a treatment strategy focusing on limited use of corticosteroids and early introduction of immunosuppressives, specifically when corticosteroids could not be tapered within 3 months.³⁶

High-dimensional flow cytometry confirmed previous findings that immune cell subsets differ between HCs and patients with SLE.¹⁰ A caveat in the previous study in cSLE that focused on immune cell subset differences is the high number of patients with low disease activity (SLEDAI ≤ 4).⁶ In our study, we preselected patients in equally sized groups with different disease activity states to overcome this caveat. Interestingly, distinct immune cell populations differed between the patients with different biological phenotypes. This indeed shows the interconnection between the biological phenotypes based on a combination of transcriptional and cytokine data and immune cells in peripheral blood, which hints towards the idea that these immune cells could also be used to reflect the identified biological phenotypes. In the ideal setting, measuring a small number of cell populations could predict a specific clinical disease course and guide treatment strategy, as has been implemented in oncology.³⁷ Possible candidates for this purpose are the CD11c+ B cell and CD11c+ CD16-DC populations identified in our study, as these cells were able to discriminate between patients in the LBP and MBP group, which may be a difficult task based on clinical characteristics. The CD11c+ B cell population has previously been associated with disease activity and proposed to be driven by an extrafollicular B cell activation route in SLE.^{38,39} Likewise, decreased blood levels of the CD11c+ CD16-DC population have been associated with active disease in SLE.⁴⁰

Better understanding of biological pathways underlying active disease, LLDAS and remission will facilitate tailored treatment strategies and improve patient outcome. Children would benefit the most as they have a more severe disease than adults with SLE.³ Multiomic data have shown their value in adult-onset SLE where the first steps are made towards stratification of patients to guide treatment choices.⁴¹ Our explorative study in children with SLE for the first time links disease activity states to biological phenotypes. We identified three clusters, in which the MBP group is the most interesting group, as from a clinical point of view, we would not have clustered these patients together. The next step is to extend patient numbers and perform machine learning techniques on combined targeted gene expression, cytokine, flow cytometry and clinical data to facilitate further understanding of underlying biological mechanisms and make a start towards stratification of patients to guide treatment choices in cSLE.

Our study has limitations. First, we have a low number of patients in our analysis. Expanding this study in a larger set of patients with different ethnic backgrounds will be our next step. The validation of the results by using a PCA, bootstrapping and a replication cohort makes us confident that expanding the cohort will produce similar results. Second, some of the analyses performed are based on multiple time points from the same individuals. Although we performed extra analyses that suggest that any underlying correlations from repeated samples from the same individual did not appear to substantively affect

the clustering, we recognise that repeated observations from the same person are correlated. Third, we have only used a combination of (targeted) transcriptomics and proteomics, while there are several other omic techniques that may also reflect the biological pathways that play a role in the pathogenesis of SLE.⁴²

In conclusion, here we show that integrating targeted transcriptomic and proteomic data in an unsupervised hierarchical clustering model leads to the identification of biological phenotypes that reliably distinguish patients with cSLE with different disease activity states, which could not be achieved by currently used clinical characteristics. This is a first step towards developing personalised treatment strategies in children with lupus based on clinical phenotype and new biomarkers. Future studies should translate and simplify these tools by reducing the number of analytes that need to be measured, which will facilitate implementation in daily clinical practice.

Acknowledgements The authors thank all the patients with cSLE and HCs for taking part in this study.

Contributors All authors were involved in drafting the article or revising it critically for important intellectual content, and all authors approved the final version to be published. Study conception and design—MJW, YMM, PDK, SK and MAV. Experimental work—MJW, SJvT, YMM, CGVH-M, HdW, AWL, PDK and MAV. Acquisition of clinical data—MJW, AWL, MJG, AM, MV and SK. Analysis and interpretation—MJW, SJvT, YMM, PDK, MAV and SK. Writing (draft manuscript)—MJW, MAV and SK. Revision of the manuscript—MJW, SJvT, YMM, HdW, CGVH-M, AWL, MJG, AM, MV, PDK, MAV and SK. MAV and SK are responsible for the overall content as the guarantors.

Funding This work was made possible by the support of the Sophia Children's Hospital Fund (B18-04), NVLE (Dutch patient organization for Lupus, APS, Scleroderma and MCTD) (BP12-1-261) and Dutch Arthritis Society (C0-19-001). The research was performed within the framework of the Erasmus Postgraduate School Molecular Medicine.

Competing interests None declared.

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

Patient consent for publication Not required.

Ethics approval This study involves human participants and was approved by the Medical Ethics Review Committee of the Erasmus University Medical Center, Rotterdam, the Netherlands (MEC-2019-0412/MEC-2005-0137). Written informed consent was obtained from all participants in compliance with the Declaration of Helsinki.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Data are available upon reasonable request.

De-identified data including R-scripts underlying this article are available to all scientists upon request.

Supplemental material This content has been supplied by the author(s). It has not been vetted by BMJ Publishing Group Limited (BMJ) and may not have been peer-reviewed. Any opinions or recommendations discussed are solely those of the author(s) and are not endorsed by BMJ. BMJ disclaims all liability and responsibility arising from any reliance placed on the content. Where the content includes any translated material, BMJ does not warrant the accuracy and reliability of the translations (including but not limited to local regulations, clinical guidelines, terminology, drug names and drug dosages), and is not responsible for any error and/or omissions arising from translation and adaptation or otherwise.

Open access This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>.

ORCID iDs

Mohamed Javad Wahadat <http://orcid.org/0000-0003-3961-1274>

Marjan A Versnel <http://orcid.org/0000-0003-0245-5386>

Sylvia Kamphuis <http://orcid.org/0000-0002-1964-352X>

REFERENCES

- Smith EMD, Lythgoe H, Midgley A, et al. Juvenile-onset systemic lupus erythematosus: update on clinical presentation, pathophysiology and treatment options. *Clin Immunol* 2019;209:108274.
- Groot N, Shaikhani D, Teng YKO, et al. Long-Term clinical outcomes in a cohort of adults with childhood-onset systemic lupus erythematosus. *Arthritis Rheumatol* 2019;71:290–301.
- Kamphuis S, Silverman ED. Prevalence and burden of pediatric-onset systemic lupus erythematosus. *Nat Rev Rheumatol* 2010;6:538–46.
- Groot N, Kardolus A, Bijl M, et al. Effects of childhood-onset systemic lupus erythematosus on academic achievements and employment in adult life. *J Rheumatol* 2021;48:915–23.
- Banchereau R, Hong S, Cantarel B, et al. Personalized immunomonitoring uncovers molecular networks that stratify lupus patients. *Cell* 2016;165:1548–50.
- Robinson GA, Peng J, Dönnnes P, et al. Disease-Associated and patient-specific immune cell signatures in juvenile-onset systemic lupus erythematosus: patient stratification using a machine-learning approach. *Lancet Rheumatol* 2020;2:e485–96.
- Panousis NI, Bertsias GK, Ongen H, et al. Combined genetic and transcriptome analysis of patients with SLE: distinct, targetable signatures for susceptibility and severity. *Ann Rheum Dis* 2019;78:1079–89.
- Huang X, Luu LDW, Jia N, et al. Multi-platform omics analysis reveals molecular signatures for pathogenesis and activity of systemic lupus erythematosus. *Front Immunol* 2022;13:833699.
- Trzupek D, Lee M, Hamey F, et al. Single-cell multi-omics analysis reveals IFN-driven alterations in T lymphocytes and natural killer cells in systemic lupus erythematosus. *Wellcome Open Res* 2021;6:149.
- Nehar-Belaid D, Hong S, Marches R, et al. Mapping systemic lupus erythematosus heterogeneity at the single-cell level. *Nat Immunol* 2020;21:1094–106.
- Guthridge JM, Lu R, Tran L-H, et al. Adults with systemic lupus exhibit distinct molecular phenotypes in a cross-sectional study. *EClinicalMedicine* 2020;20:100291.
- Rönnblom L, Leonard D. Interferon pathway in SLE: one key to unlocking the mystery of the disease. *Lupus Sci Med* 2019;6:e000270.
- Farie R, Khamashta M, Merrill JT, et al. Anifrolumab, an anti-interferon- α receptor monoclonal antibody, in moderate-to-severe systemic lupus erythematosus. *Arthritis Rheumatol* 2017;69:376–86.
- Wong CK, Lit LCW, Tam LS, et al. Hyperproduction of IL-23 and IL-17 in patients with systemic lupus erythematosus: implications for TH17-mediated inflammation in auto-immunity. *Clin Immunol* 2008;127:385–93.
- Godsell J, Rudloff I, Kandane-Rathnayake R, et al. Clinical associations of IL-10 and IL-37 in systemic lupus erythematosus. *Sci Rep* 2016;6:34604.
- Ruchakorn N, Ngamjanyaporn P, Suangtamai T, et al. Performance of cytokine models in predicting SLE activity. *Arthritis Res Ther* 2019;21:287.
- Farie R, Petri M, Zamani O, et al. A phase III, randomized, placebo-controlled study of belimumab, a monoclonal antibody that inhibits B lymphocyte stimulator, in patients with systemic lupus erythematosus. *Arthritis Rheum* 2011;63:3918–30.
- Wahadat MJ, Schonenberg-Meinema D, van Helden-Meeuwssen CG, et al. Gene signature fingerprints stratify SLE patients in groups with similar biological disease profiles: a multicentre longitudinal study. *Rheumatology (Oxford)* 2022;61:4344–54.
- Mueller YM, Schrama TJ, Ruijten R, et al. Stratification of hospitalized COVID-19 patients into clinical severity progression groups by immuno-phenotyping and machine learning. *Nat Commun* 2022;13:915.
- Park LM, Lannigan J, Jaimes MC. OMIP-069: forty-color full spectrum flow cytometry panel for deep immunophenotyping of major cell subsets in human peripheral blood. *Cytometry A* 2020;97:1044–51.
- Charrad M. NbClust: an R package for determining the relevant number of clusters in a data set. *J Stat Softw* 2014;61:1–36.
- Hennig C. Cluster-wise assessment of cluster stability. *Computational Statistics & Data Analysis* 2007;52:258–71.
- Petri M, Fu W, Ranger A, et al. Association between changes in gene signatures expression and disease activity among patients with systemic lupus erythematosus. *BMC Med Genomics* 2019;12:4.
- Heo YJ, Hwa C, Lee GH, et al. Integrative multi-omics approaches in cancer research: from biological networks to clinical subtypes. *Mol Cells* 2021;44:433–43.
- Fanourakis A, Adamichou C, Koutsovi S, et al. Low disease activity-irrespective of serologic status at baseline-associated with reduction of corticosteroid dose and number of flares in patients with systemic lupus erythematosus treated with belimumab: a real-life observational study. *Semin Arthritis Rheum* 2018;48:467–74.
- Smith EMD, Gorst SL, Al-Abadi E, et al. "it is good to have a target in mind": qualitative views of patients and parents informing a treat to target clinical trial in juvenile-onset systemic lupus erythematosus. *Rheumatology (Oxford)* 2021;60:5630–41.
- Bryant RV, Burger DC, Delo J, et al. Beyond endoscopic mucosal healing in UC: histological remission better predicts corticosteroid use and hospitalisation over 6 years of follow-up. *Gut* 2016;65:408–14.
- de Kruid MD, Lemaire LC, Giebelen IA, et al. Prednisolone dose-dependently influences inflammation and coagulation during human endotoxemia. *J Immunol* 2007;178:1845–51.
- Pascual V, Chaussabel D, Banchereau J. A genomic approach to human autoimmune diseases. *Annu Rev Immunol* 2010;28:535–71.
- Becker-Merok A, Eilertsen GØ, Nossent JC. Levels of transforming growth factor-beta are low in systemic lupus erythematosus patients with active disease. *J Rheumatol* 2010;37:2039–45.
- Kubiczkova L, Sedlarikova L, Hajek R, et al. Tgf-B-an excellent servant but a bad master. *J Transl Med* 2012;10:183.
- Smythies LE, Sellers M, Clements RH, et al. Human intestinal macrophages display profound inflammatory anergy despite avid phagocytic and bacteriocidal activity. *J Clin Invest* 2005;115:66–75.
- Wahl SM. Transforming growth factor-beta: innately bipolar. *Curr Opin Immunol* 2007;19:55–62.
- Villanueva E, Yalavarthi S, Berthier CC, et al. Netting neutrophils induce endothelial damage, infiltrate tissues, and expose immunostimulatory molecules in systemic lupus erythematosus. *J Immunol* 2011;187:538–52.
- Oke V, Gunnarsson I, Dorschner J, et al. High levels of circulating interferons type I, type II and type III associate with distinct clinical features of active systemic lupus erythematosus. *Arthritis Res Ther* 2019;21:107.
- Wahadat MJ, van den Berg L, Timmermans D, et al. LLDAS is an attainable treat-to-target goal in childhood-onset SLE. *Lupus Sci Med* 2021;8:e000571.
- van Dongen JJM, Lhermitte L, Böttcher S, et al. EuroFlow antibody panels for standardized N-dimensional flow cytometric immunophenotyping of normal, reactive and malignant leukocytes. *Leukemia* 2012;26:1908–75.
- Wu C, Fu Q, Guo Q, et al. Lupus-Associated atypical memory B cells are mtorc1-hyperactivated and functionally dysregulated. *Ann Rheum Dis* 2019;78:1090–100.
- Rincon-Arevalo H, Wiedemann A, Stefanski A-L, et al. Deep phenotyping of cd11c+ B cells in systemic autoimmunity and controls. *Front Immunol* 2021;12:635615.
- Khan SA, Nowatzky J, Jiménez-Branda S, et al. Active systemic lupus erythematosus is associated with decreased blood conventional dendritic cells. *Exp Mol Pathol* 2013;95:121–3.
- Toro-Domínguez D, Martorell-Marugán J, Martínez-Bueno M, et al. Scoring personalized molecular portraits identify systemic lupus erythematosus subtypes and predict individualized drug responses, symptomatology and disease progression. *Brief Bioinform* 2022;23:bbac332.
- Song W, Tang D, Chen D, et al. Advances in applying of multi-omics approaches in the research of systemic lupus erythematosus. *Int Rev Immunol* 2020;39:163–73.

SUPPLEMENTARY FILE

Targeted multi-omics in childhood-onset SLE reveal distinct biological phenotypes associated with disease activity; results from an explorative study
Wahadat et al.

SUPPLEMENTARY METHODS

FIGURES

Supplementary figure S1

Supplementary figure S2

Supplementary figure S3

Supplementary figure S4

Supplementary figure S5

Supplementary figure S6

Supplementary figure S7

TABLES

Supplementary table T1

Supplementary table T2

Supplementary table T3

Supplementary table T4

Supplementary table T5

REFERENCES

Supplementary Methods

Study population details

Demographics, laboratory parameters and clinical characteristics including disease activity parameters and medication usage at previous and subsequent clinical visits from time of study blood sampling were collected. Disease activity was assessed by the clinical Safety of Estrogen in Lupus National Assessment (SELENA)-SLEDAI at each visit (2). Disease flares were indicated by an increase of >3 in the clinical SELENA-SLEDAI from the previous visit and/or an intensification of treatment, respectively. Low Lupus Disease Activity State (LLDAS) was defined as described before (3).

Cytokine measurement details

Serum samples were thawed on ice and centrifuged at 1300 x g for 5 min at room temperature. Twofold dilutions were prepared in low-protein-binding plates according to the manufacturer's instructions. For TGF β 1 detection, samples were first activated with 1 N HCl and then neutralized with 1.2 N NaOH/0.5 M HEPES. Subsequently, these samples were 1:15 diluted and analyzed in the ELLA Simple Plex system (SPCKB-PS-1369/237/934/230/1108, SPCKC-PS-3649, SPCKE-PS-3426).

RT-PCR of genes from four gene modules

Blood was collected in PAXgene RNA tubes and stored in -80°C until use for whole blood RNA purification. RNA was isolated from PAXgene Blood RNA Tubes according to manufacturer's protocol (PreAnalytiX GmbH). cDNA was synthesized using the High-Capacity Reverse Transcription Kit (Applied Biosystems, Bleiswijk, The Netherlands). RT-PCR was performed on a QuantstudioTM 5 Real-Time PCR System using predesigned primer/probe sets (Applied Biosystems). Genes were selected based on four previously described gene modules (4). For calculation of

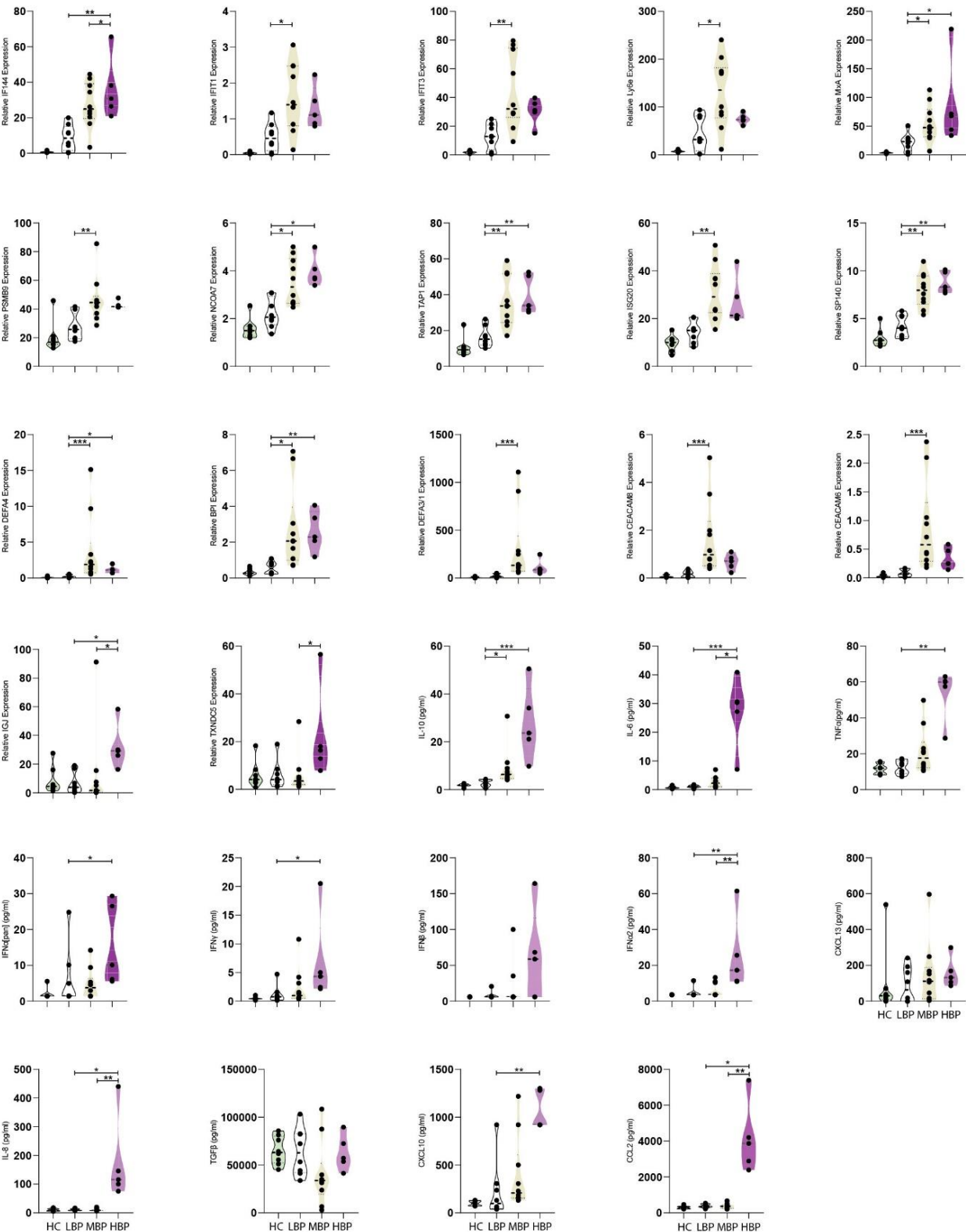
28 relative gene expressions, samples were normalized to expression of the
29 housekeeping gene Abl. Relative expression values were determined from
30 normalized CT values using the $2^{-\Delta CT}$ method (4).

31

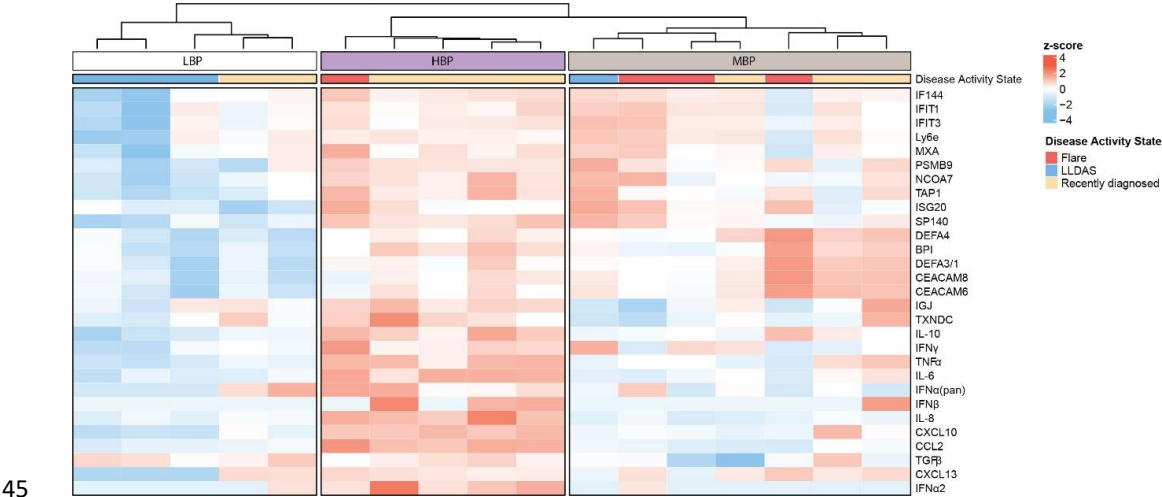
32

33

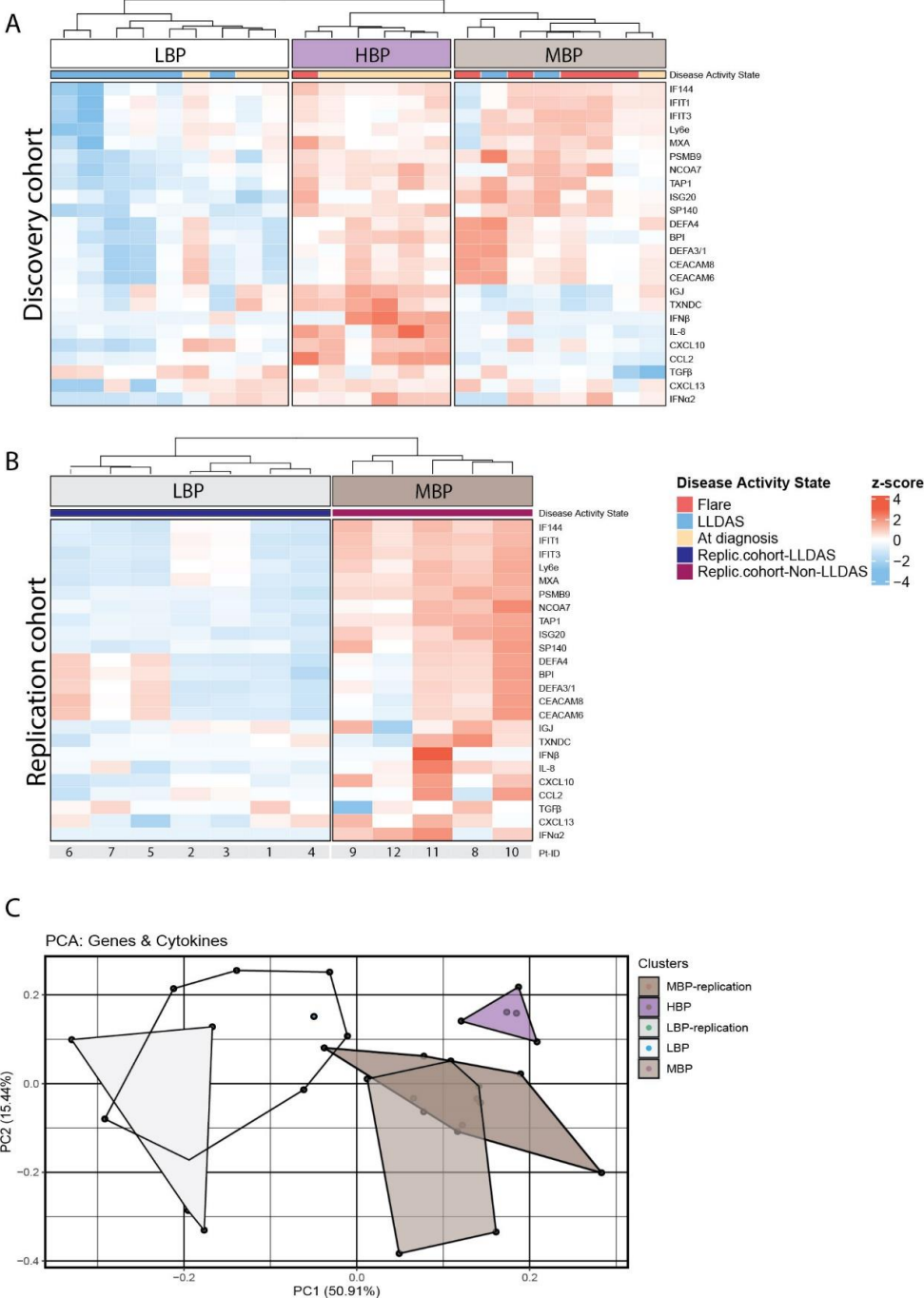
34 **Supplementary figure S1 Biological phenotypes show distinct gene expression**
35 **and cytokine levels.** Violin plots indicating different gene expression ($2^{\Delta\Delta CT}$) and
36 cytokines levels (pg/ml). Each dot represents one patient. Kruskal-Wallis test with
37 Dunn's multiple comparisons posthoc test was used for relation between the clusters.
38 HC, healthy controls, LBP, low biological phenotype; HBP, high biological phenotype;
39 MBP, mixed biological phenotype.



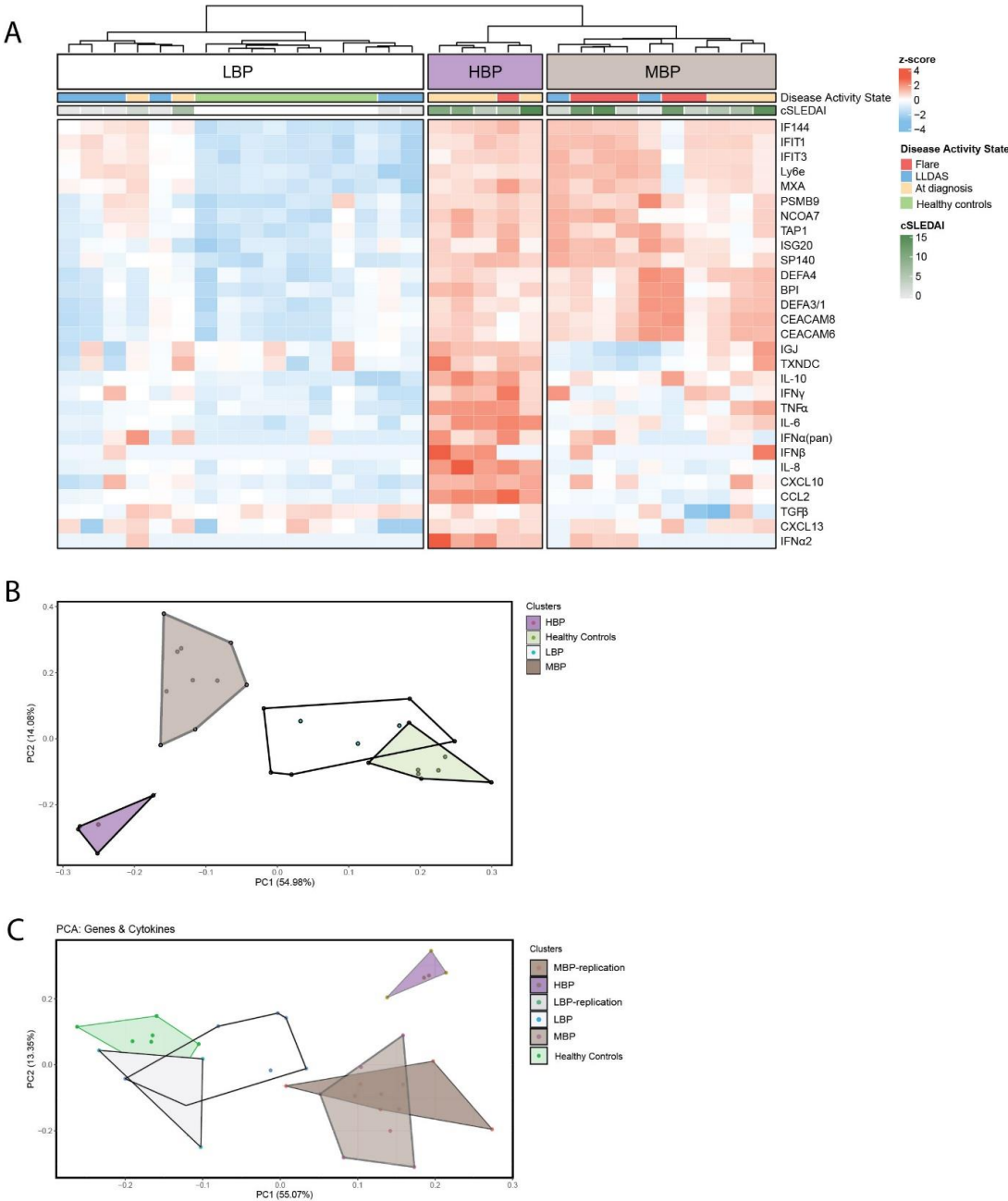
40 **Supplementary figure S2 Unsupervised hierarchical clustering remains**
41 **unchanged when only first visit samples are included.** Unsupervised hierarchical
42 clustering using Ward's agglomerative method, passing the Euclidean distance
43 between samples, using row-based log-transformed z-scores, identified 3 clusters.
44 Top banners indicate the disease activity states.



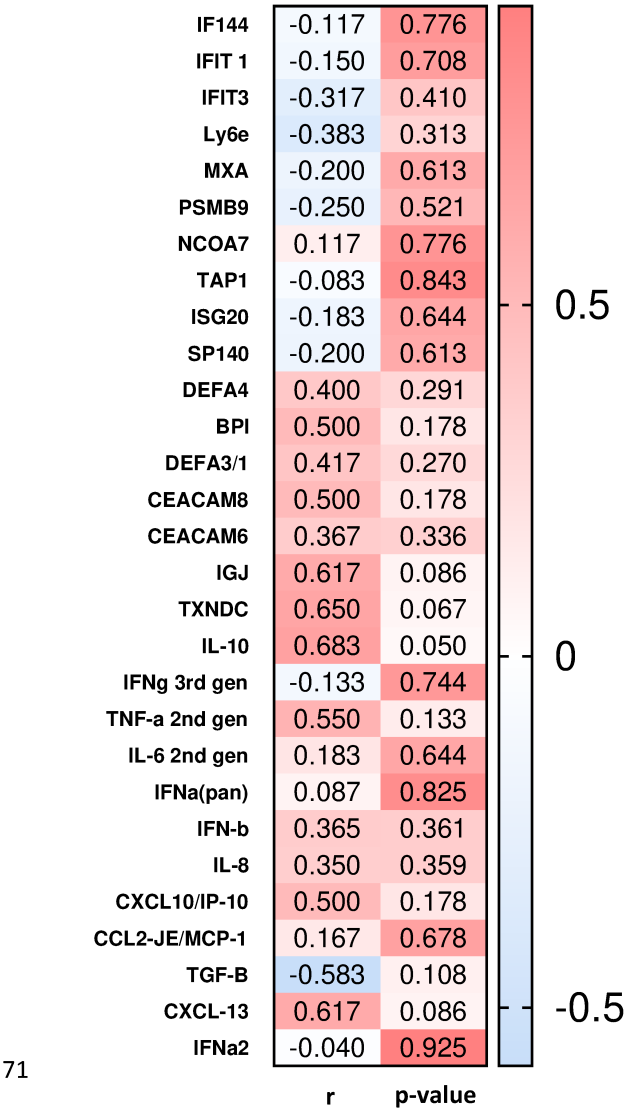
46 **Supplementary figure S3 Validation of biological phenotypes in an independent**
47 **cSLE cohort. A/B).** Unsupervised hierarchical clustering using Ward’s agglomerative
48 method and passing the Euclidean distance between samples, using row-based log-
49 transformed z-scores, identified three clusters in the discovery cohort and two
50 clusters in the replication cohort. **C).** Principal component analysis (PCA) showing
51 that there is overlap between the LBP group from the discovery cohort and LBP
52 group of the replication cohort. Patients in the MBP from the discovery cohort and
53 replication cohort show overlap. The first two components with their percentage of
54 variance are shown in parentheses.



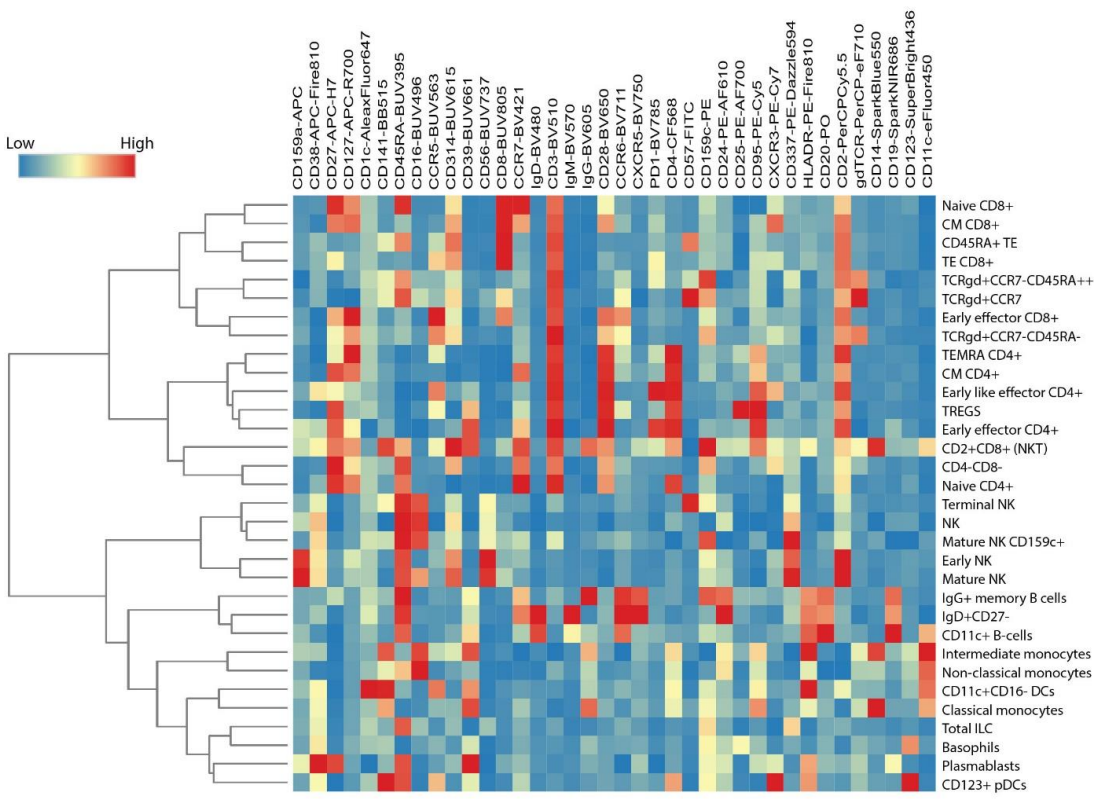
Supplementary figure S4. Unsupervised hierarchical clustering groups healthy controls within the low biological phenotype cluster of cSLE patients A) Unsupervised hierarchical clustering using Ward’s agglomerative method, passing the Euclidean distance between samples, using row-based log-transformed z-scores, identified 3 clusters. Healthy controls cluster in the LBP group. **B)** PCA shows overlap between the LBP and healthy control group. **C)** PCA shows overlap between healthy controls and patients in the LBP group from the discovery and replication cohort. Red-blue color indicates the z-scores. LBP, low biological phenotype; HBP, high biological phenotype; MBP, mixed biological phenotype.



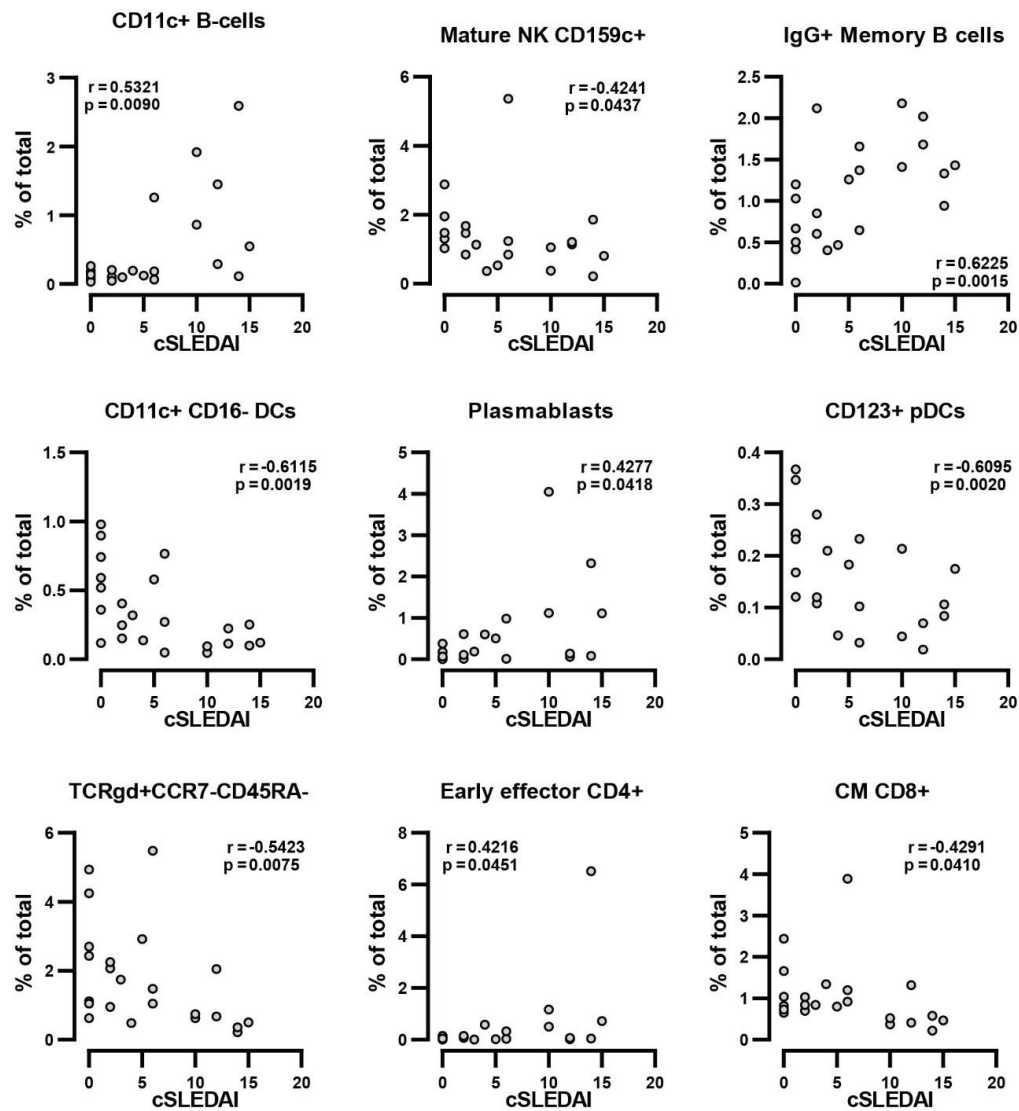
67 **Supplementary figure S5 Biological phenotypes are not influenced by the use**
68 **of corticosteroids.** Heatmap indicating the Spearman's rho and p-value between
69 prednisone use in mg/kg/day vs gene expression or cytokine levels. r represents
70 Spearman's Rho.



Supplementary figure S6 FLOWSom analysis identified 32 distinct immune cell subsets. Clustered heatmap indicating the expression of various immune cell marker. Red indicates a high expression while blue indicates a low expression.



79 **Supplementary figure S7 Clinical SELENA-SLEDAI is significantly associated**
80 **with 9 immune cell populations** Correlation plots depicting correlations between 9
81 immune cell populations and the clinical SELENA-SLEDAI (N=23). r represents
82 Spearman's rho and p indicates the p-value.



Supplementary table T1 Characteristics of replication cohort and healthy controls

ID	Gender	Age (years)	Ethnicity	Disease duration (years)	Clinical SELENA- SLEDAI	BILAG-2004	Anti-dsDNA (IU/mL) [#]	Auto-antibodies	Treatment at visit				Disease Activity State	Biological Phenotype ^{\$}
									HCQ	Prednisone (mg/kg/day)	NSAIDs	DMARD		
Pt1	Male	10	Non-White	5.44	0	0	0.9	-	Yes	.	No	No	LLDAS	LBP
Pt2	Male	13.4	Non-White	1.17	0	6	0.7	Anti-SSA	Yes	.	No	Yes	LLDAS	LBP
Pt3	Male	16.8	Non-White	4.56	0	1	41	Anti-SSA	Yes	.	No	Yes	LLDAS	LBP
Pt4	Female	17.8	White	2.81	0	1	0.4	-	Yes	.	No	No	LLDAS	LBP
Pt5	Female	16.7	White	3.93	2	10	4.1	-	Yes	.	No	Yes	LLDAS	LBP
Pt6	Female	18.1	Non-White	2.05	0	2	2.9	-	Yes	0.16	No	Yes	LLDAS	LBP

*Active refers to patients with a cSELENA-SLEDAI >4 who are not at diagnosis or have a flare
#Measured with a fluorometric enzyme immunoassay, cut-off is 10 IU/mL
**DMARD refers to mycophenolate mofetil, azathioprine, cyclophosphamide, methotrexate
\$LBP; Low Biological Phenotype, HBP; High Biological Phenotype, MBP; Mixed Biological Phenotype

Pt7	Female	17.6	White	2.42	0	6	0.1	-	Yes	.	No	Yes	LLDAS	LBP
Pt8	Female	16.4	White	0.1	5	11	0.9	Anti-SSA	Yes	0.17	No	Yes	Active*	MBP
Pt9	Female	17.8	White	3.48	5	10	10	Anti-Sm	Yes	.	No	Yes	Active*	MBP
Pt10	Female	16.5	White	0.73	4	6	0.5	Anti-SSA/B	Yes	0.12	Yes	Yes	Active*	MBP
Pt11	Female	16.2	White	0.64	8	11	93	Anti-RNP, Anti-Sm	Yes	0.08	No	Yes	Active*	MBP
Pt12	Female	16.9	White	2.18	4	5	9.5	AntiSSA/B, Anti-Sm	Yes	0.05	No	Yes	Active*	MBP
HC1	Female	17.9	White	-	-	-	-	-	-	-	-	-	-	LBP
HC2	Female	17.1	White	-	-	-	-	-	-	-	-	-	-	LBP
HC3	Male	18.3	White	-	-	-	-	-	-	-	-	-	-	LBP
HC4	Male	16.5	White	-	-	-	-	-	-	-	-	-	-	LBP
HC5	Female	14	White	-	-	-	-	-	-	-	-	-	-	LBP
HC6	Female	14.7	White	-	-	-	-	-	-	-	-	-	-	LBP
HC7	Male	12.4	White	-	-	-	-	-	-	-	-	-	-	LBP
HC7	Female	12.4	White	-	-	-	-	-	-	-	-	-	-	LBP
HC8	Female	17.9	White	-	-	-	-	-	-	-	-	-	-	LBP

Supplementary table T2. A positive logFC illustrates higher number of immune cell subsets in cSLE patients compared to HC, while a negative value indicates a higher number of immune cell subsets in the HC. The bold line indicates the threshold of 0.1.

HC vs cSLE			
	<i>logFC</i>	<i>p-value</i>	<i>FDR</i>
<i>Central memory CD8+</i>	-1.74711	7.20E-08	2.30E-06
<i>Mature NK CD159c+</i>	-1.63086	2.63E-07	4.21E-06
<i>TCRgd+CCR7-CD45RA-</i>	-1.84315	1.39E-06	1.49E-05
<i>NK</i>	-1.09615	0.000176	0.001411
<i>Terminal NK</i>	-2.05502	0.000638	0.003103
<i>CD123+ pDCs</i>	-1.29871	0.000664	0.003103
<i>CD11c+ B cells</i>	2.515655	0.000679	0.003103
<i>early like effector CD4+</i>	1.666455	0.000931	0.003723
<i>Early effector CD4+</i>	3.789902	0.001646	0.005854
<i>Plasmablasts</i>	3.038351	0.003022	0.009388
<i>CD4-CD8-</i>	0.921128	0.003259	0.009388
<i>TEMRA CD4+</i>	-0.7998	0.003521	0.009388
<i>CD11c+CD16- DCs</i>	-1.06455	0.010298	0.024734
<i>Total ILC</i>	-1.11639	0.010821	0.024734
<i>Non-classical monocytes</i>	-0.84832	0.060216	0.122393
<i>Early effector CD8+</i>	-1.21033	0.061197	0.122393
<i>Naive CD4+</i>	0.531037	0.078319	0.144305
<i>CD2+CD8+ (NKT)</i>	0.711644	0.081171	0.144305
<i>CD45RA+ Terminal effector</i>	1.318206	0.088535	0.149111
<i>IgD+CD27-</i>	0.476732	0.148027	0.23457
<i>TREGs</i>	0.50859	0.155947	0.23457
<i>Intermediate monocytes</i>	0.802366	0.161267	0.23457
<i>TCRgd+CCR7-CD45RA++</i>	-0.78278	0.24337	0.333149
<i>Naïve CD8+</i>	0.297481	0.256112	0.333149
<i>Basophils</i>	-0.38973	0.260273	0.333149
<i>Terminal effector CD8+</i>	-0.35758	0.47405	0.582145
<i>Central memory CD4+</i>	-0.15357	0.495708	0.582145
<i>IgG+ memory B cells</i>	0.356027	0.509377	0.582145
<i>Early NK</i>	0.291065	0.580484	0.640535
<i>Classical monocytes</i>	-0.14209	0.62904	0.670976
<i>Mature NK</i>	0.082463	0.917493	0.94709
<i>TCRgd+CCR7-</i>	0.065349	0.983709	0.983709

Supplementary table T3. A positive logFC illustrates higher number of immune cell subsets in the HBP group compared to LBP group, while a negative value indicates a higher number of immune cell subsets in the LBP group. The bold line indicates the threshold of 0.1.

HBP vs LBP			
	<i>logFC</i>	<i>p-value</i>	<i>FDR</i>
<i>CD11c+CD16- DCs</i>	-2.54453	1.36E-07	4.34E-06
<i>Early effector CD4+</i>	3.032646	1.68E-06	2.69E-05
<i>CD11c+ B cells</i>	2.80641	2.54E-06	2.71E-05
<i>Non-classical monocytes</i>	-2.44421	2.15E-05	0.000172
<i>Terminal NK</i>	-3.19247	0.000272	0.001493
<i>TCRgd+CCR7-CD45RA-</i>	-1.99941	0.00028	0.001493
<i>TCRgd+CCR7-</i>	-3.79634	0.002806	0.012827
<i>Total ILC</i>	-1.64627	0.004236	0.016945
<i>Mature NK CD159c+</i>	-1.18301	0.006739	0.02396
<i>Plasmablasts</i>	2.517105	0.009603	0.03073
<i>Early effector CD8+</i>	-2.55299	0.01336	0.038865

<i>Intermediate monocytes</i>	-2.10203	0.015138	0.040369
<i>Classical monocytes</i>	-0.97833	0.01767	0.043496
<i>Mature NK</i>	-1.48163	0.019892	0.045467
<i>Early NK</i>	-1.41406	0.025771	0.054978
<i>CD123+ pDCs</i>	-1.1826	0.046643	0.093287
<i>CD2+CD8+ (NKT)</i>	0.868515	0.058235	0.109619
<i>Central memory CD8+</i>	-0.83291	0.076628	0.136228
<i>Naive CD4+</i>	0.566811	0.098094	0.165211
<i>TEMRA CD4+</i>	-0.68883	0.106759	0.165652
<i>CD4-CD8-</i>	0.562335	0.108709	0.165652
<i>early like effector CD4+</i>	0.569636	0.325099	0.472871
<i>TREGs</i>	-0.4511	0.341199	0.474711
<i>IgG+ memory B cells</i>	0.671926	0.416964	0.555952
<i>NK</i>	-0.31752	0.487388	0.623857
<i>Terminal effector CD8+</i>	0.362942	0.551063	0.678231
<i>Central memory CD4+</i>	0.111494	0.750093	0.888999
<i>Naive CD8+</i>	-0.10091	0.792288	0.905472
<i>Basophils</i>	-0.13687	0.854158	0.942519
<i>TCRgd+CCR7-CD45RA++</i>	-0.14647	0.925646	0.987355
<i>IgD+CD27-</i>	-0.03022	0.966962	0.993822
<i>CD45RA+ Terminal effector</i>	-0.04186	0.993822	0.993822

Supplementary table T4. A positive logFC illustrates higher number of immune cell subsets in the HBP group compared to MBP group, while a negative value indicates a higher number of immune cell subsets in the MBP group. The bold line indicates the threshold of 0.1.

HBP vs MBP			
	<i>logFC</i>	<i>p-value</i>	<i>FDR</i>
<i>TCRgd+CCR7-</i>	-4.64398	0.000351	0.011237
<i>Non-classical monocytes</i>	-2.4343	0.000786	0.012582
<i>Classical monocytes</i>	-1.24328	0.005736	0.061188
<i>Plasmablasts</i>	2.307618	0.012353	0.075124
<i>Total ILC</i>	-1.3371	0.014008	0.075124
<i>Terminal NK</i>	-2.04466	0.018816	0.075124
<i>Early NK</i>	-1.75968	0.020078	0.075124
<i>TCRgd+CCR7-CD45RA-</i>	-1.79162	0.020948	0.075124
<i>Intermediate monocytes</i>	-1.37744	0.021129	0.075124
<i>Early effector CD8+</i>	-2.44316	0.025614	0.081965
<i>Mature NK</i>	-1.71238	0.038231	0.10636

<i>CD4-CD8-</i>	0.841503	0.039885	0.10636
<i>CD2+CD8+ (NKT)</i>	0.872938	0.051328	0.126346
<i>CD11c+CD16- DCs</i>	-1.07787	0.056967	0.13021
<i>Naive CD4+</i>	0.65101	0.136929	0.292114
<i>early like effector CD4+</i>	0.82462	0.170632	0.332529
<i>Central memory CD4+</i>	0.48435	0.176656	0.332529
<i>NK</i>	0.488394	0.193316	0.343674
<i>Mature NK CD159c+</i>	-0.61621	0.22644	0.381372
<i>CD11c+ B cells</i>	0.569034	0.448662	0.717858
<i>TEMRA CD4+</i>	-0.339	0.485386	0.739635
<i>Terminal effector CD8+</i>	-0.58677	0.558822	0.812831
<i>IgD+CD27-</i>	-0.27533	0.591283	0.822654
<i>CD45RA+ Terminal effector</i>	-0.5539	0.642539	0.856719
<i>CD123+ pDCs</i>	-0.27541	0.705776	0.863231
<i>TREGs</i>	-0.2265	0.71051	0.863231
<i>Early effector CD4+</i>	0.290938	0.751142	0.863231
<i>Central memory CD8+</i>	-0.16058	0.755327	0.863231
<i>Basophils</i>	-0.16103	0.819215	0.903961
<i>IgG+ memory B cells</i>	-0.07282	0.92338	0.969227
<i>Naive CD8+</i>	-0.03214	0.959364	0.969227
<i>TCRgd+CCR7-CD45RA++</i>	-0.02574	0.969227	0.969227

Supplementary table T5. A positive logFC illustrates higher number of immune cell subsets in the LBP group compared to MBP group, while a negative value indicates a higher number of immune cell subsets in the MBP group. The bold line indicates the threshold of 0.1.

LBP vs MBP			
	<i>logFC</i>	<i>p-value</i>	<i>FDR</i>
<i>CD11c+CD16- DCs</i>	1.466663	0.0008	0.025602
<i>CD11c+ B cells</i>	-2.23738	0.002541	0.040661
NK	0.80591	0.02366	0.244704
<i>Early effector CD4+</i>	-2.74171	0.030588	0.244704
<i>CD123+ pDCs</i>	0.907191	0.045984	0.294297
<i>Central memory CD8+</i>	0.672321	0.069403	0.370152
<i>Terminal NK</i>	1.147809	0.084308	0.385409
<i>Mature NK CD159c+</i>	0.566798	0.130962	0.523849
<i>Central memory CD4+</i>	0.372856	0.203745	0.637633
<i>Terminal effector CD8+</i>	-0.94971	0.216119	0.637633
<i>TCRgd+CCR7-CD45RA-</i>	0.621976	0.219186	0.637633
<i>IgG+ memory B cells</i>	-0.74475	0.27453	0.710676
<i>TEMRA CD4+</i>	0.349824	0.288712	0.710676
<i>TCRgd+CCR7-</i>	-0.84764	0.383948	0.870212
<i>CD4-CD8-</i>	0.279167	0.429992	0.870212
<i>Classical monocytes</i>	-0.26495	0.436739	0.870212
<i>Total ILC</i>	0.309179	0.520564	0.870212
<i>Intermediate monocytes</i>	-0.24511	0.552167	0.870212
<i>IgD+CD27-</i>	0.310412	0.552199	0.870212
<i>Early NK</i>	-0.34562	0.5531	0.870212
<i>TREGs</i>	0.224608	0.597906	0.870212
<i>CD45RA+ Terminal effector</i>	-0.51204	0.598271	0.870212
<i>early like effector CD4+</i>	0.254984	0.638872	0.888865
<i>Mature NK</i>	-0.23075	0.736046	0.970168
<i>Naïve CD4+</i>	0.0842	0.811962	0.970168
<i>Naïve CD8+</i>	0.068766	0.828235	0.970168
<i>Plasmablasts</i>	-0.20949	0.85335	0.970168
<i>TCRgd+CCR7-CD45RA++</i>	0.120722	0.877862	0.970168
<i>Early effector CD8+</i>	0.109827	0.879215	0.970168
<i>Basophils</i>	-0.02416	0.963336	0.985693
<i>Non-classical monocytes</i>	0.009906	0.974538	0.985693
<i>CD2+CD8+ (NKT)</i>	0.004424	0.985693	0.985693

References

1. Petri M, Orbai AM, Alarcon GS, Gordon C, Merrill JT, Fortin PR, et al. Derivation and validation of the Systemic Lupus International Collaborating Clinics classification criteria for systemic lupus erythematosus. *Arthritis Rheum*. 2012;64(8):2677-86.
2. Petri M, Kim MY, Kalunian KC, Grossman J, Hahn BH, Sammaritano LR, et al. Combined oral contraceptives in women with systemic lupus erythematosus. *N Engl J Med*. 2005;353(24):2550-8.
3. Franklyn K, Lau CS, Navarra SV, Louthrenoo W, Lateef A, Hamijoyo L, et al. Definition and initial validation of a Lupus Low Disease Activity State (LLDAS). *Ann Rheum Dis*. 2016;75(9):1615-21.
4. Wahadat MJ, Schonenberg-Meinema D, van Helden-Meeuwsen CG, van Tilburg SJ, Groot N, Schatorje EJJ, et al. Gene signature fingerprints stratify SLE patients in groups with similar biological disease profiles: a multicenter longitudinal study. *Rheumatology (Oxford)*. 2022.
5. Mueller YM, Schrama TJ, Ruijten R, Schreurs MWJ, Grashof DGB, van de Werken HJG, et al. Stratification of hospitalized COVID-19 patients into clinical severity progression groups by immunophenotyping and machine learning. *Nat Commun*. 2022;13(1):915.
6. Park LM, Lannigan J, Jaimes MC. OMIP-069: Forty-Color Full Spectrum Flow Cytometry Panel for Deep Immunophenotyping of Major Cell Subsets in Human Peripheral Blood. *Cytometry A*. 2020;97(10):1044-51.