OPEN ACCESS

TRANSLATIONAL SCIENCE

Single-cell resolution of longitudinal blood transcriptome profiles in rheumatoid arthritis, systemic lupus erythematosus and healthy control pregnancies

Hilde Julie T Lien [ID] ,[1] Tina T Pedersen,[1,2] Bente Jakobsen,[2] Arnar Flatberg,[1,3] Konika Chawla,[1,4] Pål Sætrom,[1,5] Mona H Fenstad[1,6]

ABSTRACT

Objectives Comparative longitudinal analyses of cellular composition and peripheral blood gene expression in Rheumatoid arthritis (RA), systemic lupus erythematosus (SLE) and healthy pregnancies.
Methods In total, 335 whole blood samples from 84 RA, SLE and healthy controls before pregnancy, at each trimester, 6 weeks, 6 months and 12 months post partum were analysed. We combined bulk and single cell RNA analyses for cell-type estimation, validated by flow cytometry, before combining this in a cell-type adjusted analysis for an improved resolution of unrecognised gene expression changes associated with RA and SLE pregnancies.
Results Patients were well regulated throughout pregnancy, and few had pregnancy complications. In SLE, the interferon signature was augmented during pregnancy, and the pregnancy signature was continued post partum. An altered cell type composition strongly influences the profile. In the pregnancy signature, transcripts involved in galactosylation potentially altering the effector functions of autoantibodies became more evident. Several genes in the adjusted RA signature are expressed in mucosal associated invariant T cells.
Conclusion We found distinct RA, SLE and pregnancy signatures, and no expression patterns could be attributed to medication or disease activity. Our results support the need for close postpartum follow-up of patients with SLE. Gene expression patterns in RA were closer to healthy controls than to SLE, and primarily became evident after cell-type adjustment. Adjusting for cell abundance unravelled gene expression signatures less associated with variation in cell-composition and highlighted genes with expression profiles associated with changes in specialised cell populations.

WHAT IS ALREADY KNOWN ON THIS TOPIC

⇒ There is a rebalancing of major proinflammatory and anti-inflammatory immune cells during pregnancy.
⇒ Rheumatoid arthritis (RA) and systemic lupus erythematosus (SLE) are affected differently by pregnancy, but the disparate disease-related and pregnancy-related responses are not well characterised.
⇒ Few studies have been conducted on these patients during pregnancy and cohorts are generally small.

INTRODUCTION

Rheumatoid arthritis (RA) and systemic lupus erythematosus (SLE) are autoimmune disorders affecting women more often than men, commonly in reproductive years. Although their clinical presentation is different, RA and SLE are genetically the most closely related of the major rheumatic diseases[1] as they share components of the type I IFN signalling pathway linked to treatment response[2–4] and disease activity.[5] One difference between the two is how they respond to pregnancy; while RA often improves, SLE is prone to flare. However, correct medication, and close follow-up, diminishes these patterns.[6 7]

Multiple mechanisms contribute to immunological tolerance of the fetus during pregnancy. Many of the cellular changes take place locally at the implantation site, but there are also systemic adaptions, accounting for the immune-modulating effects of pregnancy on disease activity.[8] More research involving comparative analyses is needed, as understanding how pregnancy-related immune-modulation differs in RA and SLE can potentially lead to better treatment options.

Distinct blood transcriptional profiles have been identified separately for both susceptibility to and severity of disease in non-pregnant SLE and RA.[9 10] These profiles are from bulk-samples, with poorly resolved cellular origin. Single cell RNA sequencing (scRNA-seq) provides an unsupervised and unbiased method for classifying and clustering cell types by the genes from each individual cell in a sample.[11–13]

We provide data from a cohort of women with universal accessibility to healthcare that have been followed closely throughout their pregnancies, are well regulated and have low disease activity. We analysed samples both by bulk- and scRNA-seq, and ran a cell-type estimation, validated by flow cytometry, before combining this in a cell-type adjusted analysis. This novel comparative approach gives an improved resolution of expression profiles in RA and SLE pregnancies.

eular   Lien HJT, *et al*. *Ann Rheum Dis* 2024;**83**:300–311. doi:10.1136/ard-2023-224644   **BMJ**

WHAT THIS STUDY ADDS
⇒ We provide comparative data from a cohort of healthy controls and well-regulated patients with RA and SLE followed before, during and after pregnancy.
⇒ We combine bulk and single cell RNA analyses in a novel approach for cellular resolution and a better understanding of the inherent differences between the groups.
⇒ RA is generally closer to healthy than SLE in gene expression profile; altered cell type composition strongly influence the SLE whole blood gene expression profile, whereas the RA whole blood gene expression profile appears to be more influenced by cellular dysregulation.
⇒ In contrast to healthy and RA, the expression levels for the largest pregnancy related gene clusters did not return to prepregnancy levels in patients with SLE post partum, indicating systemic long-term effects of pregnancy for patients with SLE.
⇒ Our results indicate possible roles of pregnancy-related decreases of smaller, more specialised immune cell populations in systemic disease modification.

HOW THIS STUDY MIGHT AFFECT RESEARCH, PRACTICE OR POLICY
⇒ A renewed focus on how cellular burden confuses results from genomic research, hiding changes in smaller cell populations and gene expression changes within cells.
⇒ A cellular resolution of observed gene expression differences can potentially lead to better monitoring, follow-up and treatment prepregnancy, during gestation and post partum.
⇒ Our results support the need for increased awareness of the postpartum period, also in patients with SLE.

METHODS

The following sections summarise the main methods; full details are provided in online supplemental methods.

Cohort and study design

In total, 335 whole blood (WB) samples were collected from 84 SLE, RA and healthy control (HEALTHY) pregnancies (figure 1A,B, tables 1 and 2). RA and SLE patient information and samples were from the Norwegian multicentre prospective quality (RevNatus) registry and biobank, respectively. Patients with RA were classified as seropositive (SPRA) or seronegative (SNRA). Information and samples from healthy controls were from Biobank1. Samples were collected before (T0), in each trimester (T1, T2, T3), and at three time points after pregnancy (T4, T5, T6) (figure 1); The majority of patients were recruited before pregnancy; the controls were recruited in their first trimester. Pregnancies with adverse pregnancy outcomes (table 1) were not excluded from further analyses. There were too few pregnancies with complications for us to analyse these data separately.

Gene expression profiles from WB

WB samples were collected in PAX-gene RNA tubes (Qiagen). RNA was isolated using PAX-gene blood RNA-kits (Qiagen). RNA-sequencing (RNA-seq) libraries were prepared using Illumina TruSeq Stranded Total RNA with Ribo-Zero Globin kit (Illumina, San Diego, California, USA), sequenced using Illumina HiSeq4000. Gene counts were quantified by the Salmon pseudoalignment method and analysed for differential expression using the voom/limma linear modelling framework.

scRNA-seq and flow cytometry of peripheral blood mononuclear cells

Peripheral blood mononuclear cells (PBMCs) were isolated using Lymphoprep density gradient on T3 and T4 samples from four patients with RA and four patients with SLE following the manufacturer's protocol. Isolated PBMCs were resuspended in 10% DMSO, frozen at −80°C for 24 hours and cryopreserved until sequencing. Samples were thawed and split in two aliquots, one for scRNA-seq and one for flow cytometry. ScRNA aliquots from the same disease and time point were combined to create four mixed samples (T3 RA, T3 SLE, T4 RA and T4 SLE), scRNA-seq libraries were generated by ×10 microdroplet-based technology; flow cytometric analyses were done in parallel on the 16 individual samples. Panels for deep immunophenotyping of B-and T cell subpopulations were adapted from EuroFlow, TRANSIM-MUNOM and the Human Immunology Project (online supplemental table S1, online supplemental figure S1).[14–16]

Combined analysis of mRNA-seq and scRNA-seq

Results from scRNA-seq and WB RNA-seq were combined for cell-type abundance estimation using CIBERSORTx[17] with default settings and permutations set to 100. The total estimated percentage of each cell fraction was calculated, and sample-specific estimates from the five cell types estimated to contribute more than 10% of the total cellular burden were added to the linear model design matrix. Subsequent normalisation, statistical analyses, sample and gene clustering, and functional enrichment analyses were then done as described above.

Statistical analyses

Cell-type estimation-deconvolution was done in CIBERSORTx, scRNA-seq data analysed in UCSC Cellbrowser 0.7.9 using uniform manifold approximation and projection (UMAP) statistics and Louvain clustering. All other analyses were done in R, using the ggPlot2 package, ClusterGenomics package[18] and gProfiler2 package.[19] Empirical Bayes moderated t-statistic adjusted with the Benjamini-Hochberg (BH) method were used to calculate significance of the differentially expressed genes (DEGs). For statistical analyses on CIBERSORTx results, a Tukey HSD test was used and for correlation with scRNA-seq data a Pearson's product-moment correlation was run.

Patient and public involvement

Continuous dialogue with the patients has been achieved through the Norwegian National Advisory Unit on Pregnancy and Rheumatic Diseases (NKSR). Their reference group has involved user representatives who give useful feedback to the researchers. NKSR contributes to relevant education and are also active in disseminating research in conferences and on social media, making it accessible for researchers, patients, clinicians and other users.

RESULTS

WB mRNA analysis reveals three distinct signatures

To identify genes with expression patterns related to pregnancy and disease state, we sequenced total RNA (RNA-seq) from 335 WB samples collected from a total of 84 SLE, RA and healthy control pregnancies. The samples were collected before, in each trimester and up to three time points after pregnancy (figure 1).

Analysing gene expression differences between each of the groups, both overall and at each time point, revealed over 7800 significantly DEGs (BH adjusted p<0.05) (online supplemental table 2 and 3, online supplemental figure S2). These grouped
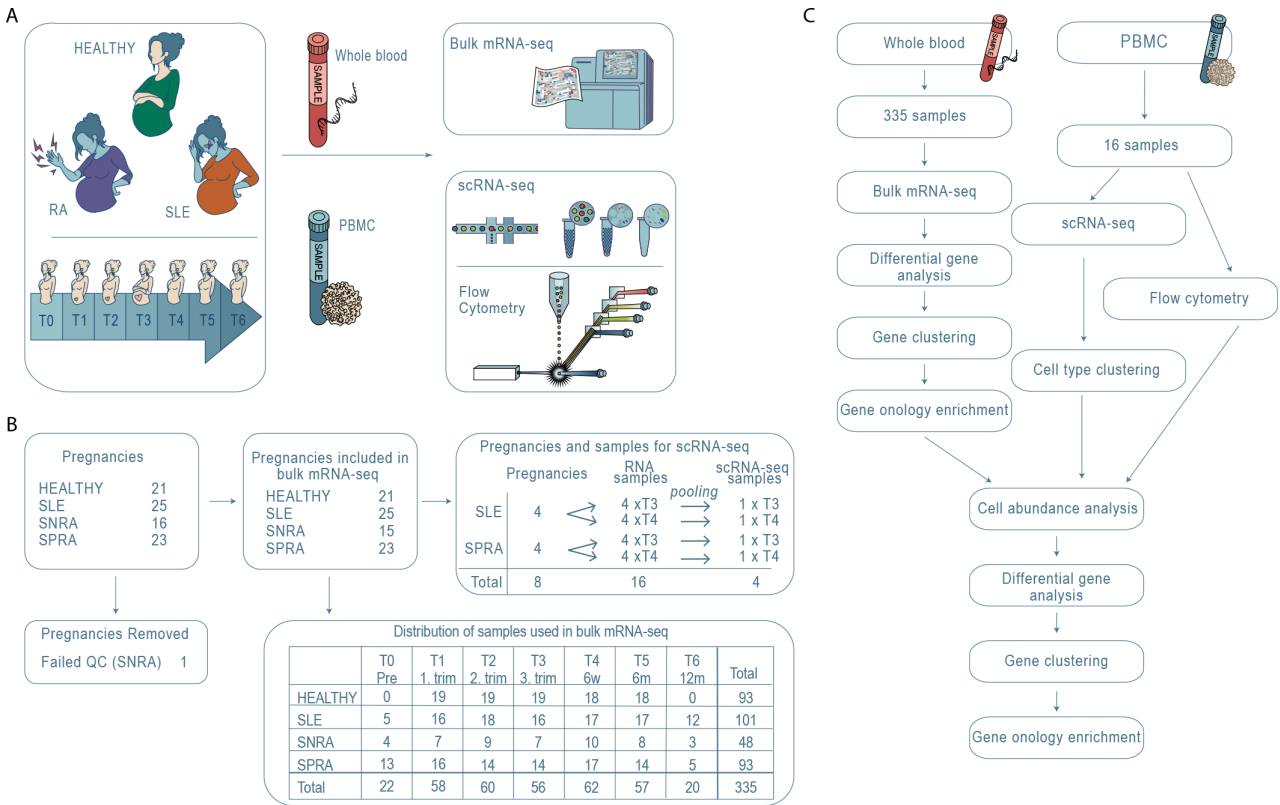
Figure 1 Study design. (A) Overview of study design with cohorts, sample times, sample types and methods. (B) Overview of samples. (C) Analysis flow chart. PBMC, peripheral blood mononuclear cell; RA, rheumatoid arthritis; SLE, systemic lupus erythematosus; SNRA, seronegative RA; SPRA, seropositive RA.

into 14 distinct clusters (G1–G14) with expression patterns related to time, pregnancy and diagnosis (figure 2A–C). The gene expression differences could not be attributed to an effect of difference in medication (disease-modifying antirheumatic drugs (DMARDs) with or without prednisolone) nor disease activity (BH-adjusted $p > 0.05$ for all genes) (online supplemental figure S3).

Pregnancy signature

The largest clusters, G5 and G14, contained genes, respectively, downregulated and upregulated during pregnancy (T1–T3), on average bottoming and peaking at T2. Using the Human Protein Atlas[20] to examine the DEGs with the highest absolute log fold change ('strongest DEGs') in these clusters, showed both clusters representing biological processes important for placentation, including immunomodulation that increases tolerance and reduces inflammation, angiogenesis and vascular remodelling. Notable regulators of these processes were upregulated (G14: *IDO, IL10, FLT1*) and downregulated (G5: *IL5RA, PTGDR2, NKX3-1, SOCS5*). The most upregulated G14 transcripts were long non-coding RNAs (lncRNAs) that are not functionally characterised yet but may be involved in immune cell proliferation.[21] Further investigating these clusters through gene ontology enrichment showed pathways significantly enriched (hypergeometric test, BH adjusted $p < 0.05$) for G14 genes were the olfactory signalling pathway and G-protein coupled receptor

Table 1 Clinical parameters

Disease group	Healthy	SLE	SNRA	SPRA
Participants, n	21	23	13	19
Clinical parameters				
Disease duration,* years	–	7 (1–18)	5 (1–11)	4 (1–17)
Age,† years	30 (23–36)	31 (22–43)	31 (25–36)	33 (26–39)
Body mass Index,† kg/m²	24 (18–37)	24 (19–28)	26 (19–30)	27 (22–40)
Pregnancy duration,* ‡ weeks	40 (36–42)	39 (29–42)	38 (29–41)	39 (35–42)
Primipara	8	9	5	9
Adverse pregnancy outcomes§	3	3	4	3
Birth weight,* kg	3.6 (2.3–4.1)	3.3 (1.1–4.2)	3.1 (1.2–3.7)	3.7 (2.5–4.2)

*Median (range) as indicated.
†Mean (range).
‡Pregnancies with a duration minimum of 24 weeks.
§Preterm birth (<week 37), low birth weight (<2500 g), eclampsia or pre-eclampsia.
SLE, systemic lupus erythematosus; SNRA, seronegative RA; SPRA, seropositive rheumatoid arthritis.

Table 2 Medication status and disease activity

Time point	LAI(P)/DAS 28*	Creatinine, µmol/L*	Prednisolone, %†	DMARDs, %‡
RA				
Pre	2.7 (1.7–4.7) (0.9)	65 (50–80) (8)	29	65
T1	2.6 (1.7–6.1) (1.1)	48 (41–62) (6)	17	65
T2	2.1 (1.7–4.1) (0.8)	45 (36–59) (5)	17	57
T3	2.1 (1.1–3.6) (0.6)	45 (35–60) (5)	19	67
6 weeks	2.3 (1.7–3.9) (0.7)	66 (48–78) (7)	22	59
6 months	2.4 (1.7–4.5) (0.8)	64 (45–69) (7)	27	82
12 months	2.6 (1.7–3.5) (0.8)	60 (53–76) (9)	38	50
SLE				
Pre	0 (0–0.1) (0.0)	61 (53–88) (13)	0	60
T1	0 (0–0.4) (0.1)	44 (37–69) (8)	44	100
T2	0 (0–0.2) (0.1)	46 (31–56) (6)	44	94
T3	0 (0–0.2) (0.1)	42 (29–61) (7)	50	100
6 weeks	0 (0–0.2) (0.0)	60 (51–99) (12)	53	100
6 months	0 (0–0.4) (0.1)	57 (42–73) (9)	47	94
12 months	0 (0–0.1) (0.0)	64 (43–76) (10)	50	83

*Median (range) (SD).
†2.5–0 mg.
‡Methotrexate, sulfasalazine and TNFα-inhibitors for RA and hydroxychloroquine and azathioprine for SLE. One patient with RA used hydoxychloroquine 6 months post partum (T5).
DAS28, Disease Activity Score-28; DMARDs, disease-modifying antirheumatic drugs; LAI(P), Lupus Activity Index (in Pregnancy); RA, rheumatoid arthritis; SLE, systemic lupus erythematosus.

signalling (figure 2D, online supplemental table 4), suggested to play a role in chemotaxis and organ construction during embryogenesis, as well as involvement in vascular remodelling during pregnancy.[22] Focusing on cell-type expression through The Genotype-Tissue Expression (GTEx) project[23] and Database of Immune Cell Expression[24] revealed the most downregulated G5 transcripts to stem from innate cytotoxic and immune regulatory cells. Among these are natural killer (NK)-cell receptors in the NKG2 group (*KLRC2, KLRC3, GNLY, KLRF1, FGFBP2*), which bind non-classical major histocompatibility-molecules and are important in conveying feto-maternal immune tolerance, cytotoxic markers (*GNLY, KLRF1, FGFBP2*), gamma-delta T (gdT)-cell receptors (*TRDC, TRDV2*), T-cell receptors enriched in regulatory T-cells (Treg) (*TRAV8-2, TRAV8-6*), and mucosal associated invariant T (MAIT) cells (*TRAV1-2*).

Additional clusters with clear pregnancy-related profiles were G13 and G9. G13 was upregulated in all pregnancies and included the interferon regulated gene (IRG) *IFI27*, which had the highest fold change in the dataset overall. This profile was mirrored by downregulated genes in G9. The G13 and G9 clusters included fewer, but more highly DEGs, including upregulated (G13) Toll-like receptors (*TLR1,2,4,5,6,8*) and B-cell activating factor (*BAFF/TNFSF13B*) and downregulated (G9) immunoglobulin heavy and light lambda and kappa chain genes (*IGH*, 6 genes; *IGL*, 14 genes; *IGK*, 8 genes; online supplemental table 2).

SLE signature

Expression levels in SLE predominantly followed similar temporal fluctuations as the other two groups, but with a higher (G9–G14) or lower (G1–G5) relative expression (figure 2A,B). This difference was not counteracted by pregnancy. Instead, for the two largest pregnancy related clusters (G5, G14) the expression levels in patients with SLE, in contrast to the healthy and RA groups, did not return to prepregnancy levels post partum (T4–T6) (figure 2C).

The genetic profile of SLE consists of both down and upregulated genes, primarily clustered in G4 and G11 respectively.

Downregulated genes in G4 include several transcripts enhanced in MAIT cells (eg, *SLC4A10, CFH, IL23R, RORC*) in addition to metalloproteases (*ADAM12, ADAM23*). Also of interest are the neutrophil transcripts in G7. Those that differ the most between SLE and healthy in the first trimester (T1) include genes involved in lipid and steroid hormone metabolism (*ABCA1, APOBR*) and inflammation or endoplasmic reticulum stress (*GRN, PRKCD*).

The SLE IFN signature genes in G11 (eg, *SIGLEC1, IFI44, IFI44L, USP18, RSAD2, ISG15*) show the highest differential expression over all in our analysis (online supplemental table 2). To investigate this signature further, we calculated interferon scores from our normalised count matrix using genes from two IFN signature models; a 28 gene IFN score (IFN 28) developed and validated through NanoString Technology[25] and a module-based IFN score[26] (online supplemental table 5a–d). The scores showed the same pattern as G11 (figure 3A–D). Most of the DEGs from the two models overlapped with G11, but a small subset of genes overlapped with G13 (figure 3E). Compared with G11, cluster G13 was more related to pregnancy than to SLE. Comparing the IFN scores for the three groups, RA and healthy had similar scores overall and at any given time point (p>0.05, Tukey's range test), while SLE differed from the other two groups both overall and at every time point (p<0.05, Tukey's range test).

RA signature

Gene expression patterns in patients with RA were closer to healthy controls than to patients with SLE. The eight genes differentially expressed between healthy, and RA are predominantly grouped in G4 and G5. These are found in T-cells (*SLC4A10, SNHG16, LRRC8B, NKX318, RPS26, HLA-H*) and monocytes (*NBPF26, NAAA*) and are involved in proliferation and tolerance.

SNRA and SPRA diverged in several clusters at T0 and T6 with opposite trajectories during pregnancy (online supplemental figure S4). Thus, gene expression changes driven by pregnancy overrode disease-group related differences between SNRA and SPRA.

PBMC analysis reveals distinct SLE and RA cell populations

To determine the cellular composition in the patients' blood, PBMCs were analysed by single cell RNA sequencing (scRNA-seq) and flow cytometry. Samples from four patients with RA and four patients with SLE were used, with two samples from each patient, one from the third trimester (T3) and one from 6 weeks post partum (T4). Both analyses revealed distinct SLE and RA cell populations.

Single cell RNA sequencing

Clustering analyses and UMAP visualisation generated 15 main clusters, which were further expanded and annotated to 20 cell types based on known marker genes (figure 4A, table 3, online supplemental figure S5). Quantifying the number of cells per cell type for the RA and SLE samples showed that RA samples had more T-cells, SLE more myeloid cells, and that the differences were enhanced in pregnancy. For both myeloid and T-cells, differences between RA and SLE were largely driven by IFN-induced clusters in the monocyte (SC-M2) and naïve T-cell (SC-T3) compartments of the SLE samples. In contrasts, RA had high levels of classical monocytes (SC-M1) and metabolically active naïve T-helper cell (SC-T1), and these cell types also showed opposite trends between RA and SLE during and after pregnancy (online supplemental figure S6, online supplemental table 6).
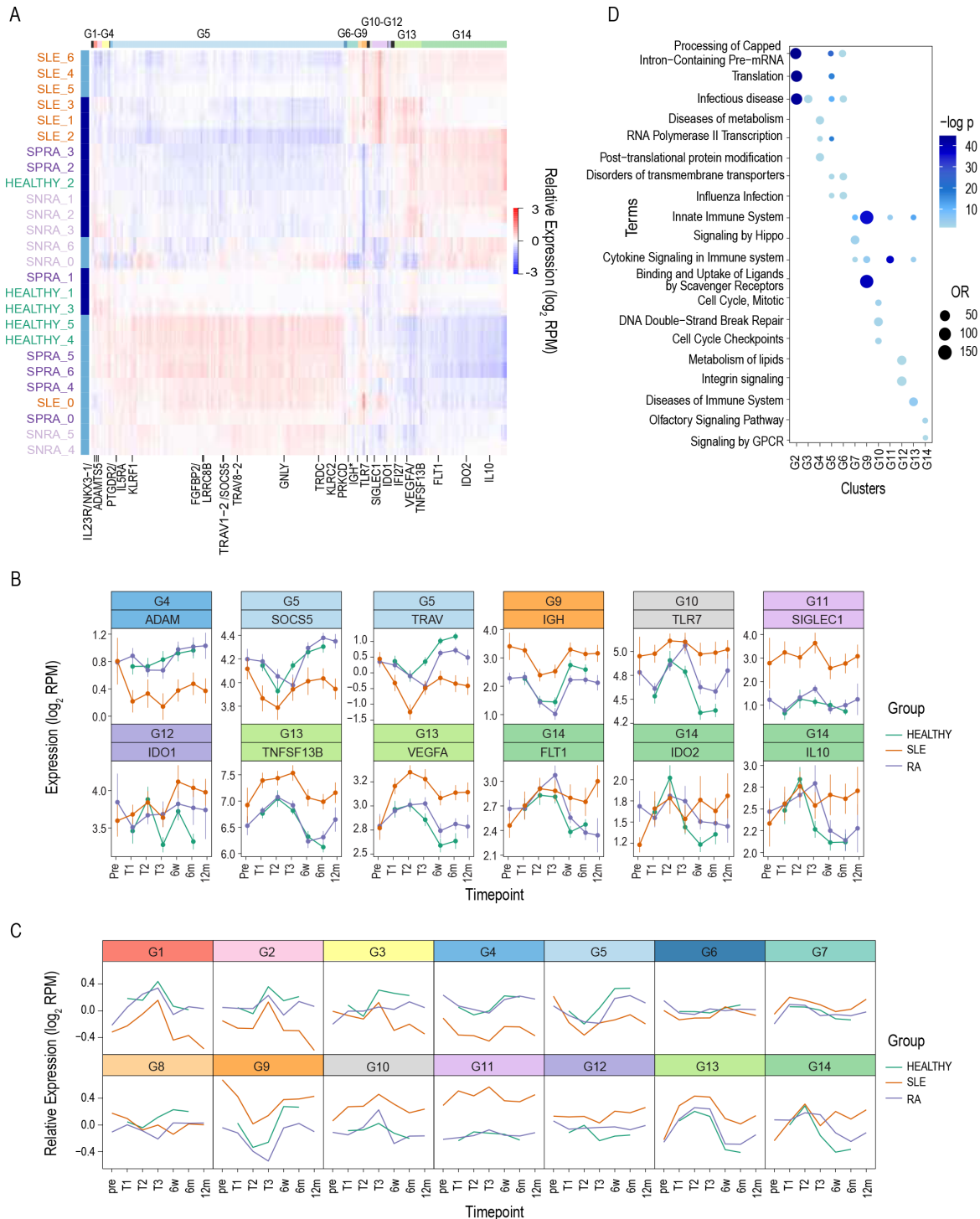
Figure 2 Whole blood mRNA sequencing results. (A) Heatmap of relative expression levels for differentially expressed genes from the whole blood mRNA-seq analysis. Samples and genes are along rows and columns, respectively. Samples are grouped and labelled by disease group and sample time (GroupTime); the left colour bar shows samples taken during (dark blue) and outside (light blue) pregnancy. Gene expression values are the gene's mean expression among samples for the corresponding GroupTime; values higher and lower compared with the gene's overall median expression are in red and blue, respectively. Samples and genes were clustered; robust gene clusters (G1–G14) are named, and colour coded on top, outliers (G0) are shown in black. Selected genes are shown on the bottom. IGH represents *IGHA1*, *IGHA2*, *IGHG1* and *IGHG3*(B) Individual gene expression profiles for 12 highlighted genes in (A). The profiles show mean (dots) gene expression (log₂ reads per million (RPM)) at each sample time for healthy (green), SLE (orange) and RA (purple) participants; vertical lines are SD. (C) Relative gene expression profiles for each individual gene cluster, labelled and colour coded as in (A). In each panel, lines show the median relative expression of the genes from the corresponding cluster at each sample time for healthy (green), SLE (orange) and RA (purple). (D) top five significant REACTOME pathways hypergeometric test, BH adjusted p<0.05) enriched for genes in clusters G1–G14. Pathways are represented by their second level term in the REACTOME pathway hierarchy (see online supplemental table 4 for complete results). Dots represent significant terms with colour showing statistical significance (minus log adjusted p value) and size showing the enrichment OR. The OR was calculated with our own normalised samples as background. BH, Benjamini-Hochberg; RA, rheumatoid arthritis; SLE, systemic lupus erythematosus.
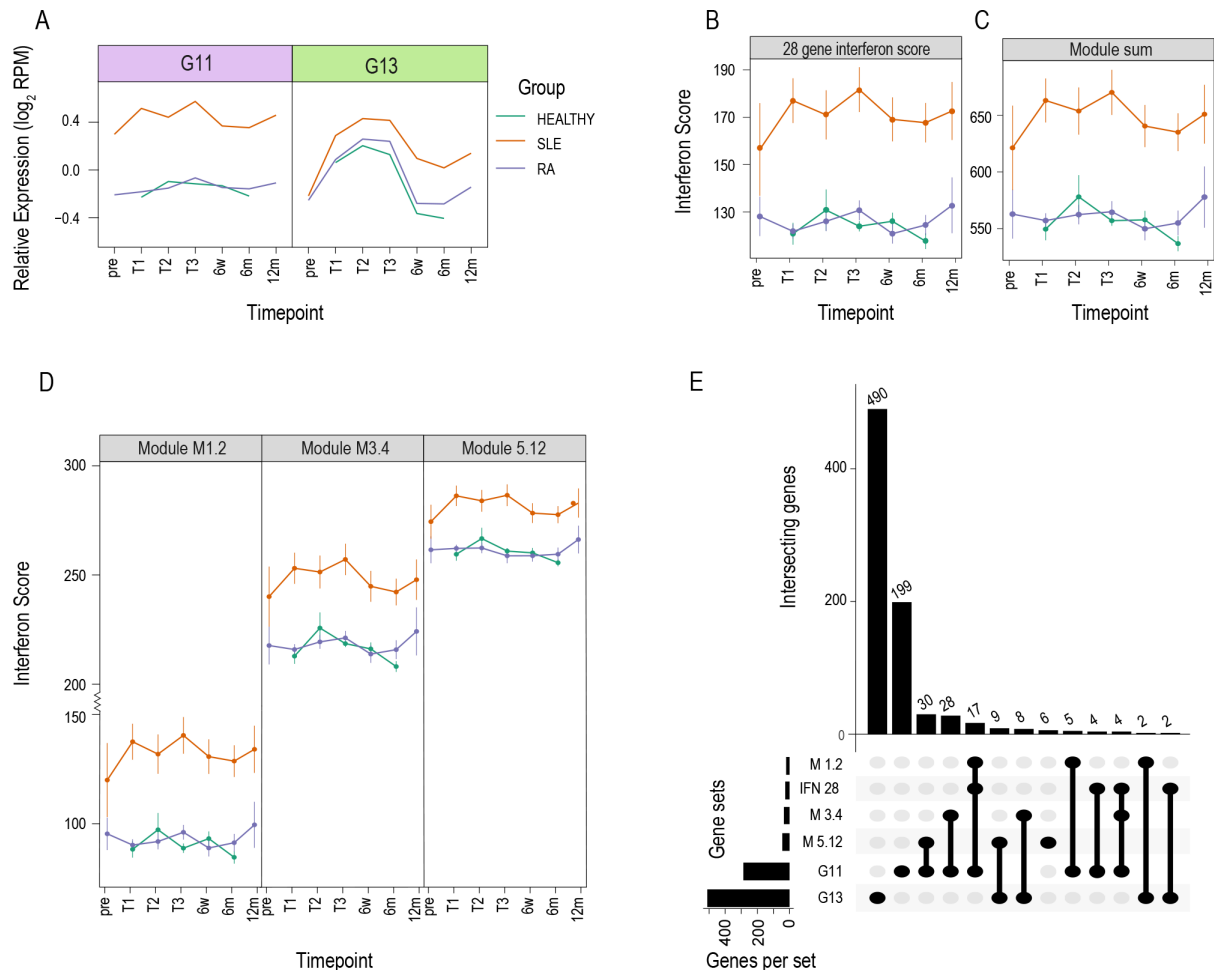
Figure 3 Interferon score. (A) An enhanced plot of the relative gene expression profiles of G11 and G13 from figure 2C; see figure 2C for details. (B) Interferon score calculated from 27 genes in a NanoString Validated 28 gene set. One gene was not found significant in our data (Benjamini-Hochberg adjusted $p>0.05$). The profiles show mean (dots) gene expression ($\log_2$ reads per million (RPM)) at each sample time for Healthy (green), SLE (orange) and RA (purple) participants; vertical lines are SD. (C) Interferon scores calculated from the sum of Module M1.2, M3.4 and M5.12 containing gene sets used in a module-based approach. (D) Interferon score shown individually for the three modules used in (C). (C, D) The profiles are labelled, and colour coded as in (B). (E) Comparison of genes in the gene sets used in (A–D). Horizontal bars show number of genes per set. Vertical bars show unique genes per set and the number of genes shared between different sets. Dots mark sets included in each of the vertical bars. RA, rheumatoid arthritis; SLE, systemic lupus erythematosus.

In both groups, the cell abundance of Cytotoxic T-and NK-cells (SC-T8, SC-T5, SC-NK1 and SC-NK2) was lower in T3 than in T4 (online supplemental table 6). Separate IFN-induced clusters formed in the monocyte (SC-M2) and naïve T-cell (SC-T3) compartments, and in SLE the SC-B2 population decreases, while the SC-PC1 population increases post partum. Differences for B-cells were small, but plasma cell (SC-PC1) levels were markedly (>4 times) higher in SLE post partum compared with T3.

Flow cytometry analysis
Adding flow cytometry in parallel to scRNA-seq allowed analysis of more cells from each PBMC sample and the possibility of classifying the cells using conventional markers for T-helper and B-cell subpopulations (online supplemental figure S1, online supplemental tables 1 and 7). The flow cytometry results showed the same overall patterns as the scRNA-seq results (figure 4B). There was no observed difference in T-regulatory cells between groups and the expected shift to a lower Th1/Th2-ratio during pregnancy was seen in both disease-groups, but more marked in RA (figure 4C, online supplemental table 7). Consistent with the

scRNA-seq data, plasma cells were significantly higher in SLE post partum (figure 4D).

Cell abundance analysis
Combining the scRNA-seq and WB mRNA-seq data allowed estimating cell-type abundances in all the WB samples (see the Methods section). The results showed the same overall patterns as both scRNA-seq and flow cytometry (figure 4B). And as these WB estimates correlated clearly with scRNA-seq results in corresponding samples (figure 4E; Spearman rank correlation 0.46, $p=2.2\times10^{-5}$), we went on to identify cell types that differed in abundance throughout and after pregnancy and between groups ($p<0.05$, Tukey's range test)(figure 4F, online supplemental table 8). Several T-cell types as well as SC-B1, SC-PC1 and SC-DC1 had a significant decrease in estimated abundance during pregnancy, bottoming out at T2-T3. Cell types SC-M1, SC-M2 and SC-T2 had the opposite signature. For the cell types with time-related fluctuations in abundance, healthy and RA samples generally showed similar fluctuations and abundances. In contrast, while the SLE samples also showed similar fluctuations, the abundance differed significantly to the healthy and
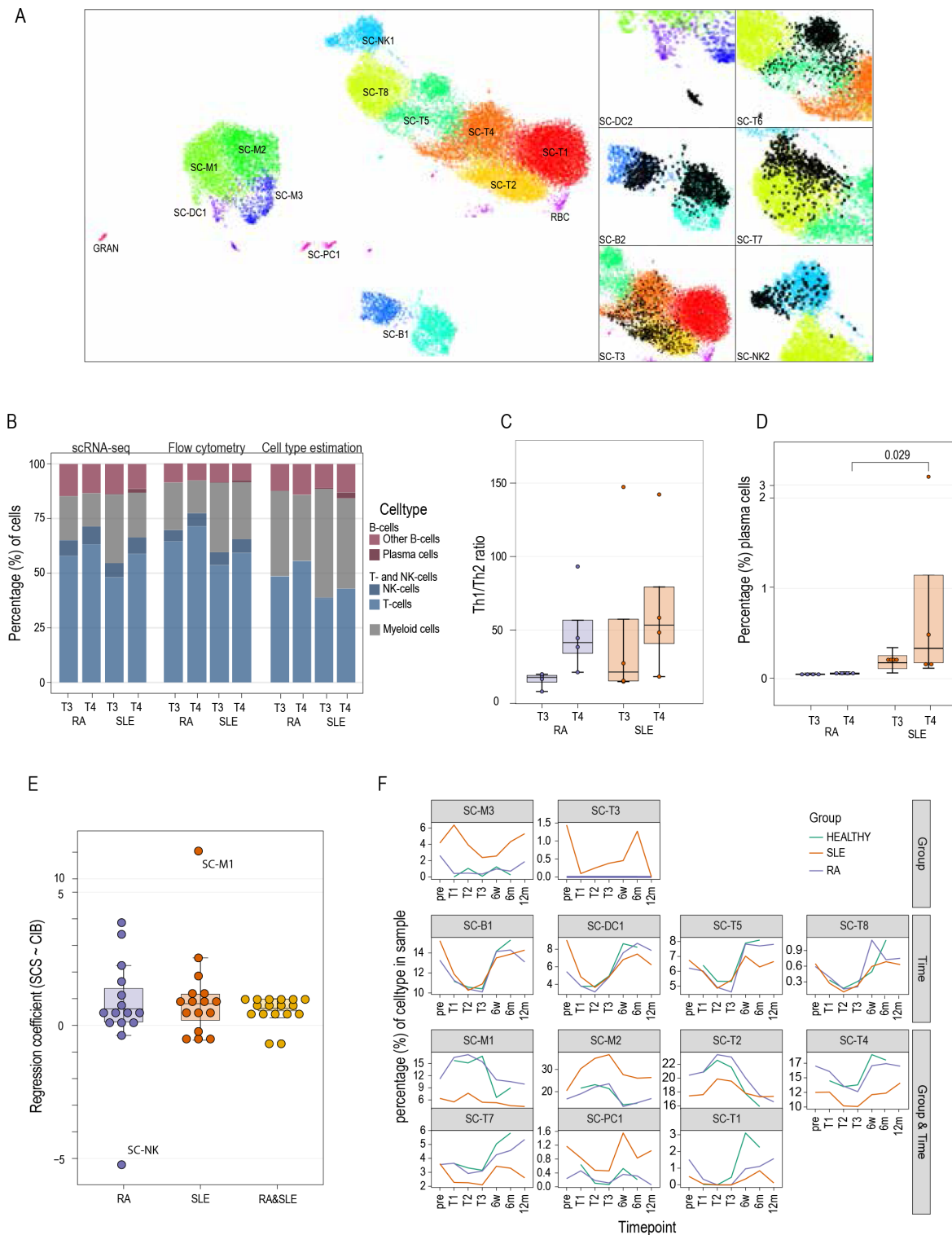
Figure 4 PBMC scRNA-seq, flow cytometry and cell type estimation results. (A) UMAP clustering of PBMC scRNA-seq expression profiles; each dot is one cell. (left) Automatically identified and (right) manually annotated clusters with their assigned cell type labels (see table 3, online supplemental figure S5). For the manually annotated clusters, each black dot represents a manually reannotated cell. (B) Relative abundance of B-cells (red), T- and NK-cells (blue) and Myeloid cells (grey) for RA and SLE in the third trimester (T3) and 6 weeks post partum (T4), estimated by scRNA-seq, flow cytometry and whole blood mRNA-seq cell type deconvolution. (C) Boxplot for the Th1/Th2 ratio found by flow cytometry for RA and SLE in the third trimester (T3) and 6 weeks post partum (T4). (D) boxplot of the percentage of plasma cells in live cells found by flow cytometry. The difference between RA and SLE at T4 was found to be significant (Wilcoxon rank sum test, p=0.025). (E) Associations between mRNA-seq cell type estimates and scRNA-seq results. Dots are regression coefficients from linear models based the mRNA-seq estimate and scRNA-seq result for one cell type from the RA (purple) and SLE (orange) samples, and RA and SLE (yellow) samples combined. (F) Estimated percentages of scRNA-seq cell types in whole blood samples. Cell types are grouped by whether they are significant between disease groups (group), throughout time (time) or both. Only significant cell types are shown (p<0.05, Tukey's range test). PBMC, peripheral blood mononuclear cell; RA, rheumatoid arthritis; SLE, systemic lupus erythematosus; UMAP, uniform manifold approximation and projection.

Table 3	UMAP clusters and associated cell types		
Cluster	Dominating cell type	Dominating group*	Dominating time point*
T-cells and NK-cells			
SC-T1	Naive T-helper cells, high metabolic activity	RA	T3 RA, T4 SLE
SC-T2	Naive T-helper cells	SLE	T3
SC-T3	IFN-induced T, T-regulatory cells	SLE	T3
SC-T4	Memory T-helper cells	RA	T4
SC-T5	Memory T (helper and cytotoxic) cells	SLE	T4
SC-T6	MAIT cells	RA	SLE T4
SC-T7	Gamma/delta T-cells	RA	T4
SC-T8	Memory T-cytotoxic	RA T3, SLE T4	T4
SC-NK1	CD16 positive NK-cells	SLE	RA T4
SC-NK2	CD16neg/CD56bright NK-cells	SLE T4	
Myeloid cells			
SC-M1	Classical monocytes	RA	T3 RA, T4 SLE
SC-M2	IFN induced monocytes	SLE	T3
SC-M3	Non-classical monocytes	SLE	T3
SC-DC1	Myeloid dendritic cells		
SC-DC2	Plasmacytoid dendritic cells		
B-cells			
SC-B1	Memory B-cells		
SC-B2	Naive B-cells, CD69 positive	SLE	T3
SC-PC1	Plasma cells	SLE	T4
Artefacts			
GRAN	Granulocytes		
RBC	Red blood cells		

*Based on OR, see online supplemental table 6.
NK, natural killer; RA, rheumatoid arthritis; SLE, systemic lupus erythematosus; UMAP, uniform manifold approximation and projection.

RA samples in some cell types. Cell types with lower abundance in SLE included SC-M1 and three SC-T groups (2, 4, 7), whereas the IFN-induced SC-M2, SC-M3, SC-PC1, and SC-T3 had higher abundance in SLE. Consistent with the flow and scRNA-seq results, and in contrast to healthy and RA, SLE had a distinct surge in plasma cells (SC-PC1) post partum.

The only cell type where the estimated abundance in RA differed from healthy controls were SC-T1 post partum, where healthy had a sharp surge compared with RA and SLE. However, this difference is due to SNRA and as there are more SPRA samples this difference is not significant when the two subtypes are combined. Furthermore, SNRA and SPRA only differed significantly in the estimated proportion of SC-T7, where the proportions in SPRA and SNRA were more like healthy and SLE, respectively (online supplemental table 9).

Cell adjustment

The cell abundance analysis revealed large changes in cell-type composition throughout pregnancy and between groups (figure 4) suggesting that many genes could be identified as DEGs due to such changes alone. To identify genes whose expression patterns could be related to other factors than cell-type composition, we reanalysed the WB gene expression data while correcting for the estimated levels of the most abundant cell types. Correcting for the five cell types each estimated to contribute to more than 10% of the total cell burden (SC-T2, SC-T4, SC-B1, SC-M1, SC-M2) resulted in less than 1000 cell-type adjusted DEGs (Adj-DEGs; online supplemental table 3a, b, online supplemental figure S2B). As in the original analysis, the majority of Adj-DEGs

differed between SLE and healthy controls. In the original analysis, more than twice as many genes were found to be downregulated as upregulated in SLE compared with healthy controls, but after cell adjustment this evened out (online supplemental table 3a, online supplemental figure S2). Investigating the effects of individual cell-type adjustment on DEGs between SLE and healthy revealed that SC-M2 could explain the majority of the original DEGs over all time points, whereas SC-T2 could explain the original DEGs 6 weeks post partum (online supplemental table 3b). Finally, in contrast to DEGs between SLE and healthy, the number of DEGs between RA and healthy were similar in the original and adjusted analyses. These results are consistent with there being less differences in cell-type compositions between RA and healthy than between SLE and healthy.

The Adj-DEGs are grouped into six individual clusters (AdjG1–AdjG6) with expression patterns related to time, disease state and pregnancy (figure 5). They had a 40%–70% overlap with preadjustment clusters, except for AdjG3 with under 20% overlap with any preadjustment cluster (online supplemental table 10).

Pregnancy signature

As in the original analysis, Adj-DEGs consisted of both upregulated (AdjG3, AdjG4) and downregulated (AdjG1, AdjG6) genes, which peaked or bottomed out at T2 on average. However, AdjG3 peaked at T3 and for AdjG1 and AdjG6, RA bottomed out at T3 (figure 5A,B). Whereas no pathways were enriched for the AdjG3 genes, the AdjG4 cluster was enriched for functions related to the innate immune system, neutrophil degranulation, platelet activation and G protein-coupled receptor signalling, AdjG1 for translation pathways, including responses to amino acid deficiency, nonsense-mediated decay and selenocysteine synthesis, and AdjG6 for complement activation and heme scavenging (figure 5C, online supplemental table 4, online supplemental figure S7).

To further interpret these results, we investigated both Adj-DEGs' tissue and cell-type expression profiles in public resources (Human Protein Atlas[20]), The GTEx project,[23] and Database of Immune Cell Expression[24] and within our scRNA-seq data.

Focusing on the Adj-DEGs with the highest absolute log fold change ('strongest Adj-DEGs') showed that many of these come from smaller cell populations and may highlight biology that is less commonly described. For genes upregulated during pregnancy, the strongest Adj-DEGs were *FOLR3*, *TMEM158*, and *ST6GALNAC4* (AdjG3) and *AC073172.1*, *ANPEP*, and *ST6GALNAC2* (AdjG4) (figure 5D, online supplemental table 2). Folate receptor gamma (*FOLR3*) is expressed in neutrophils and classical monocytes and its expression is closely correlated to that of *TNFSF13B* (also known as B-cell activating factor (*BAFF*)).[20] In our analyses, *FOLR3* was differentially expressed between SNRA and SPRA patients at all time points, but in the original analyses it was grouped among the outliers (G0) as its expression profile was not found similar enough to other genes to be placed in a cluster. Transmembrane protein 158 (*TMEM158*, also known as Ras-induced senescence protein 1 (*RIS1*)) is primarily expressed in brain and endometrium but is also distinctly expressed in classical monocytes.[20] The gene takes part in the Ras signalling pathway involved in cell death and survival; a process closely linked to TLR-recognition of endogenous nuclear antigens.[27] ST6 N-acetylgalactosaminide alpha-2,6-sialyltransferase 4 (*ST6GALNAC4*) has broad tissue expression and is enriched in plasmacytoid dendritic cells (DCs) in PBMCs.[20] We found both *TMEM158* and *ST6GALNAC4* to
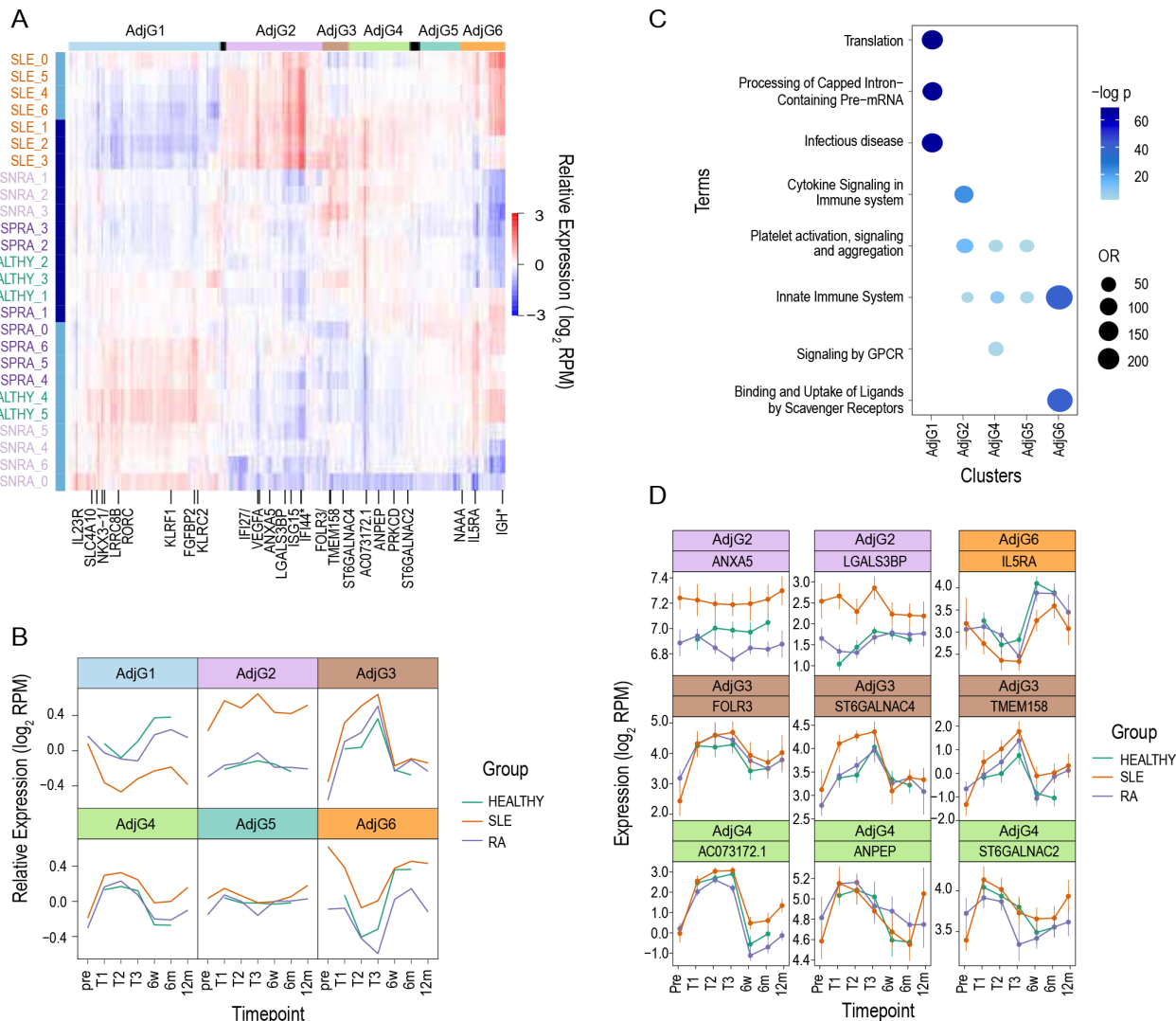
Figure 5 Differentially expressed genes after cell-type adjustment. (A) Heatmap of relative expression levels for differentially expressed genes from the cell-type adjusted whole blood mRNA-seq analysis; see figure 2A for details. IFI44 represents both *IFI44* and *IFI44L*. IGH represents *IGHA1*, *IGHA2*, *IGHG1* and *IGHG3* (B) Relative gene expression profiles shown for each individual gene cluster, labelled and colour coded as in (A); see figure 2C for details. (C) Top five significant REACTOME pathways (hypergeometric test, BH adjusted p<0.05) enriched for genes in clusters AdjG1-AdjG6; see figure 2D for details. (D) Individual gene expression profiles for nine highlighted genes in (A). BH, Benjamini-Hochberg.

be upregulated in SLE compared with healthy and SPRA in the adjusted analyses. lncRNA AC073172.1 (also known as RP11-531H8.2) is expressed in WB and has been found to have a bioinformatic association to the LPS-response process and the TLR-signalling pathway and has been hypothesised to synergise with transcription factors binding to promoters regulating sex steroid functions.[28] However, this has not been validated by laboratory experiments. We found AC073172.1 to be upregulated in SLE compared with SPRA, especially post partum. Membrane alanyl aminopeptidase (*ANPEP*) is expressed in intestine and pancreas and is enriched in neutrophils in WB.[20] Alpha-N-acetylgalactosaminide alpha-2,6-sialyltransferase 2 (*ST6GALNAC2*) has broad tissue expression and is enriched in neutrophils in WB.[20] We found both *ANPEP* and *ST6GALNAC2* to be upregulated in SPRA compared with SNRA at T0. Consistent with the pathway analyses, AdjG4 also contained other transcripts enriched in neutrophils in WB, including genes involved in steroid and lipid metabolism (*CYP27A1*), ER-stress (*MRVI1*, *UBE2J1*) and galactosylation (*GALNT14*).

For genes downregulated during pregnancy, the strongest Adj-DEGs were *RPL10P9*, *NSFP1*, and *NKX3-1* (AdjG1) and *HDC*, CLC, and *IGLV1-44* (AdjG6) (online supplemental table 2). Both ribosomal protein L10 pseudogene 9 (*RPL10P9*) and N-ethylmaleimide-sensitive factor pseudogene 1 (*NSFP1*) are pseudogenes without known functions; both have broad but low tissue expression. NK3 homeobox 1 (*NKX3-1*) is primarily expressed in prostate and salivary glands, but also shows distinct expression in eosinophils and neutrophils.[20] *NKX3-1* was down-regulated in SLE and SPRA compared with healthy and is one of several AdjG1 genes regulating the FLT1-VEGF pathway.[29] Histidine decarboxylase (*HDC*) is highly specific to basophils whereas Charcot-Leyden crystal galectin (*CLC*) is highly specific to basophils and eosinophils and is a biomarker for eosinophilic and type two inflammation.[30] Both HDC and CLC were upregulated in SPRA compared with SNRA at T0. Immunoglobulin lambda variable 1–44 (*IGLV1-44*) is expressed in B-cells and plasma cells and was one of 31 differentially expressed IGL-chain and IGH-chain transcripts—all of which were grouped in

AdjG6, comparable to G9 in the original analysis. Other genes downregulated during pregnancy (in AdjG1) included genes expressed by innate cytotoxic cells (G4 and G5 in the original analysis).

SLE signature

There were 519 Adj-DEGs differing between SLE and healthy controls (online supplemental tables 2 and 3, online supplemental figures S8 and S9), the majority of which were either downregulated during pregnancy (AdjG1; 231 genes) or generally upregulated compared with healthy and RA (AdjG2; 154 genes). As in the original analyses, genes up-regulated in SLE (AdjG2, comparable to G11) were highly enriched for IFN-signalling functions (online supplemental table 4). Focusing on the genes unique to the adjusted analysis indicated that these genes were mainly expressed in monocytes, plasma cells and DCs (online supplemental figure S10). As the analysis was adjusted for changes in monocyte levels, this indicates that these genes could be dysregulated in monocytes of patients with SLE, although we cannot exclude effects of differences in abundance of other cell types not included in the adjustment.

RA signature

After adjustment, 27 genes were dysregulated in RA compared with healthy controls (online supplemental figure S11, online supplemental tables 2 and 3), including the eight genes from the original analysis. The majority of the genes (20/27) were downregulated in RA overall or in the third trimester compared with healthy controls.

Several of the genes were also downregulated in SLE, showing a common disease profile (*EIF3G, NKX3-1, RPS26*). For some genes, however, the expression profile for SLE mostly followed the healthy controls while expression levels for RA differed (*NBPF26, LRRC8B, DEF6, CNR1*). For one gene (*NAAA,* primarily expressed in monocytes) the expression level was upregulated in RA but downregulated in SLE.

Most of the genes had decreased expression during pregnancy and 15 of the 27 genes clustered in AdjG1 (online supplemental figure S11, online supplemental table 2). Based on our single cell data these were generally expressed in T-cells and NK-cells (online supplemental figure S12) and are involved in proliferation, adaptive T-cell response and homoeostasis. Three of these genes are specific to or associated with MAIT cells (*GZMK, NCR3, SLC4A10*).[20] Several genes also have an important role in placental development (*NKX3-1, SNHG16, NRCAM*).[29 31 32] For the other clusters, the genes were either non-specific or found in the other immune cells involved in basic immunological processes (antigen binding/lysis) and tolerance. Dysregulation of *DEF6* (AdjG6) has been linked to autoimmunity[33] and *TAFA2* might also be an immune regulator.[34] The divergence between SNRA and SPRA at T0 and T6, with opposite trajectories during pregnancy (online supplemental figure S13) was also seen after adjustment.

DISCUSSION

We provide data from women followed at a rheumatological outpatient clinic connected to a national competence centre and with universal accessibility to healthcare.[35] The gene expression differences seen in our cohort first reflect patients with SLE differing from other participants and second reflect pregnancy samples differing from prepartum and postpartum samples. We also find clear differences between patients with RA and healthy. A summary of our main findings can be found in online supplemental table 13. No gene expression differences could be attributed to differences in medication or disease activity (online supplemental figure S2).

In healthy pregnancies systemic response to foreign antigens is reduced[36 37] as part of a shift towards tolerance. This immune signature is evident in our primary clusters G13 and G9 and adjusted clusters AdjG3 and AdjG4. In the adjusted analyses, part of this immune signature is found as transcripts involved in galactosylation; a post-translational protein modification altering the binding capacity and effector functions of immunoglobulins. In our dataset, these transcripts are increased in pregnancy, with a higher overall expression in SLE. Others have shown an increase in galactosylation of IgG during pregnancy in healthy and RA. An attenuation of this increase has been associated with recurrent pregnancy loss. In non-pregnant patients with RA, decreased IgG galactosylation is seen with increased disease activity and inflammation.[8]

As the humoral immune response is repressed, there is a tight balance in maintaining the pregnant immune system's defence against bacterial agents and tolerance of the fetus.[38] We find lower expression of transcripts enriched in specialised immune cells known to accumulate in the decidua and help keep this balance (Treg, MAIT, NKbright, gdT). Our findings are consistent with evidence suggesting recruitment from peripheral blood to the decidua by chemotactic factors produced by the placenta.[39–41] This mechanism may contribute to the disease-modifying effect of pregnancy and needs to be explored in further studies.

The differences found between patients with SLE and the other participants are mainly driven by the SLE IFN-signature, which was not counteracted by pregnancy. Expression of interferon related genes declined during pregnancy in healthy controls, but increased slightly in RA. In our material both the pDCs (SC-DC2) and ISG15-secreting plasma cells (SC-PC1) were more abundant in SLE than in RA in the PBMC samples, but these smaller subpopulations could not be described and verified in the larger mRNA cohort through our cell abundance estimation. We show an expansion of IFN-regulated T-cells (SC-T3) and monocytes (SC-M2) and non-classical monocytes (SC-M3) in SLE compared with RA and healthy controls. These are the main cell populations contributing to the IFN-profile. In pregnancy, there is a partial reversal of the expansion of non-classical monocytes (SC-M3), while the IFN-regulated monocytes (SC-M2) become even more dominant in the cellular composition of patients with SLE. Others have shown a similar pattern, demonstrating how an increase in monocytes, together with suppression of T-effector cells, are part of the systemic adaptation to gestation in general.[10]

The gene expression profile of SLE in pregnancy did not return to prepregnancy levels 6 weeks, 6 months or 12 months post partum (T4–T6). Although a direct association to disease activity was not detected, our finding is in line with research from the RevNatus registry, showing higher disease activity at 6 and 12 months post partum[6] and highlighting the need for tight follow-up and disease control in the first year after delivery. Among the most upregulated genes, particularly after pregnancy for patients with SLE (G14, T4), are lncRNAs. Although this might reflect general cellular turnover and nuclear activity, these lncRNAs may participate in the pathogenesis of SLE through endogenous TLR-responses and IFN I activation.[42] Among the most downregulated genes were metalloproteinases, one of them being ADAM12, which is known to be produced in the syncytiotrophoblast and is an important factor for angiogenesis and placental growth. Reduction of ADAM12 is associated with pre-eclampsia and thromboembolic complications.[43 44]

Gene expression patterns in patients with RA were closer to healthy controls than to patients with SLE. This agrees with previous findings.[45 46] Moreover, in contrast to SLE, the estimated cell type compositions in RA were similar to those in healthy (p>0.05). However, we show that cell-type adjustment is important to elucidate the RA signature, which indicates that RA to a higher extent than SLE, appears to be driven by cellular dysregulation, rather than altered cell-type composition. The majority of RA-signature genes are expressed in T-cells and NK-cells, and several were associated with MAIT cells. This association was not visible until after adjustment. Dysregulation within blood and synovial fluid MAIT cells in patients with RA have also been found by others,[47] but it has not been studied extensively. Furthermore, we found the RA-signature genes were mainly downregulated. We did identify differences between SNRA and SPRA, but there were too few samples to verify these thoroughly. In the future, larger cohorts of the RA subgroups are needed.

Single cell library preparation is still an expensive procedure where only a few samples can be prepared simultaneously. To mitigate effects of individual variation, samples were pooled, generating an average pattern rather than cell population distributions per individual. In our protocol, 99% of the cells in the sequenced samples were living cells (online supplemental figure S14). However, we cannot control for possible differences in survival between cell types during sampling and preparation and this could potentially skew cell clustering. The cell populations not detected trough cell abundance estimation were small and therefore not relevant for the adjustment procedure. However, confirmation of the fluctuations in these smaller populations could not be confirmed in the larger cohort. It has been shown that T-cells and NK-cells demand a high number of cells in order to produce a high classification accuracy.[48 49] The major NK clusters were visible in our UMAP, although only one was annotated by the software. This is also true for the T-cell populations, where several visible clusters were not annotated. Combining the software's annotation with a manual layer provided results similar to the reference studies and indicates that our clustering is sufficient. Currently, no consensus on reference materials for guided UMAP clustering for immunological cells exist, making it challenging to compare studies. Fortunately, data are rapidly accumulating, and our results from PBMCs will be a part of this.

Using scRNA-seq analysis to guide handling of bulk mRNA data provided a robust approach. Our cell-type adjusted dataset solidified patterns seen in the unadjusted dataset and removed some of the signal connected to cellular metabolism and turnover. It also highlighted the signal from new candidate genes and smaller immune cell populations, like MAIT cells' role in the RA-signature. We, therefore, see the cell-adjusted dataset as a valuable added resource in expanding the knowledge both for cell type and gene expression changes in healthy and rheumatic pregnancy.

Author affiliations
[1]Department of Clinical and Molecular Medicine, Norwegian University of Science and Technology, Trondheim, Norway
[2]Norwegian National Advisory Unit on Pregnancy and Rheumatic Diseases, Department of Rheumatology, St. Olavs hospital, Trondheim University Hospital, Trondheim, Norway
[3]Genomics Core Facility, HF, Sentral Stab, St. Olavs hospital, Trondheim University Hospital, Trondheim, Norway
[4]BioCore - Bioinformatics Core Facility, HF, Sentral stab, St. Olavs hospital, Trondheim University Hospital, Trondheim, Norway
[5]Department of Computer Science, Norwegian University of Science and Technology, Trondheim, Norway
[6]Department of Immunology and Transfusion Medicine, St. Olavs hospital, Trondheim University Hospital, Trondheim, Norway

Competing interests None declared.

Patient and public involvement Patients and/or the public were involved in the design, or conduct, or reporting, or dissemination plans of this research. Refer to the Methods section for further details.

Patient consent for publication Not applicable.

Ethics approval This study involves human participants and this study was approved by the regional ethics committee REK-nr: 2015/2014-1 and the Scientific board of RevNatus, the Norwegian multicentre prospective quality registry, following women with inflammatory rheumatic disease from preconception until 1-year post partum. The biobank has been approved by REK (REK 2012/1044). RevNatus registry has been approved by Datatilsynet. The research followed the Declaration of Helsinki for research ethics. Participants gave informed consent to participate in the study before taking part.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Data are available in a public, open access repository. Our data can be found at Gene Expression Omnibus repository (GEO), accession number GSE235508.

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.

ORCID iD
Hilde Julie T Lien http://orcid.org/0009-0008-6256-6188

REFERENCES

1 Eyre S, Orozco G, Worthington J. The genetics revolution in rheumatology: large scale genomic arrays and genetic mapping. *Nat Rev Rheumatol* 2017;13:421–32.

2 Baechler EC, Batliwalla FM, Karypis G, *et al*. Interferon-inducible gene expression signature in peripheral blood cells of patients with severe lupus. *Proc Natl Acad Sci U S A* 2003;100:2610–5.

3 Mavragani CP, La DT, Stohl W, *et al*. Association of the response to tumor necrosis factor antagonists with plasma type I interferon activity and interferon-beta/alpha ratios in rheumatoid arthritis patients: a post hoc analysis of a predominantly hispanic cohort. *Arthritis Rheum* 2010;62:392–401.

4 Raterman HG, Vosslamber S, de Ridder S, *et al*. The interferon type I signature towards prediction of non-response to rituximab in rheumatoid arthritis patients. *Arthritis Res Ther* 2012;14:R95.

5 Crow MK, Ronnblom L. Type I interferons in host defence and inflammatory diseases. *Lupus Sci Med* 2019;6:e000336.

6 Götestam Skorpen C, Lydersen S, Gilboe I-M, *et al*. Disease activity during pregnancy and the first year postpartum in women with systemic lupus erythematosus. *Arthritis Care Res (Hoboken)* 2017;69:1201–8.

7 Ostensen M, Brucato A, Carp H, *et al*. Pregnancy and reproduction in autoimmune rheumatic diseases. *Rheumatology (Oxford)* 2011;50:657–64.

8 Förger F, Villiger PM. Publisher correction: immunological adaptations in pregnancy that modulate rheumatoid arthritis disease activity. *Nat Rev Rheumatol* 2020;16:184.

9 Sumitomo S, Nagafuchi Y, Tsuchida Y, *et al*. Transcriptome analysis of peripheral blood from patients with rheumatoid arthritis: a systematic review. *Inflamm Regen* 2018;38:21.

10 Hong S, Banchereau R, Maslow B-SL, *et al*. Longitudinal profiling of human blood transcriptome in healthy and lupus pregnancy. *J Exp Med* 2019;216:1154–69.

11 Jaitin DA, Kenigsberg E, Keren-Shaul H, *et al*. Massively parallel single-cell RNA-Seq for marker-free decomposition of tissues into cell types. *Science* 2014;343:776–9.

12 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.

13 Stubbington MJT, Rozenblatt-Rosen O, Regev A, *et al*. Single-cell transcriptomics to explore the immune system in health and disease. *Science* 2017;358:58–63.

14 Maecker HT, McCoy JP, Nussenblatt R. Standardizing immunophenotyping for the human immunology project. *Nat Rev Immunol* 2012;12:191–200.

15 Pitoiset F, Cassard L, El Soufi K, *et al*. Deep phenotyping of immune cell populations by optimized and standardized flow cytometry analyses. *Cytometry A* 2018;93:793–802.

16 van Dongen JJM, van der Burg M, Kalina T, *et al*. Euroflow-based flowcytometric diagnostic screening and classification of primary immunodeficiencies of the lymphoid system. *Front Immunol* 2019;10:1271.

17 Newman AM, Steen CB, Liu CL, *et al*. Determining cell type abundance and expression from bulk tissues with digital cytometry. *Nat Biotechnol* 2019;37:773–82.

18 Nilsen G, Borgan O, Liestøl K, *et al*. Identifying clusters in genomics data by recursive partitioning. *Stat Appl Genet Mol Biol* 2013;12:637–52.

19 Kolberg L, Raudvere U, Kuzmin I, *et al*. Gprofiler2 -- an R package for gene list functional enrichment analysis and namespace conversion toolset G:profiler. *F1000Res* 2020;9:ELIXIR-709.

20 The human protein Atlas. Version: 22.0. n.d. Available: https://www.proteinatlas.org

21 de Lima DS, Cardozo LE, Maracaja-Coutinho V, *et al*. Long noncoding RNAs are involved in multiple immunological pathways in response to vaccination. *Proc Natl Acad Sci U S A* 2019;116:17121–6.

22 Itakura S, Ohno K, Ueki T, *et al*. Expression of golf in the rat placenta: possible implication in olfactory receptor transduction. *Placenta* 2006;27:103–8.

23 GTEx Portal. The genotype-tissue expression (GTEx) project version: V8. n.d. Available: https://gtexportal.org/home

24 The DICE (database of immune cell expression, expression quantitative trait Loci (eQTLs) and Epigenomics) project. 2014. version. 2022. Available: https://dice-database.org/

25 Kim H, de Jesus AA, Brooks SR, *et al*. Development of a validated interferon score using nanostring technology. *J Interferon Cytokine Res* 2018;38:171–85.

26 Chaussabel D, Quinn C, Shen J, *et al*. A modular analysis framework for blood genomics studies: application to systemic lupus erythematosus. *Immunity* 2008;29:150–64.

27 Schmit K, Michiels C. TMEM proteins in cancer: a review. *Front Pharmacol* 2018;9:1345.

28 Ma Y, Ma L, Cao Y, *et al*. Construction of a ceRNA-based lncRNA-mRNA network to identify functional lncRNAs in polycystic ovarian syndrome. *Aging (Albany NY)* 2021;13:8481–96.

29 Murthi P, Brouillet S, Pratt A, *et al*. An EG-VEGF-dependent decrease in homeobox gene NKX3.1 contributes to cytotrophoblast dysfunction: a possible mechanism in human fetal growth restriction. *Mol Med* 2015;21:645–56.

30 Aegerter H, Smole U, Heyndrickx I, *et al*. Charcot-leyden crystals and other protein crystals driving type 2 immunity and allergy. *Curr Opin Immunol* 2021;72:72–8.

31 Marcellin L, Santulli P, Gogusev J, *et al*. Endometriosis also affects the decidua in contact with the fetal membranes during pregnancy. *Hum Reprod* 2015;30:392–405.

32 Yu Z, Zhang Y, Zheng H, *et al*. LncRNA SNHG16 regulates trophoblast functions by the miR-218-5p/LASP1 axis. *J Mol Histol* 2021;52:1021–33.

33 Serwas NK, Hoeger B, Ardy RC, *et al*. Human DEF6 deficiency underlies an immunodeficiency syndrome with systemic autoimmunity and aberrant CTLA-4 homeostasis. *Nat Commun* 2019;10:4555.

34 Tom Tang Y, Emtage P, Funk WD, *et al*. TAFA: a novel secreted family with conserved cysteine residues and restricted expression in the brain. *Genomics* 2004;83:727–34.

35 Meissner Y, Strangfeld A, Costedoat-Chalumeau N, *et al*. European network of pregnancy registers in rheumatology (Eunep)-an overview of procedures and data collection. *Arthritis Res Ther* 2019;21:241.

36 Gomez-Lopez N, Romero R, Hassan SS, *et al*. The cellular transcriptome in the maternal circulation during normal pregnancy: a longitudinal study. *Front Immunol* 2019;10:2863.

37 Lima J, Cambridge G, Vilas-Boas A, *et al*. Serum markers of B-cell activation in pregnancy during late gestation, delivery, and the postpartum period. *Am J Reprod Immunol* 2019;81:e13090.

38 Raffetseder J, Lindau R, van der Veen S, *et al*. MAIT cells balance the requirements for immune tolerance and anti-microbial defense during pregnancy. *Front Immunol* 2021;12:718168.

39 Kaipe H, Raffetseder J, Ernerudh J, *et al*. MAIT cells at the fetal-maternal interface during pregnancy. *Front Immunol* 2020;11:1788.

40 Tilburgs T, Roelen DL, van der Mast BJ, *et al*. Evidence for a selective migration of fetus-specific CD4+CD25bright regulatory T cells from the peripheral blood to the decidua in human pregnancy. *J Immunol* 2008;180:5737–45.

41 Wu X, Jin L-P, Yuan M-M, *et al*. Human first-trimester trophoblast cells recruit CD56brightCD16- NK cells into decidua by way of expressing and secreting of CXCL12/stromal cell-derived factor 1. *J Immunol* 2005;175:61–8.

42 Crow MK, Olferiev M, Kirou KA. Identification of candidate predictors of lupus flare. *Trans Am Clin Climatol Assoc* 2015;126:184–96.

43 El-Sherbiny W, Nasr A, Soliman A. Metalloprotease (ADAM12-S) as a predictor of preeclampsia: correlation with severity, maternal complications, fetal outcome, and doppler parameters. *Hypertens Pregnancy* 2012;31:442–50.

44 Spencer K, Cowans NJ, Stamatopoulou A. ADAM12s in maternal serum as a potential marker of pre-eclampsia. *Prenat Diagn* 2008;28:212–6.

45 Pathi A, Wright M, Smed MK, *et al*. The rheumatoid arthritis gene expression signature among women who improve or worsen during pregnancy: a pilot study. *J Rheumatol* 2021;48:985–91.

46 Goin DE, Smed MK, Pachter L, *et al*. Pregnancy-induced gene expression changes in vivo among women with rheumatoid arthritis: a pilot study. *Arthritis Res Ther* 2017;19:104.

47 Koppejan H, Jansen DTSL, Hameetman M, *et al*. Altered composition and phenotype of mucosal-associated invariant T cells in early untreated rheumatoid arthritis. *Arthritis Res Ther* 2019;21:3.

48 Ding J, Smith SL, Orozco G, *et al*. Characterisation of CD4+ T-cell subtypes using single cell RNA sequencing and the impact of cell number and sequencing depth. *Sci Rep* 2020;10:19825.

49 Crinier A, Milpied P, Escalière B, *et al*. High-dimensional single-cell analysis identifies organ-specific signatures and conserved NK cell subsets in humans and mice. *Immunity* 2018;49:971–86.