SUPPLEMENTARY METHODS

Patients and healthy control subjects

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 one year postpartum. The RevNatus registry provides data on demographic variables, disease activity, medication, laboratory status, pregnancy outcome, self-reported health status and lactation [1]. Analyses were performed on 335 whole blood samples collected from a total of 84 SLE, RA, and healthy control pregnancies. Patients were recruited from 2015 to 2019 at the rheumatology outpatient clinic connected to the Norwegian National Advisory Unit on Pregnancy and Rheumatic Diseases (NKSR), St. Olavs Hospital. NKSR collaborate with patient representatives according to "The European League Against Rheumatism's recommendations for the inclusion of patient representatives in research". A clinical diagnosis of RA or SLE was determined by a rheumatologist, using the International Statistical Classification of Disease and Related health Problems, 10th Revision (ICD-10). Disease activity was assessed by Disease Activity Score-28–C-reactive protein 3 (DAS28-CRP3) for RA [2] and the Lupus Activity Index in Pregnancy (LAI-P) for SLE [3]. Healthy women (n=21) were recruited early in the first trimester of their pregnancy from local general practitioners' offices or via the St. Olavs hospitals website. Three of these pregnancies registered adverse pregnancy outcomes (APO) defined as: Preterm birth (<week37), low birthweight (<2500g), eclampsia or preeclampsia. Similar numbers were seen in the disease groups (Table 1). None of the APO were due to either premature rupture of membranes (PROM), infection or genetic aberrations. Data from pregnancies with APO were not excluded from further analyses. A control analysis was performed where data from the healthy who registered with PE were excluded, and no significant changes to the results were observed. Samples were further processed and stored by Biobank1 ®. All participants gave informed consent and could withdraw at any time. Samples from participants who withdrew, had a miscarriage, or did not get pregnant are included in the analysis.

Gene expression profiles from whole blood

Processing of blood samples

Peripheral whole blood samples were collected in PAX-gene RNA tubes (Qiagen) and stored at Biobank1 ®. Total RNA was extracted using PAX-gene blood RNA-kits (Qiagen). RNA concentration was measured using Qubit® RNA HS Assay Kit on a Qubit® 3.0 Fluorometer (Thermo Fisher Scientific Inc., Waltham, MA, USA). Integrity was assessed using Agilent RNA 6000 Pico Kit on a 2100 Bioanalyzer instrument (Agilent Technologies, Santa Clara, CA, USA). All analyses were done according to the manufacturers' instructions.

Library preparation and mRNA sequencing (Illumina TruSeq Stranded Total RNA with Ribo-Zero Globin)

RNA sequencing libraries were prepared using Illumina TruSeq Stranded Total RNA with Ribo-Zero Globin kit (Illumina, San Diego, CA, USA) according to the manufacturer's instructions. In brief, ~1300±500 ng total RNA was used as starting material. The first step involved the removal of globin-encoding mRNA in addition to ribosomal RNA (rRNA), using biotinylated target-specific oligos combined with Ribo/Globin-Zero RNA removal beads. Following purification, the RNA was fragmented using divalent cations at 94°C for 3 min. First and second strand cDNAs were synthesized using random oligonucleotides and SuperScript II, followed by DNA polymerase I and RNase H. Exonuclease/polymerase was used to produce blunted overhangs. Illumina SR adapter/index oligonucleotides were ligated to the cDNA after 3' end adenylation. DNA fragments were enriched by

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

13 cycles of PCR reaction. The libraries were purified using the AMPure XP (Beckman Coulter, Inc., Indianapolis, IN, USA), quantitated by qPCR using KAPA Library Quantification Kit (Kapa Biosystems, Inc., Wilmington, MA, USA) and validated using Agilent High Sensitivity DNA Kit on a Bioanalyzer (Agilent Technologies, Santa Clara, CA, USA). The size range of the DNA fragments were measured to be in the range of 220-1100 bp and peaked around 320 bp.

Prior to sequencing, the libraries were quantified (KAPA Library Quantification Kit (Illumina/ABI Prism), normalized and pooled. Quantitated libraries were further diluted to 3 nM and subject to clustering by a cBot Cluster Generation System on four HiSeq4000 flowcells (Illumina Inc. San Diego, CA, USA), according to the manufacturer's instructions. Finally, single end sequencing was performed for 75 cycles on an Illumina HiSeq4000 instrument, in accordance with the manufacturer's instructions (Illumina, Inc., San Diego, CA, USA). FASTQ files were created with bcl2fastq V2.20 (Illumina, Inc., San Diego, CA, USA).

FASTQ files were quality controlled with fastqc (v0.11.9) then filtered and trimmed by fastp (v0.20.0). Trimmed sequences were aligned to the genome reference using STAR (v2.7.3) and quality metrics were extracted with picard CollectRNASeqMetrics (v2.21.5).

Normalisation, linear modelling, and differential gene expression

Transcript counts were generated using quasi alignment (Salmon v1.3.0) to the GRCh38 transcriptome reference sequences. Transcript counts were imported into the R statistical software and aggregated to gene expression counts using the tximport (v1.14.0) Bioconductor package for downstream statistical analysis.

edgeR [4] was used to scale the raw library sizes, by calculating normalisation factors, and to create a DGEList object to store the gene expression count matrix. Gene expression counts were normalised and analysed for differential expression through the *limma* package [5]. Briefly, we modelled the gene expression as a linear model containing separate coefficients for each combination of disease status (HEALTHY, SLE, SNRA, SPRA) and sample time (T0-T6), corrected for the global intra-individual gene expression correlation. Specifically, based on two variables describing sample time, given as an integer between 0-6, and sample group, given as HEALTHY, SNRA, SPRA or SLE, we created a combined GroupTime factor variable encoding all 26 combinations of group and time in our data (Figure 1B, Figure 2A); HEALTHY_1 and SLE_6 are example values representing samples from healthy donors at T1 and SLE patients at T6, respectively. The design matrix was then created by the following call to model.matrix():

```
design <- model.matrix(~0 + GroupTime)
```

Normalisation was done through voom() from the *limma* package [6], transforming the count matrix from counts per million (cpm) to the logarithmic domain. Normalisation was followed by removal of all genes with an average cpm less than one. The intra-individual gene expression correlation was estimated by duplicateCorrelation(), using the patient identifier as the block variable. lmFit() was used to fit the linear model.

Contrasts were created by makeContrasts() such that sample groups were compared overall and at each individual timepoint (Supplementary Table 11). Overall group comparisons were achieved by subtracting the two groups' mean estimates across all timepoints. Individual timepoint comparisons were achieved by subtracting the two groups' estimates at the same timepoint. Empirical Bayes smoothing was used on the standard errors and empirical Bayes moderated t-statistics (the eBayes() function) were used to calculate significance of the differentially expressed genes (DEGs). P-values were adjusted for multiple testing for each contrast; genes with an Benjamini-Hochberg adjusted p-

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

value <0.05 were used for further analysis and clustering. Significant differentially expressed genes were examined for tissue and cell type expression through Human Protein Atlas [22, 23]), The Genotype-Tissue Expression (GTEx) project [24, 25], and Database of Immune Cell Expression [26, 27].

Gene clustering and functional enrichment analysis

Genes significant in at least one contrast were clustered based on their average expression within the samples of each GroupTime. Each gene's expression was normalized by subtracting its mean expression prior to clustering; values higher than 3 or lower than -3 were set to 3 or -3 respectively. GroupTimes were clustered by hierarchical clustering (the hclust() function in R); genes were clustered through the *clusterGenomics* package (the part() function with minSize=40, minDist = 0.5 and B=100), which identifies significant subclusters within the data [7]. Each gene cluster was further analysed by functional enrichment analysis, through the *gProfiler2* package to find over-representation of biological terms and pathways [8]. In addition to the default statistics, a calculation of Odds ratio, with the genes from our filtered normalised gene count matrix as background, was added to the analysis. Terms with a term size over 2000 were removed.

Differential gene expression related to disease activity or medication

The effect of disease activity on gene expression was modelled by combining the GroupTime variable with the disease-specific activity values, Disease Activity Score-28 (DAS28) or Lupus Activity Index (in Pregnancy) (LAI(P)), for RA and SLE, respectively. Disease scores were discretized into no (Das28<=2.6), low (Das28 2.7 – 3.2), and moderate activity (Das28 3.3 -5.1) for RA and no (LAI(P)=0) and low activity (LAI(P) >0-0.5) for SLE. The highest LAI(P) score is 0.38 so discretion into moderate activity does not exist for SLE patients.

Similarly, the effect of disease activity combined with medication was modelled by combining the GroupTime variable with disease-specific medication variables capturing four combinations of disease activity and medication (see Figure S3).

The resulting factor variables, GroupTimeDas28, GroupTimeLAIP, and GroupTimeMD, encoded all combinations of group, time, and disease-specific activity or medication values. SLE and RA were analysed separately for disease activity by subsetting the count matrix on the respective samples and creating disease-specific design matrices by the following calls to model.matrix:

```
design.das28 <- model.matrix(~0 + GroupTimeDas28)
design.laip <- model.matrix(~0 + GroupTimeLAIP)
```

Contrasts comparing different overall disease activity levels were created by makeContrasts() by subtracting the mean estimates for different activity levels across all timepoints. Normalization and statistical analyses were then done as described above.

SLE and RA were jointly analysed for medication by using the following to model.matrix to create a design matrix:

```
design.MD <- model.matrix(~0 + GroupTimeMD)
```

Contrasts comparing different overall disease activity and medication levels were created by makeContrasts() by subtracting the mean estimates for different activity and medication levels across all timepoints. Normalization and statistical analyses were then done as described above.

Calculation of Interferon score

To calculate interferon scores, genes from two gene sets were chosen; genes in a 28 gene NanoString validated interferon score[9] and genes from modules M1.2 M3.4 and M5.12 from the module-based

interferon scoring method [10]. We used genes for the modules as given by Lambers, et al. (2019)[11] Genes from these sets found significant in our data are listed in Supplementary Table 5a-d. Expression from our normalised count matrix was summarised per sample for all significant genes for both gene sets. Module genes were both summed as a total of all modules and in separate submodules. Mean and Standard deviation per score were then calculated from all samples pr GroupTime. Tukey's range test was used to find significant fluctuation differences over time and between groups.

Single cell RNA sequencing (scRNA-seq) and Flow cytometry of PBMCs

Processing of blood samples

Peripheral whole blood samples (PBMC) were collected in VACUETTE® NH sodium heparin tubes (Greiner bio-one, Austria). PBMCs were isolated using the density gradient Lymphoprep™ following the manufacturer's protocol. Isolation was followed by resuspension in 1mL RPMI (50% FBS), and 1 mL freezing media (20% DMSO, 40% FBS) was then added dropwise. The sample was split in two 1 mL cell tubes and frozen in a Corning® CoolCell™ LX Cell Freezing Container (Corning, Inc.) at -80 degrees for 24 hours before being transferred to liquid nitrogen freezer tanks until sequencing. For single cell sequencing samples from 4 RA patients and 4 SLE patients were thawed. Two samples from each patient were used, one from the 3rd trimester (T3) and one from 6 weeks postpartum (T4). The samples were transferred from liquid nitrogen to a water bath (37^{o}C) and thawed for 10 minutes. Thawed samples were transferred to falcon tubes containing 8 mL warm RPMI and the cell tubes were washed with 1 mL RPMI which were then also added to the falcon tube. The samples were washed twice by centrifugation at 500g for 10 minutes at room temperature (RT) and resuspended in 2 mL RPMI. Counting was performed on the MOXI Z Mini Automated Cell Counter (ORFLO Technologies) following the manufacturer's protocol. 25 000 cells from each of the 8 patients were pooled for each disease timepoint, generating 4 pooled samples for sequencing (Figure 1B). Flow cytometric analyses were done in parallel on single samples with material from the same sample vials.

scRNA-seq

Single-cell RNA capture and library preparation were performed from four cell suspensions, using the Chromium Next GEM Single Cell 3' cDNA Kit v3 (10X Genomics, CA; USA), according to the manufacturer's instructions.

In brief, the single-cell suspensions were counted using the Countess automated cell counter (Thermo Fisher Scientific, CA, USA). After counting, cells were loaded onto the 10X Chromium chip using the recommended loading volumes and run on the 10X Chromium V3 controller (10X Genomics, CA; USA). Immediately following GEM generation on the chip, the Gel Bards were dissolved, primers released, and cells lysed, Enabling production of barcoded, full length cDNAs through a cDNA synthesis step. After cDNA synthesis, the GEWs were broken, and pooled fractions recovered. Barcoded, full-length cDNAs were amplified via 14 PCR cycles, to generate sufficient mass for library construction. Followed by an AMPure XP beads (Beckman Coulter, Inc., IN, USA) clean up step. Enzymatic fragmentation and size selection were used to optimize cDNA amplicon size. P5, P7, i7 sample index, and TruSeq read2 (read2 primer sequence) were added via end repair, A-tailing, Adaptor Ligation, and PCR. The final libraries were purified using the AMPure XP (Beckman Coulter, Inc., Indianapolis, IN, USA). Quantitated by qPCR using KAPA Library Quantification Kit (Kapa Biosystems, Inc., Wilmington, MA, USA) and validated using Agilent High Sensitivity DNA Kit on a Bioanalyzer (Agilent Technologies, Santa Clara, CA, USA).

Libraries were normalized and pooled to 2.3 pM and subjected to clustering on a NextSeq500 high output flowcell (Illumina, San Diego, CA, USA). Finally, sequencing was performed for 92 cycles (using the following read geometry: 28 bp for read 1 (cell barcode and UMI), 8 bp for read 2 (sample index)

and 56 bp for read 3 (transcript) on a NextSeq550 instrument (Illumina, Inc. San Diego, CA, USA), according to the manufacturer's instructions. Base calling was done on the NextSeq 500 instrument by RTA 2.4.6. FASTQ files were generated using bcl2fastq2 Conversion Software v2.20.0.422 (Illumina, Inc. San Diego, CA, USA).

Flow cytometry

Panels for deep immunophenotyping of B-and T cell subpopulations were adapted from EuroFlow, TRANSIMMUNOM and the Human Immunology Project and optimised at the department of immunology and transfusion medicine St.Olavs hospital [12-14]. Antibodies used are described in supplementary table 1. The tubes chosen were designed to describe B-cell subpopulations, T-regulatory cells and CD4 positive T-cell polarisation. Gating was performed according to gating strategies described in reference publications [12-14], using the software Infinicyt (Cytognos) (Supplementary figure S1). Gating was performed on merged files as a bulk analysis, and the gates for each file were then controlled manually. An average of 790 000-870 000 cells were analysed per tube per sample, giving a lower limit of detection (LLOD) for populations of 2,5E-5/0,0025% of total PBMCs.

Gene expression profiles from single cell sequencing of PBMC

Single cell sequencing Cell Ranger/UMAP/Cell clustering

Alignment and gene expression quantification of samples was performed using the STARsolo option of the STAR (v2.7.3) aligner against the GRCh38 reference genome. Acceptable cell barcodes were downloaded from the 10x Genomics website (3M-february-2018.txt) and barcodes representing a droplet with missing cells were classified using the EmptyDrops R package. The gene expression count matrix was imported to scanpy (v1.5.1) for downstream analysis. Cell doublets were identified with Scrublet [15] and removed. Cells with read count between 1000-25000, mitochondrial gene fraction < 20%, and >400 genes expressed was kept for further analysis, and genes detected in <20 cells were filtered out. The gene expression count matrix was normalized to total size factors derived by the computeSumFactors() method of the *scran* R package and log-transformed. The 3000 most highly variable genes were used for Principal Component Analysis (PCA) and the top 20 principal components were used as input to the uniform manifold approximation (UMAP) method and clustered using the louvain method [16]. Marker genes between clusters were identified with the non-parametric Wilcoxon rank-sum test using the scanpy function "rank_genes_groups". The derived results and data were exported to the UCSC Cellbrowser [17] for further visual aid and analysis. The UMAP showed clustering that was not annotated by the software, but manually by genetic markers (Figure 3A, supplementary figure S4). The predominant cell type and short name for each cluster is given in Table 3. The clusters are partially overlapping, but no cell is annotated to more than one cluster.

Cell abundance estimations

We used the scRNA-seq data and CIBERSORTx to impute cell type fractions in the WB RNA-seq data[18]. Prior to creating the CIBERSORTx Signature Matrix, results from the UMAP analysis were cross-referenced with the processed scRNA-seq data to label each cell with its assigned cell type. As the resulting matrix was too big for CIBERSORTx, cells from cell types containing more than 1500 assigned cells were randomly subsampled such that each cell type had a maximum of 1500 cells. The resulting subsampled matrix was used as the Signature Matrix input and the WB RNA-seq count matrix as the mixture file input to CIBERSORTx. The number of permutations for significance analysis was set to 100.

The total percentage of each cell fraction, across all the samples, were calculated from the CIBERSORTx results (Supplementary Table 9). Tukey’s range test was used to find significant fluctuation differences

over time and between groups (Supplementary Table 8). CIBERSORTx cell type estimates from samples sequenced by scRNA-seq (Supplementary table 12) were compared with the scRNA-seq results by computing the average ratio per cell type per sample pool and correlating these averages with the scRNA-seq cell type estimates. One of the 16 samples used in single cell sequencing had no corresponding sample in bulk mRNA analysis as this sample failed during isolation. For RA in T3 we therefor used 3 samples to create average estimates. Pearson's product-moment correlation was used for computing correlations.

Cell adjustment

Cell types estimated by CIBERSORTx to contribute on average with more than 10% of the total cellular burden were extracted from the CIBERSORTx results. The sample-specific estimates from the resulting five cell types (CT1-CT5) were added to the linear model design matrix, in decreasing order of abundance, by the following call to model.matrix():

```
design <- model.matrix(~0 + GroupTime + CT1 + CT2 + CT3 + CT4 + CT5)
```

Subsequent normalization, statistical analyses, sample and gene clustering, and functional enrichment analyses were then done as described above. To investigate the effects of individual cell type adjustments, we also ran separate analyses where we iteratively added the five cell type estimates to the design matrix (i.e., CT1, CT1 + CT2, and so on), and for each iteration, ran the statistical analyses and quantified the number of significant genes per contrast.

References

1. 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(1):241. PubMed ID: PMC6854642
2. Prevoo ML, van 't Hof MA, Kuper HH, et al. Modified disease activity scores that include twenty-eight-joint counts. Development and validation in a prospective longitudinal study of patients with rheumatoid arthritis. Arthritis Rheum. 1995;38(1):44-8. doi:10.1002/art.1780380107
3. Buyon JP, Kalunian KC, Ramsey-Goldman R, et al. Assessing disease activity in SLE patients during pregnancy. Lupus. 1999;8(8):677-84. doi:10.1191/096120399680411272
4. Robinson MD, McCarthy DJ, Smyth GK. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics. 2010;26(1):139-40. PubMed ID: PMC2796818
5. Ritchie ME, Phipson B, Wu D, et al. limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Res. 2015;43(7):e47. PubMed ID: PMC4402510
6. Law CW, Chen Y, Shi W, et al. voom: Precision weights unlock linear model analysis tools for RNA-seq read counts. Genome Biol. 2014;15(2):R29. PubMed ID: PMC4053721
7. Nilsen G, Borgan O, Liestol K, et al. Identifying clusters in genomics data by recursive partitioning. Stat Appl Genet Mol Biol. 2013;12(5):637-52. doi:10.1515/sagmb-2013-0016
8. 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. PubMed ID: PMC7859841
9. Kim H, de Jesus AA, Brooks SR, et al. Development of a Validated Interferon Score Using NanoString Technology. J Interferon Cytokine Res. 2018;38(4):171-85. PubMed ID: PMC5963606
10. Chaussabel D, Quinn C, Shen J, et al. A modular analysis framework for blood genomics studies: application to systemic lupus erythematosus. Immunity. 2008;29(1):150-64. PubMed ID: PMC2727981

11. Lambers WM, de Leeuw K, Doornbos-van der Meer B, et al. Interferon score is increased in incomplete systemic lupus erythematosus and correlates with myxovirus-resistance protein A in blood and skin. Arthritis Research & Therapy. 2019;21(1):260. doi:10.1186/s13075-019-2034-4
12. 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. PubMed ID: PMC6585843
13. 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(8):793-802. doi:10.1002/cyto.a.23570
14. Maecker HT, McCoy JP, Nussenblatt R. Standardizing immunophenotyping for the Human Immunology Project. Nat Rev Immunol. 2012;12(3):191-200. PubMed ID: PMC3409649
15. Wolock SL, Lopez R, Klein AM. Scrublet: Computational Identification of Cell Doublets in Single-Cell Transcriptomic Data. Cell Syst. 2019;8(4):281-91.e9. PubMed ID: PMC6625319
16. Traag VA, Waltman L, van Eck NJ. From Louvain to Leiden: guaranteeing well-connected communities. Sci Rep. 2019;9(1):5233. PubMed ID: PMC6435756
17. Speir ML, Bhaduri A, Markov NS, et al. UCSC Cell Browser: visualize your single-cell data. Bioinformatics. 2021;37(23):4578-80. PubMed ID: PMC8652023
18. Newman AM, Steen CB, Liu CL, et al. Determining cell type abundance and expression from bulk tissues with digital cytometry. Nat Biotechnol. 2019;37(7):773-82. PubMed ID: PMC6610714