Check for updates

RESEARCH ARTICLE

REVISED **Single-cell multi-omics analysis reveals IFN-driven alterations in T lymphocytes and natural killer cells in systemic lupus erythematosus** [version 2; peer review: 2 approved, 1 approved with reservations]

Dominik Trzupek, Mercede Lee, Fiona Hamey, Linda S. Wicker, John A. Todd [ID], Ricardo C. Ferreira [ID]

JDRF/Wellcome Diabetes and Inflammation Laboratory, Wellcome Centre for Human Genetics, Nuffield Department of Medicine, NIHR Oxford Biomedical Research Centre, University of Oxford, Oxford, UK

Abstract

Background: The characterisation of the peripheral immune system in the autoimmune disease systemic lupus erythematosus (SLE) at the single-cell level has been limited by the reduced sensitivity of current whole-transcriptomic technologies. Here we employ a targeted single-cell multi-omics approach, combining protein and mRNA quantification, to generate a high-resolution map of the T lymphocyte and natural killer (NK) cell populations in blood from SLE patients.

Methods: We designed a custom panel to quantify the transcription of 534 genes in parallel with the expression of 51 surface protein targets using the BD Rhapsody AbSeq single-cell system. We applied this technology to profile 20,656 T and NK cells isolated from peripheral blood from an SLE patient with a type I interferon (IFN)-induced gene expression signature (IFN^{hi}), and an age- and sex- matched IFN^{low} SLE patient and healthy donor.

Results: We confirmed the presence of a rare cytotoxic CD4$^+$ T cell (CTL) subset, which was exclusively present in the IFN^{hi} patient. Furthermore, we identified additional alterations consistent with increased immune activation in this patient, most notably a shift towards terminally differentiated CD57$^+$ CD8$^+$ T cell and CD16$^+$ NK^{dim} phenotypes, and the presence of a subset of recently-activated naïve CD4$^+$ T cells.

Conclusions: Our results identify IFN-driven changes in the composition and phenotype of T and NK cells that are consistent with a systemic immune activation within the IFN^{hi} patient, and underscore

Open Peer Review

Approval Status ✓ ✓ ?

	1	2	3
version 2 (revision) 12 May 2022		✓ view	
version 1 14 Jun 2021	✓ view	? view	? view

1. **Nancy J. Olsen** [ID], Penn State MS Hershey Medical Center, Hershey, USA

2. **Camilla Jandus** [ID], University of Geneva, Geneva, Switzerland

   Mara Cenerenti [ID], University of Geneva, Geneva, Switzerland

3. **Parth Narendran**, University of Birmingham, Birmingham, UK

Any reports and responses or comments on the article can be found at the end of the article.

the added resolving power of this multi-omics approach to identify rare immune subsets. Consequently, we were able to find evidence for novel cellular peripheral biomarkers of SLE disease activity, including a subpopulation of CD57+ CD4+ CTLs.

Keywords

Single-cell RNA-sequencing (scRNA-seq), multi-omics, BD Rhapsody, AbSeq, cytotoxic CD4+ T cells (CTLs), systemic lupus erythematosus (SLE), type I interferon (IFN), biomarker

Corresponding authors: John A. Todd (jatodd@well.ox.ac.uk), Ricardo C. Ferreira (ricardo.ferreira@well.ox.ac.uk)

Author roles: Trzupek D: Data Curation, Formal Analysis, Methodology, Software, Validation, Visualization, Writing – Review & Editing; **Lee M**: Investigation, Writing – Review & Editing; **Hamey F**: Software, Writing – Review & Editing; **Wicker LS**: Funding Acquisition, Project Administration, Supervision, Writing – Review & Editing; **Todd JA**: Conceptualization, Funding Acquisition, Project Administration, Supervision, Writing – Original Draft Preparation, Writing – Review & Editing; **Ferreira RC**: Conceptualization, Data Curation, Formal Analysis, Investigation, Methodology, Project Administration, Supervision, Validation, Visualization, Writing – Original Draft Preparation, Writing – Review & Editing

REVISED **Amendments from Version 1**

We have revised the manuscript in several ways according to the reviewers suggestion to focus our results and discussion. The main changes to the original version were as follows:

- Edited figures to increase font size and cluster number / name for increased legibility (specifically Figure 2, Figure 4 and Figure 5). We also provide better resolution files of every figure.

- Revised and shortened Discussion section to improve focus and context of the results.

- Edited the text to reflect change of status of the *NCT04680637* trial cited in the text - now initiated and in the recruiting phase.

- Renamed the CD4+ Teff population as CD4+ Tconv cells in the figures and text to avoid confusion with the presence of cells with a naive phenotype in this population.

- Minor additional edits in the text to correct typographical errors and improve readability.

Any further responses from the reviewers can be found at the end of the article. The recent responses to the reviewer's are still awaiting moderator approval.

Abbreviations

CTL, cytotoxic T lymphocyte; FACS, fluorescence-activated cell sorting; IFN, type I interferon; IFN[hi], SLE patient with a detectable type I interferon-inducible gene expression signature in blood; ISG, type I interferon-stimulated gene; NK, natural killer; PBMC, peripheral blood mononuclear cells; PMA, phorbol 12-myristate 13-acetate; scRNA-seq, single-cell RNA-sequencing; SLE, systemic lupus erythematosus; T1D, type 1 diabetes; T_{CM}, central memory T cell; Tconv, conventional T cell; T_{EM}, effector memory T cell; T_{EMRA}, terminally differentiated effector memory T cell; T_{FH}, follicular helper T cell; Treg, regulatory T cell; T1D, type 1 diabetes; UMAP, Uniform Manifold Approximation and Projection; UMI, unique molecular identifier.

Introduction

Systemic lupus erythematosus (SLE) is a common autoimmune disease, characterised by chronic immune activation including a systemic type I interferon (IFN)-induced transcriptional signature in blood (henceforth referred to as the IFN signature), and the development of anti-nucleic acid autoantibodies[1,2]. The immunological complexity of SLE is underscored by the extensive heterogeneity of the clinical manifestations, which impair the management of the condition and is reflected in the paucity of tailored therapeutic options available to the patients. In particular, the presence of the IFN signature has been shown to be associated with increased disease severity[3,4]. These findings have been widely replicated and establish the pathogenic role of the IFN signalling pathway in SLE[5,6]. Furthermore, the detection of an IFN signature in blood can be used to stratify patients with chronic immune activation, and several surrogate biomarkers - for example soluble SIGLEC-1 (sSIGLEC-1) and IL-2 concentrations and a population of CD25-FOXP3+ Tregs - have been proposed to identify patients with active disease[7–9]. However, the exact mechanisms underlying the pathological chronic immune activation are only partly identified, particularly the contribution of the adaptive immune system to this condition. The recent advent of high-throughput single-cell RNA-sequencing (scRNA-seq) has provided a new tool to dissect the extensive heterogeneity of the human immune system and to identify disease-specific alterations. Nevertheless, these methods still present significant technical limitations, which include – but are not limited to – high cost and low sensitivity, especially for genes with low expression levels, that can significantly reduce the power of the analyses. In SLE, a recent scRNA-seq study did demonstrate the applicability of the approach to dissect the heterogeneity of the immune response at the single-cell level[10]. Nevertheless, the limited sensitivity of the whole-transcriptome method precluded the identification of disease-specific alterations in the T cell compartment over-and-above the previously identified IFN-induced gene expression signature.

Recently, we have developed a targeted single-cell multi-omics approach based on the BD Rhapsody system combined with the AbSeq proteomics technology, which allows us to immunophenotype the human immune system with high resolution and in a large number of cells in a more affordable way[11]. This approach, combining mRNA and protein quantification provides superior sensitivity and resolution compared to traditional whole-transcriptomic methods, which are critical factors for the single-cell characterisation of relatively quiescent cells with low RNA content such as T cells. Using this approach, we have identified a rare subset of activated T_{H1} CD4+ T cells with a distinct cytotoxic signature in blood from an SLE patient with a high IFN-induced gene expression signature (IFN[hi])[11]. To replicate and extend these findings, we here applied this technology to three additional selected blood samples: (i) an IFN[hi] SLE patient with concomitant high circulating IL-2 levels; (ii) an age- and sex-matched IFN[low] SLE patient; and (iii) an age- and sex- matched healthy donor. To further increase the sensitivity of this system to dissect the heterogeneity of T and NK cell populations, we designed a custom mRNA panel containing 565 different probes targeting 534 different genes, including a selection of genes with high cell-type discrimination potential identified in other T and NK cell datasets, thereby providing increased resolution for the clustering of these immune populations. In addition, we developed a panel of 51 AbSeq antibodies, targeting key T- and NK-cell enriched surface markers, which provide additional functional insight into the identified cell subsets. This resource provides a high-resolution map of the peripheral T and NK populations in SLE, containing ~21,000 high-quality profiled single cells, and provides a proof-of-principle framework for the application of this multi-omics approach to elucidate the cellular mechanisms contributing to autoimmune diseases.

Methods
Ethics statement
All samples and information were collected with written and signed informed consent from all participants. The study conformed to the principles of the Helsinki Declaration and was approved by the NHS Peterborough and Fenland local research ethics committee (05/Q0106/20).

Subjects

Study participants comprised two SLE patients and one healthy donor, recruited from the Cambridge BioResource (CBR). These three donors were selected from a cohort of 490 healthy donors and 41 SLE patients from the CBR, which were previously recruited to explore novel biomarkers of the IFN signature and disease activity in SLE[8,9]. Based on the blood biomarkers previously measured in this extended cohort, we selected three patients to profile in this study: an IFN^{hi} SLE patient with concomitant high circulating IL-2 levels, one age- and sex-matched IFN^{low} SLE patient and one age- and sex-matched healthy donor. Baseline demographics and characteristics are summarised in Table 1. The SLE patients were recruited for this study outside their regular clinic visits and represent population-based patients with good disease management that were otherwise well at the time of venesection.

Cell preparation and fluorescence-activated cell sorting (FACS)

Cryopreserved peripheral blood mononuclear cells (PBMCs) were thawed at 37°C and resuspended drop-by-drop in X-VIVO15 (Lonza) with 1% heat-inactivated, filtered human AB serum (Sigma). After washing in PBS, cells were incubated with Fixable Viability Dye eFluor 780 (eBioscience) for 10 minutes at room temperature. Cells were then washed with PBS + 2% FBS and incubated with human Fc block (BD Biosciences) and with the following fluorochrome conjugated antibodies: CD25-PE (Mouse anti-human monoclonal antibody, clone M-A521; BD Biosciences, cat# 555432), CD127-PE cy7 (Mouse anti-human monoclonal antibody, clone eBioRDR5; eBioscience; cat# 25-1278-42), CD4-FITC (Mouse anti-human monoclonal antibody, clone OKT4; Biolegend, cat# 317408), CD56-BV711 (Mouse anti-human monoclonal antibody, clone HCD56; Biolegend, cat# 318336), CD8-AF700 (Mouse anti-human monoclonal antibody, clone HIT8a; Biolegend, cat# 300920) and CD3-BV510 (Mouse anti-human monoclonal antibody, clone UCHT1; Biolegend, cat# 300448). Following incubation for five minutes, each sample was split in two 5ml FACS tubes (Falcon) and labelled with the respective sample barcoding antibody (sample multiplexing kit; BD Biosciences). Each donor was barcoded with two barcodes (six barcodes in total) to increase the number of sample barcodes and improve the identification of cell multiplets. The two barcodes corresponding to the same donor were collapsed into a single sample at the analysis stage. Following incubation for 30 minutes at room temperature, cells were washed two times and resuspended in PBS + 1% FBS for cell sorting at 4°C in a BD FACSAria Fusion sorter (BD Biosciences). The following T and NK cell subsets were sorted from each donor: (i) CD3$^-$CD56^{hi} (NK56^{br}); (ii) CD3$^-$CD56$^+$ (NK56^{dim}); (iii) CD3$^+$CD8$^+$ (CD8$^+$ T cells); (iv) CD3$^+$CD4$^+$ CD127^{hi}CD25$^{-/low}$ (CD4$^+$ CD127^{hi} Tconvs); (v) CD3$^+$CD4$^+$ CD127^{low}CD25$^{-/low}$ (CD4$^+$ CD127^{low} Tconvs); and (vi) CD3$^+$CD4$^+$ CD127^{low}CD25^{hi} (CD4$^+$ Tregs).

Plasma soluble IL-2RA (sCD25) measurements

Soluble IL-2RA (sCD25) concentrations were measured in plasma samples from a cohort of SLE patients (N=41) and healthy donors (N=141), which included the three donors profiled in this study, using the BD OptEIA Human ELISA Kit (BD Biosciences; cat# 559104) according to the manufacturer's instructions. Assay readout was performed using Europium-labelled streptavidin (Perkin Elmer, cat# 1244-360) combined with time-resolved fluorescence spectroscopy using DELFIA reagents (DELFIA Assay buffer, cat# 1244-106 and DELFIA Enhancement Solution, cat# 1244-104, Perkin Elmer). Plasma samples were diluted 1:20 and measured in duplicate in Corning 96-well clear polystyrene microtiter plates (Sigma, cat# CL3795-100EA) using a VICTOR X4 multilabel plate reader (Perkin Elmer). Each 96-well plate contained a recombinant sCD25 protein standard curve with a detection range of 31.25–500pg/ml.

AbSeq staining and cell capture

Following cell sorting, each cell subset was resuspended to a final concentration of 200,000 cells/ml. To capture similar numbers of cells from each subset and each donor, 100 μl of each subset (corresponding to 20,000 cells) from the same donor

Table 1. Baseline characteristics of study participants.

DonorID	Age range (years)	Sex	IFN signature (AU)[1]	sSIGLEC-1 (ng/ml)[1,2]	IL-2 (fg/ml)[3]	sCD25 (pg/ml)[4]
SLE1 (IFN^{lo})	50–55	F	-2.2	16.1	109.3	5,879.8
SLE2 (IFN^{hi})	50–55	F	7.5	24.7	672.9	12,363.3
HD	50–55	F	-5.2	4.8	108.9	3,193.7

Baseline demographic and molecular biomarkers characteristics for the study participants. Molecular biomarkers of disease activity, including the type I interferon transcriptional signature, the plasma concentration of soluble SIGLEC-1 and IL-2 were measured as described previously[8,9]. [1]Median IFN signature measured in this cohort of SLE patients was 0.06 AU (range = -11.43 - 14.36 AU). [2]Median sSIGLEC-1 concentration measured in this cohort of SLE patients was 11.0 ng/ml (range = 2.4 - 53.3 ng/ml). [3]Median IL-2 concentration measured with this assay was 93.9 fg/ml (range = <10 - 1,504.9 fg/ml) and 123.6 fg/ml (range = 19.3 - 2,527.0 fg/ml) in a cohort of healthy donors (N = 194) and SLE patients (N = 41), respectively. [4]Median sCD25 concentration was 3,205.3 pg/ml (range = 1,622.0 - 12,875.6 pg/ml) and 3,254.6 pg/ml (range = 1,270.1 - 21,282.0 pg/ml) in a cohort of healthy donors (N = 141) and SLE patients (N = 41), respectively. AU, arbitrary units; HD, healthy donor; IFN, type I interferon; N/A, not available; sCD25, soluble CD25; SLE, systemic lupus erythematosus; sSIGLEC-1, soluble SIGLEC-1.

were then pooled together. For the NK56[br] subset, cell numbers obtained from sorting were limited (ranging from 4,300 to 10,000), and therefore all cells were taken to the final donor pool. Cells were then washed and resuspended in 300 µl X-VIVO15 + 10% heat-inactivated, filtered human AB serum. Half of the volume from each donor (corresponding to ~55,000 cells) was pooled together in a 5 ml FACS tube and washed with PBS + 2% FBS. Cells were stained with a master mix of 51 oligo-conjugated AbSeq antibodies (BD Bioscience; see extended data[12]) for 45 minutes on ice.

The other half of the volume from each donor was incubated in round-bottom 96-well plates at 37ºC for 90 minutes with a PMA and ionomycin cell stimulation cocktail (eBioscience), in the absence of protein transport inhibitors. Cells were harvested into FACS tubes, washed with PBS + 2% FBS and incubated with the oligo-conjugated AbSeq antibodies, as detailed above for the resting cells.

Following incubation with AbSeq antibodies, cells were washed three times with cold BD sample buffer (BD Biosciences) to remove any unbound antibody and counted. Samples were then resuspended in 620 µl of cold BD sample buffer at the desired cell concentrations – ranging from 20 to 30 cells/µl for an estimated capture rate of 10,000-15,000 single-cells – and immediately loaded on a BD Rhapsody cartridge (BD Biosciences, cat# 633733; and cartridge reagent kit, BD biosciences, cat# 633731) for single-cell capture. Resting and *in vitro* stimulated samples were each loaded on a separate BD Rhapsody cartridge.

cDNA library preparation and sequencing

Single-cell capture and cDNA library preparation was performed using the BD Rhapsody Express Single-cell analysis system (BD Biosciences), using the BD Rhapsody cDNA (BD Biosciences, cat# 633663) and Targeted amplification (BD Biosciences, cat# 633664) kits, according to the manufacturer's instructions. cDNA was amplified for 11 cycles following the exact protocol instructions using the pre-designed Human T-cell Expression primer panel (BDiosciences, cat# 633751) containing 259 primer pairs, together with a custom designed primer panel containing 306 primer pairs (BD Biosciences, cat# 633784). Our combined panel contained 565 primer pairs targeting 534 different genes (see extended data[12]). Panel design was based on previous literature search combined with data mining of available T and NK cell datasets to identify a panel of highly variable genes that contribute highly to the clustering of the identified T and NK cell subsets. This panel was specifically enriched for lineage-defining transcription factors and other key functional markers of Th cell differentiation to increase clustering resolution.

The resulting PCR1 products were purified using AMPure XP magnetic beads (Beckman Coulter, cat#A63881) and the respective mRNA and AbSeq/Sample tag products were separated based on size selection, using different bead ratios (0.7X and 1.2X, respectively). The purified mRNA and sample tag PCR1 products were further amplified (10 cycles as per protocol instructions), and the resulting PCR2 products purified by size selection (0.8X and 1.2X for the mRNA and sample tag

libraries, respectively). The concentration, size and integrity of the resulting PCR products was assessed using both Qubit (High Sensitivity dsDNA kit; Thermo Fisher, cat# Q32854) and the Agilent 4200 Tapestation system (High Sensitivity D1000 screentape; Agilent, cat#5067-5584 and High sensitivity D1000 Reagents; Agilent, cat# 5067-5585). The final products were normalised to 2.5 ng/µl (mRNA), 1.1 ng/µl (Sample tag) and 0.5 ng/µl (AbSeq) and underwent a final round of amplification (six cycles) using indexes for Illumina sequencing to prepare the final libraries. Final libraries were quantified using Qubit and Agilent Tapestation and pooled (~39/58/3% mRNA/AbSeq/ Sample tag ratio) to achieve a final concentration of 5 nM. Final pooled libraries were spiked with 15% PhiX control DNA to increase sequence complexity and sequenced (75 bp paired-end) on a HiSeq 4000 sequencer (Illumina).

Data analysis and quality control (QC)

The FASTQ files obtained from sequencing were analysed following the BD Biosciences Rhapsody pipeline (BD Biosciences), as previously described[11], using as alignment reference the pre-designed Human T-cell expression panel containing 259 amplicon sequences (mapping to 259 different genes; BD Biosciences), a custom supplemental panel containing 306 amplicon sequences (mapping to an additional 275 genes), and an AbSeq panel of 51 oligo sequences. Reads with the same cell label, same unique molecular identifier (UMI) sequence and mapping to the same gene were collapsed into a single molecule. The recursive substitution error correction (RSEC) UMI adjustment algorithm developed by BD Biosciences was used to obtain the adjusted UMI counts for further processing and analysis.

A total of 10,973 (resting) and 14,509 (stimulated) putative cells were called by the Rhapsody data processing pipeline, of which 583 and 1,427, respectively, corresponded to pseudo-cells containing only very low mRNA and protein counts – due to low-level ambient mRNA contamination. From the 10,390 (resting) and 13,082 (stimulated) captured cells, 321 (3.9%) and 547 (4.2%), respectively, were assigned as cell multiplets based on their sample barcode information and filtered out from the analysis. Additionally, we excluded cells with highest (>40,000) UMI counts. Most cells identified as undetermined by the Rhapsody pipeline could be confirmed to be of low quality apart from the subset of NK56[br] cells identified as undetermined due to low affinity of antibodies used for multiplexing towards these cells. All remaining undetermined cells not mapping to the identified NK56[br] cluster were excluded from the analysis. Additionally, small subsets of cells of clearly distinct clusters were removed due to being identified as either apoptotic cells (characterised by lower distribution of mRNA UMI counts and higher expression of all AbSeq antibodies – due to higher levels of unspecific antibody binding on apoptotic cells), or non-T cell contaminants (small numbers of B cells and monocyte contaminants). From the initial 10,390 (resting) and 13,082 (stimulated) captured cells, we ended up with a final set of 9,568 and 11,088 cells, respectively, after QC, corresponding to a final processed dataset of 20,656 high-quality single-cells. All processed scRNA-seq data files containing the gene expression matrixes (UMI per cell) and sample tag identifiers for the resting, stimulated and integrated datasets are available as underlying data (Dataset S1)[12].

Data normalisation was performed using different methods for the protein and mRNA data: centred log-ratio (CLR) normalisation for protein, computed independently for each feature; and typical log normalisation for the mRNA data. We used Uniform Manifold Approximation and Projection (UMAP) for dimensionality reduction, increasing the number of PCA dimensions to 30, based on the Seurat elbow plot. Differential expression analysis was performed using a tailored hurdle model from the MAST package[13], and integration of data from multiple experiments was performed using a combination of canonical correlation analysis (CCA) and identification of mutual nearest neighbours (MNN), implemented in Seurat 3.0[14]. For clustering of the resting dataset, we increased the default clustering resolution parameter value to 1.1 to obtain a more fine-grained set of clusters. Although we observed relatively stable cluster assignment around this value of the clustering parameter, the optimal parameter may depend on the aim of the analysis and on the desired granularity of the resulting cell clusters.

Results

Single-cell multi-omics provides a high-resolution map of T and NK cell populations in blood

Previously, we have demonstrated that a targeted single-cell multi-omics approach, based on the BD Rhapsody system (using a pan-immune panel containing 399 probes), provides a high-resolution single-cell map of the human peripheral immune system using PBMCs[11]. To further refine the resolution of the T and NK cell populations, we have developed a custom targeted mRNA panel containing 565 primer pairs targeting 534 genes, with an emphasis on key lowly-expressed, lineage-defining, transcripts. Of the 565 primers included in this custom panel, only a subset of 260 overlapped with the pan-immune panel we used previously[11], thereby providing increased resolution to identify the functional T and NK cell subsets. We have applied this custom panel, in combination with 51 oligo-conjugated AbSeq antibodies, to profile the T and NK cell populations in PBMCs in SLE (Figure 1A). We were interested in investigating alterations in these immune
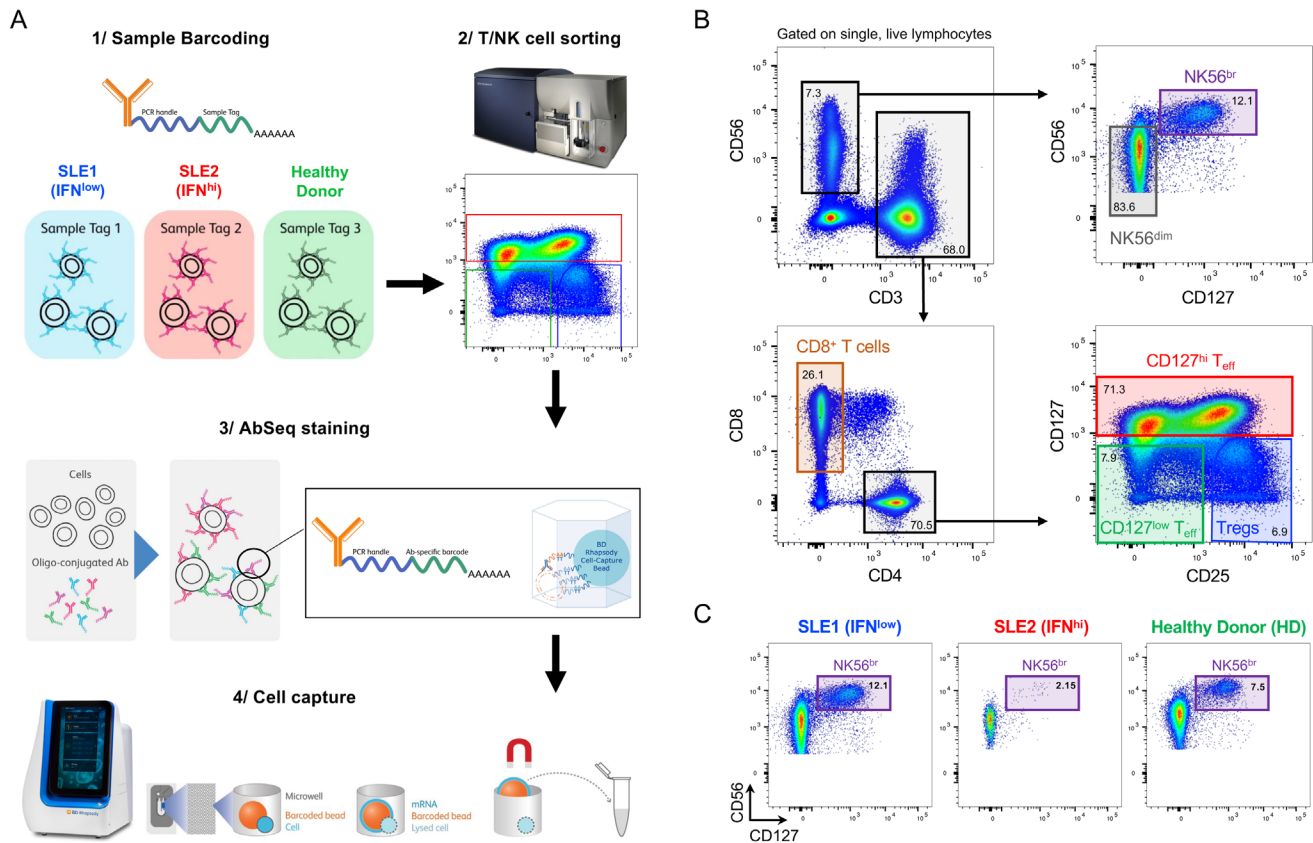
Figure 1. Characterisation of the peripheral T and NK cell populations in SLE using a single-cell multi-omics approach. (**A**) Summary of the experimental workflow based on the BD Rhapsody single-cell RNA-sequencing system combining the quantification of mRNA and surface protein targets (AbSeq). Experiment was carried out using samples from three donors: (i) an SLE patient with low expression of the type I interferon (IFN)-induced gene expression signature (IFN^low; depicted in blue); (ii) an SLE patient with a previously detected IFN-induced gene expression signature (IFN^hi; depicted in red); and (iii) a healthy donor (HD; depicted in green). Cells isolated from each donor were barcoded with oligo-conjugated antibodies and pooled together for cell capture and library preparation. (**B**) Gating strategy used for the isolation of the six T and NK cell populations assessed in this study. (**C**) Frequency of NK56^br cells, defined as CD3^- CD127^+CD56^hi NK cells, in each of the three donors.

compartments in SLE patients with a recent IFN response, so we specifically selected one independent IFN[hi] patient sample in which we had also previously measured elevated sSIGLEC-1, IL-2 and sCD25 concentrations, as well an age- and sex-matched IFN[low] patient and a healthy donor[8,9]. To reduce experimental variation, we used oligo-conjugated bar-code antibodies to identify each of the three donors and pool the cells into a single experiment (Figure 1A). Furthermore, to increase the number of T and NK cells profiled and therefore increase the ability to detect new subpopulations of cells and biomarkers, we employed a cell-subset pre-enrichment strategy by sorting six immune subsets of interest: (i) CD3[−] CD56[hi] NK cells (NK56[br]); (ii) CD3[−] CD56[+] NK cells (NK56[dim]); (iii) CD3[+] CD4[−]CD8[+] T cells (CD8[+] T cells); (iv) CD3[+] CD4[+]CD8[−] CD127[hi]CD25[−/low] T conventional cells (CD4[+] CD127[hi] Tconvs); (v) CD3[+] CD4[+]CD8[−] CD127[low]CD25[−/low] T conventional cells (CD4[+]CD127[low] Tconvs); and (vi) CD3[+]CD4[+]CD8[−]CD127[low]CD25[hi] T cells (CD4[+] Tregs; Figure 1B). To obtain a similar representation of each of these immune subsets we aimed to pool equal numbers (~20,000) of cells sorted from each subset, with the exception of the more limiting NK56[br] subset where we aimed to sort ~10,000 cells per donor. We noted a significant depletion of NK56[br] cells in the IFN[hi] patient (sorted 4,300 cells – compared to 10,000 for the IFN[low] patient and healthy donor), which may indicate

a mobilization of the circulating NK56[br] cells to tissues because of the IFN response in this patient (Figure 1C).

To obtain a more comprehensive map of the assessed subsets, we obtained single-cell transcriptomic and proteomic profiles from both resting and *in vitro* stimulated cells. In total we profiled 20,656 single cells with high quality, corresponding to 9,568 resting and 11,088 *in vitro* stimulated cells. For the resting cells, unsupervised clustering identified 15 distinct subsets, with a clear demarcation of cells from the main CD4[+], CD8[+] and NK cell populations (Figure 2A, 2B; underlying data – Dataset S2)[12]. This improved functional differentiation of the T cell compartment is illustrated by the mutually exclusive pattern of expression of the CD45RA and CD45RO protein isoforms, reflecting the transition of a naïve to a memory phenotype (Figure 2C). Furthermore, expression of the CD45RA/RO isoforms on the CD8[+] T cells identifies a subset of effector memory CD8[+] T cells within cluster 4 marked by the re-expression of CD45RA, which is compatible with a terminally differentiated effector memory T cell (TEMRA) phenotype. Although we observed good integration of the data from the three donors in this experiment, we observed some marked alterations within the relative distribution of IFN[hi] patient subsets, manifested in the donor-specific enrichment of four

Figure 2. Single-cell multi-omics provides a high-resolution map of the peripheral T and NK cell populations. (**A**) Uniform Manifold Approximation and Projection (UMAP) plot depicting the clustering of the isolated resting cells (N = 9,568 cells passing QC) from the three donors assessed in this study. (**B**) Dot plot depicts the expression of select key differentially expressed markers within the identified clusters. The size of the dots represents the frequency of the marker in the respective cluster and the colour represents the relative expression level. Protein markers (AbSeq) are shown on the left section of the plot and mRNA markers on the right section. (**C**) UMAP plot depicting the overlaid expression of the CD45RA (white to green) and CD45RO (white to red) protein isoforms on the identified cell clusters. (**D**) UMAP plot depicting the donor-specific clustering. (**E**) Compositional analysis depicting the relative frequency of cells from each donor in each of the identified clusters. A small fraction of the NK56[br] cells in cluster 7 could not be tagged with the barcoding antibodies, and therefore donor-specific identity could not be assigned for those cells (labelled as N/A on the plot). HD, healthy donor; IFN, type I interferon-induced gene expression signature; SLE, systemic lupus erythematosus.

of the identified subsets: two CD4+ T cell subsets (clusters 10 and 12), the CD8+ T$_{EM}$/T$_{EMRA}$ subset (cluster 4) and one CD16+ NK^{dim} (cluster 3) subset (Figure 2D, 2E). These observations support the presence of subtle yet defined IFN-induced alterations in the composition of rare immune subsets that can be identified using this targeted single-cell multi-omics approach.

A subset of CD4+ T cells with cytotoxic profile is specifically present in blood from IFN^{hi} patients

Among the four IFN^{hi}-enriched clusters identified in this study, one corresponded to a CD4+ T cell subset (cluster 10) with a clearly-defined cytotoxic profile. Because of this cytotoxic transcriptional profile this subset clustered together with the effector memory CD8+ T cell subsets, despite its elevated

CD4 expression and concomitant lack of CD8 expression observed by AbSeq (Figure 3A). Differential expression analysis confirmed a distinct cytotoxic profile in comparison with other CD4+ T cells, including elevated expression of canonical T cell cytotoxic markers such as *PRF1, GZMB, NKG7* and *CCL5* (Figure 3B, 3C; underlying data – Dataset S3)[12]. Notably, we had previously identified a CD4+ T cell subset with a similar cytotoxic profile using this targeted multi-omics approach that was also specifically enriched in a different IFN^{hi} SLE patient[11]. The T-cell centric custom panel used in this study provides additional resolution of this subset and lends further support to its specific enrichment in patients. Of the differentially expressed markers that compose the cytotoxic signature associated with this CD4+ CTL subset, we identified the expression of a surface protein, CD57 (encoded by *B3GAT1*; Figure 3D).
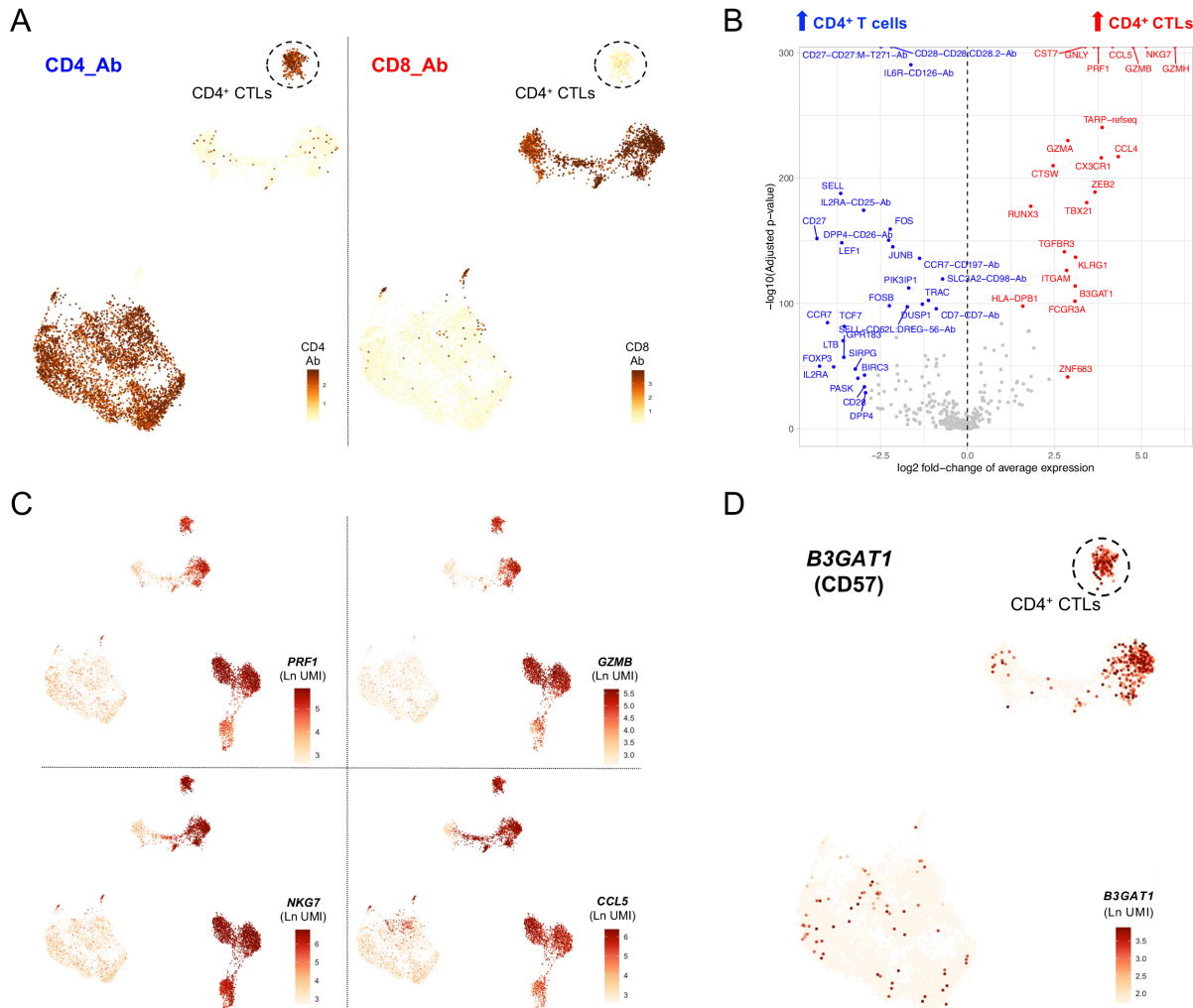
Figure 3. A subset of CD57+ CD4+ T cells with cytotoxic profile is specifically enriched in the IFN^{hi} SLE patient. (**A**) Feature plots depicting the expression of CD4 (left) and CD8 (right) at the protein level (AbSeq) on the identified T cell clusters. (**B**) Volcano plot depicts the differential expression between the CD4+ CTLs (cluster 10) and the remaining identified CD4+ T cell clusters. (**C**) Feature plots depict the expression of four canonical cytotoxic cytokine genes: *PRF1, GZMB, NKG7* and *CCL5*. (**D**) Feature plot depicting the expression of *B3GAT1*, encoding for the surface receptor CD57 on the identified T cell clusters. Ab, AbSeq antibody; CD4+ CTL, cytotoxic CD4+ T cell.

Similarly, additional genes encoding for surface proteins typically expressed on cytotoxic CD8[+] T cells, but not in CD4[+] T cells, including CX3CR1 and KLRG1, were specifically upregulated in these CD4[+] CTLs (Figure 3D). In addition to cell surface markers, CD57[+] CD4[+] CTLs were characterised by the expression of key transcription factors associated with T cell cytotoxic function, including *ZNF683* (Hobit), *IKZF1* (IKAROS), *TBX21* (T-Bet), *ZEB2, RUNX3, ID2, PRDM1* (BLIMP1) and *BHLHE40* (Figure 3B; underlying data – Dataset S3)[12]. These observations suggest that a combination of the CD4 and CD57 markers could be sufficient to identify this subset of circulating CD4[+]CD57[+] CTLs, and that the frequency of these cells could therefore represent a novel cellular biomarker of recent, and probably flaring, disease activity in SLE patients.

In vitro stimulation polarises the functional differentiation of T cells

To further increase the resolution of the functional T cell subsets identified with this multi-omics approach, we also profiled cells following a short (90 minutes) *in vitro* stimulation with PMA + ionomycin. The induction of cytokine expression and upregulation of effector function genes increased further the separation of the subpopulations (Figure 4A, 4B; underlying data – Dataset S4)[12]. This functional separation was particularly evident in CD4[+] T cells, where we were able to identify the different Th subsets, marked by the expression of the respective lineage-defining transcription factors and canonical effector cytokines (Figure 4C). We were also able to integrate both the resting and *in vitro* stimulated datasets, thereby providing increased cell
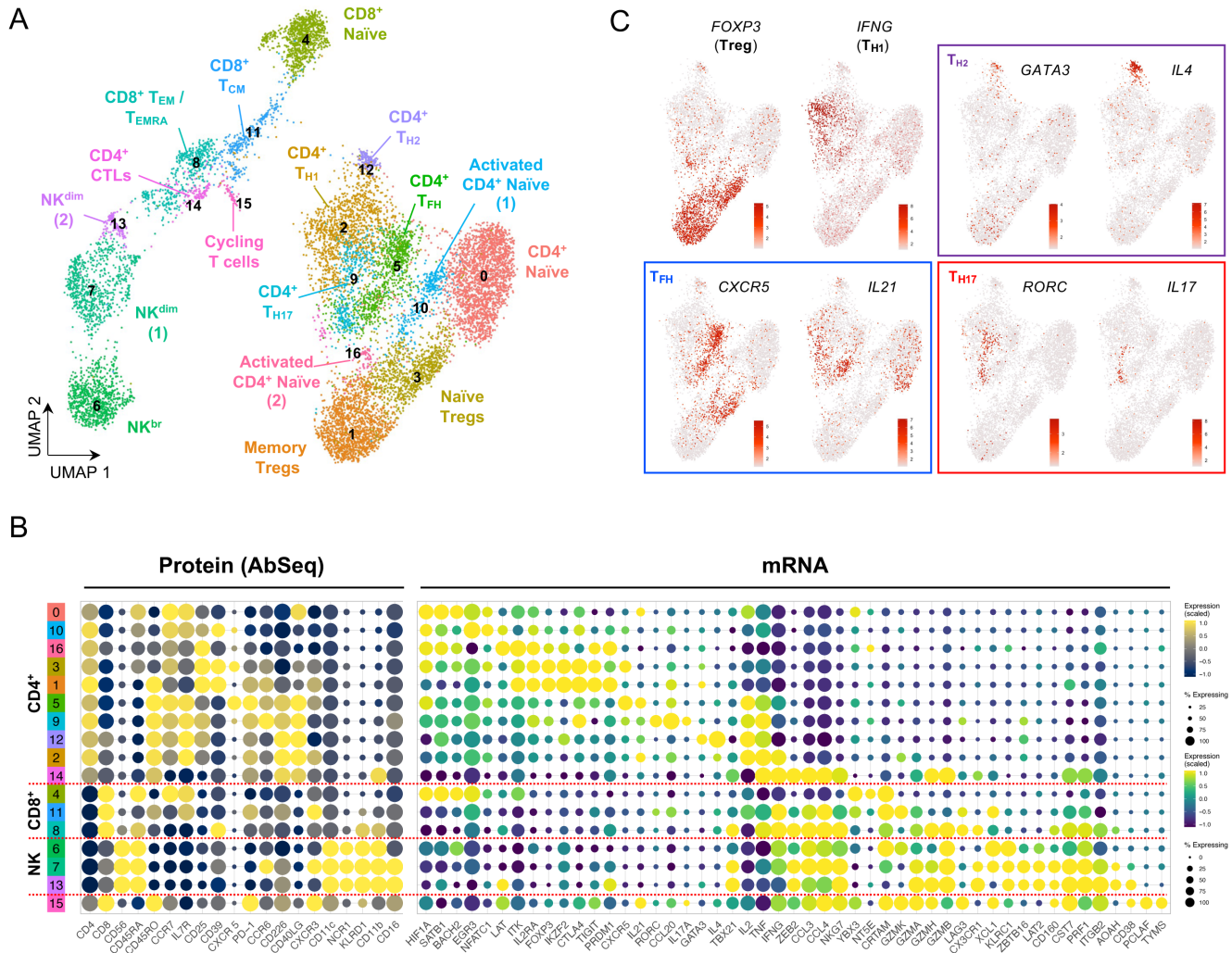
Figure 4. *In vitro* stimulation polarizes the functional differentiation of CD4[+] T cell subsets. (**A**) UMAP plot depicting the clustering of the isolated T and NK cells following short *in vitro* stimulation (90 min) with PMA + ionomycin (N = 11,088 cells passing QC). (**B**) Dot plot depicts the expression of select key differentially expressed markers within the identified *in vitro* stimulated clusters. The size of the dots represents the frequency of the marker in the respective cluster and the colour represents the relative expression level. Protein markers (AbSeq) are shown in the left section of the plot and mRNA markers on the right section. (**C**) Feature plots depicting the expression of key lineage defining transcription factors and respective effector cytokine.

numbers from each of the identified subsets (Figure 5A, 5B; underlying data – Dataset S5)[12]. The resulting clusters from the integrated analysis generally contained an informative representation of cells from both resting and *in vitro* stimulated conditions, with only a few condition-specific clusters remaining - mostly resulting from the strong induction of cytokine and effector gene expression following *in vitro* stimulation (Figure 5C). In agreement with the efficient integration of the two datasets, the compositional analysis of the integrated dataset yielded similar results to the resting cells (Figure 5D). Only a few exceptions were detected, including the split between the resting and *in vitro* stimulated CD4+ CTL clusters (clusters 14 and 17, respectively), which was due to the differential expression of increased levels of T$_{H1}$-like pro-inflammatory cytokines, such as IFN-γ, upon stimulation (Figure 5B). Nevertheless, both clusters of CD4+ CTLs were almost exclusively detected within the IFN^{hi} SLE patient (Figure 5D). In addition to this alteration in the CD4+ T cell compartment, we also observed a notable shift in the relative composition of the CD8+ T cell population within the IFN^{hi} SLE patient, with a strong reduction in the proportion of naïve CD45RA+ CD8+ T cells (cluster 5) and a concomitant increase in the proportion of

terminally differentiated CD8+ T$_{EM}$/T$_{EMRA}$ (cluster 11) cells (Figure 5E).

CD16+ NK^{dim} cells from the IFN^{hi} patient display an activated profile

In addition to the T cell compartment, NK cells isolated from the three donors clustered in four subsets: one cluster of CD16− NK56^{br} cells and three clusters of CD16+ NK^{dim} cells. Using the combined transcriptomic and proteomic profile of these cells, we were able to classify the identified NK cell clusters according to a gradient of activation, reflected by the acquisition of an increasingly cytolytic function - from the least cytotoxic NK56^{br} cells (cluster 3) to the more activated CD16+ NK^{dim} cells (clusters 7 and 8; Figure 6A, based on NK cell clusters shown in Figure 5A). Among the conventional CD16+ NK^{dim} cells, we identified a separation into two subsets (clusters 7 and 8), which was driven by a gradient of expression of markers usually expressed in terminally-differentiated cells, such as *LAG3* and *B3GAT1* (CD57), and the transcription factor *PRDM1* (encoding BLIMP-1) in cluster 8 (Figure 6B, 6C; underlying data – Dataset S6)[12]. We observed marked donor-specific differences in the relative distribution
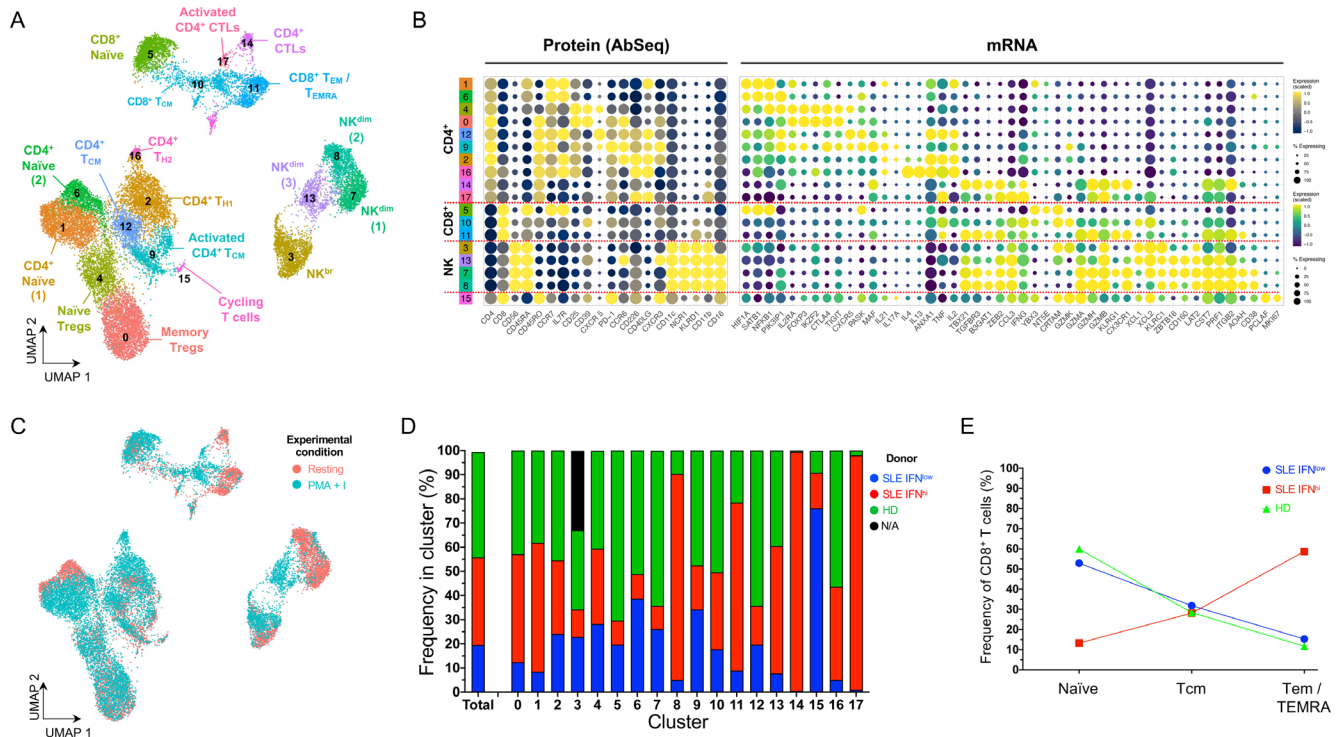

Figure 5. Integrated resting and *in vitro* stimulated data reveal subtle IFN-induced alterations in the composition of the peripheral T and NK cell populations. (A) UMAP plot depicting the clustering of integrated (resting + *in vitro* stimulated) T and NK datasets (N = 20,656 cells passing QC). **(B)** Dot plot depicts the expression of select key differentially expressed markers within the identified integrated clusters. The size of the dots represents the frequency of the marker in the respective cluster and the colour represents the relative expression level. Protein markers (AbSeq) are shown on the left section of the plot and mRNA markers on the right section. **(C)** UMAP plot depicts the treatment-specific clustering of the resting (red) and *in vitro* stimulated (teal) cells. **(D)** Compositional analysis depicting the relative frequency of cells from each donor in each of the identified clusters. A small fraction of the NK56^{br} cells in cluster 3 could not be tagged with the barcoding antibodies, and therefore donor-specific identity could not be assigned for those cells (labelled as N/A on the plot). **(E)** Relative distribution of the three identified CD8+ T cell subsets (among total CD8+ T cells) in each of the three donors assessed in this study. HD, healthy donor; IFN, type I interferon-induced gene expression signature; SLE, systemic lupus erythematosus; Tcm, central memory T cell; Tem, effector memory T cell; TEMRA, terminally differentiated effector memory T cell.
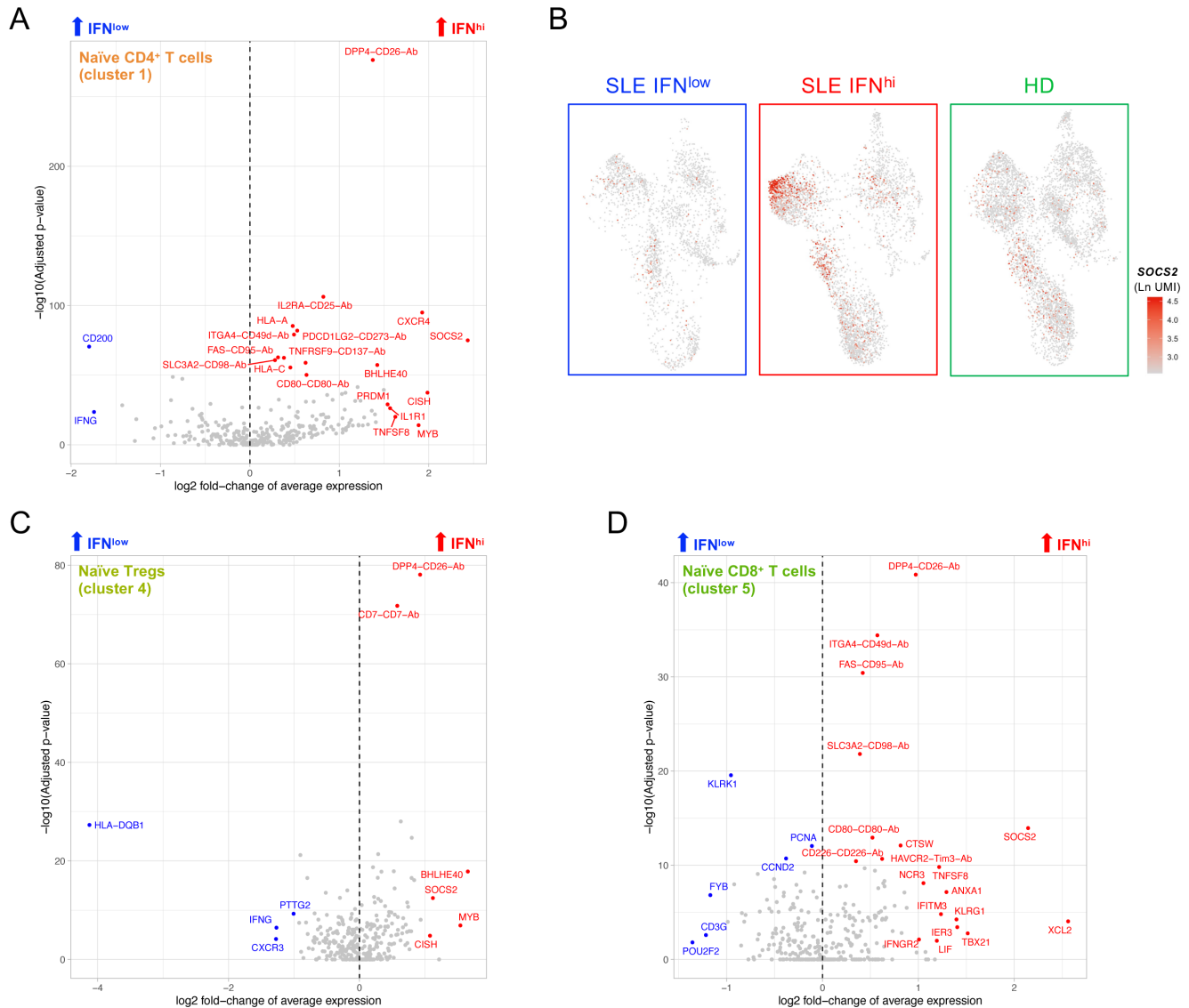
Figure 6. CD16⁺ NKᵈⁱᵐ cells from the IFNʰⁱ SLE patient display a highly differentiated phenotype. (A) Dot plot depicts the expression of select key differentially expressed markers within the identified NK cell clusters from the integrated analysis shown in Figure 5A. The size of the dots represents the frequency of the marker in the respective cluster and the colour represents the relative expression level. Protein markers (AbSeq) are shown on the left section of the plot and mRNA markers on the right section. **(B)** Volcano plot depicts the differential expression between the two CD16⁺ NKᵈⁱᵐ subsets (clusters 7 and 8) identified in the integrated analysis. **(C)** Feature plots depicting the expression of key differentially expressed markers between the CD16⁺ NKᵈⁱᵐ subsets. Markers specifically upregulated in cluster 7 are highlighted in blue and markers specifically upregulated in cluster 8 are highlighted in red.

of these NKᵈⁱᵐ subsets, with the more differentiated cluster 8 being almost exclusively detected in the IFNʰⁱ patient (Figure 5D). These data are consistent with the highly differentiated CD8⁺ T cell phenotype observed in this patient, and together support an increased differentiation of peripheral CD8⁺ T and NK cells in SLE patients with a recent or active IFN response.

Naïve T cells from the IFNʰⁱ SLE patient display a signature of recent cytokine activation

Further supporting the systematic alterations induced by the inflammatory responses in patients with active disease, we identified a distinct signature of recent activation elicited by cytokine signalling within the CD4⁺ naïve T cells (cluster 1, as defined in Figure 5A) from the IFNʰⁱ patient (Figure 7; underlying data – Dataset S7)[12]. This signature was consistent with the IFNʰⁱ-specific recently-activated naïve CD4⁺ T cell cluster identified in the resting dataset (cluster 12; Figure 2A, 2B), and was

marked by the strong upregulation of several genes associated with regulation of cytokine signalling, including *SOCS2*, *CISH*, *BHLHE40*, *CXCR4*, *PRDM1*, *DPP4* and *MYB*, as well as the upregulation of IL-2RA at the protein level (Figure 7A, 7B). Although this cytokine-induced recent activation signature was more prominent within the naive CD4⁺ T cells, it could also be detected in both naïve Tregs (cluster 4; Figure 7C) and naïve CD8⁺ T cells (cluster 5; Figure 7D), albeit with smaller differences, suggesting that peripheral naïve T cells are sensitive biomarkers of recent cytokine activation. Importantly, this signature was consistent with the high plasma concentrations of IL-2 and sCD25 detected in this IFNʰⁱ SLE patient (Table 1), indicating that an increased pro-inflammatory cytokine milieu in patients with active disease could be responsible for the observed transcriptional signature in naïve T cells. Together these findings support that gradual activation-induced alterations of the peripheral immune system are associated with the chronic immune activation in IFNʰⁱ patients, which can be

Figure 7. A subset of naïve CD4+ T cells from the IFNhi SLE patient displays a cytokine-induced activation signature. (**A**) Volcano plot depicts the differential expression between the IFNhi SLE patient and the two other donors in cells mapping to the naïve CD4+ T subset (cluster 1 of the integrated data shown in Figure 5A). (**B**) Feature plots depict the expression of the cytokine-induced signalling gene *SOCS2* in the three donors profiled in this study. (**C, D**) Volcano plots depict the differential expression between the IFNhi SLE patient and the two other donors in cells mapping to the CD4+ naïve Treg (cluster 4; **C**) and CD8+ naïve T cell (cluster 5; **D**) subsets, respectively.

robustly characterised due to the increased functional resolution of the T and NK cells provided by the targeted multi-omics approach used in this study.

Discussion

The clinical management of SLE patients has been limited by the insufficient knowledge of the pathophysiological mechanisms that drive this complex disease and the limited available biomarkers to stratify patients. The recent advent of scRNA-seq technologies has promised to improve our understanding of the disease mechanisms, and shed new light on the specific contribution of the adaptive immune system in SLE[10].

Nevertheless, the current technical challenges and reduced sensitivity of scRNA-seq technologies based on whole-transcriptome profiling, where genes such as those encoding transcription factors that are expressed at low levels are difficult to detect, have not yet provided novel actionable insights to mechanistic-based patient stratification. The combination of protein quantification with traditional scRNA-seq technologies has been also implemented, and shown to significantly increase the power to discriminate immune populations[15,16]. However, these approaches have so far been mostly restricted to total PBMC samples, thereby reducing the number and diversity of the T and NK cells captured in these studies. To circumvent

these limitations, in the present study, we have employed a targeted single-cell multi-omics approach providing both targeted mRNA and protein quantification, combined with pre-purification by flow sorting of the T and NK cell subsets, significantly increasing the number of relevant cells profiled and our confidence in assigning cell clusters to specific, and functionally relevant, sub-populations. This targeted approach – using a custom designed panel to assess the expression of 534 signature T and NK cell genes – not only provides increased sensitivity for the detection of the selected target transcripts, but also significantly decreases the costs of sequencing, thereby allowing to simultaneously quantify the expression of 51 surface expressed proteins. Here we show that this approach provides a high-resolution map of the peripheral T and NK cell populations in SLE in both resting and *in vitro* stimulated conditions. This increased resolution is particularly evident within the CD4+ T cell populations, where we are able to identify the different Th lineages in an informative way.

By applying this multiparametric deep phenotyping approach, we identified a subset of CD4+ CTLs showing hallmarks of highly differentiated T cells with a cytotoxic profile, which was specifically expanded in the IFN^{hi} SLE patient. This observation was consistent with our previous findings using this same technology on an independent IFN^{hi} SLE patient[11], and further supports that the presence of a rare subset of CD4+ CTLs in blood, marked by the co-expression of surface markers usually expressed on cytotoxic CD8+ T cells, including CD57, CX3CR1 and KLRG1, is a feature of SLE patients with active disease. The description of a similar subset of differentiated CD4+ T cells with cytotoxic profile has been well established in humans[17]. Although the frequency of these cells, defined as perforin-expressing CD4+ T cells, is very low in healthy donors (usually <2% of CD4+ T cells in adults), they are significantly expanded and can reach frequencies of 5-20% of CD4+ T cells in patients with chronic viral infection[17]. More recently, scRNA-seq studies have further refined the characterisation of this CD4+ T cell subset, including CD57 expression, in the context of both viral infection[18], anti-tumour responses[19], and autoimmunity (primary Sjogren's syndrome)[20]. In addition, a recent study has reported an increased frequency of CD4+ CTLs in super centenarians[21]. Taken together, these data strongly implicate CD4+ CTLs with antiviral response and chronic inflammation. Given that chronic IFN signalling is a central pathogenic factor in both primary Sjogren's syndrome as well as in SLE, these findings could therefore provide a mechanistic rationale underlying the increased frequency of CD57+ CD4+ CTLs in blood of patients with active IFN responses. In support of this hypothesis, plasmacytoid dendritic cell activation, which is known to be one of the main sources of IFN production, has been shown to induce a pro-inflammatory cytokine profile in CD4+ T cells[22]. In the present study, we show that the expression of CD57 and CX3CR1 on the surface of CD4+ T cells could represent specific markers for the identification of CD4+ CTLs by flow cytometry. We therefore hypothesize that the frequency of CD57+ CD4+ T cells in blood represents a novel cellular biomarker of recent or active IFN responses in patients with systemic autoimmunity, which could have informative clinical application for the monitoring of disease activity in SLE.

We note that in addition to the increased resolution of this targeted single-cell system, another key aspect of this study was the cell-subset pre-enrichment strategy, which allowed us to significantly increase the number of rare CD4+ T cell subsets. In particular, increasing the CD4+ CD127^{low} Tconv and Treg populations, which are highly enriched for activated CD4+ T cells (such as the CD4+ CTLs), allowed us to identify rare subsets that could otherwise be difficult to robustly discern in more heterogeneous peripheral immune cell populations. Similar to the CD4+ CTLs, we also observed an expansion of terminally-differentiated CD8+ T$_{EM}$/T$_{EMRA}$ cells with cytotoxic profile in the IFN^{hi} patient. An expansion of CD8+ CTLs has been previously reported in SLE patients with active disease at the single-cell level[10]. Although the expanded cytotoxic T cell cluster identified by Nehar-Belaid and colleagues displays hallmarks of CD8+ T cells, it is possible that it also encompasses the rarer subset of CD4+ CTLs, which as we show here display a very similar transcriptional profile to the CD8+ CTLs and may therefore not be discernible using whole-transcriptome single-cell technologies. Consistent with this putative pathogenic role of CD8+ CTLs in autoimmunity, a recent report has described a positive correlation between the frequency of a subset of CD57+ CD8+ T cells and the rate of C-peptide decline during the first two years after diagnosis in type 1 diabetic (T1D) patients[23]. In agreement with the CD8+ T$_{EM}$/T$_{EMRA}$ subset identified in our study, the CD57+ CD8+ T cells described in recently diagnosed T1D patients display a classical cytotoxic profile marked by the expression of GZMB and CD16, and were shown to be enriched for autoreactive TCR clones, suggesting a pathogenic function in autoimmunity[23]. Our data add to a deeper phenotypic characterisation of these potentially pathogenic CD57+ CD8+ T cells, as well as their CD57+ CD4+ CTL counterparts, thereby providing a robust tool to characterise such disease-specific T cell subsets, and to identify markers that can be used to monitor their frequency using high-throughput methods, such as flow cytometry, in large numbers of autoimmune patients.

The expression of CD57 in both CD8+ T cells and NK cells has been previously shown to be a marker of terminal differentiation, and the frequency of cytotoxic CD57+ cells has been shown to increase with age and to expand in a number of autoimmune diseases as well as in cancer[24]. Although the role of CD57 in CD4+ T cells is less well defined, it is plausible that it plays a similar role in all T cells, and is also associated with the terminal differentiation of a population of CD4+ CTLs. In addition, an expansion of CD57+ T and NK cells is also a feature of chronic viral infections, including human cytomegalovirus, human immunodeficiency virus, hepatitis C virus, and Epstein-Barr virus infections[24]. Given that the IFN gene expression signature in SLE is specifically induced by type I, but not type II, interferons, one hypothesis for the specific expansion of CD57+ cells in IFN^{hi} SLE patients is that the anti-viral type I IFN signalling pathway is directly linked with the increased differentiation of peripheral T and NK cells. These findings provide an additional cellular link between the chronic IFN signature detected in blood from approximately 50% of SLE patients and support the heightened activation profile in IFN^{hi} SLE patients.

A notable absence in this study was the identification of a distinctive IFN-induced gene expression signature in any of the defined immune subsets, which was the main source of biological variation driving the SLE-specific alterations characterised by single-cell analysis[10]. There are two main factors that could explain this observation: first is the targeted nature of the mRNA panel used in this study, which contains only a small number of IFN-stimulated genes (ISGs). Nevertheless, we did measure the expression of several canonical ISGs, such as *OAS1*, *ISG15*, *IRF7*, *IFIH1*, *IFITM2* and *IFITM3*, and they were not significantly differentially expressed between the donors profiled in this study. Furthermore, given the widespread alterations induced by IFN stimulation, a significant number of the immune genes included in our panel should also be systematically regulated by an IFN response. A more plausible second explanation is that the expression of the ISGs is typically much less pronounced within T and NK cells compared to myeloid cell populations. Since only two SLE patients were profiled in this study, both of whom were recruited outside their regular clinic visits and had no evidence of flaring or active disease, it is likely that no discernible IFN signature could be identified at the time of sample collection in the peripheral T and NK cells of these patients. Instead, our findings suggest that more subtle, and long-lasting, immune alterations in the peripheral T and NK cell populations persist even during periods of clinical remission that are not associated with an active expression of ISGs. We believe that the application of this technology to a larger number of patients with a larger range of stages of disease activity will provide a clearer identification of the disease-specific IFN signature previously described in the T and NK cell populations using single-cell technologies[10].

The major limitation of the current study is that only two patients and one healthy donor were profiled with this technology. The main aim of this study was to demonstrate the potential of applying a targeted approach - combined with a cell-subset pre-selection strategy - to dissect the heterogeneity of specific immune cell populations in more detail. By identifying a small set of patient samples with extreme phenotypic differences, we wanted to establish the applicability of this approach as a complementary tool to traditional whole-transcriptome scRNA-seq methods. This dataset therefore represents a proof-of-concept study and resource demonstrating the potential of this technology to immunophenotype the T and NK cell populations in SLE. These current findings therefore warrant further replication in larger, clinically well-characterised, cohorts; ideally in a longitudinal cohort to investigate subtle changes in the distribution of these immune populations over the course of the disease and the association with the underlying IFN response, as well as other clinical markers of disease activity. Given its targeted nature, this approach is highly scalable and could be applied to larger numbers of clinical samples and to other diseases. These features are particularly well-suited to investigate the cellular mechanisms associated with response to therapy in large clinical trials. One such example is the use of ultra-low dose IL-2 therapy (three cycles of one million units of aldesleukin administered every other day for two weeks followed by a two-week interval), which has shown initial promise and a good safety profile in SLE[25]. More recently, these findings have been extended in a randomised

clinical trial using the same dosing regimen[26]. However, one key conclusion from these studies is that response to IL-2 therapy is variable between patients, underscoring the clinical heterogeneity of SLE. In this study, He and colleagues reported that, compared to placebo, ultra-low dose IL-2 was very efficacious to induce clinical remission in patients with lupus nephritis and in patients with cutaneous rashes[26]. One potential confounding aspect in these trials is the diverse range of disease activity at the time of treatment, which can severely compromise the drug's mechanism of action and consequently the attainment of primary clinical endpoints. For example, we have observed that the concentration of IL-2 in circulation and the frequency of Tregs were increased in SLE patients with increased disease activity[9], and hypothesised that low-dose IL-2 treatment may be more effective for the maintenance of clinical remission rather than treating patients with acute or flaring disease - where additional exogenous IL-2 may be functionally redundant or proinflammatory. Indeed, the IFN[hi] patient investigated here had high plasma IL-2 and sCD25 concentrations. These findings underscore the urgent need to better understand the mechanism of action of current drugs and to identify better biomarkers associated with clinical response and non-response to therapy. We note that an IL-2 mutein, AMG592 (efavaleukin alfa), is currently being investigated in SLE (registration number: NCT03451422). A much larger follow-up phase 2b study (registration number: NCT04680637) has recently initiated recruitment of a cohort of 320 SLE patients and comparing treatment efficacy between patients with high IL-2 levels and those with normal physiological IL-2 concentrations will be of interest.

Conclusion

In conclusion, we have identified distinct alterations in the composition of the T and NK cell compartment in a patient with a discernible peripheral IFN signature, and high concentration of circulating IL-2 and sCD25. Most notably we identified an expansion of a subset of CD4[+] CTLs and a concomitant increased proportion of CD8[+] and NK cells with a terminally differentiated profile. Although an active IFN response could not be detected in these adaptive immune cells, these results suggest that additional subtle alterations in the cell composition could be more long-lasting than the prototypical IFN transcriptional signature and may represent more stable biomarkers of recent disease activity. More widely, this current study provides a framework for the application of a targeted single-cell multi-omics methodology to investigate the cellular mechanisms of human complex disorders as well as the mechanism of response to current therapies, including ultra-low dose IL-2. This information is currently largely lacking from most clinical studies and could provide critical insight for the design of future trials, including patient selection and treatment strategy, which are critical to increase the success rate of clinical trials in such clinically heterogeneous diseases as SLE.

Data availability
Underlying data
Open Science Framework (OSF). Supporting material for 'Single-cell multi-omics analysis reveals IFN-driven alterations in T lymphocytes and Natural Killer cells in systemic lupus erythematosus'. https://doi.org/10.17605/OSF.IO/EDCTN[12].

This project contains the following underlying data:

- Dataset S1: Processed gene expression data matrixes (mRNA and AbSeq UMI counts per cell) and sample tag assignment per cell for the resting and stimulated datasets.

- Dataset S2: Differentially expressed genes in the resting cell clusters.

- Dataset S3: Differentially expressed genes in resting CD4[+] CTLs (cluster 10 from the resting analysis) compared to remaining CD4[+] T cell clusters.

- Dataset S4: Differentially expressed genes in the *in vitro* stimulated cell clusters.

- Dataset S5: Differentially expressed genes in the integrated resting and *in vitro* stimulated cell clusters.

- Dataset S6: Differentially expressed genes between the CD16[+] NK[dim] cells in clusters 7 and 8 from the integrated analysis.

- Dataset S7: Differentially expressed genes in CD4[+] naïve T cells (cluster 1 from the integrated analysis) between the IFN[hi] SLE patient and the two IFN[low] donors.

Extended data

Open Science Framework (OSF). Supporting material for 'Single-cell multi-omics analysis reveals IFN-driven alterations in T lymphocytes and Natural Killer cells in systemic lupus erythematosus'. https://doi.org/10.17605/OSF.IO/EDCTN[12].

This project contains the following extended data:

- XLSX file containing targeted mRNA and protein (AbSeq) panels used in this study

Data are available under the terms of the Creative Commons Zero "No rights reserved" data waiver (CC0 1.0 Public domain dedication).

References

1. Arbuckle MR, McClain MT, Rubertone MV, *et al.*: **Development of Autoantibodies before the Clinical Onset of Systemic Lupus Erythematosus.** *N Engl J Med.* 2003; **349**(16): 1526–1533.
   PubMed Abstract | **Publisher Full Text**

2. Tsokos GC: **Systemic Lupus Erythematosus.** *N Engl J Med.* 2011; **365**(22): 2110–2121.
   PubMed Abstract | **Publisher Full Text**

3. Bennett L, Palucka AK, Arce E, *et al.*: **Interferon and Granulopoiesis Signatures in Systemic Lupus Erythematosus Blood.** *J Exp Med.* 2003; **197**(6): 711–723.
   PubMed Abstract | **Publisher Full Text** | **Free Full Text**

4. 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**(5): 2610–2615.
   PubMed Abstract | **Publisher Full Text** | **Free Full Text**

5. Banchereau J, Pascual V: **Type I Interferon in Systemic Lupus Erythematosus and Other Autoimmune Diseases.** *Immunity.* 2006; **25**(3): 383–392.
   PubMed Abstract | **Publisher Full Text**

6. Elkon KB, Wiedeman A: **Type I IFN system in the development and manifestations of SLE.** *Curr Opin Rheumatol.* 2012; **24**(5): 499–505.
   PubMed Abstract | **Publisher Full Text**

7. Rose T, Grützkau A, Hirseland H, *et al.*: **IFNα and its response proteins, IP-10 and SIGLEC-1, are biomarkers of disease activity in systemic lupus erythematosus.** *Ann Rheum Dis.* 2013; **72**(10): 1639–1645.
   PubMed Abstract | **Publisher Full Text**

8. Oliveira JJ, Karrar S, Rainbow DB, *et al.*: **The plasma biomarker soluble SIGLEC-1 is associated with the type I interferon transcriptional signature, ethnic background and renal disease in systemic lupus erythematosus.** *Arthritis Res Ther.* 2018; **20**(1): 152.
   PubMed Abstract | **Publisher Full Text** | **Free Full Text**

9. Ferreira RC, Dopico XC, Oliveira JJ, *et al.*: **Chronic Immune Activation in Systemic Lupus Erythematosus and the Autoimmune PTPN22 Trp[620] Risk Allele Drive the Expansion of FOXP3[+] Regulatory T Cells and PD-1 Expression.** *Front Immunol.* 2019; **10**: 2606.
   PubMed Abstract | **Publisher Full Text** | **Free Full Text**

10. Nehar-Belaid D, Hong S, Marches R, *et al.*: **Mapping systemic lupus erythematosus heterogeneity at the single-cell level.** *Nat Immunol.* 2020; **21**(9): 1094–1106.
    PubMed Abstract | **Publisher Full Text** | **Free Full Text**

11. Trzupek D, Dunstan M, Cutler AJ, *et al.*: **Discovery of CD80 and CD86 as recent activation markers on regulatory T cells by protein-RNA single-cell analysis.** *Genome Med.* 2020; **12**(1): 55.
    PubMed Abstract | **Publisher Full Text** | **Free Full Text**

12. Trzupek D, Ferreira R: **Supporting material for 'Single-cell multi-omics analysis reveals IFN-driven alterations in the peripheral T cell compartment of SLE patients'.** *OSF.* 2021.
    http://www.doi.org/10.17605/OSF.IO/EDCTN

13. Finak G, McDavid A, Yajima M, *et al.*: **MAST: a flexible statistical framework for assessing transcriptional changes and characterizing heterogeneity in single-cell RNA sequencing data.** *Genome Biol.* 2015; **16**: 278.
    PubMed Abstract | **Publisher Full Text** | **Free Full Text**

14. Stuart T, Butler A, Hoffman P, *et al.*: **Comprehensive Integration of Single-Cell Data.** *Cell.* 2019; **177**(7): 1888–1902.e21.
    PubMed Abstract | **Publisher Full Text** | **Free Full Text**

15. Stoeckius M, Hafemeister C, Stephenson W, *et al.*: **Simultaneous epitope and transcriptome measurement in single cells.** *Nat Methods.* 2017; **14**(9): 865–868.
    PubMed Abstract | **Publisher Full Text** | **Free Full Text**

16. Peterson VM, Zhang KX, Kumar N, *et al.*: **Multiplexed quantification of proteins and transcripts in single cells.** *Nat Biotechnol.* 2017; **35**(10): 936–939.
    PubMed Abstract | **Publisher Full Text**

17. Appay V, Zaunders JJ, Papagno L, *et al.*: **Characterization of CD4[+] CTLs Ex Vivo.** *J Immunol.* 2002; **168**(11): 5954–5958.
    PubMed Abstract | **Publisher Full Text**

18. Patil VS, Madrigal A, Schmiedel BJ, *et al.*: **Precursors of human CD4[+] cytotoxic T lymphocytes identified by single-cell transcriptome analysis.** *Sci Immunol.* 2018; **3**(19): eaan8664.
    PubMed Abstract | **Publisher Full Text** | **Free Full Text**

19. Oh DY, Kwek SS, Raju SS, *et al.*: **Intratumoral CD4 + T Cells Mediate Anti-tumor Cytotoxicity in Human Bladder Cancer.** *Cell.* 2020; **181**(7): 1612–1625.e13.
    PubMed Abstract | **Publisher Full Text** | **Free Full Text**

20. Hong X, Meng S, Tang D, *et al.*: **Single-Cell RNA Sequencing Reveals the Expansion of Cytotoxic CD4 + T Lymphocytes and a Landscape of Immune Cells in Primary Sjögren's Syndrome.** *Front Immunol.* 2021; **11**: 594658.
    PubMed Abstract | **Publisher Full Text** | **Free Full Text**

21. Hashimoto K, Kouno T, Ikawa T, *et al.*: **Single-cell transcriptomics reveals expansion of cytotoxic CD4 T cells in supercentenarians.** *Proc Natl Acad*

Sci U S A. 2019; **116**(48): 24242–24251.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

22.   Caielli S, Veiga DT, Balasubramanian P, *et al.*: **A CD4 $^+$ T cell population expanded in lupus blood provides B cell help through interleukin-10 and succinate.** *Nat Med.* 2019; **25**(1): 75–81.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

23.   Yeo L, Woodwyk A, Sood S, *et al.*: **Autoreactive T effector memory differentiation mirrors β cell function in type 1 diabetes.** *J Clin Invest.* 2018; **128**(8): 3460–3474.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

24.   Nielsen CM, White MJ, Goodier MR, *et al.*: **Functional Significance of CD57 Expression on Human NK Cells and Relevance to Disease.** *Front Immunol.* 2013; **4**: 422.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

25.   He J, Zhang X, Wei Y, *et al.*: **Low-dose interleukin-2 treatment selectively modulates CD4$^+$ T cell subsets in patients with systemic lupus erythematosus.** *Nat Med.* 2016; **22**(9): 991–993.
PubMed Abstract | **Publisher Full Text**

26.   He J, Zhang R, Shao M, *et al.*: **Efficacy and safety of low-dose IL-2 in the treatment of systemic lupus erythematosus: a randomised, double-blind, placebo-controlled trial.** *Ann Rheum Dis.* 2020; **79**(1): 141–149.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

Open Peer Review

Current Peer Review Status: ✓ ✓ ?

Version 2

Reviewer Report 16 May 2022

✓   **Camilla Jandus** [ID]
Department of Pathology and Immunology, Faculty of Medicine, University of Geneva, Geneva, Switzerland

Mara Cenerenti [ID]
Department of Pathology and Immunology, Faculty of Medicine, University of Geneva, Geneva, Switzerland

I very much appreciate the efforts of the authors to address my concerns.

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: CD4 T cells, cytotoxicity, cancer, Innate lymphoid cells

We confirm that we have read this submission and believe that we have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

Version 1

Reviewer Report 27 April 2022

?   **Parth Narendran**
Institute of Immunology and Immunotherapy, University of Birmingham, Birmingham, UK

This is a nice piece of work using RNA and protein profiling of T and NK cells in 2 SLA patients with high and low interferon signatures and a non-SLA control patient.

The authors accepted that the patient numbers are low, and that this is essentially a proof-of-concept study. There is certainly merit in increasing the patient numbers to validate this work, but this does not detract from the proof-of-concept findings of the current analysis.

My main comment relates to the disease state of the IFNhi patient and the implications for biomarker discovery. Does the IFN signature relate to current SLA disease activity or to future prognosis? The context here is that there are established clinical scoring criteria for scoring disease activity in SLE (e.g., Thanou *et al*., 2019[1]), therefore if the findings of this study (ideally replicated in a larger cohort) apply to current disease state, then the question arises as to whether it adds anything more to what is obtained through simple clinical scoring. These are questions that can be explored through larger cohort analysis. If however the multi-omics analysis brings additional information about prognosis, this would be a more intriguing outcome of this analysis.

Minor comments relate to the length of the discussion which can be more focused.

References
1. Thanou A, James JA, Arriens C, Aberle T, et al.: Scoring systemic lupus erythematosus (SLE) disease activity with simple, rapid outcome measures.*Lupus Sci Med*. 2019; **6** (1): e000365 PubMed Abstract | Publisher Full Text

Is the work clearly and accurately presented and does it cite the current literature?
Yes

Is the study design appropriate and is the work technically sound?
Yes

Are sufficient details of methods and analysis provided to allow replication by others?
Yes

If applicable, is the statistical analysis and its interpretation appropriate?
Yes

Are all the source data underlying the results available to ensure full reproducibility?
Yes

Are the conclusions drawn adequately supported by the results?
Yes

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Clinical endocrine disorders and immunology of type 1 diabetes

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have

significant reservations, as outlined above.

Author Response 04 May 2022

Ricardo Ferreira, University of Oxford, Oxford, UK

We appreciate the Reviewer's feedback and suggestions. We detail below the replies to the specific concerns raised by the reviewer:

- 1/ My main comment relates to the disease state of the IFNhi patient and the implications for biomarker discovery. Does the IFN signature relate to current SLA disease activity or to future prognosis? The context here is that there are established clinical scoring criteria for scoring disease activity in SLE (e.g., Thanou *et al*., 2019[1]), therefore if the findings of this study (ideally replicated in a larger cohort) apply to current disease state, then the question arises as to whether it adds anything more to what is obtained through simple clinical scoring. These are questions that can be explored through larger cohort analysis. If however the multi-omics analysis brings additional information about prognosis, this would be a more intriguing outcome of this analysis.

Disease activity scores in SLE have been widely implemented and reflect the high degree of clinical heterogeneity of the disease. The manuscript that the reviewer highlights (Thanou et al 2019) provides a very good summary of the diversity of different disease activity scores in SLE and, reassuringly, demonstrates good concordance between them. The need to harmonise the disease activity score between the patients is particularly important for treatment purposes, but often overlooks the heterogeneity of clinical manifestations between patients with active disease.

Because of this challenge, the search for clinical biomarkers in SLE has been a topic of intense research. The best available biomarker in SLE is currently the detection of an IFN signature in blood, which has been shown to be correlated with disease activity scores - albeit not very strongly. In the past we have investigated how different biomarkers such as soluble SIGLEC-1, soluble CD25 or IL-2 concentrations were correlated with both disease activity scores and the classical IFN signature, and have shown that patients with high IFN signature scores are associated with increased disease activity and renal disease involvement (see Oliveira J et al, Arthritis Res Ther. 2018; doi: 10.1186/s13075-018-1649-1). However, it is clear that there is vast clinical heterogeneity, and that no marker is currently sufficient to properly stratify SLE patients.

The main aim of applying single-cell multi-omics technologies in SLE is exactly to better understand the different biological mechanisms that underpin the disease and whether we can better stratify patients from treatment. In this proof-of-principle study, we wanted to investigate whether a more quantitative targeted single-cell multi-omics approach would reveal distinct alterations in the cellular composition of the T and NK cell compartment in SLE patients. For this we deliberately opted to select two extreme patients in terms of the IFN signature score that we had previously quantified. Of note, as we remark in the manuscript, both SLE patients profiled in this study were recruited outside of their regular clinic visits and were therefore assumed to have good disease management at the time of sampling. This strongly suggests that the cellular alterations that we report in this study, as well as the IFN signature in the IFNhi patient profiled in this study may not be associated with active disease activity. This is not unexpected and

> *underscores the relatively low correlation between the transcriptional IFN signature and disease activity score.*
>
> *Our proof-of-concept study in a very small cohort of patients suggests that our targeted single-cell multi-omics approach can detect substantial alterations in the T and NK cell subsets of SLE patients, and therefore supports the application of this approach in a much larger (and clinically characterised) cohort of SLE patients. As the reviewer rightly observes (and as stated above in the replies to reviewers #1 and #2), the ideal experimental setting would be to apply this approach on a prospective longitudinal cohort with ancillary disease activity information. This experimental design would allow to investigate whether the cellular alterations identified by the single-cell multi-omics analysis are indeed prognostic or reflective of flares of disease activity. Our study shows that such analysis is warranted and may provide better insight into the different pathogenic mechanisms that underlie SLE.*
>
> > ○ 2/ Minor comments relate to the length of the discussion which can be more focused.
>
> *We agree that the discussion is long and requires focussing. We have edited the discussion to reduce the size and improve the focus.*
>
> **Competing Interests:** No competing interests were disclosed.

Reviewer Report 01 April 2022

?

Camilla Jandus (iD)
Department of Pathology and Immunology, Faculty of Medicine, University of Geneva, Geneva, Switzerland

Mara Cenerenti (iD)
Department of Pathology and Immunology, Faculty of Medicine, University of Geneva, Geneva, Switzerland

This study describes the profiling of immune cell subpopulations in SLE, using a combined transcriptomic and proteomic methodology, down to the single-cell level.

The main weakness of the study is the very limited number of subjects, which dampens the enthusiasm of the reviewer. Conclusions are drawn on 1 donor per study population (1 patient IFNhi, 1 patient IFNlo, 1 HD, with only limited clinical characteristics provided). It would have been highly recommended to validate the findings obtained by the transcriptomic/Proteomic analyses

in additional samples from IFNhi and IFNlo SLE patients, by a flow cytometry based methodology.

Moreover, while the authors repeatedly stress the fact that they are utilizing a novel method combining protein and mRNA quantification, this type of analysis has also been largely used by others, eg., CITE-seq, on unsorted single immune cells, with good resolution and identification of discrete, rare immune cell states (such as CD4 CTLs).
Finally, while the inclusion of a "stimulated" condition appears very interesting to the reviewer, the rationale to use PMA/Ionomycin instead of aCD3/CD28 activation (at least for T cell subsets) is unclear. This would have provided a more physiological activation, that would better reflect the in vivo events occurring in patients.

Minor comments:
- By considering NK cells as CD56+CD127+ CD3- cells, the authors are including helper ILC3s in their sorting. This might represent a confounding parameter for the results' interpretation.
  Similarly, CD25-CD127+ CD4 T cells will also include Naive CD4 T cells, while in Fig 1B these cells are named Teff.

- Though nicely represented, the readability of several Figure panels is difficult (numbering of clusters Fig 2A, gene names and scale bar Fig 2B, blurry volcano plot Fig 3B,...).

- The discussion is very long, and highly speculative based on the findings obtained in the small patients cohort. I suggest refinement, specifically the sections arguing on the novelty of the methodology.

- The trial NCT04680637 has initiated, the authors should update its status.

Is the work clearly and accurately presented and does it cite the current literature?
Partly

Is the study design appropriate and is the work technically sound?
Partly

Are sufficient details of methods and analysis provided to allow replication by others?
Yes

If applicable, is the statistical analysis and its interpretation appropriate?
Yes

Are all the source data underlying the results available to ensure full reproducibility?
Yes

Are the conclusions drawn adequately supported by the results?
Yes

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: CD4 T cells, cytotoxicity, cancer, Innate lymphoid cells

We confirm that we have read this submission and believe that we have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however we have significant reservations, as outlined above.

Author Response 04 May 2022
Ricardo Ferreira, University of Oxford, Oxford, UK

We appreciate the Reviewer's feedback and suggestions. We detail below the replies to the specific concerns raised by the reviewer:

- 1/ The main weakness of the study is the very limited number of subjects, which dampens the enthusiasm of the reviewer. Conclusions are drawn on 1 donor per study population (1 patient IFNhi, 1 patient IFNlo, 1 HD, with only limited clinical characteristics provided). It would have been highly recommended to validate the findings obtained by the transcriptomic/Proteomic analyses in additional samples from IFNhi and IFNlo SLE patients, by a flow cytometry based methodology.

We agree that the main limitation of this study is indeed the very small sample size. Unfortunately, we do not have access to further clinical samples with ancillary clinical data about disease activity that can be used to validate these findings independently using either our single cell platform or a flow cytometry based methodology. We have included in the manuscript that some of our findings are consistent with recent findings in Sjogren's syndrome, a related systemic autoimmune disease associated with a type I IFN response, where the authors have identified a similar CD4+ CTL population through scRNA-seq, and validated by flow cytometry by identifying a fraction of GZMB+ CD4+ T cells (see ref 20 – Hong X et al; Front Immunol 2021; doi: 10.3389/fimmu.2020.594658). Furthermore, we also cite two other studies where the authors have identified a similar subset of cytotoxic CD4+ T cells in blood by scRNA-seq in the context of bladder cancer (ref 19 – Oh DY et al; Cell 2020; doi: 10.1016/j.cell.2020.05.017) and in a cohort of supercentenarians (ref 21 – Hashimoto et al; PNAS 2019; doi: 10.1073/pnas.1907883116) and further validated the results using orthogonal methods, including flow cytometry. Given these results and the known reproducibility between FACS- and AbSeq-based protein quantification methods (see Mair F et al; Cell Rep 2020; doi: 10.1016/j.celrep.2020.03.063), we are very confident of the reproducibility of this phenotype and the presence of the cell subset in blood.

We note that the main motivation of this study was to provide a proof-of-principle for the use of a targeted scRNA-seq approach to better characterise the heterogeneity of the T and NK cell compartments. We opted to investigate these immune compartments in the context of SLE given our previous experience and data showing substantial alterations in the peripheral immune system of SLE patients with active disease. As stated above in the reply to Reviewer 1 (point #1), the key outstanding question is whether these immune alterations are maintained during the course of the disease, and particularly how they are regulated by increased IFN production during active disease. This study demonstrates that this approach could be well-suited to investigate these questions in a well-characterised longitudinal cohort of SLE patients. Moreover, it provides insight into specific T and NK cell subsets that could be altered in patients with active

disease and that should be further monitored in SLE.

Nevertheless, we note that despite the sample size limitation, the increased resolution provided by this single-cell immunophenotyping approach, combined with the pre-selection strategy, provided a large number of T and NK cells profiled in this study. Furthermore, all samples were processed in parallel using sample barcodes, thereby eliminating most causes of technical variation that usually affect scRNA-seq experiments. These factors increase our confidence in the results reported in this study.

We state these limitations in the discussion and have now reduced the discussion section to provide increased focus on these limitations. We also state more clearly that this is a proof-of-principle study aiming at providing a resource for researchers to investigate the T and NK cell heterogeneity in SLE at the single-cell level.

- ○ 2/ Moreover, while the authors repeatedly stress the fact that they are utilizing a novel method combining protein and mRNA quantification, this type of analysis has also been largely used by others, eg., CITE-seq, on unsorted single immune cells, with good resolution and identification of discrete, rare immune cell states (such as CD4 CTLs).

As the reviewer correctly points out, parallel assessment of mRNA and protein at the single-cell level is not novel. However, at the time of publication of our study these methods have been limited to whole-transcriptome scRNA-seq (ie. CITE-seq) on unsorted immune cells. The main novelty of this method is the combination of a cell subset pre-enrichment approach (by sorting of specific immune cell populations), combined with the design of a custom targeted mRNA panel. This approach substantially differs from the standard whole-transcriptome approaches. In this study we wanted to highlight these differences and strengths of the targeted approach. Importantly, whole-transcriptome methods tend to have very large sequencing requirements for the mRNA libraries, compared to our targeted method. Previously, we have shown that this targeted approach results in increased power to quantify the mRNA targets present in our custom panel (see Genome Med. 2020 Jun 24;12(1):55. doi: 10.1186/s13073-020-00756-z.), leading to improved clustering resolution. This has important implications on the use of the protein quantification on both methods: for traditional whole-transcriptome scRNA-sequencing, protein quantification (ie. CITE-seq) is therefore often restricted to a small panel of lineage-discriminating markers that are mostly used for cluster annotation purposes. In contrast, the reduced sequencing requirements in this targeted approach means that we can devote a much larger proportion of the budget to sequence the protein libraries, resulting in the use of much larger protein panels (including mostly proteins associated with cell function), which can be used to identify more subtle functional changes associated with protein expression changes.

Therefore, the two methods are complementary in their approach. We believe that the targeted approach could be well-suited for a wide range of researchers, as a hypothesis-based approach to immunophenotype larger numbers of cells from specific cell populations in a more cost-effective manner. The custom nature of the mRNA and protein panels makes this a more flexible approach, although it does have the inherent limitation of the target marker selection, which does not replace the discovery-based nature of the traditional whole-transcriptome scRNA-seq methods.

We agree with the reviewers that given the points above we should refrain from claiming novelty, and instead have edited the discussion to focus more on the strengths of the method and differences with more standard combined mRNA/protein quantification approaches at the single-cell level.

- ○ 3/ Finally, while the inclusion of a "stimulated" condition appears very interesting to the reviewer, the rationale to use PMA/Ionomycin instead of aCD3/CD28 activation (at least for T cell subsets) is unclear. This would have provided a more physiological activation, that would better reflect the in vivo events occurring in patients.

The use of the PMA+Ionomycin activation was driven by previous experience in the lab, showing that very short stimulation (only 30-45 min instead of the 4-5h usually used for cytokine quantification at the protein level) induces a very distinct polarisation of the functional T_H lineages. This is mostly driven by the robust transcription of cytokine genes (as well as other activation-induced genes), which can be used to better dissect the heterogeneity of the sorted immune cell populations compared to the ex vivo (unstimulated) condition. Furthermore, as the reviewer points, since one of the motivations of this study was the characterisation of NK cells in SLE – particularly the rarer NK56br subset – we felt that a T-cell specific activation would not be best suited.

In our experience, activation with aCD3/CD28 provides a less robust induction of cytokine genes and requires much longer stimulations times as compared to PMA+Ionomycin. Nevertheless, we agree that for other studies, an aCD3/CD28 activation would be more suited and would yield very interesting data. We believe that this targeted scRNA-seq method is particularly well-suited to investigate more subtle changes associated with different immune perturbations and could therefore be adapted to a number of different experimental models and immune cell populations of interest.

- ○ 4/ By considering NK cells as CD56+CD127+ CD3- cells, the authors are including helper ILC3s in their sorting. This might represent a confounding parameter for the results' interpretation. Similarly, CD25-CD127+ CD4 T cells will also include Naive CD4 T cells, while in Fig 1B these cells are named Teff.

One of the main advantages of single-cell methods over standard bulk assays, is that they are much more tolerant of cell heterogeneity. Indeed, single-cell methods will be able to identify small proportions of distinct non-expected "contaminant" cell populations, as long as they are sufficiently represented in the cell mixture. In contrast, in bulk cell assays, the output of the experiment may well be affected by varying degrees of contamination with different cell types.

Because of this, in this study, we deliberately delineated broader gates to define our target cell populations, to avoid potentially losing any specific immune cell population. In the specific case of ILC3 cells, we did not observe any specific cluster of cells sorted from the CD56+CD127+ CD3- gate (designated NK56br gate in this study), with a profile matching ILC3 cells. This does not exclude the possibility that a small number of ILC3 cells may have been sorted in that gate, but it does ensure that the numbers are negligeable and would not affect the interpretation of our

results.

Regarding the CD25-CD127+ CD4+ T cell gate, we agree that the designation of "CD4+ CD127hi Teff" cells is confusing. As the reviewer rightly points out, there is definitely a substantial proportion of naïve cells sorted from this gate (as annotated on the respective UMAP plots). We have therefore edited the text and figure to now refer to this population as "CD127hi Tconv" cells to more clearly differentiate from the converse CD127lowCD25hi Treg gate. Similarly, we changed the designation of the cells sorted from the "CD127low Teffs" to "CD127low Tconv" for consistency (as this gate was also shown to contain a small fraction of naïve CD4+ T cells).

- ○ 5/ Though nicely represented, the readability of several Figure panels is difficult (numbering of clusters Fig 2A, gene names and scale bar Fig 2B, blurry volcano plot Fig 3B,…).

We have provided higher resolution files for our figures to increase the readability of the panels. We have also increased the font of the cluster names to improve legibility.

- ○ 6/ The discussion is very long, and highly speculative based on the findings obtained in the small patients cohort. I suggest refinement, specifically the sections arguing on the novelty of the methodology.

As detailed above, we have now edited the discussion to shorten it and to refrain from claiming novelty of the methodology, but rather to more clearly detail the differences.

- ○ 7/ The trial NCT04680637 has initiated, the authors should update its status.

We have edited the text in the revised version to reflect the change of status in the NCT04680637 trial.

Competing Interests: No competing interests were disclosed.

Reviewer Report 20 December 2021

https://doi.org/10.21956/wellcomeopenres.18624.r47452

Nancy J. Olsen iD

Division of Rheumatology, Department of Medicine, Penn State MS Hershey Medical Center, Hershey, PA, USA

This is an interesting and detailed study of the immune cell features in patients with systemic lupus erythematosus (SLE). The authors have shown the strength of combined transcriptomic and proteomic profiling in single-cell analyses. The findings suggest that standard lymphocyte subsets, such as CD4 and NK cells, are in fact very complex and that subsets of these cells, even rare ones, have functional features that likely contribute to disease pathogenesis. Strengths of this work include the detailed analyses, the very clearly presented data in multiple figures, and the fact that the proteomic results reinforce functional implications of the various phenotypes defined by flow cytometry.

The major weakness, which is discussed by the authors, is that this is a very limited analysis in terms of the patients sampled: one with high IFN signature, one with low IFN signature, and one healthy subject. The other available data on the two patients, levels of IL2 and circulating sCD25 are consistent with active disease. However, the IFN signature genes were not elevated as expected in the IFNhi individual. The explanation offered is that the IFN levels were determined at a separate time, prior to the sample taken for this study. This might imply that the cellular changes are more prolonged than the signature expression, but in the absence of longitudinal data, that is speculative. In addition, clinical features of the patients to confirm active disease are not described. The results will require further validation in larger patient samples, but the data presented suggest that this will be of interest for future studies. One additional note is that the NCT04680637 trial mentioned has now been started and is in recruiting phase; so that change might be made.

Is the work clearly and accurately presented and does it cite the current literature?
Yes

Is the study design appropriate and is the work technically sound?
Yes

Are sufficient details of methods and analysis provided to allow replication by others?
Yes

If applicable, is the statistical analysis and its interpretation appropriate?
Yes

Are all the source data underlying the results available to ensure full reproducibility?
Yes

Are the conclusions drawn adequately supported by the results?
Yes

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Autoimmune disease, systemic lupus erythematosus

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

Author Response 05 Jan 2022

Ricardo Ferreira, University of Oxford, Oxford, UK

We appreciate the reviewer's feedback. We agree that the more relevant follow up question pertains to the duration of the IFN signature and how this tracks with the cellular alterations that we observe in this study. This can only be addressed in a longitudinal cohort, and hopefully this study will provide the rationale for monitoring CD4+ CTLs in future studies.

We also note that although we do not observe a transcriptional IFN signature in the IFNhi patient, this is only true for the assessed T and NK cell subsets we profiled in this study. We cannot comment on other myeloid populations, where the IFN signature is usually more pronounced and more likely to be detected for a longer period of time.

We also appreciate the reminder about the status of the NCT04680637 trial, and will certainly edit the manuscript accordingly in the next version.

Competing Interests: No competing interests were disclosed.

Author Response 04 May 2022

Ricardo Ferreira, University of Oxford, Oxford, UK

Update (4 May 2022):

We have now edited the text in the revised version to mention that the NCT04680637 trial has started and initiated recruitment.

Competing Interests: No competing interests were disclosed.