



OPEN ACCESS

EDITED BY

Atsushi Nomura,
Juntendo University, Japan

REVIEWED BY

Iris Karina Madera-Salcedo,
Instituto Nacional de Ciencias Médicas y
Nutrición Salvador Zubirán (INCMNSZ),
Mexico
Leonie Taams,
King's College London, London

*CORRESPONDENCE

Sebastien Viatte
✉ sebastien.viatte@manchester.ac.uk

[†]These authors have contributed
equally to this work and share
first authorship

RECEIVED 10 November 2022

ACCEPTED 24 April 2023

PUBLISHED 05 May 2023

CITATION

Mulhearn B, Marshall L, Sutcliffe M,
Hannes SK, Fonseka C, Hussell T,
Raychaudhuri S, Barton A and
Viatte S (2023) Automated clustering
reveals CD4⁺ T cell subset imbalances
in rheumatoid arthritis.
Front. Immunol. 14:1094872.
doi: 10.3389/fimmu.2023.1094872

COPYRIGHT

© 2023 Mulhearn, Marshall, Sutcliffe, Hannes,
Fonseka, Hussell, Raychaudhuri, Barton and
Viatte. This is an open-access article
distributed under the terms of the [Creative
Commons Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The
use, distribution or reproduction in other
forums is permitted, provided the original
author(s) and the copyright owner(s) are
credited and that the original publication in
this journal is cited, in accordance with
accepted academic practice. No use,
distribution or reproduction is permitted
which does not comply with these terms.

Automated clustering reveals CD4⁺ T cell subset imbalances in rheumatoid arthritis

Ben Mulhearn^{1,2†}, Lysette Marshall^{1†}, Megan Sutcliffe¹,
Susan K. Hannes³, Chamith Fonseka^{3,4}, Tracy Hussell⁵,
Soumya Raychaudhuri^{1,3}, Anne Barton^{1,6}
and Sebastien Viatte^{1,5,6*}

¹Versus Arthritis Centre for Genetics and Genomics, Centre for Musculoskeletal Research, Manchester Academic Health Science Centre, The University of Manchester, Manchester, United Kingdom, ²Department of Life Sciences, University of Bath, Bath, United Kingdom, ³Division of Rheumatology, Inflammation, and Immunity, Department of Medicine, Brigham and Women's Hospital and Harvard Medical School, Boston, MA, United States, ⁴Research, eGenesis, Cambridge, MA, United States, ⁵Lydia Becker Institute of Immunology and Inflammation, Faculty of Biology, Medicine and Health, The University of Manchester, Manchester, United Kingdom, ⁶NIHR Manchester Musculoskeletal Biomedical Research Centre, Central Manchester NHS Foundation Trust, Manchester Academic Health Science Centre, Manchester, United Kingdom

Background: Despite the report of an imbalance between CD4⁺ T helper (Th) cell subsets in rheumatoid arthritis (RA), patient stratification for precision medicine has been hindered by the discovery of ever more Th cell subsets, as well as contradictory association results.

Objectives: To capture previously reported Th imbalance in RA with deep immunophenotyping techniques; to compare hypothesis-free unsupervised automated clustering with hypothesis-driven conventional biaxial gating and explore if Th cell heterogeneity accounts for conflicting association results.

Methods: Unstimulated and stimulated peripheral blood mononuclear cells from 10 patients with RA and 10 controls were immunophenotyped with a 37-marker panel by mass cytometry (chemokine receptors, intra-cellular cytokines, intra-nuclear transcription factors). First, conventional biaxial gating and standard definitions of Th cell subsets were applied to compare subset frequencies between cases and controls. Second, unsupervised clustering was performed with FlowSOM and analysed using mixed-effects modelling of Associations of Single Cells (MASC).

Results: Conventional analytical techniques fail to identify classical Th subset imbalance, while unsupervised automated clustering, by allowing for unusual marker combinations, identified an imbalance between pro- and anti-inflammatory subsets. For example, a pro-inflammatory Th1-like (IL-2⁺ T-bet⁺) subset and an unconventional but pro-inflammatory IL-17⁺ T-bet⁺ subset were significantly enriched in RA (odds ratio=5.7, $p=2.2 \times 10^{-3}$; odds ratio=9.7, $p=1.5 \times 10^{-3}$, respectively). In contrast, a FoxP3⁺ IL-2⁺ HLA-DR⁺ Treg-like subset was reduced in RA (odds ratio=0.1, $p=7.7 \times 10^{-7}$).

Conclusion: Taking an unbiased approach to large dataset analysis using automated clustering algorithms captures non-canonical CD4⁺ T cell subset imbalances in RA blood.

KEYWORDS

precision medicine, rheumatoid arthritis, CD4⁺ T cells, mass cytometry, heterogeneity

1 Introduction

Current treatment strategies for rheumatoid arthritis (RA) standardise treatments across patient groups, but as a heterogeneous disease, differences will exist in the underlying immune mechanisms. As a consequence, not all patients will respond similarly to the same drug with only 60-70% having a response to any biologic drug (1). Given that an increase of pro-inflammatory and/or a decrease of anti-inflammatory CD4⁺ T cell subsets has been reported in some patients, these cell types could represent therapeutic targets and aid patient stratification for precision medicine. However, which CD4⁺ T cell subset is associated with the disease, and in what proportion of patients, remains unclear.

The study of T cells in RA was revolutionised after the Th1/Th2 paradigm was proposed by Mossman et al. in 1986 (2). This paradigm was described as a dichotomy between type 1 T helper cells (Th1), characterised by interferon- γ (IFN- γ) and tumour necrosis factor (TNF), and type 2 T helper cells (Th2). RA pathogenesis predominantly involves Th1 cytokines, and the Th1 immune response is antagonised by Th2 cytokines [reviewed in (3)]. The Th1/Th2 paradigm was modified further when pathogenic IL-17-producing T helper cells (Th17) and regulatory T cells (Tregs) were discovered to play critical roles in initiating and regulating autoimmunity, respectively [reviewed in (4)]. Increases in Th1 and Th17 cells which antagonise, and are antagonised by, Th2 and Treg cells has been the focus of much research in autoimmune diseases including RA (5, 6). More recently, interferon-producing Th1 memory cells were found to be associated with RA using novel techniques of analysing high-dimensional single-cell data (7). A meta-analysis found that circulating Treg cells, as defined by both FOXP3 and CD25, were decreased in RA patients in 9 combined studies (8). However, other studies found no decrease in Treg frequency or function in RA (9). To add to the complexity of the paradigm, there are an ever-growing list of T helper cells [Th9 cells (10); Th22 cells (11); T follicular helper cells (Tfh) (12); peripheral helper T cells (Tph) (13)]. Critically, it has now been shown that many of these T cell types exhibit plasticity. It was previously thought that the mutually exclusive expression of a master transcription factor determined the fate of T cells, with T-bet, GATA3, ROR γ t and FoxP3 determining the fate of Th1, Th2, Th17 and Treg cells respectively. However, co-expression of these transcription factors may temporarily alter the effector function of T cell subsets [reviewed in (14)] and many of these T cells might also be transitional and therefore appear only in very low frequencies.

The plasticity of T cell subsets and heterogeneity of RA may also explain why clinical trials targeting specific cytokines, for instance IL-17, have not reached their primary endpoint (15). Stratified medicine may allow clinicians to identify RA patients who have predominantly IL-17-driven disease to improve the design of such trials, as researchers have successfully shown in psoriatic arthritis (16). Exploring peripheral blood immunophenotypes by flow cytometry in RA has been shown to mirror findings in the synovial compartment (17) and still represents the most widely used and cost-effective technique to enumerate immune cell populations. The advent of mass cytometry (CyTOF) has vastly increased the number of cellular markers which can be detected simultaneously (18) and together with the development of high-throughput automated methods of data analysis (19–23), rare and novel cell types involved in the aetiology of RA have been recently identified (7, 13, 24, 25). Automated cell clustering algorithms allow for unbiased marker combinations and therefore the hypothesis-free discovery of unanticipated cell subsets, relevant to disease and precision medicine (7, 22, 26, 27). We postulate that Th cell plasticity and overlap between subsets and definitions explain conflicting associations and lack of reproducibility in small sample sizes. Therefore, we developed a 37 marker T cell mass cytometry panel to encompass most definitions for the most studied CD4⁺ T cell subsets to date (Th1, Th2, Th17 and Treg).

The aims of this study are first, to confirm the T helper cell subset imbalance previously described in RA; second, to explore whether this imbalance is detectable in a small sample size, if the standard definition of Th subsets is relaxed in favour of Th cell plasticity (unbiased marker combinations) and third, to test whether innovative techniques (mass cytometry and unsupervised clustering algorithm) facilitate the identification of pro- and anti-inflammatory CD4⁺ T cell subsets over standard techniques (flow cytometry and manual bi-axial gating) in a small sample size.

2 Methods

2.1 Patient and public involvement

The Versus Arthritis Centre for Genetics and Genomics in Manchester has a Research User Group (RUG) of patients with various rheumatologic conditions, which includes RA. Members of the RUG meet regularly to review the research carried out in the Centre. They highlighted the importance of understanding basic

mechanisms of disease from a patient's perspective. The RUG supports research of basic disease mechanisms to identify biomarkers for stratification into treatment response categories. Some of the following comments were made: "the sooner anyone with RA can get on the correct treatment, the better"; "Personalising from onset would be perfect"; "If only a blood test is needed, then even better".

2.2 Patient selection and sample collection

A cohort of 10 RA patients stabilised on therapy and 10 healthy volunteers were recruited from the National Repository (North West Ethics committee approval MREC 99/8/84) at the Manchester Royal Infirmary (NIHR portfolio ID 7881). Peripheral Blood Mononuclear Cells (PBMC) were isolated from 18 ml of blood by density gradient centrifugation using Ficoll-Paque plus (GE Healthcare Life Sciences) and cryopreserved at -150°C .

2.3 Mass cytometry antibody panel

We adapted a previously published mass cytometry T cell panel (13) to encompass surface (chemokine receptors), intracytoplasmic (cytokines) and intra-nuclear markers (transcription factors) used to define Th and Treg subsets (Table 1).

2.4 T cell enrichment and stimulation

Thawed PBMCs were rested at 37°C for one hour prior to enrichment of $\text{CD}3^{+}$ T cells by positive selection using magnetic cell separation (MACS) with CD3 MicroBeads (Miltenyi Biotec). T cell receptor stimulation was achieved with Dynabeads Human T-Activator CD3/CD28 beads (Fisher Scientific UK Ltd) at a concentration of 1 bead per 2 cells, in the presence of brefeldin A and monensin (both Fisher Scientific UK Ltd) and incubated for 4.5 hours at 37°C .

2.5 Mass cytometry staining protocol

$\text{CD}3^{+}$ T cells were incubated with cisplatin prior to incubation with an extracellular antibody cocktail (detailed protocol in Supplementary Methods). Subsequently, cells were fixed, permeabilised and stained for intracellular antigens for 30 minutes on ice. A solution of Intercalator-Ir was added to each well. Samples were run on Helios mass cytometers by the Longwood Medical Area CyTOF Core Facility at the Dana-Farber Cancer Institute, Boston, USA.

2.6 Data pre-processing

The Nolan Normalizer MATLAB plugin (28) was used to normalise the signal in each channel to the signals from the Maxpar Calibration beads.

2.7 Traditional gating analysis

Traditional biaxial gating was carried out using FlowJo V.10.8 (BD Biosciences) to identify well-described $\text{CD}4^{+}$ Th and Treg subsets using standard definitions (29) (example gating strategy in Supplementary Figure 1). The percentages of positive cells are represented as a proportion of the parent T cell population ($\text{CD}4^{+}$ or $\text{CD}8^{+}$). Geometric mean intensity is used to quantify cytokine expression on an individual cell level. The Mann-Whitney U test was used to compare RA and Healthy Controls (HC). All p -values are unadjusted (tests are not independent).

2.8 Automated clustering workflow

The CyTOF clustering workflow from Nowicka et al. (20) was modified to include an extended quality control approach and a different statistical framework for association testing (see below). The analysis was performed in R (v4.1.0) and all plots were produced with *ggplot2* (v3.3.5) (30), unless stated otherwise. The parent population for the automated analysis was $\text{CD}3^{+}$ cells, with cleaned $\text{CD}4^{+}$ and $\text{CD}8^{+}$ populations combined.

2.8.1 Quality control steps

Normalised cytometric data was manually inspected using FlowJo V.10.8 (BD Biosciences) to ensure that at least 10 cell events were identified in a group in at least 3 samples. IL-5 was excluded due to the absence of any positive cell groups (Supplementary Figure 2). Furthermore, only samples with at least 1000 live single T cells were included in subsequent analyses. Cell events with extreme expression were excluded per individual marker. After the removal of unsuccessful markers and extreme events, analysis was performed on data transformed with the inverse hyperbolic sine (arcsinh) function (cofactor = 5). Finally, to identify outlying samples and potential batch effects, multidimensional scaling (MDS) and principal component analysis (PCA) plots were produced from median expression of panel markers in each sample.

2.8.2 Clustering with FlowSOM/ConsensusClusterPlus

The dataset was downsampled to an equal number of randomly selected cells (n) from each sample, where n was equal to the number of cells in the smallest sample. The automated clustering steps were performed with *FlowSOM* (v2.0.0) (19) and *ConsensusClusterPlus* (v1.56.0) (23) using agglomerative hierarchical consensus clustering. Cells were clustered into 20 populations using Euclidean distance and average linkage (Supplementary Figure 3). Heatmaps of median marker expression for each cluster were generated with *heatmap* (v1.0.12) (31). For visualisation purposes, these values represent the median of the arcsinh-transformed, 0 – 1 scaled expression data. For 0 – 1 scaling, the 0.01 and 0.99 quantiles for each marker were used as the lower and upper boundary, respectively.

TABLE 1 The antigen target, heavy metal and clone of each conjugated antibody used in the CyTOF T cell panel.

Target	Epitope location	Marker significance	Label	Clone
CD8a	Extracellular	Cytotoxic T cell	141Pr	RPA-T8
HLA-DR	Extracellular	Activation marker	143Nd	L243
IL-2	Intracellular	T cell cytokine	144Nd	MQ1-17H12
CD4	Extracellular	T helper cell	145Nd	RPA-T4
CD127 (IL-7Ra)	Extracellular	Cytokine receptor	147Sm	A019D5
CD278/ICOS	Extracellular	Checkpoint protein	148Nd	C398.4A
CD200	Extracellular	Myeloid cell regulator	149Sm	OX-104
IL-22	Intracellular	Th2 cytokine	150Nd	22URTI
IL-5	Intracellular	Th2 cytokine	151Eu	TRFK5
TNF	Extracellular	Inflammatory cytokine	152Sm	Mab11
CD62L (L-selectin)	Extracellular	Adhesion molecule	153Eu	DREG-56
CD38	Extracellular	Lymphocyte activator	154Sm	HIT2
CD279 (PD-1)	Extracellular	T cell suppressor	155Gd	EH12.2H7
CD134 (OX40)	Extracellular	Co-stimulatory molecule	156Gd	ACT35
IFN- γ	Intracellular	Th1 cytokine	158Gd	B27
FoxP3	Intracellular	Treg marker	159Tb	259D/C7
Tbet	Intracellular	Th1 transcription factor	160Gd	4B10
CD197 (CCR7)	Extracellular	Chemokine receptor involved in lymph node homing	161Dy	G043H7
CD152 (CTLA-4)	Extracellular	Checkpoint protein	162Dy	14D3
CD183 (CXCR3)	Extracellular	Th1 chemokine receptor	163Dy	G025H7
IL-17A	Intracellular	Th17 cytokine	164Dy	N49-653
CD45RO	Extracellular	Memory T cell marker	165Ho	UCHL1
IL-10	Intracellular	Regulatory cytokine	166Er	JES3-9D7
Gata3	Intracellular	Th2 transcription factor	167Er	TWAJ
ROR gamma	Intracellular	Th17 transcription factor	168Er	600214
CD25 (IL-2R)	Extracellular	IL-2 receptor	169Tm	2A3
CD28	Extracellular	Costimulatory molecule	170Er	CD28.2
CD185 (CXCR5)	Extracellular	T follicular helper cell chemokine receptor	171Yb	RF8B2
IL-21	Intracellular	Regulatory cytokine	172Yb	3A3-N2
IL-4	Intracellular	Th2 cytokine	173Yb	MP4-25D2
CD154 (CD40L)	Extracellular	Activated T cell marker	174Yb	24-31
Perforin	Intracellular	Cytotoxic T cell effector protein	175Lu	B-D48
CD196 (CCR6)	Extracellular	Th17 chemokine receptor	176Yb	11A9
Iridium - CellID		DNA binder to identify cell events	191Ir	
Iridium - CellID		DNA binder to identify cell events	193Ir	
Cisplatin		DNA binder to assess cell viability	194Pt	
CD3	Extracellular	Pan-T cell marker	209Bi	UCHT1

Target location indicates if the antibody was used in the extracellular or intracellular antibody cocktail. All antibodies were purchased from Fluidigm.

2.8.3 Cluster visualisation with t-SNE

The similarity of single cells in two-dimensional space was visualised with the dimensionality reduction technique t-stochastic neighbour embedding (t-SNE), implemented with *Rtsne* (v0.15) (32). Clusters were annotated manually by their surface phenotype.

2.9 Statistical analysis

To identify clusters with differential representation in RA samples compared to healthy controls, we used mixed-effects modelling of associations of single cells (MASC) (7) for each cluster, including age and sex of donors as fixed-effects covariates (justification in the quality control section), sample ID as a random-effect covariate and the case-control status as the contrast term. Clusters that are significantly enriched or depleted in RA are defined as those with a *P* value ≤ 0.05 after Bonferroni correction (*n* = number of clusters tested).

2.10 Data sharing

All mass cytometry data is freely available at <https://flowrepository.org/id/FR-FCM-Z5RM> (33).

3 Results

3.1 Study design and subject characteristics

PBMCs from 10 RA patients and 10 healthy controls (Table 2) were left unstimulated or stimulated with anti-CD3/CD28 beads (*n* = 40 samples) prior to deep immunophenotyping by mass cytometry.

TABLE 2 Subject characteristics of cases (RA) and healthy controls.

	Cases (n = 10)	Controls (n = 10)	<i>P</i> value
Age (mean ± SD)	61 ± 14	46 ± 8	0.02
Female	7	6	1
RF-positive	7	—	—
ACPA-positive	6	—	—
DAS28 (mean ± SD)	3.4 ± 2.2	—	—
Glucocorticoids	2	0	—
Methotrexate	5	0	—
Other DMARD*	3	0	—
Biologics	1	0	—

*Sulfasalazine or hydroxychloroquine
 RF, rheumatoid factor; ACPA, anti-citrullinated protein antibody; DAS, disease activity score; DMARD, disease-modifying anti-rheumatic drug.
P values calculated for age and gender by Mann-Whitney U and Fisher's exact test, respectively.

Manual gating with traditional statistical analysis fails to identify large differences between blood from patients with RA and healthy controls.

Sequential biaxial gating of CyTOF data was used to identify commonly described CD4⁺ and CD8⁺ T cell subsets as defined by expression of surface markers, transcription factors and/or intracellular cytokines. Conventional statistical analysis with Mann-Whitney *U* found no significant differences in the proportions of Th CD4⁺ subsets between RA and HC (Figure 1); only Treg (defined as CD4⁺ FoxP3⁺) were markedly decreased in RA. No CD8⁺ T cell subsets were found to be differentially abundant in RA (Figure 2).

3.2 CyTOF automated clustering pipeline quality control steps

Low quality cells with extreme expression values were excluded after individual appraisal of expression distributions for each marker. This step removed 0.34% (4,146/1,230,364) of cells from the full dataset, with between 0.09% and 1.4% of cells removed from each sample. Diagnostic MDS and PCA plots suggest a differential trend between RA and HC (MDS2 in Figure 3A, PC1 and PC2 in Figure 3B), indicating slight differences in marker expression between the two groups (Supplementary Figure 4). Data structure of this kind is likely caused by study design or technical artefacts (e.g. batch effects), which requires correction, when the potential confounder is identified. As age is known to associate with the frequency of several cell types (e.g. naïve T cell populations), we studied the effect of age on data structure: Figure 3C shows that samples largely cluster by age group, similarly to previous reports (34), suggesting that the structure in the data is attributable to participant age rather than technical batch effects. Therefore, we included age as a covariate for statistical association testing.

3.3 Automated clustering identifies well-described T cell subsets

Automated clustering was performed on downsampled quality-controlled CyTOF data; analyses of unstimulated and stimulated T cell datasets were performed separately. For both datasets, T cell subsets were identified from a parent population of CD3⁺ T cells, including both CD4⁺ and CD8⁺ populations. Within the unstimulated dataset, we identified 20 distinct T cell subsets, including 11 CD4⁺ and nine CD8⁺ (Figure 4A). Amongst the CD4⁺ populations are a FoxP3⁺ subset (subset 2U; a Treg-like subset), two IL-17⁺ subsets (subsets 10U and 1U; Th17 subsets), a Th1-like subset (subset 4U) and a CD38⁺ subset (subset 19U).

Although analysis was performed independently, clusters identified in the stimulated dataset were broadly similar to those of the unstimulated dataset (Figure 4B). As expected, t-SNE plots coloured by expression level of several markers highlight an increase in the number of cells expressing TNF upon stimulation (Figures 4C, D).

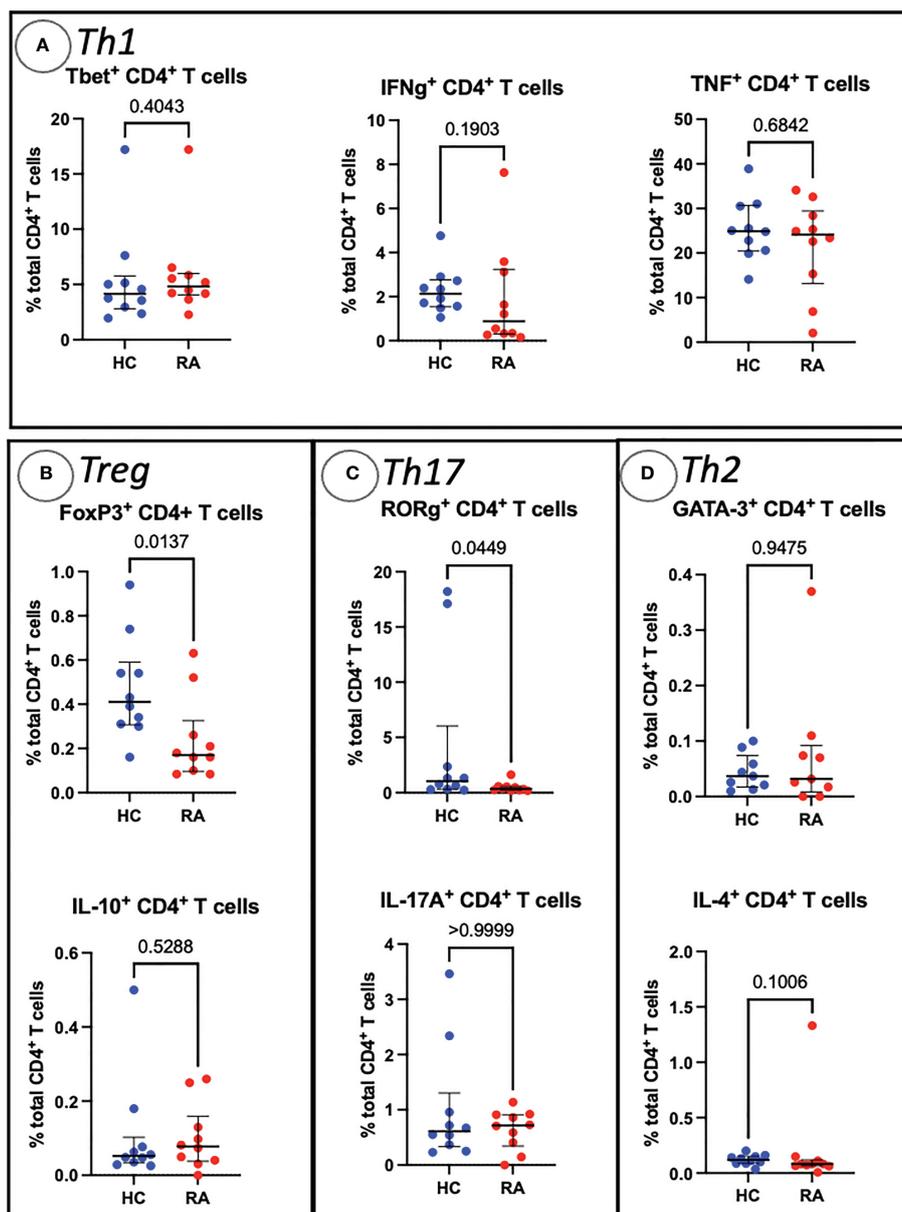


FIGURE 1
 Abundances of Th and Treg CD4⁺ T cell subsets identified by manual gating of CyTOF data shown with median, interquartile range and P value (shown above the pairwise comparisons). Points represent RA and HC samples. Parent population is labelled on the y-axis. Mann Whitney U P values are shown above the pairwise comparisons. Graphs with cytokine expression are taken from the stimulated T cell dataset and those with surface markers and transcription factors are taken from the unstimulated dataset. (A) Represents markers associated with Th1 cells, (B) Treg cells, (C) represents Th17-related markers, and (D) Th2-related markers.

3.4 MASC identifies multiple T cell clusters with differential expression between RA and healthy samples

MASC was used to identify clusters with differential abundance between RA samples and HC, adjusting for the effects of age, gender, and donor. Six unstimulated cell subsets and eight stimulated subsets were found to be differentially abundant with MASC (Figure 5). The majority were CD4⁺ and, of these, three

populations were present in both unstimulated and stimulated states, including enrichment of Th1 and Th17 subsets (CD4⁺ Tbet⁺ IL-17⁺ and CD4⁺ Tbet⁺ IL-2⁺ cell clusters) and depletion of a Treg-like subset (CD4⁺ FoxP3⁺ IL-2⁺) in RA (Table 3, Figure 6). Interestingly, a Th2 subset (CD4⁺ T_{CM} GATA3⁺) was decreased in RA (although it did not reach statistical significance).

Notably, Figures 7, 8 highlight the heterogeneity in subset proportions between individual patients (raw values in Supplementary Tables 1, 2). For example, the abundance of the

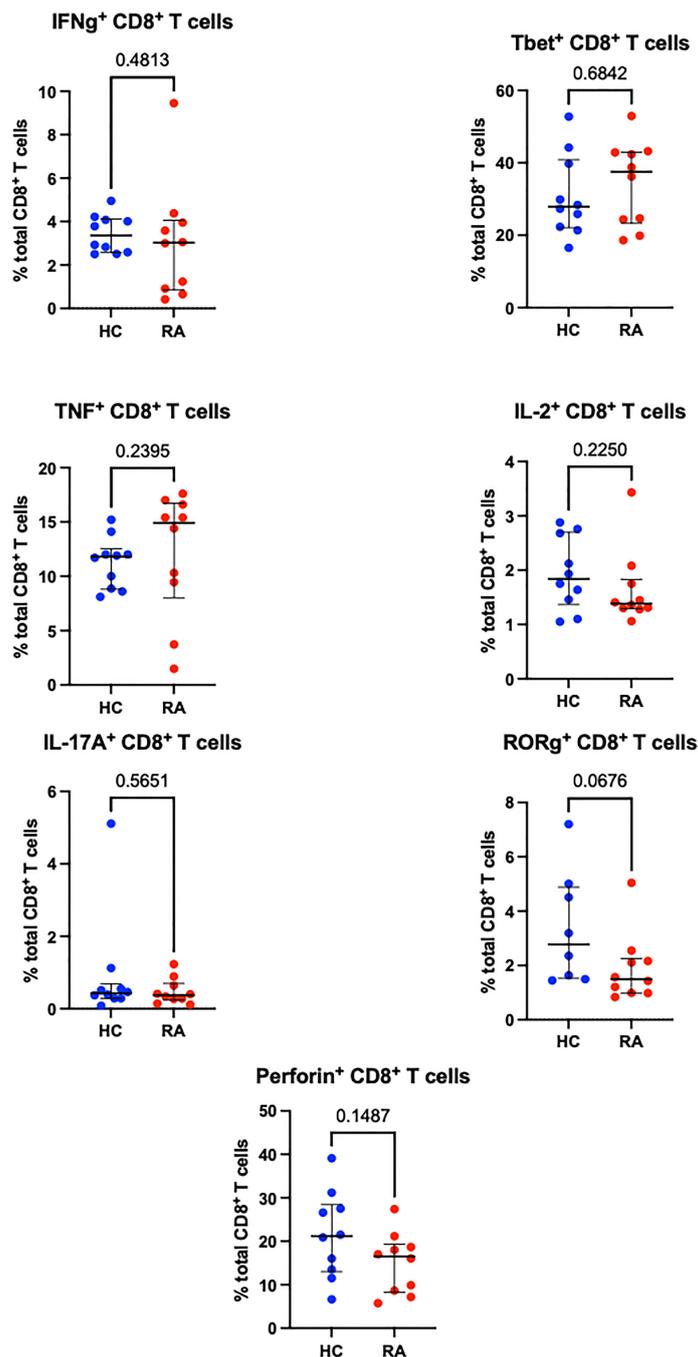
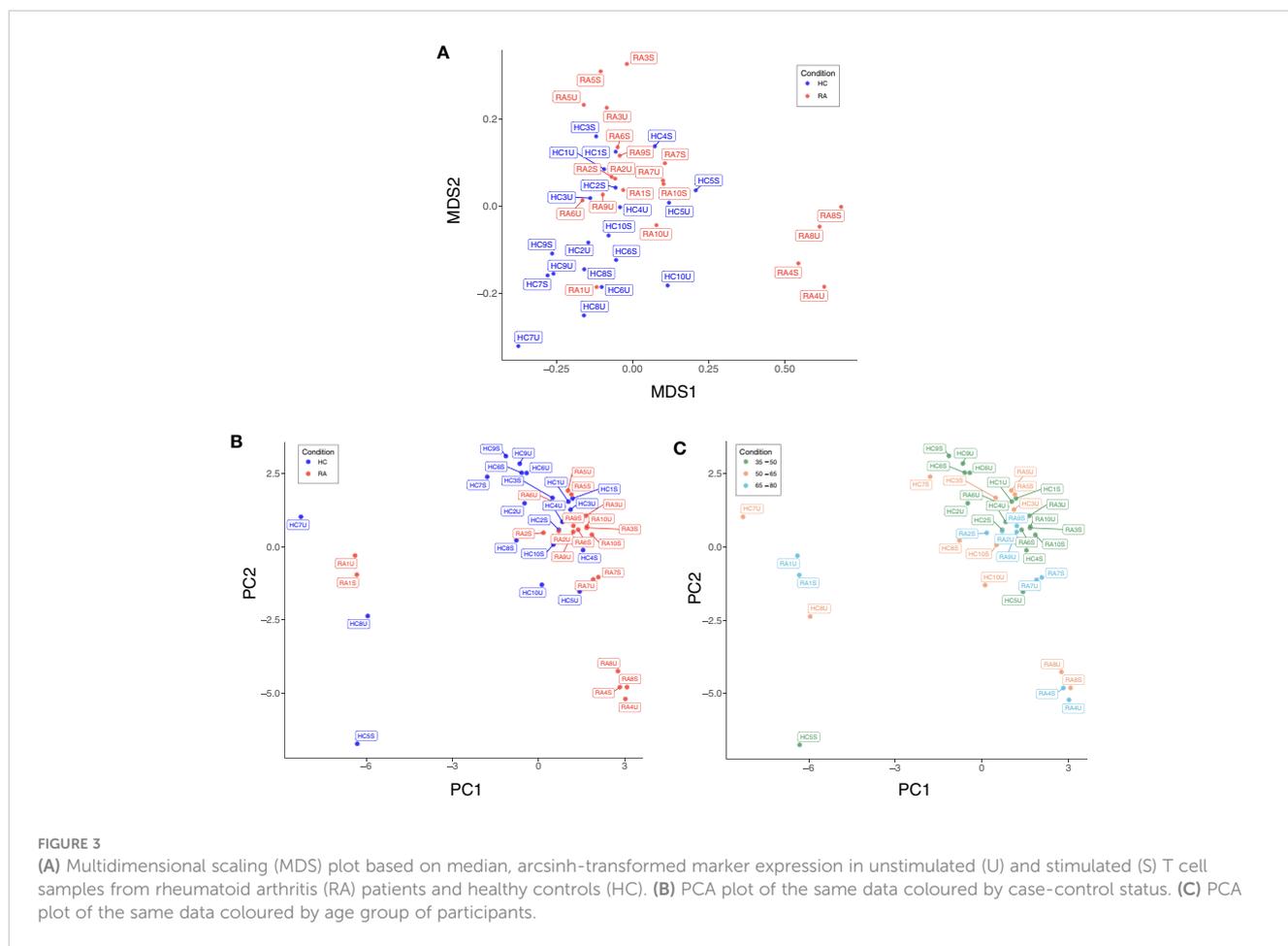


FIGURE 2
 Abundances of CD8⁺ T cell subsets identified by manual gating of CyTOF data shown with median and interquartile range, and P value. Points represent RA and HC samples. Parent population is labelled on the y-axis. Mann-Whitney U P values are shown above the pairwise comparisons. Graphs with cytokine expression are taken from the stimulated T cell dataset and those with surface markers and transcription factors are taken from the unstimulated dataset.

CD4⁺ TNF⁺ cells range from < 1% to > 7% of cells in stimulated RA samples (Figure 8, subset 8S). Before any correction, naïve T cells were found to be increased in frequency in controls compared to cases, as expected by the age difference between the two groups. This difference disappeared after statistical adjustment for age implemented in MASC,

demonstrating that the correction for the age-driven stratification observed in the PCA plots was sufficient in terms of statistical outcome.

Cell cluster phenotype, abundances, odds ratios, and P values for all cell clusters generated by the pipeline are presented in Table 3.



3.5 Agnostically defined cell clusters can be identified using conventional techniques

To provide internal validation of the automated pipeline’s significant clusters, manual gating and standard statistical analysis of agnostically defined clusters was performed. Based on the phenotypes presented in **Figure 4** and the level of expression of each marker, gates were drawn by eye around all positive cells if the cluster had low expression of a marker on the heatmap, but only around high positive cells if the marker had high expression.

Using these definitions of T cell subsets and a standard statistical approach, manual gating was able to identify a clear imbalance between pro-inflammatory Th subsets and anti-inflammatory subsets in RA. For example, out of the unstimulated dataset, manual gating was able to identify that subset 2U (CD4⁺ Tcm FoxP3⁺ IL-2⁺), a Treg-like subset, is decreased in RA ($P = 0.0004$; **Figure 9**). The same population could also be identified in the stimulated dataset (subset 2Sb, CD4⁺ Tcm FoxP3⁺ IL-2⁺, $P = 0.0002$). Pro-inflammatory Th cells were also clearly identified and statistically significantly increased in RA (CD4⁺ Tcm Tbet⁺ IL-17⁺, subset 1U, $P = 0.0172$; and subset 1S, $P = 0.0022$; **Figure 9**). Therefore, innovative approaches easily capture known CD4⁺ T cell subset imbalances in RA blood, even in a small sample size, by allowing for unconventional marker combinations.

4 Discussion

We show the superiority of innovative immunophenotyping and automated analytical techniques over hypothesis-driven conventional biaxial gating (based on canonical biomarkers) to identify a T cell subset imbalance in RA peripheral blood in small sample sizes. Our strategy allows for unusual combinations of T cell markers, for instance T-bet and IL-17, capturing the plasticity of T helper cell populations, a more difficult task to achieve with standard gating strategies.

Th17 cells are known to be phenotypically unstable (35). In our study, T-bet⁺ IL-17⁺ T cells are increased in RA blood. These cells have been termed Th1-like Th17 cells in the literature, co-produce IL-17 and IFN- γ , and are known to contribute to inflammation in the context of autoimmunity (35–37). Interestingly, IFN- γ production in Th1-like Th17 cells has been shown to be repressed by miR-146a (37) and loss of function polymorphisms in the microRNA-146a (miR-146a) gene have been associated with an increased risk of developing RA and lupus in both European and Asian genetic association studies (38, 39); taken together, these data provide some mechanistic evidence that Th1-like Th17 cells are involved in the pathogenesis of RA. The methods outlined in this paper easily detect an increase of this subset in RA.

We also found a significant decrease in the number of IL-2⁺ FoxP3⁺ T cells (Treg-like cells) in RA samples compared to

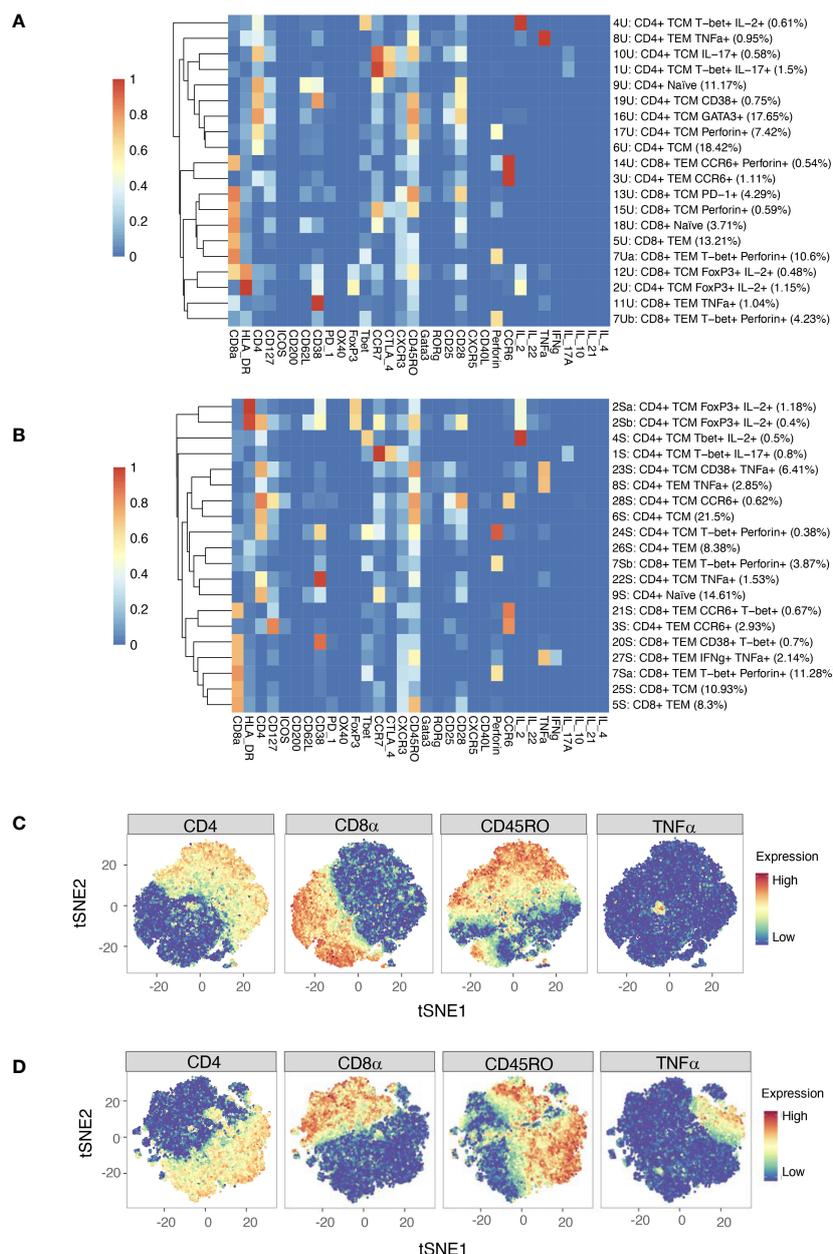


FIGURE 4

(A) Heatmap of median marker expression (arcsinh, 0–1 transformed) for 20 T cell clusters identified in unstimulated cell samples by automated clustering. Each sample was downsampled to 2746 cells ($n = 10$ RA, 10 HC). The dendrogram represents hierarchical clustering using Euclidean distance and average linkage. (B) As (A), for stimulated cell samples ($n = 10$ RA, 10 HC). Each sample was downsampled to 2224 cells. (C) t-SNE plots of the same unstimulated cells, coloured by expression level of a selection of markers. (D) As (C), for stimulated cells. Note that clustering was performed on a parent population comprising both CD4⁺ cells and CD8⁺ cells.

controls. It is well-established that IL-2 is critical in maintaining Treg function (40) and that some Tregs produce IL-2 (41). However, the relative role of paracrine versus autocrine IL-2 secretion from nearby effector T cells (42) versus IL-2⁺ FoxP3 T cells (41) is unclear, and the low level or absence of CD25 expression on the IL-2⁺ FoxP3⁺ subsets in this data appears to conflict with classical Treg definitions (reviewed by (43)). Other studies have characterised subpopulations of CD4⁺ FoxP3⁺ memory cells that produce IL-2 (44). Consistent with our findings, the authors show that these cells possess a phenotype that lies between Tregs and other T helper cells,

where they exhibit lower expression of classical Treg markers, including CD25, than their non-IL-2-producing counterparts. Therefore, rather than identifying *bona fide* Tregs, the unsupervised analysis has potentially identified IL-2-producing Treg-like subpopulations that may be more relevant for RA. Regardless, these findings highlight the utility of automated analyses for capturing non-canonical cell types, better reflecting the heterogeneity and plasticity of T cells.

Due to our small sample size, our association results could theoretically represent false positive due to a sampling bias or

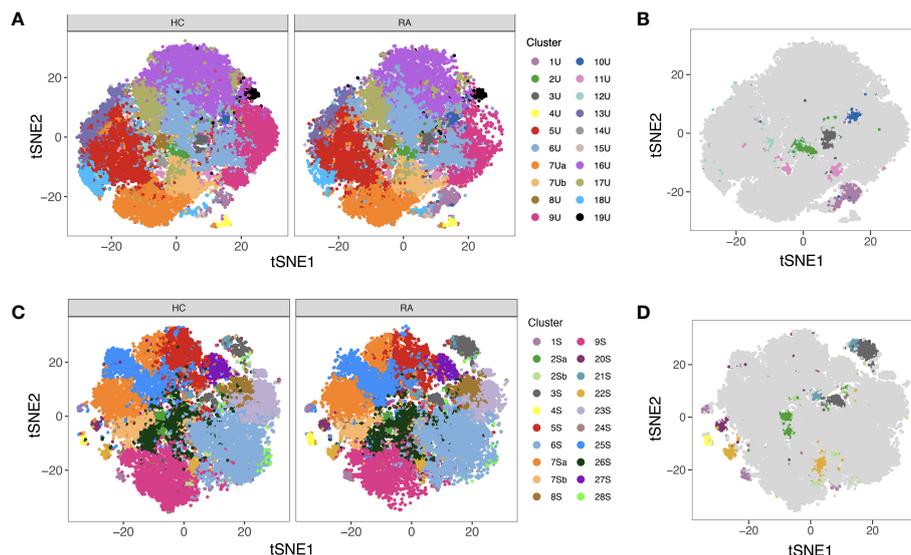


FIGURE 5
(A) tSNE plots showing unstimulated cells (one point = a single cell), each coloured by the cluster they were assigned to during automated clustering of merged CD4⁺ and CD8⁺ cells. The cells are split by those derived from healthy control (HC) samples (left) and those derived from RA samples (right) to show potential differences in cluster size between conditions. Unstimulated samples (n = 10 RA, 10 HC) were downsampled to 2746 cells each prior to clustering. **(B)** The same tSNE plot of unstimulated T cells, now showing all RA and healthy cells together. Clusters that were found to be differentially abundant in RA vs healthy controls according to MASC are shown in colour (P ≤ 0.05), whilst those not differentially abundant are shown in grey. **(C)** As A, but for stimulated cells. Stimulated samples (n = 10 RA, 10 HC) were downsampled to 2224 cells each prior to clustering. **(D)** As **(B)**, for stimulated cells. Note that cell subsets identified in both states (unstimulated and stimulated) are denoted by the same colour across subfigures.

TABLE 3 MASC results for clusters present in **A)** both unstimulated (U) and stimulated (S) samples, **B)** unstimulated samples alone, and **C)** stimulated samples alone.

Subset	Cluster number	% total RA cells	% total HC cells	OR (RA vs HC)	P value
A) Shared subsets (Unstimulated, Stimulated)					
CD4 ⁺ TCM T-bet ⁺ IL-17 ⁺ (combined Th1 and Th17 phenotype)	1U	2.6	0.4	12.2 (5.3 – 28.1)	1.5 x 10 ⁻⁴
	1S	1.4	0.2	9.7 (3.8 – 24.5)	1.5 x 10 ⁻³
CD4 ⁺ TCM FoxP3 ⁺ IL-2 ⁺ (Treg-like)	2U	0.5	1.8	0.3 (0.2 – 0.5)	4.6 x 10 ⁻³
	2Sa + 2Sb	0.4/0.09	2.0/0.7	0.1 (0.1 – 0.2)/0.03 (0.01 – 0.09)	7.7 x 10 ⁻⁷ /1.9 x 10 ⁻⁷
CD4 ⁺ TEM CCR6 ⁺ (Th17)	3U	1.6	0.6	4.4 (2.2 – 8.7)	5.1 x 10 ⁻³
	3S	4.8	1.1	9.4 (4.6 – 19.5)	7.9 x 10 ⁻⁵
CD4 ⁺ TCM T-bet ⁺ IL-2 ⁺ (Th1)	4U	0.8	0.4	3.5 (1.6 – 7.4)	6.8 x 10 ⁻²
	4S	0.8	0.2	5.7 (2.6 – 12.3)	2.2 x 10 ⁻³
CD8 ⁺ TEM	5U	15.9	10.5	1.7 (1.1 – 2.7)	3.4 x 10 ⁻¹
	5S	6.8	9.8	0.6 (0.4 – 0.9)	5.4 x 10 ⁻¹
CD4 ⁺ TCM	6U	22.8	14.0	1.5 (0.9 – 2.7)	1
	6S	22.5	20.6	0.9 (0.5 – 1.5)	1
CD8 ⁺ TEM T-bet ⁺ Perforin ⁺	7Ua + 7Ub	12.2/3.7	9.0/4.8	1.0 (0.5 – 1.9)/0.8 (0.4 – 1.5)	1
	7Sa + 7Sb	12.7/3.3	9.9/4.5	0.9 (0.5 – 1.9)/0.9 (0.4 – 2.0)	1
CD4 ⁺ TEM TNFα ⁺ (Th1)	8U	1.0	0.9	1.1 (0.6 – 1.9)	1
	8S	3.8	1.9	1.9 (0.8 – 4.7)	1

(Continued)

TABLE 3 Continued

Subset	Cluster number	% total RA cells	% total HC cells	OR (RA vs HC)	P value
CD4 ⁺ Naïve	9U	7.0	15.4	0.6 (0.3 – 1.1)	1
	9S	11.8	17.4	0.8 (0.4 – 1.4)	1
B) Unstimulated-specific subsets					
CD4 ⁺ TCM IL-17 ⁺ (Th17)	10U	0.9	0.2	5.7 (3.1 – 10.3)	4.0 x 10⁻⁴
CD8 ⁺ TEM TNFα ⁺	11U	1.5	0.6	3.7 (2.0 – 7.1)	1.4 x 10⁻²
CD8 ⁺ TCM FoxP3 ⁺ IL-2 ⁺	12U	0.5	0.5	0.2 (0.1 – 0.5)	2.5 x 10⁻²
CD8 ⁺ TCM PD-1 ⁺	13U	2.4	6.2	0.4 (0.2 – 0.7)	6.5 x 10 ⁻²
CD8 ⁺ TEM CCR6 ⁺ Perforin ⁺	14U	0.8	0.3	3.5 (1.6 – 7.6)	7.3 x 10 ⁻²
CD8 ⁺ TCM Perforin ⁺	15U	0.9	0.2	3.8 (1.6 – 8.7)	9.4 x 10 ⁻²
CD4 ⁺ TCM GATA3 ⁺	16U	15.0	20.3	0.6 (0.3 – 1.2)	1
CD4 ⁺ TCM Perforin ⁺	17U	6.9	7.9	0.7 (0.4 – 1.2)	1
CD8 ⁺ Naïve	18U	2.2	5.3	0.8 (0.4 – 1.5)	1
CD4 ⁺ TCM CD38 ⁺	19U	0.8	0.7	1.6 (1.0 – 2.7)	1
C) Stimulated-specific subsets					
CD8 ⁺ TEM CD38 ⁺ T-bet ⁺	20S	1.0	0.4	6.3 (2.3 – 17.0)	1.8 x 10⁻²
CD8 ⁺ TEM CCR6 ⁺ T-bet ⁺	21S	0.9	0.4	3.8 (1.8 – 7.9)	3.8 x 10⁻²
CD4 ⁺ TCM TNFα ⁺	22S	2.2	0.8	4.0 (1.9 – 8.5)	4.1 x 10⁻²
CD4 ⁺ TCM CD38 ⁺ TNFα ⁺	23S	3.9	9.0	0.3 (0.1 – 0.8)	4.1 x 10 ⁻¹
CD4 ⁺ TCM T-bet ⁺ Perforin ⁺	24S	0.2	0.6	0.3 (0.1 – 0.8)	4.1 x 10 ⁻¹
CD8 ⁺ TCM	25S	11.7	10.2	1.4 (0.9 – 2.0)	1
CD4 ⁺ TEM	26S	9.1	7.6	1.1 (0.5 – 2.4)	1
CD8 ⁺ TEM IFN-γ ⁺ TNFα ⁺	27S	2.3	2.0	0.9 (0.4 – 2.3)	1
CD4 ⁺ TCM CCR6 ⁺	28S	0.5	0.8	0.7 (0.4 – 1.5)	1

For each shared cell subset, the first row indicates statistics for the unstimulated subset and the second for the stimulated subset. Elements separated by “/” indicate statistics for two clusters with the same definition found in the same dataset. Here, ‘a’ and ‘b’ refer to the larger and smaller cluster, respectively. The odds ratio is shown with a 95% confidence interval. Subsets with differential abundance between RA and healthy samples are highlighted by a bold P value (α = 0.05). T_{CM}: T central memory. T_{EM}: T effector memory.

caused by multiple testing. However, this is very unlikely, as first we confirm a plausible and previously reported imbalance of Th/Treg-like subsets in RA; second the mechanistic evidence supports our findings, and finally all statistical tests have been stringently corrected for multiple testing. These elements reinforce the credibility of our main conclusion: the superiority of hypothesis-free clustering of many markers in small sample sizes over standard hypothesis-driven bi-axial gating.

The hallmark of precision medicine and the primary aim of stratified medicine initiatives in RA (45) consists of stratifying patients into treatment response categories. Our study shows large heterogeneity across patients with some with a predominantly Th1 profile, others a Th17 profile, or a Th1-like Th17 profile, while others are mainly characterised by a low IL-2⁺

Treg-like profile. As an example, stratifying patients by their IL-2⁺ Treg titres may dictate response to low-dose IL-2 therapy, as has now been successfully trialled in lupus (46) and RA (47). Furthermore, a Japanese group found that stratified medicine for psoriatic arthritis patients increased response rates significantly (16). Immunophenotyped patients received ustekinumab if Th1-dominant and secukinumab if either Th17-dominant or with a Th1/Th17-hybrid phenotype resulting in clinically significant improvement in response rates (16). The same group have good one year follow-up data (48), suggesting that the immunophenotype could be the main predictor of drug response in precision medicine. Stratifying RA patients by their Th1/Th17 phenotype, or by the Th1-like Th17 phenotype highlighted in this study, may suggest that, despite the fact that previous trials of

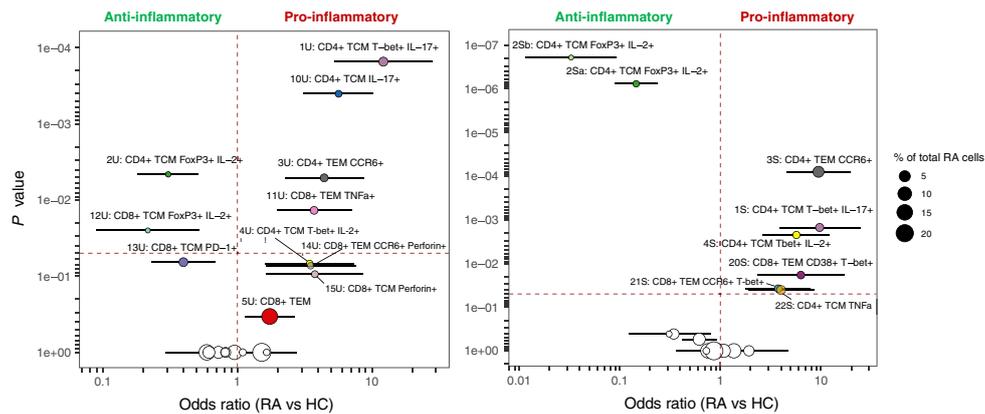


FIGURE 6 MASC results for unstimulated (left) and stimulated (right) clusters identified by automated clustering of merged CD4⁺ and CD8⁺ cells. Odds ratio (with 95% CI) is shown against P value, and the size of the point represents the % of total RA cells from the dataset that comprise each cluster. The vertical dashed line represents an odds ratio of 1, and the horizontal dashed line represents the threshold for statistical significance ($\alpha = 0.05$). Cell subsets identified in both states (unstimulated and stimulated) are denoted by the same colour in both plots.

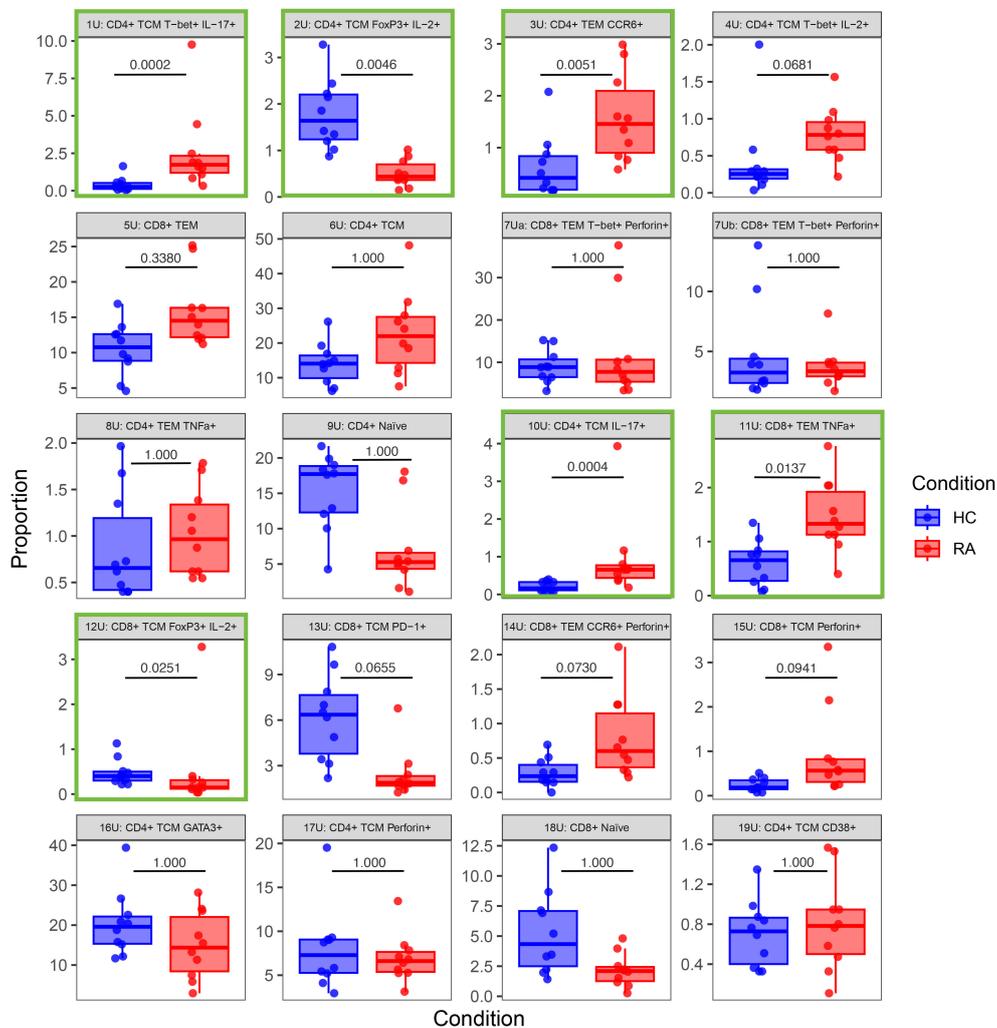


FIGURE 7 Box-and-whisker plots of the proportions (%) of 20 T cell populations identified in unstimulated RA and healthy control (HC) samples by automated clustering of merged CD4⁺ and CD8⁺ cells. The median, first and third quartiles and minimum and maximum values are shown. The P value obtained from MASC is shown for each cluster, and a green outline indicates differential abundance between RA and HC (i.e. $P \leq 0.05$). $n = 10$ HC, 10 RA.

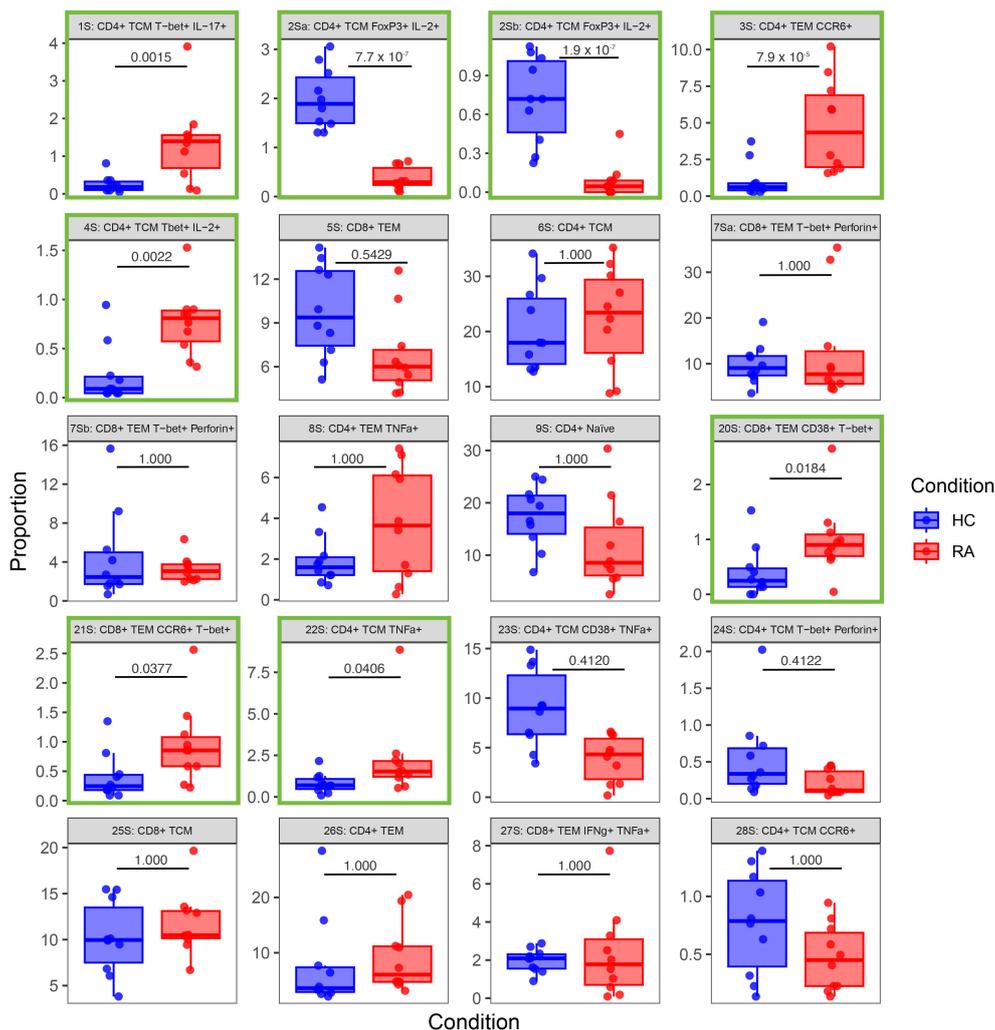


FIGURE 8 Box-and-whisker plots of the proportions (%) of 20 T cell populations identified in stimulated RA and healthy control (HC) samples by automated clustering of merged CD4+ and CD8+ cells. The median, first and third quartiles and minimum and maximum values are shown. The P value obtained from MASC is shown for each cluster, and a green outline indicates differential abundance between RA and HC (i.e. P < 0.05). n = 10 HC, 10 RA.

secukinumab in RA failed to reach their primary endpoint, it could be a viable treatment option in RA patient subgroups; further work is required to confirm this hypothesis (49).

We have shown that non-biased automated analysis of large immune datasets is successful in identifying Th cell imbalance in RA with implications for precision medicine. We now plan to expand on these findings by testing their role in patient stratification for treatment response studies.

5 Key messages

What is already known about this subject?

- Large cytometric datasets are difficult to analyse manually due to the multidimensional nature of the data.

- Manual data analysis may introduce significant bias and be underpowered to interrogate multidimensional data.

What does this study add?

- This study has shown that an unbiased automated clustering algorithm can successfully interrogate large cytometric datasets, finding differential expression of 2 rare T cell populations in RA patients (decreased IL-2+ Treg-like cells; increased Th1-like Th17 cells).

How might this impact on clinical practice?

- Understanding the underlying immunopathology of autoimmune disease is a prerequisite to finding new drug targets for treatment.

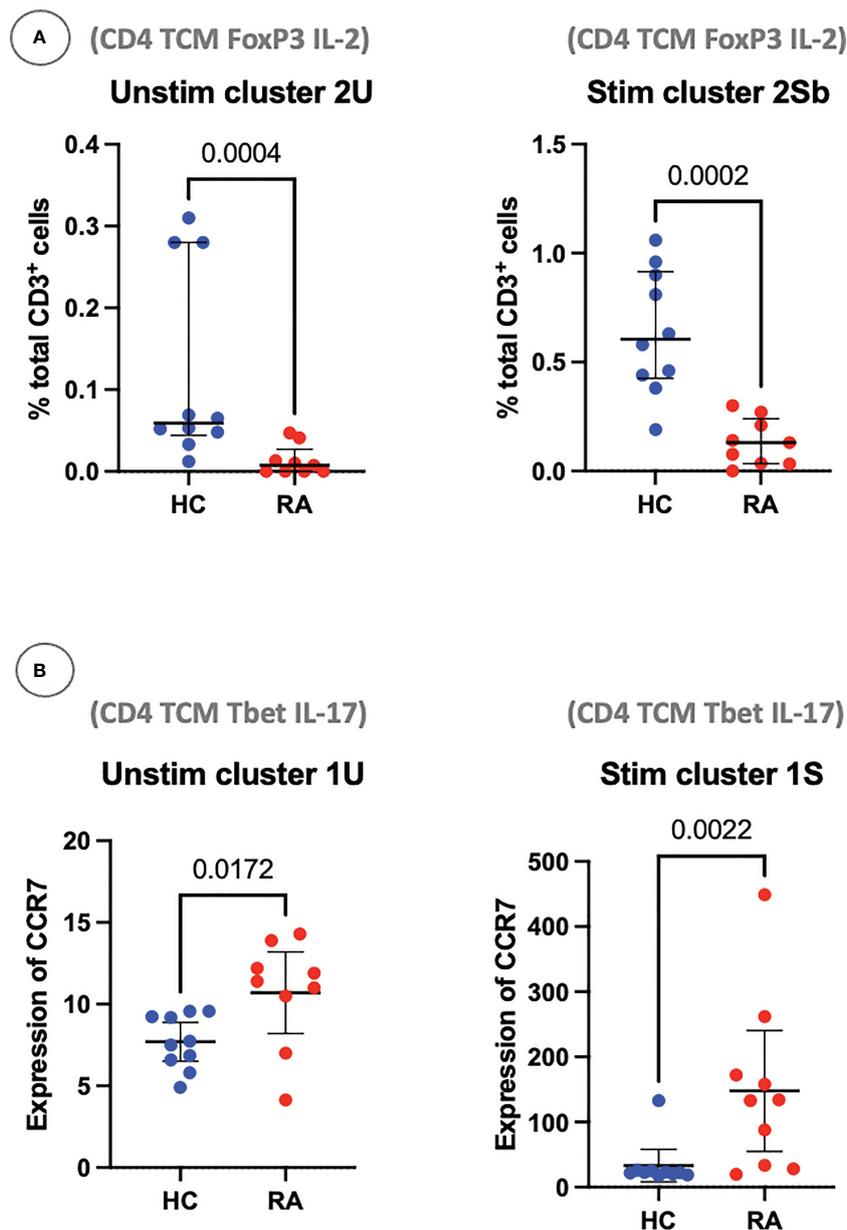


FIGURE 9
(A) Clusters 2U and 2Sb represent the abundance of manually gated subsets previously identified by the CyTOF automated clustering pipeline, shown here with median and interquartile range. Parent population is labelled on the y-axis. **(B)** Clusters 1U and 1S represent the expression of CCR7 on Th1-like IL-17⁺ cells, with median and interquartile range. Points represent RA and HC samples. Mann Whitney *U* *P* values are shown above the pairwise comparisons. TEM, T effector memory cell; TCM, T central memory cell.

- As our knowledge of T cell heterogeneity expands, discovering rare T cell populations will directly feed into the concept of precision medicine, to: ‘treat arthritis, right first time’.

number(s) can be found below: <https://flowrepository.org/>, FR-FCM-Z5RM.

Ethics statement

The studies involving human participants were reviewed and approved by North West Ethics committee approval MREC 99/8/84 Manchester Royal Infirmary (NIHR portfolio ID 7881). The patients/participants provided their written informed consent to participate in this study.

Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession

Author contributions

SV conceived the project, SV, SR and AB guided it through its stages of development, ensured progress, and contributed to the final manuscript. BM ran the laboratory experiments with help from SH. MS pre-processed the raw mass cytometry data. BM and LM ran the data analysis (BM: FlowJo; LM: R pipeline) and drafted the manuscript. CF supported the automated pipeline analysis and provided R code for MASC. TH and SR provided immunology expertise. All authors contributed to the article and approved the submitted version.

Funding

This research was supported by Versus Arthritis (grant 21754), the NIHR Manchester Biomedical Research Centre, and the Wellcome Trust Institutional Strategic Support Fund (WT ISSF) in Precision Medicine and Single Cell Research. AB is an NIHR Senior Investigator. BM is an NIHR-funded Academic Clinical Fellow.

Acknowledgments

We would like to thank Sarah Ashton (senior study coordinator) for her invaluable support, all the staff at the University of Manchester Versus Arthritis Centre for Genetics and Genomics and at the Kellgren Centre for Rheumatology, Central Manchester University NHS Foundation Trust, UK. Most

References

1. Hyrich KL, Watson KD, Silman AJ, Symmons DPM. Predictors of response to anti-TNF- α therapy among patients with rheumatoid arthritis: results from the British society for rheumatology biologics register. *Rheumatology* (2006) 45(12):1558–65. doi: 10.1093/rheumatology/ke1149
2. Mosmann TR, Cherwinski H, Bond MW, Giedlin MA, Coffman RL. Two types of murine helper T cell clone. i. definition according to profiles of lymphokine activities and secreted proteins. *J Immunol* (1986) 136(7):2348–57. doi: 10.4049/jimmunol.136.7.2348
3. Schulze-Koops H, Kalden JR. The balance of Th1/Th2 cytokines in rheumatoid arthritis. *Best Pract Res Clin Rheumatol* (2001) 15:677–91. doi: 10.1053/berh.2001.0187
4. Afzali B, Lombardi G, Lechler RI, Lord GM. The role of T helper 17 (Th17) and regulatory T cells (Treg) in human organ transplantation and autoimmune disease. *Clin Exp Immunol* (2007) 148:32–46. doi: 10.1111/j.1365-2249.2007.03356.x
5. Niu Q, Cai B, Huang Z, Shi Y, Wang L. Disturbed Th17/Treg balance in patients with rheumatoid arthritis. *Rheumatol Int* (2012) 32:2731–6. doi: 10.1007/s00296-011-1984-x
6. van Hamburg JP, Asmawidjaja PS, Davelaar N, Mus AMC, Colin EM, Hazes JMW, et al. Th17 cells, but not Th1 cells, from patients with early rheumatoid arthritis are potent inducers of matrix metalloproteinases and proinflammatory cytokines upon synovial fibroblast interaction, including autocrine interleukin-17A production. *Arthritis Rheum* (2011) 63:73–83. doi: 10.1002/art.30093
7. Fonseka CY, Rao DA, Teslovich NC, Korsunsky I, Hannes SK, Slowikowski K, et al. Mixed-effects association of single cells identifies an expanded effector CD4+ T cell subset in rheumatoid arthritis. *Sci Transl Med* (2018) 10(463):eaq0305. doi: 10.1126/scitranslmed.aq0305
8. Morita T, Shima Y, Wing JB, Sakaguchi S, Ogata A, Kumanogoh A. The proportion of regulatory T cells in patients with rheumatoid arthritis: a meta-analysis. *PLoS One* (2016) 11:e0162306. doi: 10.1371/journal.pone.0162306

importantly, we are grateful and thankful for all the patients and volunteers who took part in this study.

Conflict of interest

Authors CF and SH are currently employed by the companies eGenesis and BioNTech, respectively. They were not employed by these companies during the conception of the study or the generation of the experimental data.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fimmu.2023.1094872/full#supplementary-material>

9. Walter GJ, Fleskens V, Frederiksen KS, Rajasekhar M, Menon B, Gerwien JG, et al. Phenotypic, functional, and gene expression profiling of peripheral CD45RA+ and CD45RO+ CD4+CD25+CD127^{low} Treg cells in patients with chronic rheumatoid arthritis. *Arthritis Rheumatol* (2016) 68:103–16. doi: 10.1002/art.39408
10. Ciccia F, Guggino G, Rizzo A, Manzo A, Vitolo B, Manna MPL, et al. Potential involvement of IL-9 and Th9 cells in the pathogenesis of rheumatoid arthritis. *Rheumatol (United Kingdom)* (2015) 54(12):2264–72. doi: 10.1093/rheumatology/kev252
11. Miyazaki Y, Nakayama S, Kubo S, Nakano K, Iwata S, Miyagawa I, et al. Th22 cells promote osteoclast differentiation via production of IL-22 in rheumatoid arthritis. *Front Immunol* (2018) 9:2901. doi: 10.3389/fimmu.2018.02901
12. Deng J, Fan C, Gao X, Zeng Q, Guo R, Wei Y, et al. Signal transducer and activator of transcription 3 hyperactivation associates with follicular helper T cell. *Front Immunol* (2018) 9:1226. doi: 10.3389/fimmu.2018.01226
13. Rao DA, Gurish MF, Marshall JL, Slowikowski K, Fonseka CY, Liu Y, et al. Pathologically expanded peripheral T helper cell subset drives B cells in rheumatoid arthritis. *Nature* (2017) 542:110–4. doi: 10.1038/nature20810
14. Mirlekar B. Co-Expression of master transcription factors determines CD4+ T cell plasticity and functions in auto-inflammatory diseases. *Immunol Lett* (2020) 222:58–66. doi: 10.1016/j.imlet.2020.03.007
15. Genovese MC, Durez P, Richards HB, Supronik J, Dokoupilova E, Mazurov V, et al. Efficacy and safety of secukinumab in patients with rheumatoid arthritis: a phase II, dose-finding, double-blind, randomised, placebo controlled study. *Ann Rheum Dis* (2013) 72:863–9. doi: 10.1136/annrheumdis-2012-201601
16. Miyagawa I, Nakayama S, Nakano K, Kubo S, Iwata S, Miyazaki Y, et al. Precision medicine using different biological DMARDs based on characteristic phenotypes of peripheral T helper cells in psoriatic arthritis. *Rheumatol (United Kingdom)* (2019) 58:336–44. doi: 10.1093/rheumatology/key069

17. Ishigaki K, Shoda H, Kochi Y, Yasui T, Kadono Y, Tanaka S, et al. Quantitative and qualitative characterization of expanded CD4 + T cell clones in rheumatoid arthritis patients. *Sci Rep* (2015) 5:12937. doi: 10.1038/srep12937
18. Bendall SC, Simonds EF, Qiu P, Amir E -a. D, Krutzik PO, Finck R, et al. Single-cell mass cytometry of differential immune and drug responses across a human hematopoietic continuum. *Sci* (1979) (2011) 332:687–96. doi: 10.1126/science.1198704
19. Van Gassen S, Callebaut B, Van Helden MJ, Lambrecht BN, Demeester P, Dhaene T, et al. FlowSOM: using self-organizing maps for visualization and interpretation of cytometry data. *Cytometry Part A* (2015) 87(7):636–45. doi: 10.1002/cyto.a.22625
20. Robinson MD, Nowicka M, Krieg C, Weber LM, Hartmann FJ, Guglietta S, et al. CyTOF workflow: differential discovery in high-throughput high-dimensional cytometry datasets. *F1000Res* (2017) 6:748. doi: 10.12688/f1000research.11622.2
21. Qiu P, Simonds EF, Bendall SC, Gibbs KD, Bruggner RV, Linderman MD, et al. Extracting a cellular hierarchy from high-dimensional cytometry data with SPADE. *Nat Biotechnol* (2011) 29:886–91. doi: 10.1038/nbt.1991
22. Bruggner RV, Bodenmiller B, Dill DL, Tibshirani RJ, Nolan GP. Automated identification of stratifying signatures in cellular subpopulations. *Proc Natl Acad Sci U.S.A.* (2014) 111:E2770–7. doi: 10.1073/pnas.1408792111
23. Wilkerson MD, Hayes DN. ConsensusClusterPlus: a class discovery tool with confidence assessments and item tracking. *Bioinformatics* (2010) 26(12):1572–3. doi: 10.1093/bioinformatics/btq170
24. Zhang F, Wei K, Slowikowski K, Fonseka CY, Rao DA, Kelly S, et al. Defining inflammatory cell states in rheumatoid arthritis joint synovial tissues by integrating single-cell transcriptomics and mass cytometry. *Nat Immunol* (2019) 20:928–42. doi: 10.1038/s41590-019-0378-1
25. Stahl EA, Raychaudhuri S, Remmers EF, Xie G, Eyre S, Thomson BP, et al. Genome-wide association study meta-analysis identifies seven new rheumatoid arthritis risk loci. *Nat Genet* (2010) 42:508–14. doi: 10.1038/ng.582
26. Korsunsky I, Millard N, Fan J, Slowikowski K, Zhang F, Wei K, et al. Fast, sensitive and accurate integration of single-cell data with harmony. *Nat Methods* (2019) 16:1289–96. doi: 10.1038/s41592-019-0619-0
27. Kang JB, Nathan A, Weinand K, Zhang F, Millard N, Rumker L, et al. Efficient and precise single-cell reference atlas mapping with symphony. *Nat Commun* (2021) 12(1):5890. doi: 10.1038/s41467-021-25957-x
28. Finck R, Simonds EF, Jager A, Krishnaswamy S, Sachs K, Fantl W, et al. Normalization of mass cytometry data with bead standards. *Cytometry A* (2013) 83:A:483–494. doi: 10.1002/cyto.a.22271
29. Maecker HT, McCoy JP, Nussenblatt R. Standardizing immunophenotyping for the human immunology project. *Nat Rev Immunol* (2012) 12:191–200. doi: 10.1038/nri3158
30. Wickham H. Positioning. In: *ggplot2*. Use R!. Springer, Cham. (2016). doi: 10.1007/978-3-319-24277-4_7
31. Kolde R. Package “pheatmap”: pretty heatmaps. *Version 1012* (2019).
32. Van Der Maaten LJP, Hinton GE. Visualizing high-dimensional data using t-sne. *J Mach Learn Res* (2008) 9(nov):2579–605.
33. Spidlen J, Breuer K, Brinkman R. Preparing a minimum information about a flow cytometry experiment (MIFlowCyt) compliant manuscript using the international society for advancement of cytometry (ISAC) FCS file repository (FlowRepository.org). *Curr Protoc Cytom* (2012) 61(1):10.18.1–10.18.26. doi: 10.1002/0471142956.CY1018S61
34. Alpert A, Pickman Y, Leipold M, Rosenberg-Hasson Y, Ji X, Gaujoux R, et al. A clinically meaningful metric of immune age derived from high-dimensional longitudinal monitoring. *Nat Med* (2019) 25(3):487–95. doi: 10.1038/s41591-019-0381-y
35. Kamali AN, Noorbakhsh SM, Hamedifar H, Jadidi-Niaragh F, Yazdani R, Bautista JM, et al. A role for Th1-like Th17 cells in the pathogenesis of inflammatory and autoimmune disorders. *Mol Immunol* (2019) 105:107–15. doi: 10.1016/J.MOLIMM.2018.11.015
36. Cohen CJ, Crome SQ, MacDonald KG, Dai EL, Mager DL, Levings MK. Human Th1 and Th17 cells exhibit epigenetic stability at signature cytokine and transcription factor loci. *J Immunol* (2011) 187:5615–26. doi: 10.4049/JIMMUNOL.1101058
37. Schmolka N, Papotto PH, Romero PV, Amado T, Enguita FJ, Amorim A, et al. MicroRNA-146a controls functional plasticity in T cells by targeting NOD1. *Sci Immunol* (2018) 3:1392. doi: 10.1126/SCIIMMUNOL.AAO1392/SUPPL_FILE/AAO1392_TABLE_S2.ZIP
38. Zhang LL, Wu XX, Wang XF, Di DS, Huang Q, Liu RS, et al. Genetic variant in microRNA-146a gene is associated with risk of rheumatoid arthritis. *Ann Med* (2021) 53:824–9. doi: 10.1080/07853890.2021.1933163
39. Löfgren SE, Frostegård J, Truedsson L, Pons-Estel BA, D’Alfonso S, Witte T, et al. Genetic association of miRNA-146a with systemic lupus erythematosus in europeans through decreased expression of the gene. *Genes Immun* (2012) 13:268–74. doi: 10.1038/GENE.2011.84
40. de la Rosa M, Rutz S, Dorninger H, Scheffold A. Interleukin-2 is essential for CD4+CD25+ regulatory T cell function. *Eur J Immunol* (2004) 34:2480–8. doi: 10.1002/EJL.200425274
41. Chawla AS, Khalsa JK, Dhar A, Gupta S, Umar D, Arimbasseri GA, et al. A role for cell-autocrine interleukin-2 in regulatory T-cell homeostasis. *Immunology* (2020) 160:295–309. doi: 10.1111/IMM.13194
42. Furtado GC, De Curotto Lafaille MA, Kutchukhidze N, Lafaille JJ. Interleukin 2 signaling is required for CD4(+) regulatory T cell function. *J Exp Med* (2002) 196:851–7. doi: 10.1084/JEM.20020190
43. Sakaguchi S, Miyara M, Costantino CM, Hafler DA. FOXP3+ regulatory T cells in the human immune system. *Nat Rev Immunol* (2010) 10:490–500. doi: 10.1038/nri2785
44. Bendfeldt H, Scheel MB, Steinbrink K, Radbruch A, Herzel H, Baumgrass R. IL-2 expression in activated human memory FOXP3+ cells critically depends on the cellular levels of FOXP3 as well as of four transcription factors of T cell activation. *Front Immunol* (2012) 3:264/BIBTEX. doi: 10.3389/FIMMU.2012.00264/BIBTEX
45. Barton A, Pitzalis C. Stratified medicine in rheumatoid arthritis-the MATURA programme. *Rheumatol (Oxford)* (2017) 56:1247–50. doi: 10.1093/rheumatology/kew369
46. Humrick JY, Cacoub P, Rosenzweig M, Pitoiset F, PHAM HP, Guidoux J, Leroux D, et al. Low-dose interleukin-2 therapy in active systemic lupus erythematosus (LUPIL-2): a multicentre, double-blind, randomised and placebo-controlled phase II trial. *Ann Rheum Dis* (2022) 81(12):1685–94. doi: 10.1136/ARD-2022-222501
47. Zhang X, Miao M, Zhang R, Liu X, Zhao X, Shao M, et al. Efficacy and safety of low-dose interleukin-2 in combination with methotrexate in patients with active rheumatoid arthritis: a randomized, double-blind, placebo-controlled phase 2 trial. *Signal Transduct Target Ther* (2022) 7(1):141–9. doi: 10.1038/S41392-022-00887-2
48. Miyagawa I, Nakayamada S, Ueno M, Miyazaki Y, Ohkubo N, Inoue Y, et al. Precision medicine based on the phenotypic differences in peripheral T helper cells in patients with psoriatic arthritis: one year follow-up outcomes. *Front Med (Lausanne)* (2022) 9:934937/BIBTEX. doi: 10.3389/FMED.2022.934937/BIBTEX
49. Blanco FJ, Mörck R, Dokoupilova E, Codding C, Neal J, Andersson M, et al. Secukinumab in active rheumatoid arthritis: a phase III randomized, double-blind, active comparator- and placebo-controlled study. *Arthritis Rheumatol* (2017) 69:1144–53. doi: 10.1002/art.40070