ORIGINAL ARTICLE

immunology
Insights into immunity in health and disease

A unique human cord blood CD8$^+$CD45RA$^+$CD27$^+$CD161$^+$ T-cell subset identified by flow cytometric data analysis using Seurat

Julen Gabirel Araneta Reyes[1,2,3] | Duan Ni[1,2,3] | Brigitte Santner-Nanan[1,2,3] | Gabriela Veronica Pinget[1,3] | Lucie Kraftova[1,3,4,5] | Thomas Myles Ashhurst[6] | Felix Marsh-Wakefield[7,8] | Claire Leana Wishart[1,9,10] | Jian Tan[1,11] | Peter Hsu[12,13] | Nicholas Jonathan Cole King[1,6,9,10,14,15] | Laurence Macia[1,6,11] | Ralph Nanan[1,2,3]

[1]Charles Perkins Centre, The University of Sydney, Sydney, New South Wales, Australia

[2]Nepean Hospital, Nepean Blue Mountains Local Health District, Penrith, New South Wales, Australia

[3]Nepean Clinical School, The University of Sydney, Sydney, New South Wales, Australia

[4]Department of Microbiology, Faculty of Medicine, University Hospital in Pilsen, Charles University, Pilsen, Czech Republic

[5]Biomedical Center, Faculty of Medicine, Charles University, Pilsen, Czech Republic

[6]Sydney Cytometry Core Research Facility, Charles Perkins Centre, The University of Sydney and Centenary Institute, Sydney, New South Wales, Australia

[7]Liver Injury and Cancer Program, Centenary Institute, Sydney, New South Wales, Australia

[8]Human Cancer and Viral Immunology Laboratory, The University of Sydney, Sydney, New South Wales, Australia

[9]Viral immunopathology Laboratory, Infection, Immunity and Inflammation Research Theme, School of Medical Sciences, Faculty of Medicine and Health, The University of Sydney, Sydney, New South Wales, Australia

[10]Ramaciotti Facility for Human System Biology, The University of Sydney and Centenary Institute, Sydney, New South Wales, Australia

[11]School of Medical Sciences, Faculty of Medicine and Health, The University of Sydney, Sydney, New South Wales, Australia

[12]Kids Research, The Children's Hospital at Westmead, Sydney, New South Wales, Australia

[13]Discipline of Child and Adolescent Health, The University of Sydney, Sydney, New South Wales, Australia

[14]The University of Sydney Institute for Infectious Diseases, The University of Sydney, Sydney, New South Wales, Australia

[15]Sydney Nano, The University of Sydney, Sydney, New South Wales, Australia

Correspondence
Ralph Nanan, Nepean Clinical School, The University of Sydney, Sydney, NSW, Australia.
Email: ralph.nanan@sydney.edu.au

Abstract

Advances in single-cell level analytical techniques, especially cytometric approaches, have led to profound innovation in biomedical research, particularly in the field of clinical immunology. This has resulted in an expansion of high-dimensional data, posing great challenges for comprehensive and unbiased analysis. Conventional manual analysis is thus becoming untenable to handle these challenges. Furthermore, most newly developed computational methods lack flexibility and interoperability, hampering their accessibility and usability. Here, we

Julen Gabirel Araneta Reyes and Duan Ni should be considered joint first authors.

adapted Seurat, an R package originally developed for single-cell RNA sequencing (scRNA-seq) analysis, for high-dimensional flow cytometric data analysis. Based on a 20-marker antibody panel and analyses of T-cell profiles in both adult blood and cord blood (CB), we showcased the robust capacity of Seurat in flow cytometric data analysis, which was further validated by Spectre, another high-dimensional cytometric data analysis package, and conventional manual analysis. Importantly, we identified a unique $CD8^+$ T-cell population defined as $CD8^+CD45RA^+CD27^+CD161^+$ T cell that was predominantly present in CB. We characterised its IFN-γ-producing and potential cytotoxic properties using flow cytometry experiments and scRNA-seq analysis from a published dataset. Collectively, we identified a unique human CB $CD8^+CD45RA^+CD27^+CD161^+$ T-cell subset and demonstrated that Seurat, a widely used package for scRNA-seq analysis, possesses great potential to be repurposed for cytometric data analysis. This facilitates an unbiased and thorough interpretation of complicated high-dimensional data using a single analytical pipeline and opens a novel avenue for data-driven investigation in clinical immunology.

INTRODUCTION

The rapid development of analytical technologies at a single-cell level over recent decades has revolutionised biological and medical research, particularly in the field of immunology. Immune cell populations are well-known for their heterogeneity and tools such as flow cytometry (or fluorescence-activated cell sorting; FACS), cytometry by time-of-flight mass spectrometry (CyTOF) and single-cell RNA sequencing (scRNA-seq), facilitate an in-depth identification and characterisation of various immune cell types [1]. Conventionally, analysis of cytometric data (including flow, spectral and mass cytometry) has relied on manual analysis based on empirical gating strategies under expert supervision. This is extremely labour-intensive and tedious, as the complex cytometric data is limited to permutational visualisation of two-dimensional (2D) plots (FACS plots). These plots feature different pairs of marker combinations, which require arduous sequential inspection [2]. The possible combinations of markers from a given panel increase exponentially with the addition of extra parameters. As more and more state-of-the-art cytometric panels exceed 20 markers [3, 4], thorough manual gating analysis is becoming increasingly challenging and impractical [5]. Furthermore, such analytical workflows are inevitably subject to bias, considering their dependence on empirical knowledge and subjective selection and inspection of markers.

These limitations hamper analyses and potentially conceal novel findings.

Various computational approaches have been developed as potential solutions, including methods for dimension reduction (such as *t*-distributed stochastic neighbour embedding [6] and uniform manifold approximation and projection (UMAP) [7]), clustering (such as PhenoGraph and FlowSOM [8, 9]) and automated cell gating and classification [5, 10]. These tools all accelerate high-dimensional data analysis. Moreover, they have revolutionised cytometry-based research, transitioning from the conventional hypothesis-driven strategy that focuses on specific cell types or markers, to more unbiased and comprehensive methods, that simultaneously take all data into account. Despite their notable success, these tools still suffer significant limitations. For example, many of these computational modalities are separate, and some even require specific data formatting and processing procedures. This is not user-friendly and hampers their usability and accessibility in the broader research community. While some integrative toolkits combine these modalities and offer end-to-end analysis of cross-platform cytometric data, including normalisation, integration and clustering, such as Spectre [10] and ImmPort Galaxy [11], these are still few in number. There are also some other commercial toolkits of this kind, like OMIQ [12, 13] and Cytobank [14], but they usually require paid subscriptions, limiting their availability. Furthermore, owing to

their non-open-source nature, commercial toolkits can lag in flexible customising services, as well as community-driven support, maintenance and improvements, potentially preventing optimal usability and adaptability. Hence, there is great interest in more accessible and adaptable tools.

Computational analyses of complex cytometric data have considerably benefited clinical immunological research. On one hand, clinical immunological data is notorious for its heterogeneity, highlighting the need for computational tools for data cleaning, batch alignment, and unbiased analysis. On the other hand, given the limited availability of clinical samples (such as samples with rare disease backgrounds or longitudinal samples), expanding the markers and dimensionality of cytometry panels may help to achieve more comprehensive and efficient investigation, and represents an unprecedented opportunity for high-dimensional data analysis. Leveraging the rapidly developing computational analytical tools for clinical immunological studies is thus emerging as a promising avenue to provide more detailed insights into clinical contexts whilst maximising the values of limited clinical samples.

An area of increasing interest is the characterisation of immune profiles in umbilical cord blood (CB) compared to adult blood (AB). The striking immunological differences between CB and AB not only offer critical insights for disease pathogenesis but also provide an ideal scenario for showcasing the analytical power of computational toolkits in clinical applications.

The main components of the CB immune compartment are cord blood mononuclear cells (CBMCs), known to exhibit unique characteristics relative to peripheral blood mononuclear cells (PBMCs) from AB, due to the semi-allogeneic environment of pregnancy. Mirroring the foetal immune system [15], CBMCs feature a more naïve phenotype [16–18] and are implicated in the physiology and pathology during both pregnancy and later in life [19–22]. Hence, understanding CB immune profiles and their differences from AB provides precious insights into immunological development at different stages, as well as sheds light on the complex immunobiology of pregnancy.

Here, we developed a 20-marker antibody panel for thoroughly immunophenotyping T cells in both CB and AB. We adapted Seurat, a widely used end-to-end package originally for scRNA-seq analysis, for the resulting high-dimensional flow cytometric data analysis after primary processing in FlowJo. This workflow identified several previously underappreciated T-cell subsets in AB, validated by Spectre, another computational cytometric analytical package, and conventional manual gating, showcasing the capacity of Seurat. Importantly, using

Seurat for a comparative study of CB and AB profiles, we revealed a unique CB T-cell population, characterised as $CD8^+CD45RA^+CD27^+CD161^+$ T cells. Analysis of previously published scRNA-seq data confirmed this identified population and hinted at its possible cytotoxic and pro-inflammatory properties. Together, this represents the first application example of using Seurat as a complete flow cytometric analysis workflow and demonstrates its robust analytical performance. It emerges as a simple and easy-to-use toolkit for cytometric data analysis, particularly for its pre-existing wide scRNA-seq user community. Seurat also features as a single platform but with various supplementary tools and plugins facilitating single-cell analysis of both protein and RNA data, as well as their comparisons and cross-validation. This represents a novel unbiased discovery tool for complex single-cell data analysis in clinical immunology.

MATERIALS AND METHODS

Sample preparation and flow cytometry experiments

PBMCs from AB samples and CBMCs from CB samples (Table S1) were prepared as described in Supporting Information S1.1. For flow cytometry experiments, cells were processed, barcoded, and stained as Supporting Information S1.2 and S1.3. and analysed using the 5-laser Aurora Spectral cytometer (Cytek Biosciences, USA) on the same day.

Cytometric data was unmixed with SpectroFlo and analysed using FlowJo (BD Life Science) based on the gating strategy in Figure S1A. The $CD3^+CD4^+$ and $CD3^+CD8^+$ T-cell populations were manually gated and exported as CSV files with their scaled values for further computational analysis in RStudio using Seurat and Spectre.

High-dimensional flow cytometry data analysis

Data preparation

For analysis using Seurat 4.3.0 and Spectre 1.0.0, the exported CSV files of the gated populations of interest were loaded in RStudio (v4.1.2). Next, data tables for each sample were merged together using Spectre's *do.merge. files()* function. The data then underwent a hyperbolic arcsine (arcsinh) transformation (co-factor = 2000) using Spectre (*do.asinh*). To avoid bias and skewing due to different cellularity, the processed cytometric data was first

randomly downsampled to the same number among samples before analysis.

Analysis using Seurat

After reading in the downsampled data matrix generated in Data preparation, along the conventional Seurat analysis workflow [23–26], procedures such as QC and normalisation originally for transcriptomic data were skipped. Furthermore, to preserve the transformed flow cytometry data structure for analysis, the data scaling process was bypassed by selecting "do.scale = FALSE, do.center = FALSE" in the *ScaleData()* function. All features were selected as variable features for further analysis.

After that, principal component analysis (PCA) was performed. Based on the PCA scores, the top PCs contributing to 99% of variance were selected for the subsequent cluster analysis using the *FindNeighbors() and FindClusters()* functions and dimensional reduction with the *RunUMAP()* function. We utilised the *DotPlot()* and *VlnPlot()* functions in Seurat to visualise the expression of various markers in different clusters.

A detailed workflow for analysing flow cytometry data using Seurat is available in the Supporting Information S1.4.2.

Analysis using Spectre

For analysis with Spectre, input data were first downsampled and then analysed following [10], except a 5×5 self-organising map (SOM) was used. The final clustering numbers were also adjusted to the same as Seurat's results for comparison.

scRNA-seq analysis

Preprocessed scRNA-seq data was downloaded from Gene Expression Omnibus from a previous study [27] (GEO accession: GSE158493). Since the original dataset was grouped based on sample origins (foetal spleen, full-term umbilical CB and adult peripheral blood), the LIGER package was used to re-integrate the data with different origins to eliminate potential batch effects [28]. After that, the data were analysed with Seurat 4.3.0. [23–26]. Marker genes were identified using Seurat's *FindMarkers()* function based on its default settings, with thresholds for differentially expressed genes defined as $p < 0.05$ and log2(fold change) higher than 0.5 or lower than -0.5. Gene ontology (GO) analysis and gene set enrichment analysis (GSEA) was run with package fgsea and ClusterProfiler following their tutorials [29, 30].

Statistics

Statistical analysis was performed with GraphPad PRISM 10 or in RStudio using Seurat. Mann–Whitney U test was used for comparison between the two groups. Differences were considered to be statistically significant when $p < 0.05$.

RESULTS

Seurat is a reliable tool for high-dimensional flow cytometry data analysis

Analysis of human peripheral blood, and the heterogenous T-cell compartments therein, is of great importance considering the availability of blood samples and its ubiquitous use in medical research and clinical diagnosis. Here, we established a 20-marker antibody panel (Table 1) to comprehensively characterise the functional profiles of CD4$^+$ and CD8$^+$ T cells in human PBMCs. Details about the markers and their functions are described in Table 1. They cover various aspects of T-cell biology, including maturation, activation, migration and function.

PBMCs from healthy adult individuals were collected and processed based on the workflow in Figure 1a. To reduce antibody wastage as well as minimise the inter-sample variations caused by batch effects introduced during experimental processes, which is an intrinsic problem for clinical studies involving flow cytometry, we utilised a barcoding system leveraging the anti-CD45 labelling. In the current study, PBMC samples were stained with our 20-colour antibody panel and analysed by spectral cytometry. After demultiplexing, the resulting cytometric data were manually gated for CD4$^+$ and CD8$^+$ T-cell populations (Figure S1A).

For proof-of-concept, the CD8$^+$ T-cell data was analysed with Seurat and compared with Spectre, based on the 12 functional markers out of the total 15 surface markers, excluding lineage markers such as CD45, CD3 and CD8. The results were validated through manual gating as well.

Seurat clustered the adult CD8$^+$ T-cell compartment into 14 clusters (Se1-14) and projected them onto the 2D UMAP plot, as shown in Figure 1b. Historically, T cells are gated into four populations based on expression of CD45RA and CD27. The CD45RA$^+$CD27$^+$ population is defined as naïve, the CD45RA$^-$CD27$^+$ population as central memory (CM), the CD45RA$^-$CD27$^-$ population as effector memory (EM) and the CD45RA$^+$CD27$^-$ is effector memory that re-express CD45RA (EMRA) [31]. Comparing Seurat's results to this classification, we found

T A B L E 1 A 20-marker antibody panel for comprehensive profiling of T cells in human peripheral blood mononuclear cells (PBMCs).

Marker target	Staining	Markers/functional implications	References
CD45	Surface	Leukocytes (excludes erythrocytes and platelets)	
CD3	Surface	T lymphocytes	
CD4/CD8	Surface	Helper T cells/cytotoxic T cells	
CD45RA	Surface	High expression identifies naïve T cells	[31, 32]
CD27	Surface	Highly expressed in naïve T cells but lost in fully differentiated cells after persistent antigen stimulation	[31]
Integrin β7 (β7)	Surface	Gut-homing marker.	[33]
CD161	Surface	Type 17 responses and cytotoxic functions.	[34, 35]
PSGL-1 (CLA)	Surface	Skin-homing marker.	[36]
CXCR-5	Surface	Follicular T cell marker.	[37]
LAG-3	Surface	Inhibitory receptor upregulated in recently activated lymphocytes.	[38]
FcERI (FCER1A)	Surface	High affinity receptor for IgE.	[39]
CD49b	Surface	Integrin (α2β1) that binds to collagen and modulates T cell stimulation, cytokine production and survival.	[40, 41]
CD137	Surface	Costimulatory receptor, promoting proliferation and survival, upregulated in activated lymphocytes.	[42]
CRTH2	Surface	Type 2 responses.	[43]
CD40L	Surface	Costimulatory ligand expressed on activated T cells.	[44]
T-box protein 21 (T-bet)	Intracellular	Type 1 responses and effector $CD8^+$ cells differentiation	[45]
GATA-3	Intracellular	Type 2 responses and cytotoxic $CD8^+$ function	[46]
BCL-6	Intracellular	Involved in follicular helper T cell differentiation. Involved in granzyme B production and memory differentiation in $CD8^+$ T cells.	[47, 48]
Forkhead box protein P3 (FoxP3)	Intracellular	Transcription factor expressed by regulatory T cells.	[49]
Nur77	Intracellular	Transcription factor indicative of T cell activation.	[50]

that Se1, Se2 and Se8 were naïve, Se4, Se5 and Se7 were CM, Se11 and Se13 were EM, and Se12 was EMRA. Interestingly, Seurat clustering retrieved populations exhibiting intermediate expression levels of CD45RA and CD27, such as clusters Se3, Se6, Se9, Se10 and Se14 (Figure 1c,d), which did not necessarily fall into any of the four conventional gates mentioned above. This implies that the conventional gating strategy based only on positive or negative expression of markers is incomplete. Seurat, on the other hand, can provide a more comprehensive analysis for multiple markers and thus enable the discovery of previously unidentified subsets.

Feature markers identified by Seurat and the potential functional properties of each cluster are summarised

in Table S4. Clusters with naïve phenotypes (Se1, Se2 and Se8) expressed different levels of integrin β7, indicating differential gut-homing potentials [33]. Two CM clusters, Se4 and Se7, were also high in integrin β7, however, they could be separated based on CXCR-5 and CD161 expression, as markers for follicular T cells and cytotoxic T cells, respectively. The CM cluster Se5 uniquely featured the skin-homing marker, CLA. Cluster Se6 and Se9 both expressed intermediate levels of CD161, but Se6 also had intermediate expression of integrin β7 while Se9 did not express integrin β7 at all. Se3 exhibited a similar profile to Se6, except for the absence of CD161. In addition, their intermediate expression of CD45RA implied that they might represent the transitional state between EM and
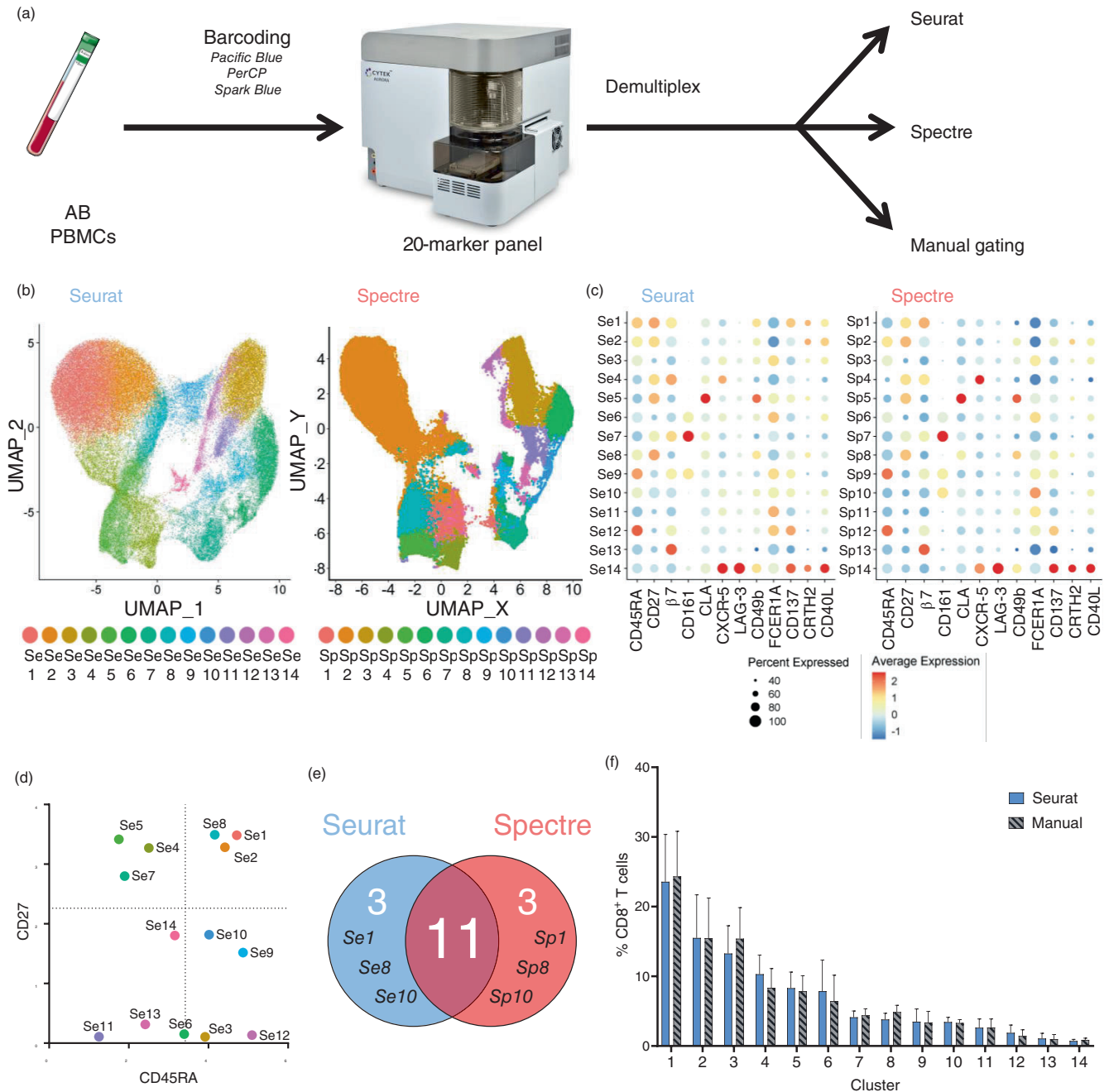
FIGURE 1 Adapting Seurat for high-dimensional flow cytometric data analysis retrieved robust results on adult blood (AB) peripheral blood mononuclear cells (PBMCs), confirmed by Spectre and manual analysis. (a) An overview of the adult blood (AB) study design. PBMCs were isolated from blood samples from healthy adult and then were first labeled with anti-CD45 antibodies with different fluorophores or their combinations for barcoding. After that, PBMCs were pooled together and stained with our 20-marker antibody panel and analysed with spectral cytometry. Next, the resulting data were first demultiplexed based on their CD45 marker signals and then subject to analysis with Seurat, Spectre and manual gating. (b) Uniform manifold approximation and projection (UMAP) plots visualising the clustering results from Seurat (left) and Spectre (right) based on the adult PBMC CD8$^+$ T-cell experiment. One colour represents one cluster. (c) Dot plots visualising the clusters identified by Seurat (left) and Spectre (right) and their marker expression profiles. The size of the dot corresponds to the percentage of cells expressing the corresponding markers and the colour gradient reflects the average normalised expression of the corresponding markers. (d) Projection of 14 clusters identified by Seurat based on the adult PBMC CD8$^+$ T-cell experiment onto the two-dimensional (2D) plot comparing their expression of CD27 and CD45RA. The dashed lines denote the average normalised expression of CD27 and CD45RA for all cells. (e) Venn diagram comparing the clustering results from Seurat and Spectre. Both methods were set to generate 14 clusters and 11 out of 14 clusters could be identified by both methods, while Se1, Se8 and Se10 could only be identified by Seurat and Sp1, Sp8 and Sp10 could only be identified by Spectre. (f) Bar chart comparing the proportions per sample within the total CD8$^+$ T-cell compartments of the clusters identified by Seurat or retrieved by manual gating. $N = 5$ per group and data are presented as mean ± s.e.m.

EMRA [51]. For EM-like and EMRA-like clusters, Se13 and Se12, but not Se11, they expressed high levels of integrin β7. Finally, Se14 was characterised by its high level of LAG-3, a marker for early T-cell activation, while Se10 might be a transitional population during the T-cell activation and maturation, considering its intermediate expression of both CD45RA and CD27. Other markers in our panel, including CD49b and FCER1A, were expressed at relatively low levels in T cells in the absence of stimulation. These subtle differences could still be detected by Seurat and confirmed with manual analysis (Figures 1c and S2A,B). For example, CD49b, a collagen-binding integrin, showed the highest level in the skin-homing population Se5, while FCER1A expression inversely correlated with CD27 expression, as exemplified in Se3, Se6, Se9, Se11 and Se12. Considering their relatively low level of expression, further research is required to evaluate whether such mild differences are of biological significance.

In parallel, we applied Spectre's workflow to analyse our dataset, manually defining 14 clusters as the final clustering output to cross-validate the findings from Seurat. A comparison of the Seurat and Spectre results revealed that their outcomes were similar, with 11 of the 14 clusters being identified by both approaches, displaying comparable cluster sizes and feature marker expression (Figures 1c,e and S2C). This indicated the robust and reliable performance of Seurat. As for their discrepancies, the Sp2 identified by Spectre was divided into three distinct clusters by Seurat, Se1, Se2 and Se8 (Figure S2D). In contrast, Se4 in Seurat analysis was split into Sp1 and Sp4 by Spectre (Figure S2D). In an independent comparative study, we also found that most cells were grouped into the same cluster by both methods, with minor discrepancies (Figure S3 and Supporting Information Materials and Methods S1.4.3). Together, these results indicated that clustering analyses from Seurat and Spectre were comparable and since they adopted different algorithms for their analyses, some differences were potentially to be expected.

Importantly, we further validated the results from Seurat with manual gating. As shown in Figure S4, all 14 Seurat clusters could be gated out according to their marker expression on 2D FACS plots. Projection of Seurat and Spectre clustering results on the 2D FACS plots showed similar distributions compared to the manual gating results (Figures S5–S7). Furthermore, the proportion of each population in the total CD8$^+$ T cells were also comparable between manual gating and Seurat clustering (Figure 1f).

Finally, we similarly applied the workflow to the gated CD4$^+$ T cells and retrieved 15 clusters with Seurat (Figure S8A,B). These subsets could also be similarly confirmed by manual gating and cross-validated with Spectre, which identified 12 out of the 15 subsets obtained by Seurat (Figure S8C,D).

Together, these results demonstrated that Seurat, a tool originally developed for scRNA-seq data analysis, is also applicable and robust for high-dimensional flow cytometric data analysis. Its analysis helps to characterise the CD8$^+$ T cells in PBMCs in more detail, retrieving novel T-cell sub-clusters.

Comparative profiling of CD8$^+$ T cells in AB and CB

Human circulatory T cells are plastic across the human lifespan and might be linked to differential disease susceptibility across human lives [52]. We next aimed to compare the profiles of the T-cell compartments from the CB and AB using Seurat.

As shown in Figure 2, PBMCs from AB and CBMCs from CB were first analysed based on our 20-parameter antibody panel with manual gating. Next, the high-dimensional data were analysed by Seurat and Spectre (Figures 3 and 4).

The proportion of CD4$^+$ and CD8$^+$ T cells among total lymphocytes showed no difference in PBMCs and CBMCs (Figure S1B). As previously shown [16], there were more naïve T cells and fewer EM T cells in both CD4$^+$ and CD8$^+$ compartments from CBMCs compared with PBMCs (Figure 2b–d). Reflective of a mature phenotype, adult PBMCs had higher proportions of CD4$^+$ CM T cells and CD8$^+$ EMRA T cells (Figure 2c,d). Our comprehensive 20-parameter antibody panel enabled an in-depth characterisation of the functional status of T cells.

As shown in Figure 2e, CD4$^+$ T cells in adult PBMCs had higher proportions of cells expressing T-bet, CLA and CXCR-5, while more CD4$^+$ T cells from CB expressed CD27, β7, and LAG-3. Similarly, CB CD8$^+$ T cells had a higher expression of CD27, β7, alongside FoxP3, whilst adult CD8$^+$ T cells had higher proportions of cells expressing CD161, CLA, CXCR-5 and T-bet (Figures 2f and S9A,B). These findings are consistent with previous reports of the higher gut-homing potential of CBMCs, whilst PBMCs are more likely to migrate to the skin [16, 52–54]. Moreover, the lower expression of T-bet and overall naïve-biased phenotype of CB T cells coincides with their reduced IFN-γ, IL-4 and IL-13 production compared to AB T cells (Figure S9C,D) [55].

Analysis of AB and CB CD8$^+$ T cells with Seurat identifies a unique CB CD8$^+$CD45RA$^+$CD27$^+$CD161$^+$ T cell subset

Conventional manual gating workflows are empirical and subject to bias. Such analysis is less likely to
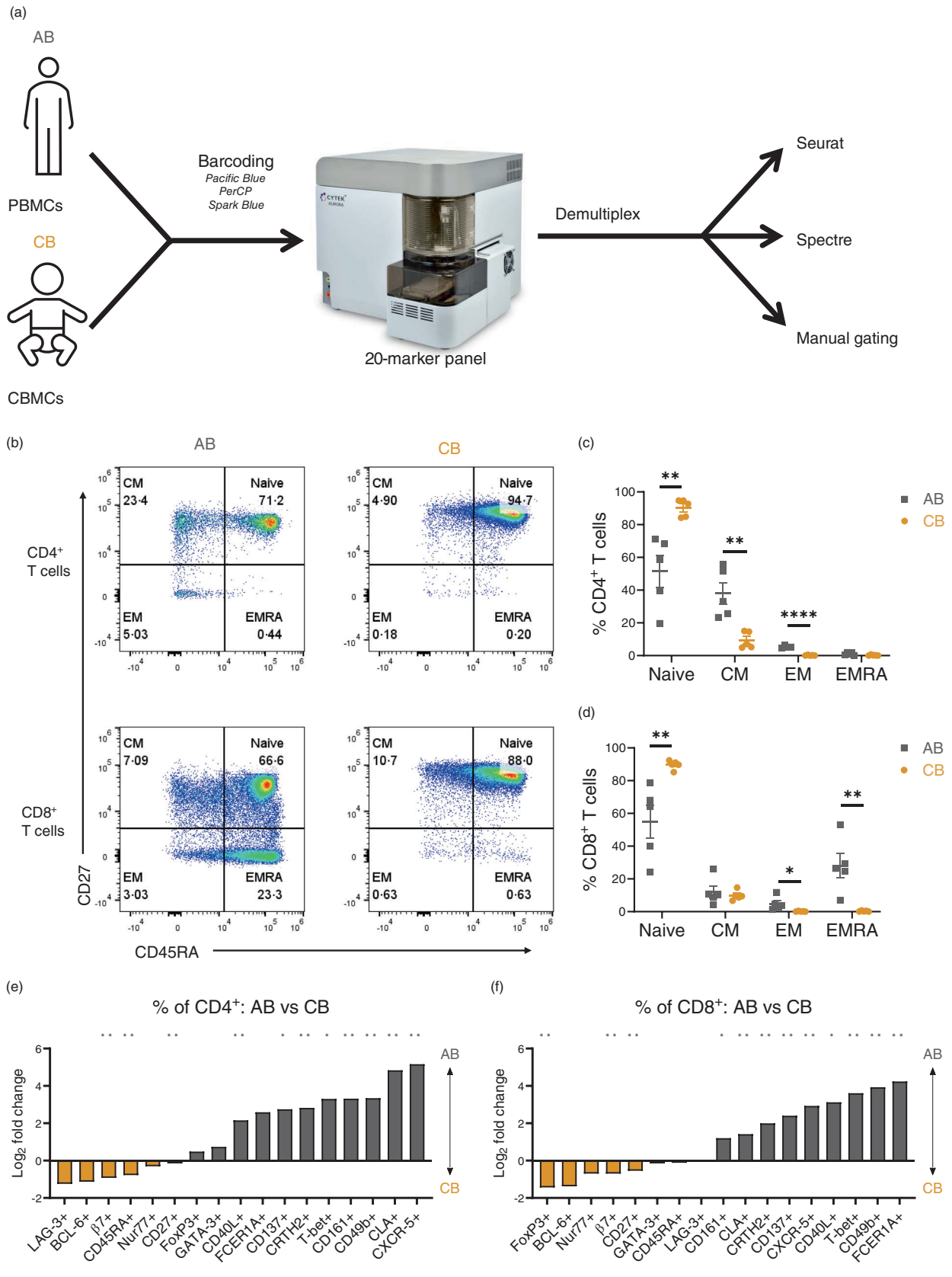
FIGURE 2   Legend on next page.

potentially reveal new cell populations. Focusing on the $CD8^+$ T-cell compartment, we, therefore, performed unsupervised PCA on the AB and CB combined dataset based on expression levels of the 12 surface markers used for clustering (Figure 3a). As expected, PBMC samples were distinctly separated from CBMC samples along the first PC (PC1, accounting for 18.0% of the variance), which was consistent with the differential expression of various T-cell functional markers in AB samples relative to CB samples (Figure 2e,f).

Next, Seurat was used for more in-depth analyses, identifying 13 sub-populations from the combined AB and CB dataset (Figure 3b). Seurat clustering distinguished three naïve (Clusters 1, 2 and 8), three CM (Clusters 3, 4 and 7), one EM (Cluster 11), and two EMRA clusters (Clusters 5 and 12), with the remaining clusters exhibiting intermediate expression levels of CD45RA and/or CD27 (Clusters 6, 9, 10, 13) (Figure 3c,d). Details about feature markers and potential functional properties of the identified $CD8^+$ T-cell subsets are summarised in Table S5. Clusters 1 and 2 shared a naïve phenotype but differed in their expression of integrin β7 and gut-homing potential, consistent with previous findings demonstrating an increased expression of gut-homing receptors in naïve CB lymphocytes [56]. The naïve Cluster 8 was also high in integrin β7 but additionally expressed CD161, suggesting a cytotoxic phenotype. Out of the three CM subsets identified by Seurat, Clusters 3 and 7 of the CM class, both had high integrin β7 expression, indicative of a migratory preference to the gut. Cluster 7 differed from Cluster 3 through its low expression of CD45RA and high expression of the follicular T-cell marker, CXCR-5. Opposed to gut-homing, the CM cluster 4 did not express integrin β7 but instead highly expressed the skin-homing marker CLA and integrin CD49b. Integrin β7 levels could also distinguish the EMRA-like Clusters 5 versus 12 as well as Clusters 6 versus 9, with Clusters 6 and 9 also expressing CD161. Finally, cluster 13 featured high levels of LAG-3 and CXCR-5.

We again compared the analyses from Seurat and Spectre based on the combined AB and CB dataset. Both methods identified 13 clusters, of which 11 were commonly shared, with comparable feature markers and cluster sizes (Figures S9E–G, S10B and S11B). Similarly, Seurat clusters in both AB and CB could also be validated through manual gating (Figures S10 and S11). These results again highlighted the robust performance of Seurat for analysing high-dimensional cytometric data.

Striking differences in the abundance of the Seurat clusters were found comparing AB and CB samples (Figure 3b and S12A). The naïve Cluster 1 was predominantly abundant in AB whilst Cluster 2 consisted primarily of CB cells. Seurat cluster analysis was able to separate these two naïve populations based on the differential expression of integrin β7 and CD27, which were enriched in the naïve population from CB (Cluster 2) [16, 56]. Consistent with previous studies comparing AB and CB, $CD8^+$ EM populations were almost exclusively found in AB [57]. This includes all the identified EM (Cluster 11) and EMRA (Clusters 5 and 12) subsets, in addition to Clusters 6, 9 and 10. Unlike these EM subsets, CM populations were present in both AB and CB. Of note, Cluster 3 was dominant in CB whilst Clusters 4 and 7 were enriched in AB. Finally, the LAG-3$^+$ Cluster 13 was equally abundant in both AB and CB.

Moreover, we also documented the differential expression of various markers between the clusters from AB and CB. Reflective of their naïve phenotype and lack of antigenic exposure, several CB $CD8^+$ T-cell populations (Clusters 1, 2, 3, 4, 7, 8, 9) exhibit higher expression of CD27 compared to their adult equivalents (Figure 3e). In contrast, the expression of CXCR-5 in Clusters 7 and 13 alongside CLA expression in Cluster 4 was higher in adult $CD8^+$ T cells compared to CB (Figure 3e).

Intriguingly, our clustering analysis identified a population (Cluster 8) that is almost exclusive to CBMCs (7.11% of $CD8^+$ T cells in CB vs. 0.19% in AB) (Figure 4a–c). This cluster is characterised as $CD8^+CD45RA^+CD27^+CD161^+$, partly overlapping with

FIGURE 2 Comparative analysis of T-cell profiles in peripheral blood mononuclear cells (PBMCs) from adult blood (AB) and cord blood mononuclear cells (CBMCs) from cord blood (CB). (a) An overview of the AB and CB study design. PBMCs and CBMCs were isolated and then were first labeled with anti-CD45 antibodies with different fluorophores or their combinations for barcoding. After that, PBMCs were pooled together and stained with our 20-marker antibody panel and analysed with spectral cytometry. Next, the resulting data were first demultiplexed based on their CD45 marker signals and then subject to analysis with Seurat and manual gating. (b) Representative flow cytometric plots of $CD4^+$ and $CD8^+$ T cells from AB and CB to identify naïve ($CD45RA^+CD27^+$), central memory (CM; $CD45RA^-CD27^+$), effector memory (EM; $CD45RA^-CD27^-$) and effector memory cells re-expressing CD45RA (EMRA; $CD45RA^+CD27^-$) subsets. (c, d) Scatter bar charts for the proportions of naïve, CM, EM and EMRA subsets within $CD4^+$ (c) and $CD8^+$ (d) T cells from AB and CB. (e, f) Bar charts for the log2(fold change) comparing AB versus CB for the proportions of populations expressing the corresponding markers among $CD4^+$ (e) and $CD8^+$ (f) T cells. The asterisks denote the populations whose proportions are significantly different between AB and CB. $N = 5$ per group and data are presented as mean, with *$p < 0.05$ and **$p < 0.01$ by unpaired Mann–Whitney U test.
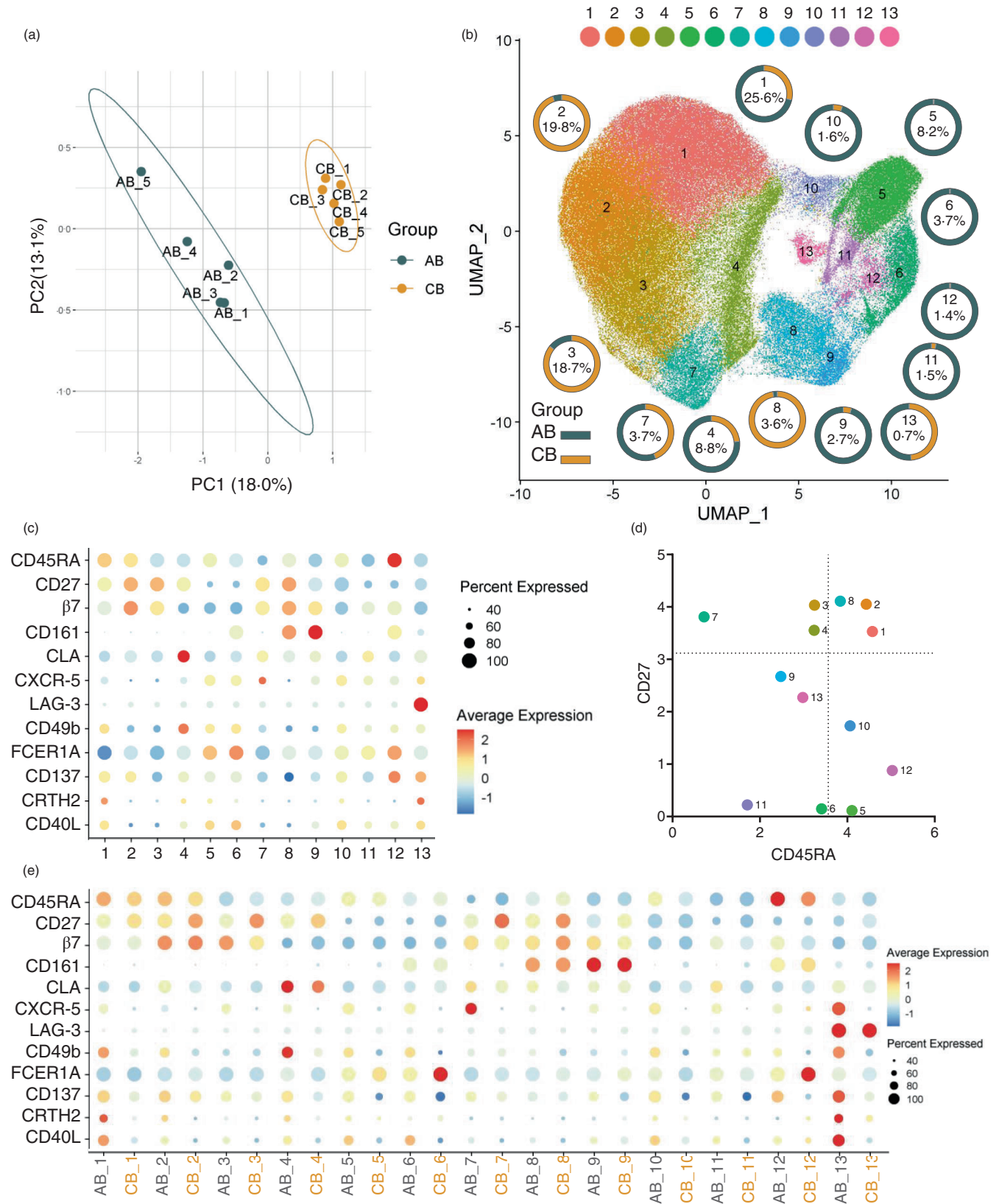
FIGURE 3 Legend on next page.

a previously reported but not fully characterised CD8$^+$CD161$^+$ T-cell population found in CB [58, 59]. Analysis based on our 20-parameter panel provided an unprecedented functional overview of the CD8$^+$CD45RA$^+$CD27$^+$CD161$^+$ T-cell subset. We discovered that this newly identified population had high integrin β7 expression (Figure 4b,d) as well as higher CLA, BCL-6, T-bet and GATA-3, but lower FoxP3 levels compared to its CD161$^-$ counterpart (Figure 4e). No differences in LAG-3, CXCR-5, Nur77, CD137 and FCER1A were found between the CD8$^+$CD45RA$^+$CD27$^+$CD161$^+$ and CD8$^+$CD45RA$^+$CD27$^+$CD161$^-$ T cells (Figure 4e). Upon stimulation with PMA and ionomycin, this subset predominantly produced IFN-γ and IL-4 but lowly expressed IL-5, IL-10 and IL-13. It only differed significantly in IL-10 production relative to the CD161$^-$ counterpart (Figures 4f and S12B). Based on these results, the CD8$^+$CD45RA$^+$CD27$^+$CD161$^+$ sub-population (Cluster 8) appeared to be a pro-inflammatory and cytotoxic T-cell subset.

In summary, using our high-dimensional antibody panel and Seurat analysis, we thoroughly profiled the CD8$^+$ T-cell compartment in AB versus CB. This revealed a unique CD8$^+$CD45RA$^+$CD27$^+$CD161$^+$ T-cell subset in CB which we characterised.

Cross-validation and further characterisation of the CD8$^+$CD45RA$^+$CD27$^+$CD161$^+$ T-cell subset using scRNA-seq

To validate our findings and further characterise the population identified by Seurat in CBMCs, we leveraged a recently published scRNA-seq dataset for naïve CD8$^+$ T cells, which analysed 18 513 cells across different developmental stages and compartments [27]. This dataset covers naïve T cells from foetal spleen, umbilical CB and adult peripheral blood. As shown in Figures 5a and S13A, the overall naïve CD8$^+$ T-cell population (sorted as CD8$^+$CD45RA$^+$CD27$^+$CCR7$^+$CD95$^-$) was further clustered into four subsets. Cluster0 expressed high levels of *RGS1*, which is linked to T-cell exhaustion [60]; cluster1 was high in *IL7R* and *SELL*, indicative of a naïve phenotype; and cluster2 was marked by *MT2A* and *RPS4Y1*, which might be potentially linked to a memory T-cell phenotype [61, 62]. Intriguingly, Cluster 3 exhibited similar features to the population we identified, characterised by its expression of *KLRB1* (gene encoding CD161) (Figure 5b). Differentially expressed gene analysis with DESeq2 [63] found 93 genes upregulated in Cluster 3 compared with the other naïve CD8$^+$ T cells, while 2 genes were downregulated (Figure 5c). Cluster 3 expressed higher levels of *GZMA* and *GZMK*, linked to cytotoxic T-cell features. Together with *MAFF*, they represented the core signature genes upregulated by CD161 in T-cell lineages reported previously [34, 58, 64, 65]. This again highlighted the close similarity, if not equivalence, between Cluster 3 and the CD8$^+$CD45RA$^+$CD27$^+$CD161$^+$ T-cell subset we identified. There were other additional inflammation-related genes upregulated, such as *ID1* [66] and *CMC1* [67]. Interestingly, *CCL5*, related to a memory phenotype, was also significantly upregulated in Cluster 3. Moreover, several chemokine receptors were also increased, such as *CXCR3* and *CXCR4*.

GO analysis was next carried out with the differentially expressed genes using ClusterProfiler [29, 30] based on the Gene Ontology pathways [68]. As shown in Figure 5d, compared with other naïve CD8$^+$ T cells, Cluster 3 was enriched for pathways related to *T-cell chemotaxis*, *lymphocyte chemotaxis* and *T-cell migration*, consistent with the differentially expressed gene analysis (Figure 5c) and our flow cytometric staining data for β7 and CLA (Figure 4e). On the other hand, this population was suppressed in pathways related to ribosomes.

The GO analyses were based only on the differentially expressed genes, and we next conducted GSEA based on

FIGURE 3 Unbiased comparative analysis of CD8$^+$ T-cell profiles in peripheral blood mononuclear cells (PBMCs) from adult blood (AB) and cord blood mononuclear cells (CBMCs) from cord blood (CB) with principal component analysis (PCA) and Seurat. (a) PCA based on the marker expression levels on CD8$^+$ T cells from AB and CB samples. (b) Uniform manifold approximation and projection (UMAP) plot visualising the clustering results from Seurat and the compositional contribution of each cluster from AB versus CB. The donut charts visualised the proportions of each cluster that are from AB (dark blue) and CB (yellow), and the numbers within denoted the proportions of the corresponding clusters within the overall CD8$^+$ T cells from both AB and CB. (c) Dot plot visualising the clusters identified by Seurat and their marker expression profiles. The size of the dot corresponds to the percentage of cells expressing the corresponding markers and the colour gradient reflects the average normalised expression of the corresponding markers. (d) Projection of 13 clusters identified by Seurat based on the AB versus CB experiment onto the two-dimensional (2D) plot comparing their expression of CD27 and CD45RA. The dashed lines denote the average normalised expression of CD27 and CD45RA for all cells. (e) Dot plot comparing the clusters from AB versus CB identified by Seurat and their marker expression profiles. The size of the dot corresponds to the percentage of cells expressing the corresponding markers and the colour gradient reflects the average normalised expression of the corresponding markers.
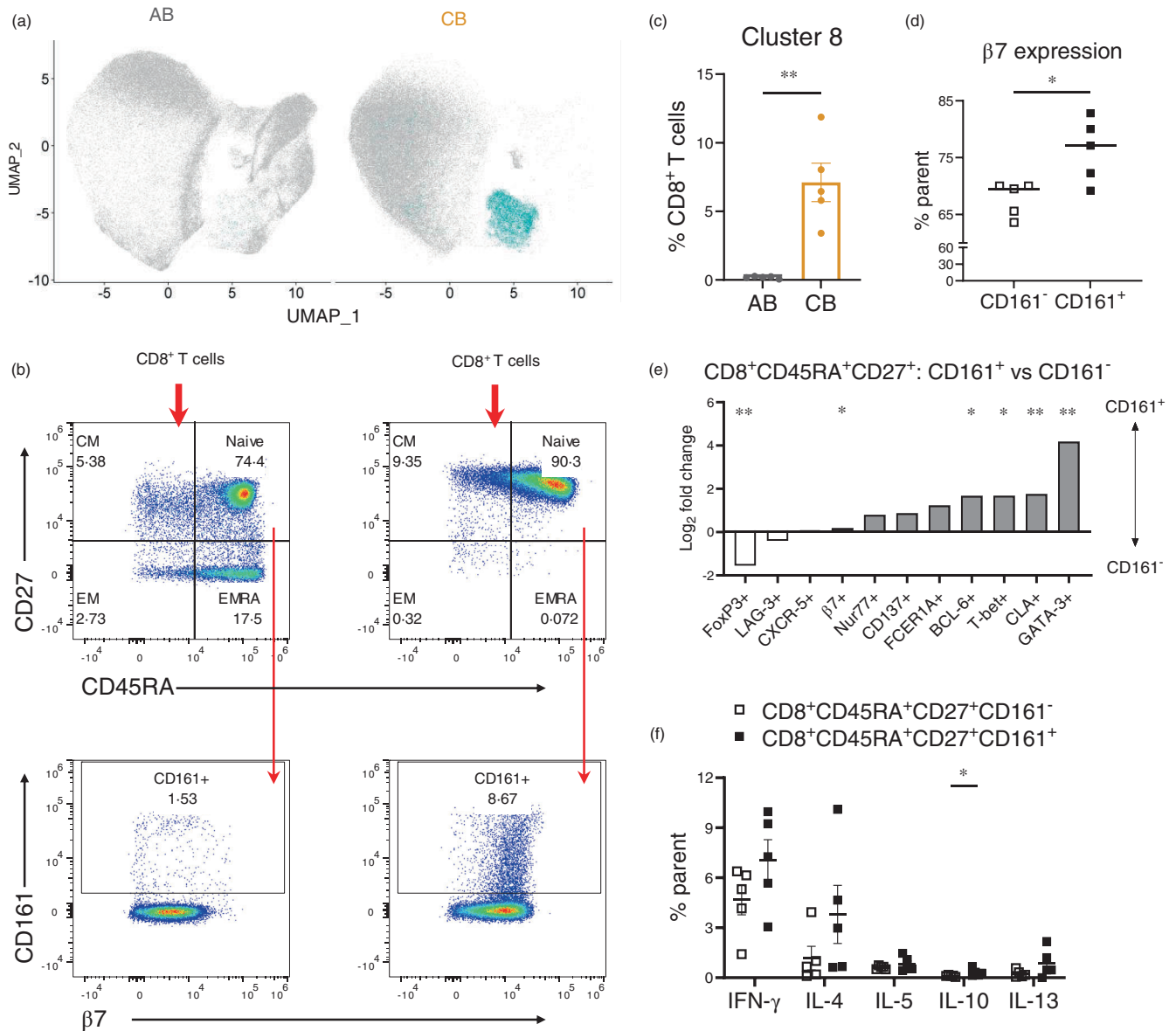
FIGURE 4 Seurat analysis identified a unique CD8+CD45RA+CD27+CD161+ T-cell population in cord blood. (a) Overlay of the newly identified CD8+CD45RA+CD27+CD161+ T-cell subset (Cluster 8) onto the uniform manifold approximation and projection (UMAP) plots for AB (left) and CB (right) CD8+ T-cell compartments. (b) Manual gating strategies to identify the CD8+CD45RA+CD27+CD161+ T-cell subset (Cluster 8) from AB (left) and CB (right) CD8+ T-cell compartments. (c) Scatter bar chart for the proportions of CD8+CD45RA+CD27+CD161+ T-cell subset (Cluster 8) in AB (black) and CB (yellow) CD8+ T-cell compartments. (d) Scatter dot plot of the proportions of cells expression integrin β7 in CD8+CD45RA+CD27+CD161+ and CD8+CD45RA+CD27+CD161− T-cell subsets. (e) Bar chart for the log2(fold change) comparing the proportions of populations expressing the corresponding markers among CD8+CD45RA+CD27+CD161+ and CD8+CD45RA+CD27+CD161− T-cell subsets. The asterisks denote the populations whose proportions are significantly different between AB and CB. (f) Scatter dot plot of the proportions of cells expressing IFN-γ, IL-4, IL-5, IL-10, and IL-13 in CD8+CD45RA+CD27+CD161+ and CD8+CD45RA+CD27+CD161− T-cell subsets. $N = 5$ per group and data are presented as mean, with *$p < 0.05$ and **$p < 0.01$ by unpaired Mann–Whitney U test.

the overall transcriptomic data using the fgsea package [69]. GSEA using the *Hallmark gene set* showed that Cluster 3 displayed enrichment in inflammatory pathways such as *TNFα signalling* via *NFκB*, *IL2-STAT5 signalling* and *IFNγ response* (Figure 5e). In addition, they

also showed enrichment in the *MAPK signalling pathway* and the *T-cell receptor signalling pathway* (Figure 5f,g).

In summary, scRNA-seq analysis from an independent dataset cross-validated the CD8+CD45RA+CD27+CD161+ population identified by our Seurat-based analysis. It also
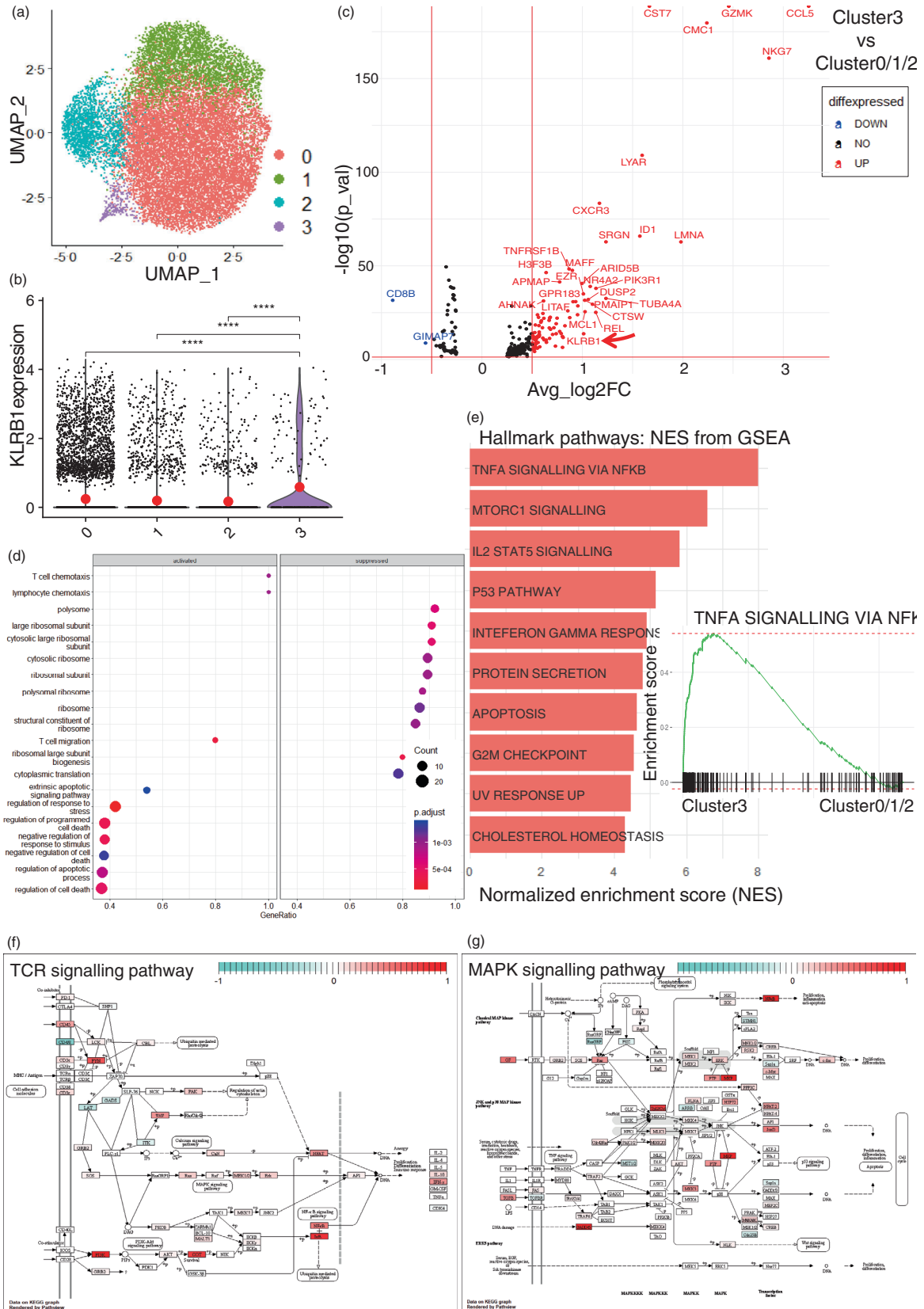
FIGURE 5 Legend on next page.

characterised it as a naïve subset but with a potential pro-inflammatory and cytotoxic profile.

DISCUSSION

There has been a significant expansion of both the size and complexity of cytometric data, especially in the field of clinical immunology. Such high-dimensional and complicated datasets cause great difficulties for conventional manual analytical strategies, inevitably hampering comprehensive and unbiased analyses and interpretation. Consequently, myriad computational toolkits have been developed, aiming to address these challenges, but their effective applications are sometimes restricted, suffering from a lack of flexibility and interoperability. Recently, packages such as Spectre [10] and tidyof [70] were developed, attempting to provide integrative, end-to-end services for cytometric data analysis. However, they have only been sporadically applied, due to them being standalone pipelines, requiring adaptation to completely new packages, and demanding significant prerequisite coding knowledge.

In the present study, we repurposed Seurat, a well-established package for scRNA-seq data analysis, for high-dimensional flow cytometric data analysis [23–26]. Comparison of Seurat and other currently available analytical packages are summarised in Table S6. Among them, Seurat has long shared great popularity within the field of single-cell analysis. It is community-driven and well-supported and has more than 20 R packages for related data processing and analysis. Therefore, it is likely to be more accessible and easier to use, particularly for broad users with previous experiences in scRNA-seq

looking to use flow cytometry to complement their investigative breadth.

Here, we showcased the robust capacity of Seurat, based on our experiments profiling the T-cell compartments in AB and CB. Overall, Seurat generated similar results to Spectre, which were also confirmed by manual gating. Importantly, with our approach, we identified a unique T-cell subset ($CD8^+CD45RA^+CD27^+CD161^+$) within CB and cross-validated its functional profiles with an independent scRNA-seq dataset using Seurat. Together, these data highlight the great potential of Seurat for cytometric data analysis. It represents a simple single platform for the unbiased analysis of both protein and RNA data at single-cell resolution. This will enable simpler comparison and cross-validation of cytometric and scRNA-seq studies and facilitate more comprehensive investigations and discoveries in clinical immunology.

A plethora of state-of-the-art mathematical algorithms or statistical models are used in the field of cytometry computational analysis. These include the FlowSOM modality used in Spectre, carrying out clustering based on a SOM method, while Seurat first ran PCAs on the overall dataset, next constructed a K-nearest neighbour (KNN) graph, similar to PhenoGraph, another single-cell analytical tool, and then adopted the Louvain algorithm to group cells together [10, 23]. Such differences might account for the discrepancies we observed when comparing the clustering outcomes from Seurat and Spectre. Detailed comparisons of these mathematical methodologies are not within the scope of the current work, but previous study has shown that FlowSOM and PhenoGraph were the top-performing unsupervised methods for mass cytometry data clustering analysis. The KNN graph model deployed in PhenoGraph excelled in

FIGURE 5 Single-cell RNA sequencing (scRNA-seq) cross-validation and characterisation of the newly identified $CD8^+CD45RA^+CD27^+CD161^+$ T-cell population. (a) Uniform manifold approximation and projection (UMAP) plot visualising the clustering results from re-analysis of the scRNA-seq data for naïve $CD8^+$ T cells from GEO: GSE158493. (b) Violin plot visualising the expression of KLRB1 (encoding CD161) across clusters identified by Seurat from scRNA-seq analysis. Red dots indicate the average expression of the corresponding group. (c) Volcano plot showing the differentially expressed genes comparing $CD8^+CD45RA^+CD27^+CD161^+$ and $CD8^+CD45RA^+CD27^+CD161^-$ T-cell subsets. The red dots indicate genes upregulated in $CD8^+CD45RA^+CD27^+CD161^+$ T cells, the blue dots indicate genes downregulated in $CD8^+CD45RA^+CD27^+CD161^+$ T cells, and the black ones are genes without significant changes. KLRB1, the gene that we used as the population-defining gene for Cluster 3 was specified. (d) Dot plot for the activated (left) and suppressed (right) pathways in $CD8^+CD45RA^+CD27^+CD161^+$ T-cell subset compared to $CD8^+CD45RA^+CD27^+CD161^-$ T-cell subset analysed by ClusterProfiler. The size of the dot corresponds to the number of genes within the corresponding pathway, the GeneRatio is calculated by the ratio between the number of genes that are significantly differentially expressed comparing $CD8^+CD45RA^+CD27^+CD161^+$ and $CD8^+CD45RA^+CD27^+CD161^-$ T-cell subsets and the number of total genes involved in the corresponding pathway, and the colour gradient reflects the adjusted p value for the comparison. (e) The top enriched gene sets of gene set enrichment analysis (GSEA) comparing $CD8^+CD45RA^+CD27^+CD161^+$ and $CD8^+CD45RA^+CD27^+CD161^-$ T-cell subsets based on the Hallmark gene sets and the representative GSEA plot of the top enriched gene set 'TNFA SIGNALING VIA NFKB'. (f, g). Visualisation of differentially expressed genes comparing $CD8^+CD45RA^+CD27^+CD161^+$ T-cell subsets relative to $CD8^+CD45RA^+CD27^+CD161^-$ T-cell subsets that are involved in T-cell receptor (TCR) signalling pathway (f) and mitogen-activated protein kinase (MAPK) signalling pathway (g).

its clustering precision, stability and robustness in identifying sub-clusters, relative to other approaches like flowMeans, DEPECHE and Xshift [71]. Since Seurat shares a similar KNN model to PhenoGraph, it is reasonable to expect it could also show robust capacity in more generic cytometric data analysis. This warrants more systematic comparisons in future work. On the other hand, differences in results from Seurat and Spectre highlight that utilising both methods in tandem will provide a more complete understanding of complex datasets.

In addition to its distinct mathematical nature, as one of the cutting-edge end-to-end analytical tools for single-cell data, Seurat has already been widely used in various research and clinical settings and is vigorously maintained and supported by its broad user community. This contributes to its easy accessibility and high user-friendliness and might reduce the coding burden as well, as users, especially those with previous experience in scRNA-seq, would not need to learn a completely new package or coding language for analysis. Moreover, the application of Seurat potentially opens more possibilities for cytometric data analysis. For example, the built-in function *FindMarkers()* in Seurat might facilitate easier marker identification, particularly for high-dimensional cytometric data, as the number of markers is expanding continuously. However, caution should still be taken to interpret its outcomes. Also, as a popular scRNA-seq analysis package, Seurat could also act as a wrapper with favourable interoperability around a wide range of complementary packages or plugins originally developed for scRNA-seq analysis, such as LIGER and Harmony for data integration [28, 72]. They might also be applicable to cytometric data, such as for batch correction, and could provide novel possibilities for high-dimensional data analyses once validated. Thus, adapting Seurat offers a single simple platform to analyse, compare and cross-validate protein and RNA, and even potentially other multi-omic single-cell data.

Previously, there were a few reports applying Seurat for protein-level single-cell analysis, such as for cellular indexing of transcriptomes and epitopes by sequencing (CITE-seq) [26] and CyTOF [73, 74], and leveraging the rPCA integration method in Seurat for spectral cytometry analysis [52, 75]. Recently, there has also been a similar attempt, adapting Scanpy, a Python-based scRNA-seq analysis package, to analyse mass cytometry data [76]. To our knowledge, the present work represents the first example of applying Seurat as a complete flow cytometric analysis workflow. Harnessing our 20-colour antibody panel and the Seurat-based analysis pipeline, we reported a unique T-cell subset in CBMCs, characterised as $CD8^+CD45RA^+CD27^+CD161^+$ T cells. This subset partly overlapped with the previously described $CD8^+CD161^+$ T cells [58]. Previously, studies first discovered the differential (low/intermediate/high) levels of CD161 on $CD8^+$ T cells in AB, which correlated with their various functional activities, including cytokine production, proliferation, and lytic activity [34]. AB $CD8^+CD161^{hi}$ T cells were predominantly mucosal-associated invariant T (MAIT) cells [65], while the $CD8^+CD161^{int}$ population represented a memory T-cell subset which was enriched in the colonic lamina propria [58]. Consistent with this, our clustering analysis found a $CD8^+CD161^+$ population predominantly existing in AB (Figure 3b), although, with the current clustering setting, both Seurat and Spectre failed to further subdivide it into $CD161^{int}$ and $CD161^{hi}$ subsets. This can be overcome by finetuning the clustering parameters depending on the particular scientific question of interest. As for our newly identified $CD8^+CD45RA^+CD27^+CD161^+$ subset, considering its naïve phenotype, it is not surprising that it is almost negligible in AB.

The role of the $CD8^+CD161^+$ T cells in CB remains elusive. Developmentally, it was found that the $CD8^+CD161^{hi}$ T cells in CB might be the progenitor for post-natal MAIT cells [59, 64, 65]. Functionally, the $CD161^{hi}$ subset produced IFN-γ and IL-17 [64, 77], while the $CD161^{int}$ subset, despite expressing markers like CD45RA and CCR7, still exhibited a preprogrammed transcriptomic profile reflective of their AB counterpart [58]. Our current clustering analysis could not further separate the population based on CD161 levels, but adjusting the clustering parameters could potentially help to differentiate them considering their intermediate to high CD161 expression (Figure 4b). The naïve phenotype of this CB-enriched subset, based on the expression of CD45RA and CD27, is similar to previous reports [58, 77], and we also confirmed its IFN-γ production. Previously, there were limited studies investigating the functional surface markers of $CD8^+CD161^+$ T cells, such as CCR6 [65]. Here, we characterised the $CD8^+CD45RA^+CD27^+CD161^+$ population as high in integrin β7 but low in CLA expression, implying a preference to gut over skin-homing. Thus, this CB population might represent the progenitors for AB $CD8^+CD161^{int}$ T cells, which are enriched within the colon [58]. $CD8^+CD161^+$ T cells in AB are involved in the response to tissue-localised inflammation triggered by intracellular and viral pathogens [58, 65], while their functional implications in CB remain elusive. Likewise, both our flow cytometry and scRNA-seq data suggested the pro-inflammatory and cytotoxic properties of CB $CD8^+CD45RA^+CD27^+CD161^+$ T cells. In addition, as CD161 contributes to prenatal immune suppression [78], this subset might be involved in maintaining tolerance in the semi-allogenic context of pregnancy.

In summary, we have adapted Seurat, a widely used scRNA-seq analysis package, for high-dimensional flow

cytometric data analysis and showcased its performance through the identification of a unique CD8⁺CD45RA⁺ CD27⁺CD161⁺ T-cell population in CB. Such a pipeline presents a novel avenue for comprehensive analysis of high-dimensional complex cytometric and multi-modal data, facilitating unbiased data-driven studies and discovery.

AUTHOR CONTRIBUTIONS

ACKNOWLEDGEMENTS

CONFLICT OF INTEREST STATEMENT

The 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.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors upon reasonable request.

ETHICS STATEMENT

This study was reviewed and approved by the Human Research Ethics Committee of the Nepean Blue Mountains Local Health District Network (HREC 03/006) according to the Declaration of Helsinki. All participants were recruited on a volunteer basis. Participants or their guardians signed a written informed consent before sample collection.

ORCID

Duan Ni https://orcid.org/0000-0002-3902-2843
Nicholas Jonathan Cole King https://orcid.org/0000-0002-3877-9772

REFERENCES

1. Ginhoux F, Yalin A, Dutertre CA, Amit I. Single-cell immunology: past, present, and future. Immunity. 2022;55(3):393–404.
2. Saeys Y, Van Gassen S, Lambrecht BN. Computational flow cytometry: helping to make sense of high-dimensional immunology data. Nat Rev Immunol. 2016;16(7):449–62.
3. Park LM, Lannigan J, Jaimes MC. OMIP-069: forty-color full spectrum flow cytometry panel for deep immunophenotyping of major cell subsets in human peripheral blood. Cytometry A. 2020;97(10):1044–51.
4. Spitzer MH, Nolan GP. Mass cytometry: single cells, many features. Cell. 2016;165(4):780–91.
5. Mair F, Hartmann FJ, Mrdjen D, Tosevski V, Krieg C, Becher B. The end of gating? An introduction to automated analysis of high dimensional cytometry data. Eur J Immunol. 2016;46(1):34–43.
6. van der Maaten L, Hinton G. Visualizing data using t-SNE. J Mach Learn Res. 2008;9(86):2579–605.
7. Becht E, McInnes L, Healy J, Dutertre CA, Kwok IWH, Ng LG, et al. Dimensionality reduction for visualizing single-cell data using UMAP. Nat Biotechnol. 2018;37:38–44.
8. Levine JH, Simonds EF, Bendall SC, Davis KL, Amir ED, Tadmor M, et al. Data-driven phenotypic dissection of AML reveals progenitor-like cells that correlate with prognosis. Cell. 2015;162(1):184–97.
9. 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 A. 2015;87(7):636–45.
10. Ashhurst TM, Marsh-Wakefield F, Putri GH, Spiteri AG, Shinko D, Read MN, et al. Integration, exploration, and analysis of high-dimensional single-cell cytometry data using Spectre. Cytometry A. 2022;101(3):237–53.
11. Bhattacharya S, Dunn P, Thomas CG, Smith B, Schaefer H, Chen J, et al. ImmPort, toward repurposing of open access immunological assay data for translational and clinical research. Sci Data. 2018;5:180015.
12. OMIQ. Flow Cytometry Software, Reimagined [Internet]. Available from: https://www.omiq.ai
13. Realize R&D Digital Transformation. Realize R&D Digital Transformation [Internet]. Available from: https://www.dotmatics.com/
14. Kotecha N, Krutzik PO, Irish JM. Web-based analysis and publication of flow cytometry experiments. Curr Protoc Cytom. 2010;53(1):9.32.1–12.20.40. https://doi.org/10.1002/0471142956.cy1017s53
15. Rackaityte E, Halkias J. Mechanisms of fetal T cell tolerance and immune regulation. Front Immunol. 2020;11:588. https://doi.org/10.3389/fimmu.2020.00588

16. Szabolcs P, Park KD, Reese M, Marti L, Broadwater G, Kurtzberg J. Coexistent naive phenotype and higher cycling rate of cord blood T cells as compared to adult peripheral blood. Exp Hematol. 2003;31(8):708–14.

17. Chen L, Cohen AC, Lewis DB. Impaired allogeneic activation and T-helper 1 differentiation of human cord blood naive CD4 T cells. Biol Blood Marrow Transplant. 2006;12(2):160–71.

18. Galindo-Albarrán AO, López-Portales OH, Gutiérrez-Reyna DY, Rodríguez-Jorge O, Sánchez-Villanueva JA, Ramírez-Pliego O, et al. CD8+ T cells from human neonates are biased toward an innate immune response. Cell Rep. 2016;17(8):2151–60.

19. Luciano AA, Yu H, Jackson LW, Wolfe LA, Bernstein HB. Preterm labor and chorioamnionitis are associated with neonatal T cell activation. PLoS One. 2011;6(2):e16698.

20. Ponsonby AL, Collier F, O'Hely M, Tang MLK, Ranganathan S, Gray L, et al. Household size, T regulatory cell development, and early allergic disease: a birth cohort study. Pediatr Allergy Immunol. 2022;33(6):e13810.

21. Collier F, Ponsonby AL, O'Hely M, Tang MLK, Saffery R, Molloy J, et al. Naïve regulatory T cells in infancy: associations with perinatal factors and development of food allergy. Allergy. 2019;74(9):1760–8.

22. Smith M, Tourigny MR, Noakes P, Thornton CA, Tulic MK, Prescott SL. Children with egg allergy have evidence of reduced neonatal CD4+CD25+CD127lo/− regulatory T cell function. J Allergy Clin Immunol. 2008;121(6):1460–1466.e7.

23. Satija R, Farrell JA, Gennert D, Schier AF, Regev A. Spatial reconstruction of single-cell gene expression data. Nat Biotechnol. 2015;33(5):495–502.

24. Butler A, Hoffman P, Smibert P, Papalexi E, Satija R. Integrating single-cell transcriptomic data across different conditions, technologies, and species. Nat Biotechnol. 2018;36(5):411–20.

25. Stuart T, Butler A, Hoffman P, Hafemeister C, Papalexi E, Mauck WM, et al. Comprehensive integration of single-cell data. Cell. 2019;177(7):1888–1902.e21.

26. Hao Y, Hao S, Andersen-Nissen E, Mauck WM, Zheng S, Butler A, et al. Integrated analysis of multimodal single-cell data. Cell. 2021;184(13):3573–3587.e29.

27. Bunis DG, Bronevetsky Y, Krow-Lucal E, Bhakta NR, Kim CC, Nerella S, et al. Single-cell mapping of progressive fetal-to-adult transition in human naive T cells. Cell Rep. 2021; 34(1):108573.

28. Welch JD, Kozareva V, Ferreira A, Vanderburg C, Martin C, Macosko EZ. Single-cell multi-omic integration compares and contrasts features of brain cell identity. Cell. 2019;177(7):1873–1887.e17.

29. Yu G, Wang LG, Han Y, He QY. clusterProfiler: an R package for comparing biological themes among gene clusters. Omics. 2012;16(5):284–7.

30. Wu T, Hu E, Xu S, Chen M, Guo P, Dai Z, et al. clusterProfiler 4.0: a universal enrichment tool for interpreting omics data. Innovation. 2021;2(3):100141.

31. Hamann D, Baars PA, Rep MHG, Hooibrink B, Kerkhof-Garde SR, Klein MR, et al. Phenotypic and functional separation of memory and effector human CD8+ T cells. J Exp Med. 1997;186(9):1407–18.

32. Okada R, Kondo T, Matsuki F, Takata H, Takiguchi M. Phenotypic classification of human CD4+ T cell subsets and their differentiation. Int Immunol. 2008;20(9):1189–99.

33. Gorfu G, Rivera-Nieves J, Ley K. Role of β7 integrins in intestinal lymphocyte homing and retention. Curr Mol Med. 2009; 9(7):836–50.

34. Takahashi T, Dejbakhsh-Jones S, Strober S. Expression of CD161 (NKR-P1A) defines subsets of human CD4 and CD8 T cells with different functional activities. J Immunol. 2006; 176(1):211–6.

35. Fergusson J, Fleming V, Klenerman P. CD161-expressing human T cells. Front Immunol. 2011;2:36. https://doi.org/10.3389/fimmu.2011.00036

36. de Jesús-Gil C, Sans-de SanNicolàs L, García-Jiménez I, Ferran M, Celada A, Chiriac A, et al. The translational relevance of human circulating memory cutaneous lymphocyte-associated antigen positive T cells in inflammatory skin disorders. Front Immunol. 2021;12:652613. https://doi.org/10.3389/fimmu.2021.652613

37. Chevalier N, Jarrossay D, Ho E, Avery DT, Ma CS, Yu D, et al. CXCR5 expressing human central memory CD4 T cells and their relevance for humoral immune responses. J Immunol. 2011;186(10):5556–68.

38. Chocarro L, Blanco E, Zuazo M, Arasanz H, Bocanegra A, Fernández-Rubio L, et al. Understanding LAG-3 signaling. Int J Mol Sci. 2021;22(10):5282.

39. Wang J, Lindholt JS, Sukhova GK, Shi MA, Xia M, Chen H, et al. IgE actions on CD4+ T cells, mast cells, and macrophages participate in the pathogenesis of experimental abdominal aortic aneurysms. EMBO Mol Med. 2014;6(7):952–69.

40. Boisvert M, Gendron S, Chetoui N, Aoudjit F. Alpha2beta1 integrin signaling augments T cell receptor-dependent production of interferon-gamma in human T cells. Mol Immunol. 2007;44(15):3732–40.

41. Boisvert M, Chetoui N, Gendron S, Aoudjit F. Alpha2beta1 integrin is the major collagen-binding integrin expressed on human Th17 cells. Eur J Immunol. 2010;40(10):2710–9.

42. Palma C, Binaschi M, Bigioni M, Maggi CA, Goso C. CD137 and CD137 ligand constitutively coexpressed on human T and B leukemia cells signal proliferation and survival. Int J Cancer. 2004;108(3):390–8.

43. Tsuda H, Michimata T, Sakai M, Nagata K, Nakamura M, Saito S. A novel surface molecule of Th2- and Tc2-type cells, CRTH2 expression on human peripheral and decidual CD4+ and CD8+ T cells during the early stage of pregnancy. Clin Exp Immunol. 2001;123(1):105–11.

44. Elgueta R, Benson MJ, de Vries VC, Wasiuk A, Guo Y, Noelle RJ. Molecular mechanism and function of CD40/CD40L engagement in the immune system. Immunol Rev. 2009;229(1):152–72. https://doi.org/10.1111/j.1600-065X.2009.00782.x

45. Yeo CJJ, Fearon DT. T-bet-mediated differentiation of the activated CD8+ T cell. Eur J Immunol. 2011;41(1):60–6.

46. Tai TS, Pai SY, Ho IC. GATA-3 regulates the homeostasis and activation of CD8+ T cells. J Immunol. 2013;190(1):428–37.

47. Nurieva RI, Chung Y, Martinez GJ, Yang XO, Tanaka S, Matskevitch TD, et al. Bcl6 mediates the development of T follicular helper cells. Science. 2009;325(5943):1001–5.

48. Liu Z, Guo Y, Tang S, Zhou L, Huang C, Cao Y, et al. Cutting edge: transcription factor BCL6 is required for the generation, but not maintenance, of memory CD8+ T cells in acute viral infection. J Immunol. 2019;203(2):323–7.

49. Mayer CT, Floess S, Baru AM, Lahl K, Huehn J, Sparwasser T. CD8+Foxp3+ T cells share developmental and phenotypic features with classical CD4+Foxp3+ regulatory T cells but lack potent suppressive activity. Eur J Immunol. 2011;41(3):716–25.

50. Liebmann M, Hucke S, Koch K, Eschborn M, Ghelman J, Chasan AI, et al. Nur77 serves as a molecular brake of the metabolic switch during T cell activation to restrict autoimmunity. Proc Natl Acad Sci U S A. 2018;115(34):E8017–26.

51. Carrasco J, Godelaine D, Van Pel A, Boon T, van der Bruggen P. CD45RA on human CD8 T cells is sensitive to the time elapsed since the last antigenic stimulation. Blood. 2006; 108(9):2897–905.

52. Wong MT, Ong DEH, Lim FSH, Teng KWW, McGovern N, Narayanan S, et al. A high-dimensional atlas of human T cell diversity reveals tissue-specific trafficking and cytokine signatures. Immunity. 2016;45(2):442–56.

53. Staton TL, Habtezion A, Winslow MM, Sato T, Love PE, Butcher EC. CD8+ recent thymic emigrants home to and efficiently repopulate the small intestine epithelium. Nat Immunol. 2006;7(5):482–8.

54. Hsu PS, Lai CL, Hu M, Santner-Nanan B, Dahlstrom JE, Lee CH, et al. IL-2 enhances gut homing potential of human naive regulatory T cells early in life. J Immunol. 2018;200(12): 3970–80.

55. Nitsche A, Zhang M, Clauss T, Siegert W, Brune K, Pahl A. Cytokine profiles of cord and adult blood leukocytes: differences in expression are due to differences in expression and activation of transcription factors. BMC Immunol. 2007; 8(1):18.

56. Crespo M, Martinez DG, Cerissi A, Rivera-Reyes B, Bernstein HB, Lederman MM, et al. Neonatal T-cell maturation and homing receptor responses to toll-like receptor ligands differ from those of adult naive T cells: relationship to prematurity. Pediatr Res. 2012;71(2):136–43.

57. Liu DD, Hong WC, Qiu KY, Li XY, Liu Y, Zhu LW, et al. Umbilical cord blood: a promising source for allogeneic CAR-T cells. Front Oncol. 2022;12:944248. https://doi.org/10.3389/fonc.2022.944248

58. Fergusson JR, Hühn MH, Swadling L, Walker LJ, Kurioka A, Llibre A, et al. CD161intCD8+ T cells: a novel population of highly functional, memory CD8+ T cells enriched within the gut. Mucosal Immunol. 2016;9(2):401–13.

59. Walker LJ, Kang YH, Smith MO, Tharmalingham H, Ramamurthy N, Fleming VM, et al. Human MAIT and CD8αα cells develop from a pool of type-17 precommitted CD8+ T cells. Blood. 2012;119(2):422–33.

60. Bai Y, Hu M, Chen Z, Wei J, Du H. Single-cell transcriptome analysis reveals RGS1 as a new marker and promoting factor for T-cell exhaustion in multiple cancers. Front Immunol. 2021;12:767070.

61. Sjöstedt E, Zhong W, Fagerberg L, Karlsson M, Mitsios N, Adori C, et al. An atlas of the protein-coding genes in the human, pig, and mouse brain. Science. 2020;367(6482): eaay5947.

62. Zhao HC, Chen CZ, Song HQ, Wang XX, Zhang L, Zhao HL, et al. Single-cell RNA sequencing analysis reveals new immune disorder complexities in Hypersplenism. Front Immunol. 2022;13:921900.

63. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. 2014;15(12):550.

64. Fergusson JR, Smith KE, Fleming VM, Rajoriya N, Newell EW, Simmons R, et al. CD161 defines a transcriptional and functional phenotype across distinct human T cell lineages. Cell Rep. 2014;9(3):1075–88.

65. Billerbeck E, Kang YH, Walker L, Lockstone H, Grafmueller S, Fleming V, et al. Analysis of CD161 expression on human CD8+ T cells defines a distinct functional subset with tissue-homing properties. Proc Natl Acad Sci U S A. 2010; 107(7):3006–11.

66. Yang Y, Liou HC, Sun XH. Id1 potentiates NF-kappaB activation upon T cell receptor signaling. J Biol Chem. 2006;281(46): 34989–96.

67. Li Y, Ren P, Dawson A, Vasquez HG, Ageedi W, Zhang C, et al. Single-cell transcriptome analysis reveals dynamic cell populations and differential gene expression patterns in control and aneurysmal human aortic tissue. Circulation. 2020; 142(14):1374–88.

68. Gene Ontology Consortium, Aleksander SA, Balhoff J, Carbon S, Cherry JM, Drabkin HJ, et al. The gene ontology knowledgebase in 2023. Genetics. 2023;224(1):iyad031.

69. Korotkevich G, Sukhov V, Budin N, Shpak B, Artyomov MN, Sergushichev A. Fast gene set enrichment analysis. bioRxiv. 2021. https://doi.org/10.1101/060012

70. Keyes TJ, Koladiya A, Lo YC, Nolan GP, Davis KL. tidytof: A user-friendly framework for scalable and reproducible high-dimensional cytometry data analysis. Bioinform Adv. 2023; 3(1):vbad071.

71. Liu X, Song W, Wong BY, Zhang T, Yu S, Lin GN, et al. A comparison framework and guideline of clustering methods for mass cytometry data. Genome Biol. 2019;20(1):297.

72. 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(12):1289–96.

73. Wang K, Yang Y, Wu F, Song B, Wang X, Wang T. Comparative analysis of dimension reduction methods for cytometry by time-of-flight data. Nat Commun. 2023;14(1):1836.

74. Run CyTOF analysis with Seurat [Internet]. [cited 2023 Jul 25]. Available from: https://tjburns08.github.io/run_cytof_with_seurat.html

75. Verstegen NJM, Hagen RR, van den Dijssel J, Kuijper LH, Kreher C, Ashhurst T, et al. Immune dynamics in SARS-CoV-2 experienced immunosuppressed rheumatoid arthritis or multiple sclerosis patients vaccinated with mRNA-1273. Elife. 2022;11:e77969.

76. Büttner M, Hempel F, Ryborz T, Theis FJ, Schultze JL. Pytometry: flow and mass cytometry analytics in Python. bioRxiv. 2022. https://doi.org/10.1101/2022.10.10.511546

77. Dusseaux M, Martin E, Serriari N, Péguillet I, Premel V, Louis D, et al. Human MAIT cells are xenobiotic-resistant, tissue-targeted, CD161hi IL-17-secreting T cells. Blood. 2011; 117(4):1250–9.

78. Halkias J, Rackaityte E, Hillman SL, Aran D, Mendoza VF, Marshall LR, et al. CD161 contributes to prenatal immune suppression of IFN-γ–producing PLZF+ T cells. J Clin Invest. 2019;129(9):3562–77.

79. Reyes JGA, Ni D, Santner-Nanan B, Pinget GV, Kraftova L, Ashhurst TM, et al. A unique human cord blood CD8+CD45RA+CD27+CD161+ T cell subset identified by flow cytometric data analysis using Seurat. bioRxiv. 2023. https://doi.org/10.1101/2023.08.01.549954

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.