

# Multimodal discovery of a pathogenic B cell-dependent T cell state in multiple sclerosis

Redwan Farooq,<sup>1,2</sup> Brian Cutler,<sup>1,2</sup> Marco Pisa,<sup>1,3</sup> Subita Balaram Kuttikkatte,<sup>1,2</sup> Jacqueline Palace,<sup>1,3</sup> Gabriele C. DeLuca,<sup>1,3</sup> Ralf Gold,<sup>4</sup> Fabian J. Theis,<sup>5,6,7</sup> Lars Fugger<sup>1,2</sup> and Kathrine E. Attfield<sup>1,2</sup>

## Abstract

B cell depleting therapies in multiple sclerosis (MS) have transformed disease management, yet the immunological mechanisms linking B cells to chronic compartmentalised CNS inflammation, the key pathological driver of disability accrual, remain poorly defined.

Here, we leverage anti-CD20 therapy as an *in vivo* perturbational probe to disentangle MS immunobiology. We combine longitudinal, high-dimensional multimodal immune profiling in patients initiating ocrelizumab with a novel machine learning pipeline to resolve treatment-induced shifts in continuous immune cell states, revealing treatment-associated modulation of shared biological processes that act across the boundaries of discretely partitioned cell types.

We identify a distinct chronically activated, proinflammatory–cytotoxic T cell state with CNS-homing properties that is selectively depleted following B cell ablation. Importantly, across independent datasets, the same T cell state is enriched in the circulation of patients with clinically aggressive relapsing disease, in CSF-enriched expanded clonotypes, including a subset with proven Epstein–Barr virus-specificity, and in chronically inflamed lesion rims in end-stage MS. Finally, we demonstrate that blockade of lymphocyte trafficking across the blood–brain barrier with the anti-integrin  $\alpha 4$  therapy natalizumab leads to enrichment of this cell state in the circulation, demonstrating a mechanistically concordant effect across distinct high-efficacy MS therapies.

Together, our findings support a B cell-dependent, pathogenic T cell state which links peripheral immune activation to CNS-compartmentalised smouldering neuroinflammation across disease stages. Direct therapeutic manipulation of this T cell state may represent a

1 key opportunity for targeting chronic neuroinflammation and facilitate a strategic shift away  
2 from broad immune cell ablation.

3  
4  
5 **Author affiliations:**

6 1 Nuffield Department of Clinical Neurosciences, University of Oxford, John Radcliffe  
7 Hospital, Oxford, OX3 9DU, UK

8 2 MRC Translational Immune Discovery Unit, Weatherall Institute of Molecular Medicine,  
9 University of Oxford, John Radcliffe Hospital, Oxford, OX3 9DS, UK

10 3 Department of Neurology, Oxford University Hospitals NHS Foundation Trust, John  
11 Radcliffe Hospital, Oxford, OX3 9DU, UK

12 4 Department of Neurology, St. Josef-Hospital, Ruhr University Bochum, Bochum, 44801,  
13 Germany

14 5 Institute of Computational Biology, Helmholtz Munich, Munich, 80939, Germany

15 6 TUM School of Life Sciences Weihenstephan, Technical University of Munich, Munich,  
16 80333, Germany

17 7 TUM School of Computing, Information and Technology, Technical University of Munich,  
18 Munich, 80333, Germany

19  
20 Correspondence to: Kathrine E. Attfield

21 Nuffield Department of Clinical Neurosciences, University of Oxford, John Radcliffe  
22 Hospital, Oxford, OX3 9DU, UK

E-mail: kate.attfield@ndcn.ox.ac.uk

**Running title:** Pathogenic B cell-dependent T cells in MS

**Keywords:** multiple sclerosis; neuroimmunology; B cells; T cells; single-cell transcriptomics; machine learning

## Introduction

Multiple sclerosis (MS) is the most common inflammatory disease of the CNS,<sup>1</sup> affecting more than 2.8 million people worldwide with a rising prevalence.<sup>2</sup> Recent evidence has challenged the historical dichotomy between relapsing–remitting MS (RRMS) and primary/secondary progressive MS (PPMS/SPMS), reframing the disease as a continuous spectrum,<sup>3</sup> with neuronal loss driven by underlying inflammation.<sup>4,5</sup> Multiple lines of evidence suggest that the MS arises in individuals with a genetically primed immune system, precipitated by specific environmental triggers<sup>6</sup>; in particular, infection with Epstein–Barr virus (EBV) appears to be a prerequisite.<sup>7–9</sup> Despite these advances, an immunological model that mechanistically unifies aetiological factors with pathological processes and therapeutic response across the evolving disease continuum remains elusive.

In the last decade, the remarkable clinical efficacy of anti-CD20 monoclonal antibodies, which deplete B cells, has supported a shift in our understanding of the immunological basis of MS. However, the exact mechanism of B cell pathogenicity in MS remains undetermined. The lack of broadly prevalent disease-specific autoantibodies in MS is highly suggestive of a predominantly non-humoral role.<sup>10–12</sup> It is well established that CD8<sup>+</sup> T cells predominate the lymphocytic infiltrate of actively inflamed MS lesions across disease stages,<sup>13–16</sup> alongside highly abundant myeloid cells comprising macrophages and microglia.<sup>6</sup> T cells co-localise with B cell infiltrates in perivascular spaces and meningeal lymphoid aggregates.<sup>17,18</sup> These observations suggest that complex dysregulated interactions between B cells, T cells and myeloid cells may drive MS.

Recent studies have supported the notion that anti-CD20 therapies act beyond the B cell compartment, with secondary effects on peripheral myeloid and memory T cell subsets.<sup>19–24</sup> However, the ability to reliably measure these effects has been restricted by analytical approaches reliant on discretely partitioned cell populations. Importantly, improved

1 methods for single-cell data analysis have revealed a continuum of cellular states  
2 exhibiting shared biological processes across cell type boundaries.

3 We performed cellular indexing of transcriptomes and epitopes by sequencing (CITE-seq)  
4 on longitudinal peripheral blood mononuclear cell (PBMC) samples from people with MS  
5 (pwMS) receiving ocrelizumab, the first humanised anti-CD20 monoclonal antibody  
6 licensed for treatment of MS. To overcome the limitations of prior studies, we developed a  
7 novel machine learning pipeline to measure treatment-induced shifts in continuous and  
8 overlapping transcriptional states, allowing us to perform a deep and unbiased *in vivo*  
9 characterisation of subtle shifts in cellular phenotypes after B cell depletion in the disease  
10 context.

11 We identify a distinct transcriptional and surface protein phenotype of B cell-dependent T  
12 cells in the blood of pwMS, which we validate in multiple independent and orthogonal  
13 single-cell datasets, tracking this cell state across tissue compartments, disease stages  
14 and therapeutic modalities. Our findings reveal a chronically activated, proinflammatory-  
15 cytotoxic and CNS-homing T cell state, which we hypothesise is driven by latent EBV  
16 infection in B cells, arises in the circulation in early-stage disease, becomes clonally  
17 expanded and resident within the CNS, and drives chronic compartmentalised  
18 inflammation, representing a unified pathogenic B-T cell axis throughout the disease  
19 course.

## 20 Materials and methods

### 21 Consent and ethical approval

22 Study participants were recruited from an internal tissue bank of pwMS being initiated on  
23 ocrelizumab treatment within Oxford University Hospitals NHS Foundation Trust (OUH),  
24 Oxford, UK since 2019. Informed consent was obtained from all participants to collect  
25 blood samples for research purposes. The study was approved by the local clinical  
26 research ethics committee (NHS REC 21/SC/0353).

### 27 Study cohorts

#### 28 Discovery cohort

29 Study participants included in the discovery cohort were pwMS being initiated on  
30 ocrelizumab treatment with no prior history of disease modifying therapy (DMT). All MS  
31 diagnoses were confirmed in accordance with the 2017 revisions of the McDonald criteria  
32 by MS neurologists within OUH, with supporting MRI scans reviewed by neuroradiologists.

Patients with RRMS were classified as having high relapse frequency ( $RR_{high}$ ) if they met clinical criteria for rapidly evolving severe (RES) MS: 2 disabling relapses within a 12-month period with radiological evidence of disease activity on serial MRI scans.

## Treatment non-responder

We identified an additional treatment-naïve participant with RRMS and suboptimal response to ocrelizumab for inclusion in this study.

## Public datasets

Publicly available single-cell sequencing datasets were identified for validation of our findings from the discovery cohort: (i) a single-cell RNA sequencing (scRNA-seq) compendium of CSF and blood mononuclear cells from pwMS, other neurological diseases and healthy controls compiled from multiple studies<sup>25–29</sup> (Synapse accession syn51730532), (ii) scRNA-seq and T cell receptor (TCR) sequencing of paired CSF and blood mononuclear cells from pwMS and healthy controls<sup>30</sup> [Gene Expression Omnibus (GEO) accession GSE133028 and GSE291328], (iii) single nucleus RNA sequencing (snRNA-seq) data of post-mortem brain tissue of pwMS and healthy control white matter<sup>31,32</sup> [GEO accession GSE180759 and European Genome-phenome Archive (EGA) accession EGAC50000000231], and (iv) CITE-seq data of PBMCs from pwMS with longitudinal samples before and during treatment with natalizumab<sup>33</sup> (GEO accession GSE144744).

## Clinical data extraction

Demographic and clinical data were extracted for all participants from electronic health records.

## Sample processing

### Sample collection and storage

Blood samples were collected from all participants prior to commencing treatment (baseline), two weeks after the first ocrelizumab infusion (early) and immediately prior to each subsequent infusion (late). No pre-treatment sample was available for the treatment non-responder participant. The dosing schedule for ocrelizumab comprised an initiation phase of two 300 mg infusions given two weeks apart, followed by extended interval 600 mg maintenance infusions, with the timing of infusions determined on an individual basis by detectable CD19<sup>+</sup> B cell repopulation ( $\geq 10$  cells/ $\mu$ L) on monthly blood monitoring. PBMCs were isolated from fresh whole blood samples using density gradient centrifugation and cryopreserved in liquid nitrogen.

## Single-cell library preparation and sequencing

PBMCs were thawed at 37°C, counted, assessed for viability and processed using Chromium Next GEM Single Cell 3' Reagent Kit v3.1 (discovery cohort) or Chromium GEM-X Single Cell 3' Reagent Kit v4 (treatment non-responder) (10x Genomics). Cells from each sample were labelled with CITE-seq antibodies from TotalSeq-A Human Universal Cocktail v1.0 (BioLegend) or TotalSeq-B Human Universal Cocktail v1.0 (BioLegend) and corresponding TotalSeq-A or TotalSeq-B anti-human hashtag antibodies (BioLegend). Longitudinal samples from each patient were pooled and 16,000–32,000 cells were loaded per well of a Chromium microfluidic chip, using 2–3 wells per participant. Library preparation was performed according to the manufacturer instructions, with a modification in Step 2.2 for samples processed using the TotalSeq-A antibodies, as described previously.<sup>33</sup> cDNA library concentrations and sizes were measured by Qubit (Thermo Fisher Scientific) and TapeStation (Agilent Technologies), respectively.

Processed cDNA libraries for RNA, antibody-derived tags (ADTs) and hashtag oligonucleotides (HTOs) were pooled and sequenced across multiple flow cells using either the Illumina NextSeq 500 platform (in house) or NovaSeq X Plus platform (Novogene UK), using read configurations for each library type as recommended by 10x Genomics and a minimum sequencing depth of 20,000 reads per cell for RNA libraries, 10,000 reads per cell for ADT libraries and 2,000 reads per cell for HTO libraries.

## Data analysis

Preprocessing, quality control, normalisation, variable feature selection, dimensionality reduction and integration of data from CITE-seq experiments performed in this study was performed using a custom computational pipeline and in accordance with published single-cell best practices<sup>34</sup>.

We developed a novel analytical pipeline to robustly link treatment-related perturbation to cellular phenotype at single-cell resolution. In brief, manifold enhancement of latent dimensions (MELD)<sup>35</sup> was used to quantify treatment effect within each donor, and a donor-stratified permutation test was used to quantify uncertainty. These were aggregated across donors to derive a perturbation score metric per cell state, representing a statistically robust, denoised measure of consensus treatment effect. Consensus non-negative matrix factorisation (cNMF)<sup>36</sup> was used to infer gene programs representing overlapping cellular phenotypes. Random forest regression models were trained to predict perturbation score from gene program usages per cell state, and Shapley additive explanations (SHAP)<sup>37</sup> values were used to quantify feature importance. Using perturbation scores and SHAP values, we derived a program effect score metric for each gene program, representing an

importance-weighted measure of treatment-related perturbation in cells with high program activity.

Differential expression analysis for paired RNA and ADT data was performed using total variational inference (totalVI).<sup>38</sup>

Cell type annotations from the discovery dataset were transferred to validation datasets using Seurat mutual nearest neighbour-based reference mapping.<sup>39</sup> Precomputed cNMF-derived gene programs from the discovery dataset were projected to validation datasets using star cell annotator (starCAT).<sup>40</sup>

Details of data analysis steps outlined in this section, along with formal definitions for perturbation and program effect scores, are provided in Supplementary Methods.

## Statistical analysis

No statistical methods were used to predetermine the number of participants in the discovery cohort. Participants were selected based on meeting our clinical inclusion criteria and availability of samples.

Statistical tests for cell type marker gene identification, perturbation scores, gene set enrichment analysis for gene programs and differential expression analyses are described in Supplementary Methods.

All other statistical tests were performed using linear models in R. Comparisons of program scores across samples with repeated measures from the same biological replicates (donors) were performed using linear mixed effects models (as implemented in lme4)<sup>41</sup> with Satterthwaite estimates of degrees of freedom (as implemented in lmerTest).<sup>42</sup> All model formulas included fixed effects for age and sex, and mixed effects models included a random effect for donor. Marginal means, confidence intervals and *P*-values for contrasts of interest were computed using emmeans.<sup>43</sup>

Multiple testing correction was performed for false discovery rate (FDR) control using the Benjamini–Hochberg method, and adjusted *P*-values are reported in the text and figures. For statistical significance, a FDR threshold of 0.05 was selected for all hypothesis tests, except in tests of significantly perturbed cell states following treatment, where a more conservative threshold of  $10^{-5}$  was selected.

# Results

## Dynamic longitudinal changes in peripheral immune cells in response to ocrelizumab

To investigate the broader effects of anti-CD20 therapy on the peripheral immune system in MS, we performed CITE-seq on PBMC samples from a discovery cohort ( $n = 14$ ) of treatment-naïve pwMS commencing ocrelizumab (Fig. 1A–B). We collected samples prior to commencing treatment (baseline), two weeks after the first ocrelizumab infusion (early) and immediately prior to each subsequent infusion (late), which were timed to coincide with the earliest detectable CD19<sup>+</sup> B cell repopulation. For each individual, we obtained between 3–5 longitudinal samples during treatment. Participants were grouped into three categories: (i) RRMS with high relapse frequency (RR<sub>high</sub>,  $n = 5$ ), (ii) RRMS with low relapse frequency (RR<sub>low</sub>,  $n = 4$ ), and (iii) PPMS ( $n = 5$ ). Clinical and demographic characteristics of all participants are shown in Supplementary Table 1.

We generated an integrated dataset of 264,512 cells comprising 5 main (level 1), 20 coarse (level 2) and 34 fine (level 3) cell types (Fig. 1C, Supplementary Figure 1A–C and Supplementary Table 2). Cell type annotations were validated by manual inspection of marker genes (Supplementary Data 1 and Supplementary Data 2) and surface proteins (Fig. 1D).

To quantify treatment effects, we used MELD<sup>35</sup> to compute perturbation scores for each transcriptional cell state (Fig. 2A), whereby negative scores represent depletion and positive scores represent enrichment relative to the donor-specific pre-treatment baseline (Fig. 1E–G). These scores represent a statistically robust, denoised measure of consensus treatment effect across donors per cell state, accounting for both biological and technical sources of noise.

As expected, we observed the strongest depletion in naïve and memory B cells at the early post-treatment time point, with a markedly weaker effect in plasma cells (including plasmablasts), which are known to downregulate surface CD20 expression. At the late post-treatment time points, the majority of the B cell compartment remained strongly depleted, except for notable enrichment of a subset of naïve B cells and unchanged plasma cell abundance relative to baseline.

The strongest treatment effect outside the B cell compartment was an enrichment in monocyte populations (particularly CD14<sup>+</sup> monocytes) at both the early and late post-treatment time points. We also observed clear shifts in the T cell compartment post-treatment, particularly at the late time points, with depletion of several memory and

unconventional populations (within CD4<sup>+</sup>, CD8<sup>+</sup>,  $\gamma\delta$  and MAIT cells) and enrichment of naive populations (especially within CD4<sup>+</sup> cells).

Taken together, these findings demonstrate broad, temporally dynamic, cell type-specific shifts in the peripheral immune system following ocrelizumab treatment, with important implications for the possible mechanisms underlying its therapeutic benefit in pwMS.

## Early repopulating B cells are immature and antigen-inexperienced

To gain a more granular understanding of the dynamics of B cell repopulation during extended interval treatment with ocrelizumab, we then performed an unbiased characterisation of the enriched subset of naive B cells observed at the late post-treatment time points (Fig. 3A–B).

We used cNMF<sup>36</sup> to infer cell-type specific gene programs, with cells assigned continuous program scores reflecting the contribution of each program to its transcriptional profile (Fig. 2B). We then correlated gene programs with treatment effects by training machine learning models to predict perturbation scores based on program scores. Using these models, we derived program effect scores, which quantify the strength and direction of association of each gene program with the consensus treatment effect across donors, allowing us to measure subtle but consistent shifts in cellular phenotypes after ocrelizumab (Fig. 2C).

We identified 14 gene programs within the B cell compartment, which we annotated based on a combination of their top-weighted genes, expression pattern across fine (level 3) cell types and gene set enrichment analysis (Supplementary Figure 2, Supplementary Figure 3, Supplementary Data 3 and Supplementary Data 4). We excluded housekeeping programs and those capturing low-level contaminating transcripts from other cell types.

We identified the *Transitional Naive* program as solely enriched during early B cell repopulation (Fig. 3C). This transcriptional program was expressed in a subset of naive B cells (Supplementary Figure 2 and Supplementary Figure 3) and the top-weighted genes were enriched for known markers of B cell immaturity (*MME*, *CD9*, *CD38*, *SOX4*, *EBF1*, *TCL1A*, *VPREB1*, *VPREB3*), supporting their identity as recently differentiated, antigen-inexperienced B cells migrating out of the bone marrow into the circulation. We confirmed enrichment of this cell state during B cell repopulation at the individual level (Fig. 3D) by comparing mean program usage per donor in naive B cells between the late post-treatment and pre-treatment time points (marginal mean difference 0.598, 95% CI [0.534, 0.661],  $P = 3.77 \times 10^{-10}$ ).

Of the remaining programs, those most strongly associated with persistent depletion at the late post-treatment time points were linked to mature memory B cell (*Homeostatic Memory, Primed Memory, Atypical Memory*) and naive B cell (*Mature Naive*) states (Supplementary Figure 2 and Supplementary Figure 3). By contrast, the *Early ASC*, *IgA<sup>+</sup> ASC* and *IgG<sup>+</sup> ASC* programs, expressed exclusively in plasma cells (Supplementary Figure 2 and Supplementary Figure 3), showed the weakest program effect scores, indicating that these cell populations are relatively spared from depletion by anti-CD20 therapies.

Our findings demonstrate a reprogramming of the B cell compartment during extended interval ocrelizumab treatment, with the early repopulating B cells between infusions harbouring an immature, antigen-inexperienced phenotype, persistent depletion of memory B cells and preservation of antibody-secreting cells (ASCs).

## **Delayed depletion of a trafficking, activated, residency-primed, granzyme K-effector T cell (TARGET) phenotype**

Next, we aimed to characterise the observed phenotypic shifts in the T cell compartment during ocrelizumab treatment (Fig. 4A–B), based on our hypothesis that B cell depletion would result in reduced activation of pathogenic T cells in the disease context.

Using cNMF, we identified and annotated 36 T cell-specific gene programs (Supplementary Figure 4, Supplementary Figure 5, Supplementary Data 3 and Supplementary Data 4). We again excluded housekeeping programs and those capturing low-level contaminating transcripts from other cell types.

We identified three gene programs, expressed predominantly in naive T cells (Supplementary Figure 4 and Supplementary Figure 5), that were strongly enriched post-treatment (Fig. 4C). The first, *Survival Signalling*, had top marker genes enriched for pro-survival and growth signalling pathways (*BCL2*, *PIM1*, *PDK1*, *IGF1R*). The second, *Naive CD4*, featured canonical marker genes of naive T cells (*IL7R*, *SELL*, *CCR7*, *TCF7*, *LEF1*, *BACH2*, *FOXO1*). The top genes for the third program, *Precursor Memory*, included genes encoding TCR co-stimulatory molecules (*CD28*, *ICOS*) and early activation markers (*CD38*, *IRF4*), suggestive of transcriptional priming for activation and naive-to-memory differentiation. Together, these programs suggest enrichment of antigen-inexperienced T cell states exhibiting tonic survival and growth signalling after ocrelizumab, consistent with reduced B cell-dependent T cell activation.

Conversely, we identified three gene programs, expressed predominantly in memory T cells (Supplementary Figure 4 and Supplementary Figure 5), that were strongly depleted post-treatment (Fig. 4C). The largest effect was observed for the *Cytotoxic Inflammatory*

program, with top marker genes indicating a distinct cytotoxic effector polarisation (*CD8A*, *EOMES*, *GZMK*, *GZMA*, *GZMM*, *GZMH*, *CST7*, *NKG7*) featuring high expression of the non-canonical granzymes K/A/M, which collectively have reduced direct cytolytic function and greater intercellular proinflammatory signalling capacity compared to granzyme B (*GZMB*) and perforin (*PRF1*). This program additionally featured genes involved in inflammatory cytokine/chemokine signalling (*IFNG*, *IL32*, *CCL3*, *CCL4*, *CCL5*, *CXCR4*) and MHC class II-related genes (*HLA-DRB1*, *HLA-DRA*, *HLA-DRB5*, *HLA-DQA1*, *HLA-DQB1*, *HLA-DPA1*, *HLA-DPB1*, *HLA-DMA*, *CD74*), which are upregulated in activated interferon- $\gamma$ -responsive T cells. The second program was annotated *Th1-like*, based on several genes associated with  $T_H1$  polarisation (*STAT4*, *JAK2*, *CXCR3*, *CD226*). However, the top marker genes of this program were additionally enriched for genes supporting adhesion/migration, including *ITGA4* (integrin  $\alpha 4$ ) and *ITGA1* (integrin  $\alpha 1$ ), along with genes linked to tissue residency, such as *PRDM1* (BLIMP-1) and *ZNF683* (HOBIT), suggesting a broader phenotype with increased capacity for endothelial transmigration into the CNS and differentiation into tissue-resident cells. The third program was an *Activation Signalling* program, featuring genes associated with signalling pathways downstream of TCR engagement, including AP-1 family genes (*FOS*, *FOSB*, *JUN*, *JUNB*), other immediate early genes (*IER2*, *EGR1*, *EGR3*, *NR4A1*, *NR4A3*), NF- $\kappa$ B pathway genes (*NFKB1*, *NFKBIA*, *NFKBIZ*, *TNFAIP3*) and T cell activation markers (*CD69*, *CD83*, *TNF*).

To delineate the overall T cell phenotype most strongly depleted following anti-CD20 therapy, we derived a composite score by summing the individual program scores for these three programs in each cell (Supplementary Figure 6A) and confirmed that this score was selectively enriched in cell states with negative perturbation scores (Supplementary Figure 6B). We hereafter describe this composite score as the trafficking, activated, residency-primed, granzyme K-effector T cell (TARGET) transcriptional signature, reflecting the inferred functional polarisation captured by the component gene programs.

To further characterise the TARGET phenotype, we analysed which cell types were enriched among cells with high scores ( $> 0.2$ ) for this program (Supplementary Figure 6C). Cells highly expressing the TARGET signature were predominantly CD8<sup>+</sup> or CD4<sup>+</sup> memory cells (81%, CD4:CD8 ratio 0.45). We therefore performed protein-level differential expression analysis to identify the cell surface markers associated with this transcriptional signature within memory T cells (Fig. 4E and Supplementary Data 5). This confirmed high expression of markers of innate-like cytotoxic effector function (*CD8*, *CLEC15A*, *CD56*, *CD57*, *NKp80*, *CD94*, *NKG2D*, *KIR2DL3*, *KIR3DL1*) and chronic activation (*IL-2R $\beta$* , *IFNGR1*, *CD83*, *CD38*, *HLA-DR/DP/DQ*, *PD-1*, *TIGIT*, *LAG-3*), as well as molecules linked to chemotaxis (*CCR5*, *CXCR3*, *CX3CR1*), adhesion (integrin  $\alpha 4$ , integrin  $\beta 7$ , integrin  $\alpha L$ , *ICAM-1*, *PECAM-1*) and tissue retention (integrin  $\alpha E$ , integrin  $\alpha 1$ , *CD101*, *CD73*, *GPR56*). Together, these findings

support the interpretation that the TARGET signature defines a subpopulation of activated circulating cytotoxic effector memory T cells that are primed to migrate into tissues and differentiate into tissue-resident memory cells.

As recent studies have identified a population of CD20<sup>dim</sup> T cells that are also rapidly depleted along with B cells by anti-CD20 therapies,<sup>44–47</sup> we sought to determine whether the post-treatment depletion of cells with high TARGET signature was attributable to low-level CD20 expression on these cells. We used protein-level CD3 and CD20 expression combined with the RNA-based cell type annotations to gate CD20<sup>dim</sup> T cells (Supplementary Figure 7A). We found that although memory T cells with a high TARGET program score were more likely to be CD20<sup>dim</sup> (odds ratio 1.3, Fisher exact test  $P = 0.0278$ ), the CD20<sup>dim</sup> subpopulation only accounted for a minor proportion (1.23%) of these cells (Supplementary Figure 7B). Moreover, direct comparison of CD20 protein and gene expression to B cell subsets revealed their expression in TARGET high cells was substantially lower than that seen in plasma cells (Supplementary Figure 7C–D), which are notably spared from the direct effects of anti-CD20 therapies.

To further investigate the relevance of low-level CD20 expression in these cells on the observed treatment effect, we assessed the temporal dynamics of TARGET signature depletion following ocrelizumab. We hypothesised that if these cells were directly targeted by ocrelizumab, they would be rapidly depleted following the first dose (at the early post-treatment time point), mirroring the time course of B cell depletion. We compared longitudinal mean program usage per donor in memory T cells (Fig. 4F) and found the TARGET signature was not significantly depleted at the early post-treatment time point compared to baseline (marginal mean difference  $-0.0168$ , 95% CI  $[-0.0417, 0.0081]$ ,  $P = 0.176$ ). However, it was significantly depleted at the late post-treatment time points compared to both pre-treatment (marginal mean difference  $-0.0616$ , 95% CI  $[-0.0865, -0.0367]$ ,  $P = 1.16 \times 10^{-4}$ ) and early post-treatment (marginal mean difference  $-0.0448$ , 95% CI  $[-0.0697, -0.0199]$ ,  $P = 1.75 \times 10^{-3}$ ) time points.

Collectively, these results suggest that B cell depletion with ocrelizumab is associated with secondary depletion of a distinct chronically activated, proinflammatory–cytotoxic and tissue-homing memory T cell phenotype. Although a minor subset of these cells exhibits low-level CD20 expression and may be directly targeted by the drug, the delayed time course of the shift in the broader cell state following treatment is most consistent with an indirect mechanism, possibly reflecting loss of B cell-dependent antigenic stimulation or cytokine support. Additionally, this finding is consistent with the clinical observation of delayed downregulation of inflammation in patients treated with anti-CD20 therapies, which was hitherto poorly understood.

## TARGET signature correlates with clinical disease activity

Given cells highly expressing the TARGET signature are depleted following treatment with ocrelizumab, we hypothesised that they may have a pathogenic role in the untreated disease context. We therefore next explored whether their pre-treatment abundance was linked to clinical disease severity. We compared mean pre-treatment TARGET program score within memory T cells between donors in different disease groups (Fig. 4F) and found higher TARGET signature in the  $RR_{high}$  group compared to the  $RR_{low}$  group (marginal mean difference 0.0375, 95% CI [0.0103, 0.0646],  $P = 0.031$ ), indicating that patients with more clinically aggressive MS have an enrichment of this cell phenotype within their circulating memory T cell repertoire.

Longitudinal samples from pwMS with breakthrough disease activity during ocrelizumab treatment are rare due to its high efficacy in RRMS.<sup>48</sup> However, we identified a treatment-naive pwMS with poor therapeutic response despite depletion of CD19<sup>+</sup> B cells below the lower limit of detection, experiencing breakthrough clinical disease activity and accumulation of disability (EDSS 2.0 to 6.0) over two years. As an additional exploratory analysis, we performed CITE-seq on longitudinal PBMC samples from this patient to determine if there was a measurable difference in the pattern of treatment-related shift in the TARGET signature compared to the discovery cohort of patients exhibiting good therapeutic response. We generated an integrated dataset comprising 24,195 cells (Supplementary Figure 8A). We used starCAT<sup>40</sup> to project the T cell-specific gene programs from our discovery dataset onto the T cells from this additional donor, allowing us to compute TARGET program scores in these cells. Intriguingly, we found an absence of the longitudinal trend of decreasing TARGET signature enrichment within memory T cells at the late post-treatment time points in this donor (Supplementary Figure 8B). This may suggest a link between their poor treatment response and either inadequate suppression of this cell state or alternative immunological mechanisms playing a stronger role in driving their disease activity.

Together, these findings support pathogenicity of the TARGET program. Cells with this transcriptional signature are enriched in the circulation of pwMS with more aggressive relapsing disease, suggesting an expanded circulating reservoir of proinflammatory T cells primed to migrate into the CNS. Concordantly, we identified a failure in downregulation of this cell state in the circulation in an individual with poor therapeutic response to ocrelizumab; although this preliminary finding requires validation in a larger cohort of pwMS with anti-CD20 treatment failure, it is consistent with the hypothesis that the treatment efficacy of ocrelizumab may be dependent on dampening this T cell phenotype.

## **TARGET signature enriched in clonally expanded memory T cells in CSF and chronic lesion rims of pwMS**

We next sought to test our hypothesis that TARGET high cells in the blood represent a circulating reservoir of cells with high pathogenic potential which migrate to the CNS and contribute to MS pathology.

First, we investigated the CNS-homing capability of circulating TARGET high cells using a published compendium of scRNA-seq studies of mononuclear cells from both peripheral blood and CSF across neurological diseases<sup>25–29</sup> (Fig. 5A). The final integrated dataset contained 178,027 cells from 30 treatment-naïve individuals (Supplementary Figure 9), with either established MS or its prodromal stage clinically isolated syndrome (CIS) ( $n = 21$ , 18 with paired PBMC/CSF) or healthy control donors ( $n = 9$ , all with paired PBMC/CSF). We identified significant enrichment of the TARGET signature within memory T cells in the CSF (Fig. 5B) in MS/CIS (marginal mean difference 0.0438, 95% CI [0.0216, 0.0661],  $P = 7.65 \times 10^{-4}$ ) but not healthy control (marginal mean difference 0.0012, 95% CI [−0.0310, 0.0784],  $P = 0.938$ ) donors. In addition, the magnitude of the inter-compartment difference, calculated as the CSF:PBMC log ratio of mean TARGET program score (Fig. 5C), was greater in patients with MS/CIS compared to healthy control donors (marginal mean difference 0.578, 95% CI [0.138, 1.02],  $P = 0.012$ ). Taken together, these findings strongly support the hypothesis that cells with a TARGET phenotype in the circulation have a high propensity to cross the blood–brain barrier (BBB). Moreover, this effect appears to be more pronounced in pwMS, possibly reflecting a greater chemotactic drive recruiting these cells to the CNS in these individuals.

A recent study using paired scRNA-seq and TCR sequencing in blood and CSF samples from pwMS and healthy control donors (Fig. 5D) demonstrated CSF-enriched clonal expansion of predominantly CD8<sup>+</sup> memory T cells, a subset of which were restricted to EBV antigens.<sup>30</sup> As these cells were shown to express chronic activation, cytotoxicity, trafficking and tissue residency markers, which overlap with the top-weighted genes of programs in the TARGET signature, we hypothesised that circulating cells with high expression of the TARGET signature may undergo clonal expansion following migration across the BBB.

To test this hypothesis, we compared TARGET program expression between expanded clonotype groups within memory T cells from this publicly available dataset. The final integrated dataset contained 108,306 cells from 15 treatment-naïve individuals (Supplementary Figure 10), with either MS/CIS ( $n = 12$ ) or healthy control donors ( $n = 3$ ). We grouped memory T cells with available TCR sequencing data into clonotypes which were disproportionately highly expanded in the CSF compared to the blood (CSF-enriched) and

those which were not (CSF-unenriched), as previously described.<sup>30</sup> We found that the TARGET signature was significantly higher in the CSF-enriched clonotypes found in both blood (marginal mean difference 0.204, 95% CI [0.147, 0.261],  $P = 7.91 \times 10^{-8}$ ) and CSF (marginal mean difference 0.150, 95% CI [0.097, 0.202],  $P = 3.43 \times 10^{-6}$ ) in MS/CIS (Fig. 5E). Specifically, the enrichment of this signature in CSF-enriched clonotypes across compartments strongly supports its interpretation as a peripherally induced state that is primed for trafficking across the BBB, with subsequent cognate antigen re-encounter driving compartmentalised clonal expansion within the CNS.

The most abundant adaptive immune cells identified in MS lesions are tissue-resident memory T ( $T_{rm}$ ) cells, which are predominantly CD8<sup>+</sup>,<sup>49</sup> and are thought to arise from circulating precursors, which fully differentiate upon encountering their cognate antigen within tissues. In pwMS,  $T_{rm}$  cells are present in the rims of chronic inflammatory lesions, suggesting a maladaptive immune response to local antigens, and may be responsible for driving persistent smouldering inflammation. Given the presence of tissue retention markers in cells with high TARGET signature in the blood, we hypothesised that this signature may represent a circulating precursor state of pathogenic dysregulated  $T_{rm}$  cells that arise in chronic MS.

To test this hypothesis, we identified two published datasets<sup>31,32</sup> of snRNA-seq from post-mortem brain tissue of pwMS with annotated samples from different intra-lesional niches, including demyelinated core, inflamed rim of both mixed active/inactive (chronic active) and inactive (chronic inactive) lesions and perilesional NAWM; both datasets also included white matter samples from healthy control donors with no known neurological disease (Fig. 5F). The final integrated dataset contained 1,417 cells from 18 individuals (Supplementary Figure 11), either SPMS ( $n = 12$ ) or healthy control donors ( $n = 6$ ). Due to the relatively low abundance of T cells in brain tissue samples, several donor-tissue type combinations had fewer than 10 cells detected, making donor-level mean aggregation of program scores unreliable. Therefore, we compared cell-level scores across tissue types using a mixed effects model with donor as a random effect to mitigate pseudoreplication. Overall, we detected higher TARGET signature in T cells from MS lesion rims and perilesional NAWM compared to both MS lesion core and white matter from healthy control donors (Fig. 5G). Statistically significant differences were detected in pairwise comparisons between lesion core and chronic active lesion rims (marginal mean difference 0.0739, 95% CI [0.0263, 0.122],  $P = 0.0241$ ), chronic inactive lesion rims (marginal mean difference 0.0663, 95% CI [0.0153, 0.117],  $P = 0.0341$ ) and perilesional NAWM (marginal mean difference 0.0704, 95% CI [0.0189, 0.122],  $P = 0.0341$ ), as well as between healthy white matter and chronic active lesion rims (marginal mean difference 0.0616, 95% CI [0.0134, 0.110],  $P = 0.0341$ ).

These validation analyses support the disease relevance of the TARGET signature from multiple perspectives. Overall, our findings suggest that circulating memory T cells with this phenotype preferentially migrate into the CNS and undergo compartmentalised clonal expansion in pwMS. Crucially, we demonstrate that the same transcriptional signature is enriched in the blood and CSF of patients with relapsing MS, as well as in chronically inflamed brain lesions from individuals with end-stage progressive disease, supporting a common axis of immune dysregulation throughout the disease course.

## TARGET signature-expressing cells are trapped in the periphery by natalizumab

Whereas most DMTs for MS, including ocrelizumab, work by depleting or functionally modulating circulating immune cells, some DMTs confer therapeutic benefit by blocking the migration of immune cells into the CNS. Within this class of DMTs, natalizumab is the drug with the highest clinical efficacy, comparable to ocrelizumab.<sup>50</sup> Natalizumab is a monoclonal antibody against integrin  $\alpha 4$ , which forms part of a heterodimeric complex with integrin  $\beta 1$  to form the cellular adhesion molecule VLA-4, which plays a critical role in lymphocyte transmigration across the BBB; natalizumab inhibits the binding of this complex to VCAM-1 on vascular endothelial cells to prevent migration into the CNS.

Given our findings supporting enhanced CNS migration of TARGET high cells and their expression of integrin  $\alpha 4$  (*ITGA4*), we sought to investigate whether these cells were enriched in the circulation of pwMS under natalizumab treatment, reflecting reduced ability to enter the CNS. We used a large published CITE-seq dataset of PBMC from pwMS<sup>33</sup> with longitudinal samples taken before and during treatment with natalizumab (Fig. 6A). The final integrated dataset contained 56,252 cells from nine pwMS with pre- and post-natalizumab samples (Supplementary Figure 12). We subsequently computed perturbation scores and TARGET program scores per cell. We noted that within the T cell compartment, there was enrichment of a broad range of CD4<sup>+</sup> and CD8<sup>+</sup> memory T cell states in the circulation following natalizumab treatment, consistent with impairment of migration into target tissues and secondary lymphoid organs (Fig. 6B–C). In particular, cytotoxic T cell states with higher perturbation scores were also enriched for cells with higher TARGET program scores. We compared the mean score per donor between the pre- and post-treatment time points (Fig. 6D) and found a significant enrichment of the TARGET signature in circulating cytotoxic memory T cells at the post-treatment time point (marginal mean difference 0.0358, 95% CI [0.0159, 0.0556],  $P = 3.19 \times 10^{-3}$ ), consistent with the prior hypothesis of reduced migration of these cells into the CNS during treatment.

The countereffect of ocrelizumab and natalizumab on the abundance of circulating TARGET phenotype cells is a key observation which strongly supports the peripheral origin of the pathogenic adaptive immune response in MS. Furthermore, this finding serves as valuable reverse translational triangulation providing novel insight into a potential common mode of therapeutic benefit across two highly effective DMTs with differing mechanisms of action.

## Discussion

Depletion of B cells in pwMS is highly efficacious at reducing disease burden, indicating that these cells or the cells that they interact with are key pathogenic drivers. Using a reverse translational study design combined with a novel analytical strategy to robustly correlate treatment effects with continuous transcriptional states across broad immune cell types, we discovered a B cell-dependent T cell phenotype (TARGET) with characteristics of chronic activation, proinflammatory–cytotoxic polarisation and tissue homing. Across multiple independent and orthogonal single-cell datasets, we provide convergent evidence to support that this cellular phenotype is enriched in the blood in highly active RRMS, migrates to the CNS with subsequent compartmentalised clonal expansion, and is enriched in chronically inflamed brain lesions in SPMS. The conflation of these findings supports a peripheral origin of this cell state in pwMS, followed by CNS compartmentalisation, and suggests a common pathogenic axis throughout the disease course.

We found a predominance of transitional naive cells within the repopulating B cell repertoire during extended interval ocrelizumab treatment, with persistent depletion of mature naive and memory B cells. Transitional naive B cells are immature and antigen-inexperienced, lacking the functional capacity to activate pathogenic T cells.<sup>11</sup> Our findings are supported by those of previous studies,<sup>51–53</sup> and this alignment serves as an important validation of the sensitivity and robustness of our novel analysis strategy.

Our study defines TARGET, a transcriptionally defined cell state seen predominantly in CD8<sup>+</sup> memory T cells exhibiting the strongest depletion within the T cell compartment following treatment with ocrelizumab. We further demonstrate that the temporal trajectory of this perturbation is consistent with a secondary effect due to loss of B cells; importantly, our findings cannot be solely accounted for by a direct drug effect on a minor subset of CD20<sup>dim</sup> T cells which may have acquired low-level CD20 surface expression via trogocytosis during recent B cell contact.<sup>54</sup> However, the precise mechanisms underlying the contraction of this cell state in the circulation following B cell depletion remains

unresolved; plausible options include apoptosis by neglect of B cell-dependent survival signals, or a reduced polarisation of cells into this state due to loss of B cell engagement.

Surface protein markers associated with the TARGET transcriptional signature in our study show strong overlap with markers delineating depleted CD8<sup>+</sup> T cell subpopulations following anti-CD20 therapy in prior work,<sup>19–22</sup> strengthening the evidence that these findings represent a convergent cell phenotype. To our knowledge, only two prior studies<sup>23,24</sup> have applied high-throughput single-cell sequencing-based approaches to profile immunologic changes after B cell depletion; these studies primarily identified changes in proinflammatory myeloid populations, along with memory and regulatory CD4<sup>+</sup> T cell subsets.

Using a published dataset of paired scRNA-seq and TCR sequencing from blood and CSF mononuclear cells, we demonstrate higher expression of the TARGET signature in memory T cell clonotypes that are disproportionately expanded in the CSF compared to the blood in pwMS. Importantly, high-throughput *in silico* and *in vitro* antigen specificity screening of the CSF-enriched CD8<sup>+</sup> TCRs revealed three clonotypes with confirmed donor-specific MHC-restricted EBV specificity in pwMS, but none specific to CNS autoantigens or other herpesviruses.<sup>30</sup> Prior studies have shown preferential reduction in circulating EBV-specific CD8<sup>+</sup> T cells following ocrelizumab<sup>22,55</sup> and enrichment during treatment with natalizumab,<sup>56</sup> mirroring shifts in the TARGET signature observed in our study under the same therapies. Combined, these findings strongly suggest that at least some TARGET phenotype cells in pwMS arise in response to chronic latent EBV infection in memory B cells.

Our study links the peripherally originating TARGET signature to CNS-resident T<sub>rm</sub> cells in MS by demonstrating that memory T cells extracted from snRNA-seq of post-mortem MS brain tissue are enriched for the same signature, supporting our hypothesis that it represents a circulating precursor state, although our data lacks direct evidence of shared clonality between these cells in blood and brain tissue within the same donor. However, this hypothesis is further strengthened by prior studies that have demonstrated the existence of a subset of circulating memory T cells skewed towards the T<sub>rm</sub> fate by co-expression of transcription factors HOBIT (*ZNF683*) and BLIMP-1 (*PRDM1*),<sup>57,58</sup> both of which feature highly in TARGET; these circulating T cells with markers of tissue residency priming represent a previously unrecognised link anchoring recent advances in T<sub>rm</sub> biology to MS pathogenesis. Recently, high-resolution spatial transcriptomics has delineated a pathogenic interferon- $\gamma$ -mediated signalling axis between CD8<sup>+</sup> T<sub>rm</sub> cells and dysfunctional inflammatory microglia within chronic active lesion rims,<sup>59</sup> highlighting a key role played by these cells in driving smouldering, compartmentalised neuroinflammation in MS.

Collectively, our study provides a conceptual framework for unifying findings across different tissue compartments and disease stages to build a compelling model of a pathogenic B–T cell axis that may be central to MS. The observation that EBV-responsive memory T cells are strongly downregulated following B cell depletion in pwMS suggest that memory B cells may act as a latent EBV reservoir driving persistent stimulation of these cells. Using therapeutic B cell depletion as a probe, our study sought to delineate the transcriptional signature of these cells, highlighting a chronically activated state, production of proinflammatory cytokines and non-canonical cytotoxic effector molecules such as granzyme K, as well as priming for migration into tissues and differentiation into resident memory cells. Subsequently, proinflammatory  $T_{rm}$  cells within the CNS may contribute to smouldering MS pathology, both directly through complement activation, and through chemokine- and cytokine-mediated recruitment and activation of both CNS-intrinsic and circulating immune cells.<sup>59,60</sup> Importantly, granzyme K-expressing, chronically activated  $CD8^+$  T cells have recently been linked to promoting inflammation in diverse contexts including chronic viral infections, autoimmune diseases, cancer, and advancing age.<sup>61,62</sup> The precise antigenic stimulus which may promote differentiation of circulating TARGET precursor states into  $T_{rm}$  cells remains uncertain; plausible options include re-encounter with EBV-infected memory B cells within the CNS compartment,<sup>17,18</sup> cross-reactivity of EBV peptides with CNS-derived antigens<sup>63</sup> or a combination of both mechanisms. Increasing compartmentalisation of these cells over time may account for the reduced efficacy of peripherally acting DMTs in patients with late-stage progressive disease. Crucially, this framework provides a plausible mechanistic rationale for the clinical consensus that early and aggressive immunomodulatory therapy is key to reducing long-term disability progression in MS.

Although our discovery cohort comprised a relatively small number of patients, this is mitigated by the orthogonal validation of our findings in independent and larger cohorts of pwMS and healthy control donors. By using data from real-world cohorts of pwMS, we sought to overcome several limitations of animal models and *in vitro* studies of pathogenic mechanisms underlying MS; specifically, these may not recapitulate the mechanisms driving human disease, nor capture the complexity of immune interactions in the disease–therapy context. Our approach enabled the unbiased identification of disease-associated cell states and generation of testable mechanistic hypotheses grounded in true disease biology.

In conclusion, we discover a peripherally originating, CNS-homing, B cell-dependent pathogenic T cell phenotype in pwMS, which reframes the mechanism of action of B cell depletion therapies and links peripheral immune activation to compartmentalised CNS inflammation. Our data is consistent with an emerging view of the disease as an aberrant

immune response against chronic latent EBV infection in memory B cells in genetically predisposed individuals.<sup>64,65</sup> Crucially, our data supports a model of dysfunctional immune interactions driving a modifiable pathogenic T cell state in pwMS, advocating a strategic shift away from broad and indiscriminate cell type ablating therapies.<sup>6</sup> Early and specific targeting of the molecular pathways enriched within these circulating T cells, in combination with CNS-penetrant therapies targeting their tissue-resident successors, may present key opportunities for developing more effective and durable therapies for pwMS.

## Data availability

Raw sequencing data (FASTQ files) generated in this study cannot be made publicly available to protect the privacy of human participants. These data are currently being deposited in EGA and will be available under controlled access for research use. Processed data (count matrices and associated metadata) necessary to reproduce the analysis have been deposited in GEO under accession GSE316688. All code and software environment specifications required to reproduce the analysis are publicly available on GitHub ([https://github.com/redwanfarooq/ocrelizumab\\_paper](https://github.com/redwanfarooq/ocrelizumab_paper)). Key software packages and versions used are listed in Supplementary Table 3.

## Acknowledgements

We are grateful to all participants who provided samples for this study and medical personnel within the MS Clinical Trials Unit (University of Oxford/OUH) for their assistance with recruiting and obtaining consent from participants, as well as collecting blood samples.

## Funding

This work was supported by funding from Aims2Cure (HMD00630), OAK Foundation (OFIL-22-254) and Medical Research Council (UK Research and Innovation) (MC\_UU\_00036/3). R.F. was supported by a Clinical Research Training Fellowship funded by the Association of British Neurologists and Guarantors of Brain (HMR04540).

## Competing interests

F.J.T. holds equity interests in RN.AI Therapeutics, Dermagnostix, and Cellarity, serves as a member of the Scientific Advisory Board of CytoReason, and Valinor Industries, and

receives consulting fees from Immunai, CytoReason, Genbio, Bioturing and Valinor Industries. All other authors declare no competing interests.

## Supplementary material

Supplementary material is available at *Brain* online.

## References

1. Jakimovski D, Bittner S, Zivadinov R, et al. Multiple sclerosis. *The Lancet*. 2024;403(10422):183-202. doi:10.1016/S0140-6736(23)01473-3
2. Walton C, King R, Rechtman L, et al. Rising prevalence of multiple sclerosis worldwide: Insights from the Atlas of MS, third edition. *Multiple Sclerosis Journal*. 2020;26(14):1816-1821. doi:10.1177/1352458520970841
3. Ganjgahi H, Häring DA, Aarden P, et al. AI-driven reclassification of multiple sclerosis progression. *Nature Medicine*. 2025;31(10):3414-3424. doi:10.1038/s41591-025-03901-6
4. Frischer JM, Bramow S, Dal-Bianco A, et al. The relation between inflammation and neurodegeneration in multiple sclerosis brains. *Brain*. 2009;132(5):1175-1189. doi:10.1093/brain/awp070
5. Lassmann H. Pathogenic Mechanisms Associated With Different Clinical Courses of Multiple Sclerosis. *Frontiers in Immunology*. 2019;9:3116. doi:10.3389/fimmu.2018.03116
6. Attfield KE, Jensen LT, Kaufmann M, Friese MA, Fugger L. The immunology of multiple sclerosis. *Nat Rev Immunol*. 2022;22(12):734-750. doi:10.1038/s41577-022-00718-z
7. Handel AE, Williamson AJ, Disanto G, Handunnetthi L, Giovannoni G, Ramagopalan SV. An Updated Meta-Analysis of Risk of Multiple Sclerosis following Infectious Mononucleosis. *PLOS ONE*. 2010;5(9):e12496. doi:10.1371/journal.pone.0012496
8. Jacobs BM, Giovannoni G, Cuzick J, Dobson R. Systematic review and meta-analysis of the association between Epstein–Barr virus, multiple sclerosis and other risk factors. *Multiple Sclerosis Journal*. 2020;26(11):1281-1297. doi:10.1177/1352458520907901
9. Bjornevik K, Cortese M, Healy BC, et al. Longitudinal analysis reveals high prevalence of Epstein-Barr virus associated with multiple sclerosis. *Science*. 2022;375(6578):296-301. doi:10.1126/science.abj8222

- 1 10. de Sèze J, Maillart E, Gueguen A, et al. Anti-CD20 therapies in multiple sclerosis:  
2 From pathology to the clinic. *Frontiers in Immunology*. 2023;Volume 14 - 2023.
- 3 11. Cencioni MT, Mattosio M, Magliozzi R, Bar-Or A, Muraro PA. B cells in multiple  
4 sclerosis — from targeted depletion to immune reconstitution therapies. *Nature Reviews*  
5 *Neurology*. 2021;17(7):399-414. doi:10.1038/s41582-021-00498-5
- 6 12. Höftberger R, Lassmann H, Berger T, Reindl M. Pathogenic autoantibodies in  
7 multiple sclerosis — from a simple idea to a complex concept. *Nature Reviews Neurology*.  
8 2022;18(11):681-688. doi:10.1038/s41582-022-00700-2
- 9 13. Babbe H, Roers A, Waisman A, et al. Clonal Expansions of Cd8+ T Cells Dominate  
10 the T Cell Infiltrate in Active Multiple Sclerosis Lesions as Shown by Micromanipulation and  
11 Single Cell Polymerase Chain Reaction. *Journal of Experimental Medicine*.  
12 2000;192(3):393-404. doi:10.1084/jem.192.3.393
- 13 14. van Nierop GP, van Luijn MM, Michels SS, et al. Phenotypic and functional  
14 characterization of T cells in white matter lesions of multiple sclerosis patients. *Acta*  
15 *Neuropathologica*. 2017;134(3):383-401. doi:10.1007/s00401-017-1744-4
- 16 15. Machado-Santos J, Saji E, Tröscher AR, et al. The compartmentalized inflammatory  
17 response in the multiple sclerosis brain is composed of tissue-resident CD8+ T  
18 lymphocytes and B cells. *Brain*. 2018;141(7):2066-2082. doi:10.1093/brain/awy151
- 19 16. Fransen NL, Hsiao CC, van der Poel M, et al. Tissue-resident memory T cells invade  
20 the brain parenchyma in multiple sclerosis white matter lesions. *Brain*. 2020;143(6):1714-  
21 1730. doi:10.1093/brain/awaa117
- 22 17. Serafini B, Rosicarelli B, Veroni C, Mazzola GA, Aloisi F. Epstein-Barr Virus-Specific  
23 CD8 T Cells Selectively Infiltrate the Brain in Multiple Sclerosis and Interact Locally with  
24 Virus-Infected Cells: Clue for a Virus-Driven Immunopathological Mechanism. *Journal of*  
25 *Virology*. 2019;93(24):10.1128/jvi.00980-19. doi:10.1128/jvi.00980-19
- 26 18. Serafini B, Rosicarelli B, Veroni C, Aloisi F. Tissue-resident memory T cells in the  
27 multiple sclerosis brain and their relationship to Epstein-Barr virus infected B cells. *Journal*  
28 *of Neuroimmunology*. 2023;376:578036. doi:10.1016/j.jneuroim.2023.578036
- 29 19. Sabatino JJ, Wilson MR, Calabresi PA, Hauser SL, Schneck JP, Zamvil SS. Anti-CD20  
30 therapy depletes activated myelin-specific CD8+ T cells in multiple sclerosis. *Proceedings*  
31 *of the National Academy of Sciences*. 2019;116(51):25800-25807.  
32 doi:10.1073/pnas.1915309116

20. Garcia A, Dugast E, Shah S, et al. Immune Profiling Reveals the T-Cell Effect of Ocrelizumab in Early Relapsing-Remitting Multiple Sclerosis. *Neurol Neuroimmunol Neuroinflamm*. 2023;10(3):e200091. doi:10.1212/NXI.0000000000200091
21. Mathias A, Pantazou V, Perriot S, et al. Ocrelizumab Impairs the Phenotype and Function of Memory CD8+ T Cells. *Neurology Neuroimmunology & Neuroinflammation*. 2023;10(2):e200084. doi:10.1212/NXI.0000000000200084
22. Abbadessa G, Lepore MT, Bruzzaniti S, et al. Ocrelizumab Alters Cytotoxic Lymphocyte Function While Reducing EBV-Specific CD8+ T-Cell Proliferation in Patients With Multiple Sclerosis. *Neurology Neuroimmunology & Neuroinflammation*. 2024;11(4):e200250. doi:10.1212/NXI.0000000000200250
23. Wei J, Moon J, Yasumizu Y, et al. Transcriptomic profiling after B cell depletion reveals central and peripheral immune cell changes in multiple sclerosis. *The Journal of Clinical Investigation*. 2025;135(11). doi:10.1172/JCI182790
24. Waede M, Kingo C, Damsbo K, et al. Phenotypic Changes in a Monocyte Cluster with High Interleukin-1 Beta Expression during Long-Term Anti-CD20 Therapy. *Annals of Neurology*. 2025;n/a(n/a). doi:10.1002/ana.78004
25. Cantoni C, Smirnov RA, Firulyova M, et al. A single-cell compendium of human cerebrospinal fluid identifies disease-associated immune cell populations. *The Journal of Clinical Investigation*. 2025;135(1). doi:10.1172/JCI177793
26. Esaulova E, Cantoni C, Shchukina I, et al. Single-cell RNA-seq analysis of human CSF microglia and myeloid cells in neuroinflammation. *Neurology Neuroimmunology & Neuroinflammation*. 2020;7(4):e732. doi:10.1212/NXI.0000000000000732
27. Pappalardo JL, Zhang L, Pecsok MK, et al. Transcriptomic and clonal characterization of T cells in the human central nervous system. *Science Immunology*. 2020;5(51):eabb8786. doi:10.1126/sciimmunol.abb8786
28. Ramesh A, Schubert RD, Greenfield AL, et al. A pathogenic and clonally expanded B cell transcriptome in active multiple sclerosis. *Proc Natl Acad Sci U S A*. 2020;117(37):22932-22943. doi:10.1073/pnas.2008523117
29. Schafflick D, Xu CA, Hartlehnert M, et al. Integrated single cell analysis of blood and cerebrospinal fluid leukocytes in multiple sclerosis. *Nat Commun*. 2020;11(1):247. doi:10.1038/s41467-019-14118-w

- 1 30. Hayashi F, Mittl K, Dandekar R, et al. Antigen specificity of clonally enriched CD8+ T  
2 cells in multiple sclerosis. *Nature Immunology*. 2026;27(3):490-502. doi:10.1038/s41590-  
3 025-02412-3
- 4 31. Absinta M, Maric D, Gharagozloo M, et al. A lymphocyte–microglia–astrocyte axis in  
5 chronic active multiple sclerosis. *Nature*. 2021;597(7878):709-714. doi:10.1038/s41586-  
6 021-03892-7
- 7 32. Lerma-Martin C, Badia-i-Mompel P, Ramirez Flores RO, et al. Cell type mapping  
8 reveals tissue niches and interactions in subcortical multiple sclerosis lesions. *Nature*  
9 *Neuroscience*. 2024;27(12):2354-2365. doi:10.1038/s41593-024-01796-z
- 10 33. Kaufmann M, Evans H, Schaupp AL, et al. Identifying CNS-colonizing T cells as  
11 potential therapeutic targets to prevent progression of multiple sclerosis. *Med*.  
12 2021;2(3):296-312 e8. doi:10.1016/j.medj.2021.01.006
- 13 34. Heumos L, Schaar AC, Lance C, et al. Best practices for single-cell analysis across  
14 modalities. *Nature Reviews Genetics*. 2023;24(8):550-572. doi:10.1038/s41576-023-  
15 00586-w
- 16 35. Burkhardt DB, Stanley JS, Tong A, et al. Quantifying the effect of experimental  
17 perturbations at single-cell resolution. *Nature Biotechnology*. 2021;39(5):619-629.  
18 doi:10.1038/s41587-020-00803-5
- 19 36. Kotliar D, Veres A, Nagy MA, et al. Identifying gene expression programs of cell-type  
20 identity and cellular activity with single-cell RNA-Seq, eds. Alfonso Valencia, Naama Barkai,  
21 Elisabetta Mereu, Berthold Göttgens. *eLife*. 2019;8:e43803. doi:10.7554/eLife.43803
- 22 37. Lundberg SM, Erion G, Chen H, et al. From local explanations to global  
23 understanding with explainable AI for trees. *Nature Machine Intelligence*. 2020;2(1):56-67.  
24 doi:10.1038/s42256-019-0138-9
- 25 38. Gayoso A, Steier Z, Lopez R, et al. Joint probabilistic modeling of single-cell multi-  
26 omic data with totalVI. *Nature Methods*. 2021;18(3):272-282. doi:10.1038/s41592-020-  
27 01050-x
- 28 39. Hao Y, Hao S, Andersen-Nissen E, et al. Integrated analysis of multimodal single-cell  
29 data. *Cell*. 2021;184(13):3573-3587.e29. doi:10.1016/j.cell.2021.04.048
- 30 40. Kotliar D, Curtis M, Agnew R, et al. Reproducible single-cell annotation of programs  
31 underlying T cell subsets, activation states and functions. *Nature Methods*.  
32 2025;22(9):1964-1980. doi:10.1038/s41592-025-02793-1

- 1 41. Bates D, Mächler M, Bolker B, Walker S. Fitting Linear Mixed-Effects Models Using  
2 Lme4. *Journal of Statistical Software*. 2015;67(1):1-48. doi:10.18637/jss.v067.i01
- 3 42. Kuznetsova A, Brockhoff PB, Christensen RHB. lmerTest Package: Tests in Linear  
4 Mixed Effects Models. *Journal of Statistical Software*. 2017;82(13):1-26.  
5 doi:10.18637/jss.v082.i13
- 6 43. Lenth RV. Least-Squares Means: The R Package lsmeans. *Journal of Statistical*  
7 *Software*. 2016;69(1):1-33. doi:10.18637/jss.v069.i01
- 8 44. Palanichamy A, Jahn S, Nickles D, et al. Rituximab Efficiently Depletes Increased  
9 CD20-Expressing T Cells in Multiple Sclerosis Patients. *The Journal of Immunology*.  
10 2014;193(2):580-586. doi:10.4049/jimmunol.1400118
- 11 45. Wilk E, Witte T, Marquardt N, et al. Depletion of functionally active CD20+ T cells by  
12 rituximab treatment. *Arthritis & Rheumatism*. 2009;60(12):3563-3571.  
13 doi:10.1002/art.24998
- 14 46. Gingele S, Jacobus TL, Konen FF, et al. Ocrelizumab Depletes CD20+ T Cells in  
15 Multiple Sclerosis Patients. *Cells*. 2019;8(1). doi:10.3390/cells8010012
- 16 47. Shinoda K, Li R, Rezk A, et al. Differential effects of anti-CD20 therapy on CD4 and  
17 CD8 T cells and implication of CD20-expressing CD8 T cells in MS disease activity.  
18 *Proceedings of the National Academy of Sciences*. 2023;120(3):e2207291120.  
19 doi:10.1073/pnas.2207291120
- 20 48. Butzkueven H, Farr P, Ozakbas S, et al. Real-World Effectiveness of Ocrelizumab in  
21 Relapsing Multiple Sclerosis: An MSBase Registry Sub-Study. *Multiple Sclerosis and*  
22 *Related Disorders*. Published online November 2025. doi:10.1016/j.msard.2025.106885
- 23 49. Smolders J, van Luijn MM, Hsiao CC, Hamann J. T-cell surveillance of the human  
24 brain in health and multiple sclerosis. *Seminars in Immunopathology*. 2022;44(6):855-867.  
25 doi:10.1007/s00281-022-00926-8
- 26 50. Barbuti E, Castiello A, Pozzilli V, et al. Comparative effectiveness, safety and  
27 persistence of ocrelizumab versus natalizumab in multiple sclerosis: A real-world, multi-  
28 center, propensity score-matched study. *Neurotherapeutics*. 2025;22(2):e00537.  
29 doi:10.1016/j.neurot.2025.e00537
- 30 51. Nissimov N, Hajiyeve Z, Torke S, et al. B cells reappear less mature and more  
31 activated after their anti-CD20-mediated depletion in multiple sclerosis. *Proceedings of*  
32 *the National Academy of Sciences*. 2020;117(41):25690-25699.  
33 doi:10.1073/pnas.2012249117

52. Rodríguez-Mogeda C, van Lierop ZYGJ, van der Pol SMA, et al. Extended interval dosing of ocrelizumab modifies the repopulation of B cells without altering the clinical efficacy in multiple sclerosis. *Journal of Neuroinflammation*. 2023;20(1):215. doi:10.1186/s12974-023-02900-z
53. Wang M, Otto C, Fernández Zapata C, et al. Comprehensive analysis of B cell repopulation in ocrelizumab-treated patients with multiple sclerosis by mass cytometry and proteomics. *iScience*. 2025;28(5):112383. doi:10.1016/j.isci.2025.112383
54. Ochs J, Nissimov N, Torke S, et al. Proinflammatory CD20+ T cells contribute to CNS-directed autoimmunity. *Science Translational Medicine*. 2022;14(638):eabi4632. doi:10.1126/scitranslmed.abi4632
55. Schneider-Hohendorf T, Wünsch C, Schulte-Mecklenbeck A, et al. B- and T cell receptor sequencing elucidates characteristics of lymphocyte depletion by ocrelizumab. *iScience*. 2025;28(8). doi:10.1016/j.isci.2025.113068
56. Schneider-Hohendorf T, Gerdes LA, Pignolet B, et al. Broader Epstein–Barr virus–specific T cell receptor repertoire in patients with multiple sclerosis. *Journal of Experimental Medicine*. 2022;219(11):e20220650. doi:10.1084/jem.20220650
57. Mackay LK, Minnich M, Kragten NAM, et al. Hobit and Blimp1 instruct a universal transcriptional program of tissue residency in lymphocytes. *Science*. 2016;352(6284):459-463. doi:10.1126/science.aad2035
58. Parga-Vidal L, Behr FM, Kragten NAM, et al. Hobit identifies tissue-resident memory T cell precursors that are regulated by Eomes. *Science Immunology*. 2021;6(62):eabg3533. doi:10.1126/sciimmunol.abg3533
59. Feng R, Spieth L, Liu L, et al. Single-cell spatial transcriptomic profiling defines a pathogenic inflammatory niche in chronic active multiple sclerosis lesions. *Immunity*. Published online October 2025. doi:10.1016/j.immuni.2025.10.003
60. Merkler D, Vincenti I, Masson F, Liblau RS. Tissue-resident CD8 T cells in central nervous system inflammatory diseases: Present at the crime scene and ...guilty. *Current Opinion in Immunology*. 2022;77:102211. doi:10.1016/j.coi.2022.102211
61. Guo CL, Wang CS, Wang XH, Yu D, Liu Z. GZMK+CD8+ T cells: Multifaceted roles beyond cytotoxicity. *Trends in Immunology*. 2025;46(8):562-572. doi:10.1016/j.it.2025.06.003
62. Xin C, Liu P, Zhan Q, Cao W. GzmK+ CD8 T cells in inflammatory diseases. *Frontiers in Immunology*. 2025;16:1661755. doi:10.3389/fimmu.2025.1661755

63. Veroni C, Aloisi F. The CD8 T Cell-Epstein-Barr Virus-B Cell Dialogue: A Central Issue  
in Multiple Sclerosis Pathogenesis. *Frontiers in Immunology*. 2021;Volume 12 - 2021.
64. Giovannoni G. EBV-specific T-cell responses are telling us something important  
about multiple sclerosis. *Brain*. 2025;148(3):692-694. doi:10.1093/brain/awaf027
65. Lindsey JW. Epstein-Barr virus and multiple sclerosis: A mechanism comes into  
focus. *Multiple Sclerosis and Related Disorders*. 2026;105:106884.  
doi:10.1016/j.msard.2025.106884

## Figure legends

**Figure 1 Longitudinal multimodal single-cell profiling of the peripheral immune repertoire during ocrelizumab treatment.** (A) Graphical summary of discovery cohort of treatment-naïve pwMS ( $n = 14$ ), grouped by clinical parameters into three categories: RRMS with high disease activity ( $RR_{high}$ ,  $n = 5$ ), RRMS with low disease activity ( $RR_{low}$ ,  $n = 4$ ) and PPMS (PP,  $n = 5$ ). Longitudinal blood samples were collected before and during treatment with ocrelizumab, at both early (two weeks after first dose) and late (6–18 months after previous dose based on individual  $CD19^+$  B cell repopulation timing) time points. CITE-seq was performed on PBMCs isolated from whole blood, to measure the transcriptome and surface expression of 160 protein markers for each cell. (B) Heatmap summarising key demographic and clinical characteristics (rows) for each individual (columns). (C) UMAP visualisation of 264,512 cells comprising the discovery dataset, coloured and labelled by cell type. (D) Dot plot showing expression of canonical immune surface protein markers (x-axis) in each coarse (level 2) cell type (y-axis); colour scale shows normalised expression scaled to unit variance and dot size shows percentage of cells in each group with expression above threshold levels manually selected for each protein based on inspection of the expression distribution across all cells. (E) UMAP visualisation of cell states coloured by treatment-related perturbation score at early and late post-treatment time points; positive values (red) represent enrichment and negative values (blue) represent depletion relative to pre-treatment baseline. (F) Bar plots showing percentage (x-axis) of states in each coarse (level 2) cell type (y-axis) statistically significantly enriched (red) or depleted (blue) at the early and late post-treatment time points. (G) Violin/box plots showing distribution of perturbation scores within each cell type, coloured by post-treatment time point. Statistical testing for perturbation scores was performed with Z-tests using an empirical permutation-based null distribution for each cell state. MS = multiple sclerosis; pwMS = people with MS; RRMS = relapsing-remitting MS; PPMS = primary progressive MS;

CITE-seq = cellular indexing of transcriptomes and epitopes by sequencing; PBMC = peripheral blood mononuclear cell; UMAP = uniform manifold approximation and projection. Figure partially created in BioRender. Farooq, R. (2026) <https://BioRender.com/bte9vv0>.

**Alt text:** Diagrams and data plots, arranged in panels A–G, showing the experimental design for the discovery cohort, participant demographics, cell type annotations and their surface protein markers, and the longitudinal effect of ocrelizumab treatment on relative abundance across continuous cell states in the circulating immune cell repertoire.

## **Figure 2 Overview of analysis strategy for characterising shifts in cellular phenotypes due to treatment.**

(A) MELD is used to estimate the empirical probability density of cells from each pre- and post-treatment time point over the joint low dimensional manifold of cell state space. Estimated likelihoods for each cell state at each time point are computed separately based on the sampling distribution of cells from each donor and used to compute relative likelihoods compared to the donor-specific pre-treatment baseline; the median across donors is used as the consensus treatment-induced perturbation signal. This signal is standardised against a donor-stratified permutation-based empirical null distribution to compute statistically robust, denoised perturbation scores. (B) cNMF is used to infer additive gene programs (representing cell identity or biological function) within major immune cell populations, with program scores assigned to each cell state for each program as a continuous measure of cellular phenotype. (C) Non-linear relationships between continuous program scores and perturbation scores across cell states are modelled using random forest regression and SHAP-based feature importances obtained from the trained models are used to calculate program effect scores, representing an importance-weighted measure of treatment effect on the cellular phenotype represented by each gene program. MELD = manifold enhancement of latent dimensions; cNMF = consensus non-negative matrix factorisation; SHAP = Shapley additive explanations.

**Alt text:** Illustrative schematic diagrams, , arranged in panels A–C, outlining the computational analysis pipeline.

**Figure 3 Repopulating B cells after ocrelizumab are skewed towards an immature, antigen-inexperienced transitional phenotype.** (A) UMAP visualisation of B cells, which are shown highlighted (dark red) in UMAP visualisation of all cells (inset), coloured and labelled by coarse (level 2) cell type. (B) UMAP visualisation of B cell states coloured by treatment-related perturbation score at early and late post-treatment time points; positive

values (red) represent enrichment and negative values (blue) represent depletion relative to pre-treatment baseline. **(C)** Bar plot showing program effect scores (x-axis) for B cell-specific gene programs (y-axis); positive scores (red) indicate that high program expression is associated with enrichment post-treatment and negative scores (blue) indicate that high program expression is associated with depletion post-treatment. **(D)** Box plots comparing mean program scores per donor (y-axis) between pre- and late post-treatment time points (x-axis) for selected gene programs, with lines connecting paired data from the same individual. Statistical testing was performed using a linear mixed effects model adjusted for age and sex, *P*-values were calculated with *t*-tests of marginal means (using Satterthwaite degrees of freedom) and adjusted for multiple testing using the Benjamini-Hochberg method. UMAP = uniform manifold approximation and projection.

**Alt text:** Data plots, arranged in panels A–D, showing longitudinal treatment-related changes in B cell states following ocrelizumab.

**Figure 4 Circulating trafficking, activated, residency-primed, granzyme K-effector T cell (TARGET) signature is less abundant following long-term B cell depletion and correlates with clinical disease activity.** **(A)** UMAP visualisation of T cells, which are shown highlighted (dark green) in UMAP visualisation of all cells (inset), coloured and labelled by coarse (level 2) cell type. **(B)** UMAP visualisation of T cell states coloured by treatment-related perturbation score at early and late post-treatment time points; positive values (red) represent enrichment and negative values (blue) represent depletion relative to pre-treatment baseline. **(C)** Bar plot showing program effect scores (x-axis) for T cell-specific gene programs (y-axis); positive scores (red) indicate that high program expression is associated with enrichment post-treatment and negative scores (blue) indicate that high program expression is associated with depletion post-treatment. **(D)** Heatmap of top 100 weighted genes per component program of the TARGET signature with selected genes highlighted corresponding to key biological processes represented by each program; colour scale shows gene weights centred and scaled to unit variance. **(E)** Volcano plot showing differentially expressed proteins which are more abundant in memory T cells with high TARGET program score, with mean log<sub>2</sub> fold change on the x-axis and Bayes factor (posterior evidence for differential expression) on the y-axis; statistically significant proteins (FDR threshold < 0.05) are highlighted in blue and selected key markers are labelled. **(F)** Box plots comparing mean TARGET program scores per donor (y-axis) between pre- and early/late post-treatment time points (left panel) or between disease groups at the pre-treatment time point (right panel); lines connect paired data from the same individual. Statistical testing was performed using linear mixed effects models (within-donor repeated

measures) or standard linear models (no within-donor repeated measures) adjusted for age and sex; *P*-values were calculated with *t*-tests of marginal means (using Satterthwaite degrees of freedom for mixed effects models) and adjusted for multiple testing using the Benjamini–Hochberg method. UMAP = uniform manifold approximation and projection; FDR = false discovery rate.

**Alt text:** Data plots, arranged in panels A–F, showing longitudinal treatment-related changes in T cell states following ocrelizumab, highlighting gene and surface protein markers associated with the TARGET transcriptional signature, as well as comparing enrichment of this signature longitudinally during ocrelizumab treatment and between groups of patients with different disease subtypes at baseline.

**Figure 5 TARGET signature is enriched among memory T cells in the CSF, CSF-enriched expanded clonotypes, and chronically inflamed brain tissue of pwMS.** (A) Graphical summary of CSF/blood validation dataset of publicly available scRNA-seq studies with paired blood and CSF samples obtained from MS/CIS (*n* = 21) or healthy control donors (*n* = 9). (B) Box plots comparing mean TARGET program scores per donor (y-axis) in each compartment (x-axis) between MS/CIS (left panel) and healthy control donors (right panel); lines connect paired data from the same individual. (C) Box plots comparing CSF:blood log ratio of mean TARGET program score per donor (y-axis) in each disease group (x-axis). (D) Graphical summary of TCR clonality validation dataset of publicly available paired scRNA-seq and TCR sequencing study from blood and CSF samples obtained from MS/CIS (*n* = 12) or healthy control donors (*n* = 3). Expanded TCR clonotypes (> 0.75% of donor-specific TCR repertoire) were grouped into CSF-enriched (CSF:blood ratio > 2) or CSF-unenriched. Three CSF-enriched expanded TCR clonotypes were determined to be EBV-specific. (E) Box plots comparing mean TARGET program scores per donor (y-axis) in each expanded TCR clonotype group (x-axis) between disease group (columns) and tissue compartment (rows); lines connect paired data from the same individual. (F) Graphical summary of brain tissue validation dataset of publicly available post-mortem brain snRNA-seq data obtained from SPMS (*n* = 12) and healthy control donors (*n* = 6). SPMS samples were taken from subregions of chronic active/inactive demyelinating white matter lesions (core, rim, perilesional NAWM) and samples of normal white matter were taken from healthy control donors. Analyses were restricted to lymphocyte clusters extracted from the full snRNA-seq datasets. (G) Violin/box plots comparing TARGET program score (y-axis) in T cells from different subregions (x-axis); text labels above the x-axis indicate the number of T cells from each subregion in the dataset. Statistical testing was performed using linear mixed effects models (within-donor repeated measures) or standard linear models (no within-donor

repeated measures) adjusted for age and sex;  $P$ -values were calculated with  $t$ -tests of marginal means (using Satterthwaite degrees of freedom for mixed effects models) and adjusted for multiple testing using the Benjamini–Hochberg method. CSF = cerebrospinal fluid; scRNA-seq = single-cell RNA sequencing; MS = multiple sclerosis; CIS = clinically isolated syndrome; TCR = T cell receptor; EBV = Epstein-Barr virus; snRNA-seq = single nucleus RNA sequencing; SPMS = secondary progressive MS; NAWM = normal-appearing white matter. Figure partially created in BioRender. Farooq, R. (2026) <https://BioRender.com/bte9vv0>.

**Alt text:** Diagrams and data plots, arranged in panels A–G, showing the experimental designs for the CSF, TCR clonality and brain tissue external validation datasets, as well as comparing enrichment of the TARGET transcriptional signature between blood and CSF of MS and healthy control donors, between CSF-enriched and CSF-unenriched clonotypes, and between different brain MS lesion subregions and healthy control white matter with statistical annotation.

**Figure 6 Cytotoxic memory T cells with TARGET phenotype are trapped in the circulation during natalizumab treatment.** (A) Graphical summary of natalizumab validation dataset of publicly available PBMC CITE-seq study with pre- and post-natalizumab samples obtained from donors with RRMS ( $n = 9$ ). (B) UMAP visualisation of T cells coloured and labelled by coarse (level 2) cell type. (C) UMAP visualisation of T cell states coloured by treatment-related perturbation score at post-treatment time point; positive values (red) represent enrichment and negative values (blue) represent depletion relative to pre-treatment baseline. Cytotoxic memory T cells are indicated with a dashed grey circle. UMAP visualisation coloured by TARGET program score is inset; upper limit of colour scale is clipped at 99<sup>th</sup> percentile. (D) Box plots comparing mean TARGET program scores per donor ( $y$ -axis) between pre- and late post-treatment time points ( $x$ -axis); lines connect paired data from the same individual. Statistical testing was performed using a linear mixed effects model adjusted for age and sex;  $P$ -value was calculated with  $t$ -test of marginal means (using Satterthwaite degrees of freedom). PBMC = peripheral blood mononuclear cell; CITE-seq = cellular indexing of transcriptomes and epitopes by sequencing; MS = multiple sclerosis; RRMS = relapsing-remitting MS; UMAP = uniform manifold approximation and projection. Figure partially created in BioRender. Farooq, R. (2026) <https://BioRender.com/bte9vv0>.

**Alt text:** Diagrams and data plots, arranged in panels A–D, showing the experimental design for the natalizumab external validation dataset, treatment-related changes in T cell states

1 following natalizumab, as well as comparing enrichment of the TARGET transcriptional  
2 signature between the pre- and post-treatment time points with statistical annotation.

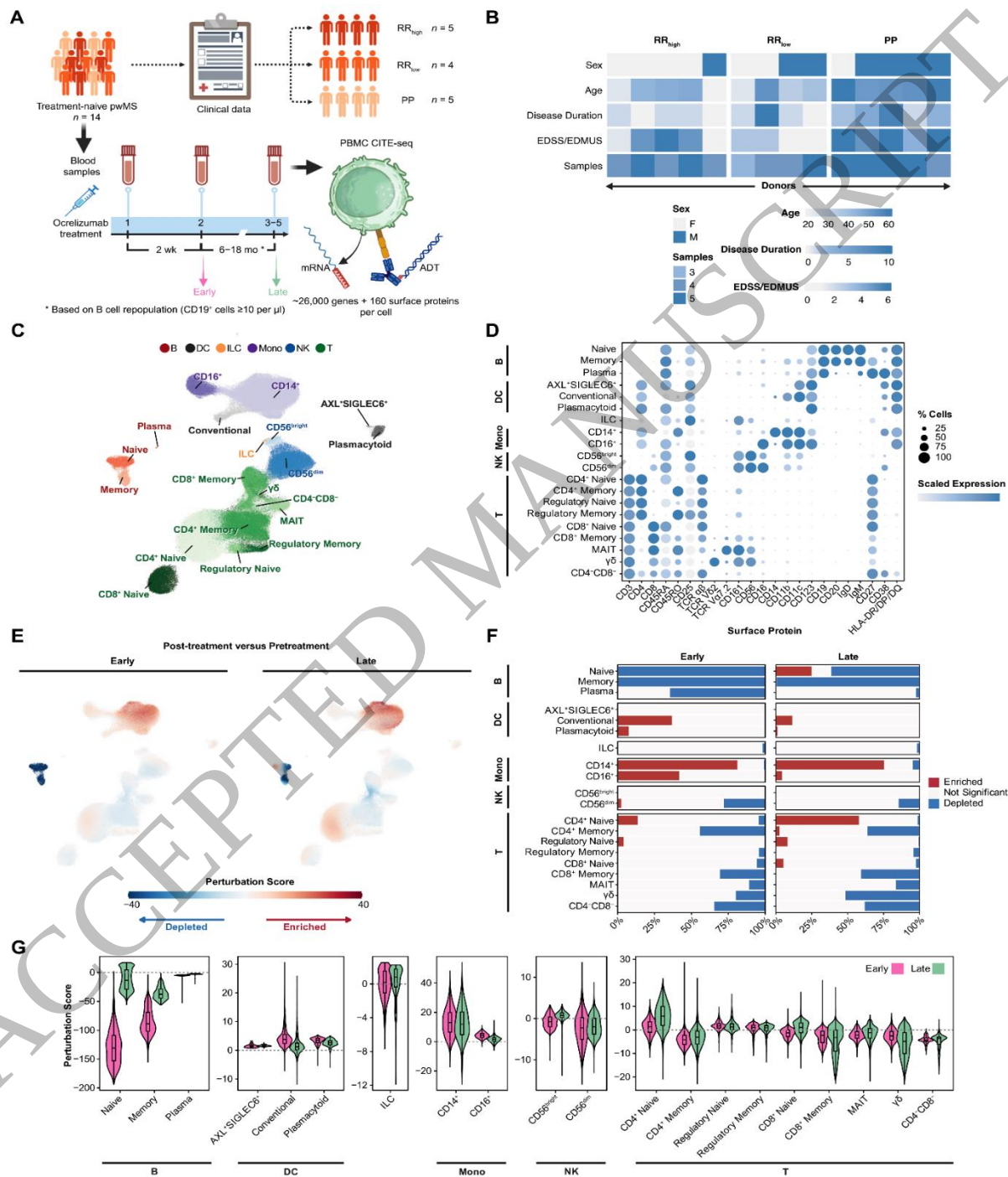


Figure 1  
180x225 mm (x DPI)

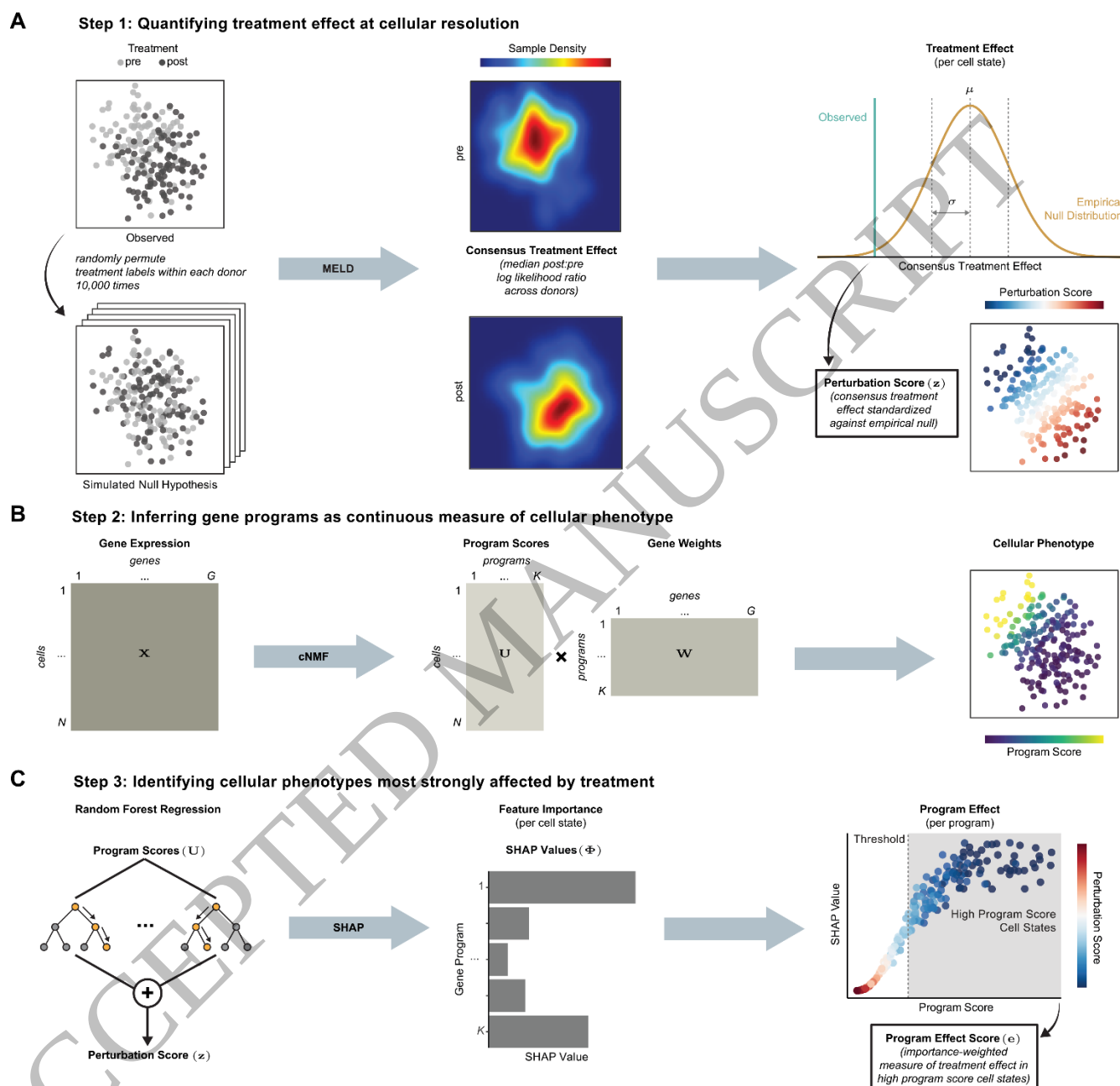


Figure 2  
180x172 mm (x DPI)

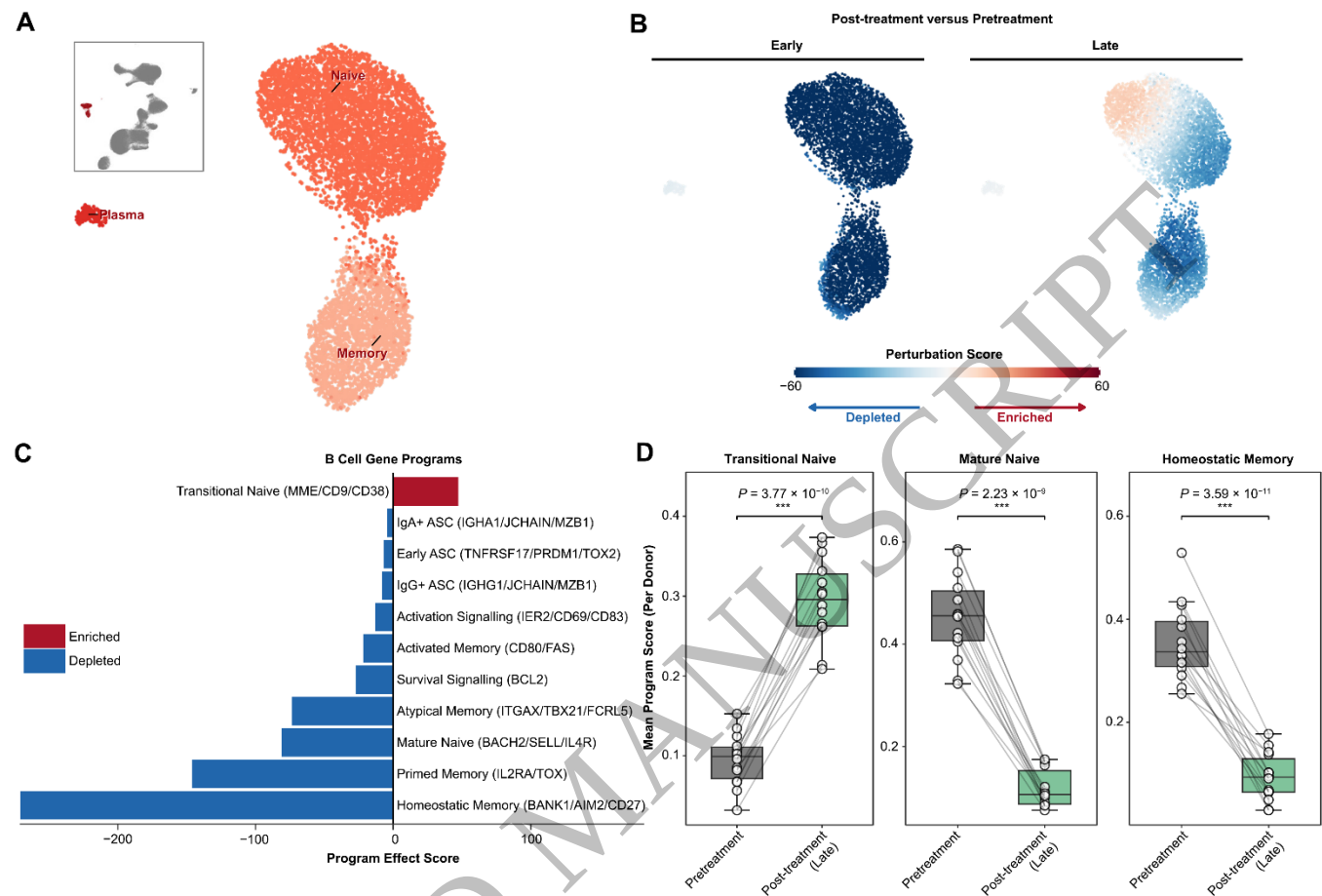


Figure 3  
180x123 mm (x DPI)



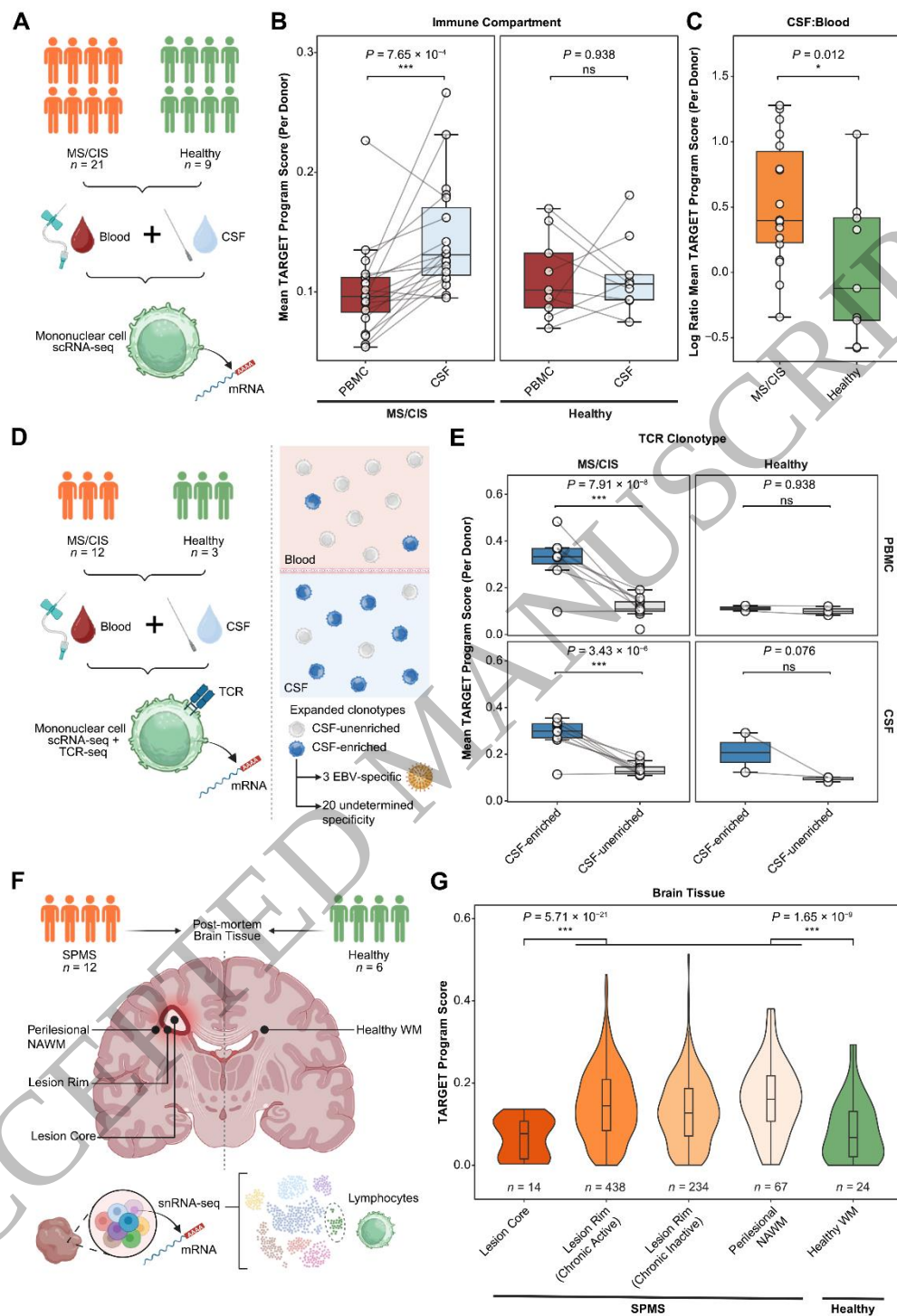


Figure 5  
135x195 mm (x DPI)

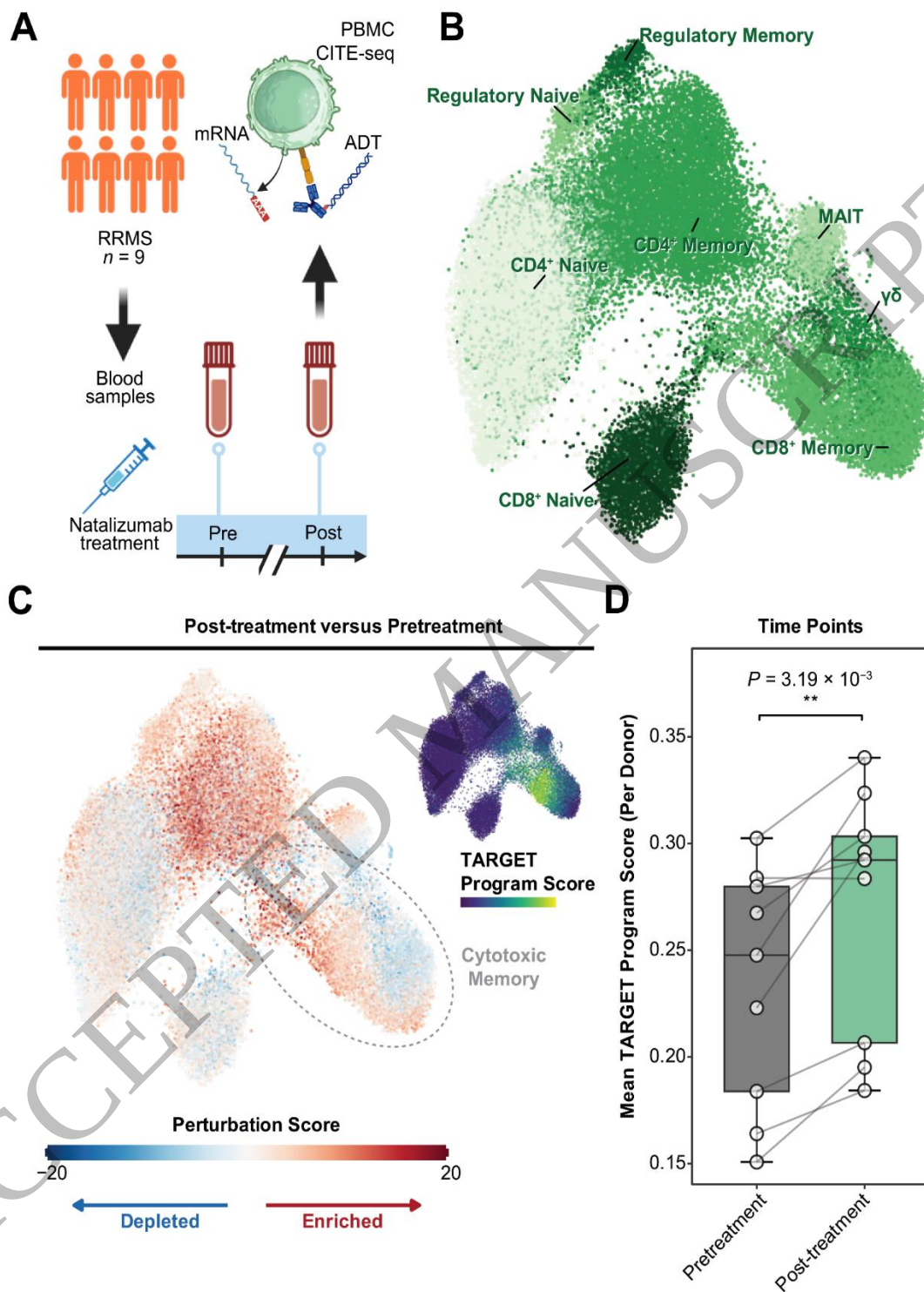


Figure 6  
187x244 mm (x DPI)