

Longitudinal single cell profiling of epitope specific memory CD4⁺ T cell responses to recombinant zoster vaccine

Received: 1 April 2024

Accepted: 25 February 2025

Published online: 08 March 2025


 Check for updatesXiaomin Wen^{1,7}, Alex K. Hu², Scott R. Presnell^{1,2}, Emily S. Ford^{3,4},
David M. Koelle^{1,3,4,5,6} & William W. Kwok^{1,3} 

Vaccination leads to rapid expansion of antigen-specific T cells within in the first few days. However, understanding of transcriptomic changes and fates of antigen-specific T cells upon vaccination remains limited. Here, we investigate the fate of memory CD4⁺ T cells upon reactivation to recombinant zoster vaccine for shingles at cellular and transcriptional levels. We show that glycoprotein E-specific memory CD4⁺ T cells respond strongly, their frequencies remain high, and they retain markers of cell activation one year following vaccination. Memory T cells with the most dominant TCR clonotype pre-vaccination remain prevalent at year one post-vaccination. These data implicate a major role for pre-existing memory T cells in perpetuating immune repertoires upon re-encountering cognate antigens. Differential gene expression indicates that cells post-vaccination are distinct from cells at baseline, suggesting committed memory T cells display transcriptional changes upon vaccination that could alter their responses against cognate immunogens.

Antigen-specific T cells that target the vaccine immunogen expand rapidly in the first few days after vaccination. This is followed by contraction of the antigen-specific T cell population and subsequent generation of memory T cells that are essential for protective immunity¹. Recent advances in DNA and RNA-sequencing technology have allowed for investigation of changes in T cell receptor (TCR) repertoires pre- and post-vaccination^{2–4}. As expected, expansion of specific TCR clonotypes followed by contraction was observed after vaccination. Most of these studies showed that vaccination increased the diversity of the TCR repertoires, and common CDR3 motifs were observed amongst individuals with identical HLA^{3–7}. A subsequent drop in clonal diversity has also been reported⁸. Primary vaccination was noted to preferentially expand naïve T cells of high clonal diversity but not naïve T cells of low clonal diversity and pre-existing memory T cells^{9,10}. However, studies investigating transcriptomic changes

of antigen-specific T cell responses upon vaccination in humans remain limited, with most studies focusing on the generation of memory CD8⁺ T cells from naïve T cells upon primary vaccination^{11–13}. These studies show that memory precursor CD8⁺ T cells have transcriptomic signatures associated with cellular metabolism and proliferation pathways¹¹, whereas memory CD8⁺ T cells have transcriptomic signatures that resemble naïve T cells and an epigenomic signature that resembles effector T cells^{12,13}. A study focused on circulating follicular helper (cTfh) CD4⁺ T cell differentiation demonstrated an increased inflammatory gene signature 7 days post-vaccination in adults over 65 years old versus younger adults between 18 and 35 years old¹⁴. Despite these results, the fate of memory T cells upon revaccination has not been sufficiently examined. Other than an increase in frequency of antigen-specific memory T cells, it remains unclear whether revaccination induces additional transcriptional changes in memory T cells that

¹Center for Translational Immunology, Benaroya Research Institute at Virginia Mason, Seattle, WA, USA. ²Center for Systems Immunology, Benaroya Research Institute at Virginia Mason, Seattle, WA, USA. ³Department of Medicine, University of Washington School of Medicine, Seattle, WA, USA. ⁴Fred Hutchinson Cancer Center, Seattle, WA, USA. ⁵Department of Global Health, University of Washington, Seattle, WA, USA. ⁶Department of Laboratory Medicine and Pathology, University of Washington, Seattle, WA, USA. ⁷Present address: AstraZeneca Pharmaceuticals, Gaithersburg, MD, USA.

 e-mail: bkwok@benaroyaresearch.org

influence their immune responsiveness. It is also unclear whether memory T cells will lose their self-renewal capacity upon reactivation and transition into terminally differentiated CD28⁺CD4⁺ T cells¹⁵.

For an investigation of memory CD4⁺ T cell responses to vaccine challenge, the current study examines the CD4⁺ T cell response to recombinant zoster vaccine (RZV) SHINGRIXTM in adults over 50 years old. RZV is a subunit vaccine that contains the recombinant varicella zoster virus (VZV) glycoprotein E (gE) together with the adjuvant AS01_B, which consists of monophosphoryl lipid A (MPL) and saponin QS-21¹⁶, and is a 2-dose regimen administered at least 60 days apart. Prior studies demonstrated that this vaccine elicits strong T cell responses against VZV and is highly effective in protecting against herpes zoster^{16,17}. Since most adults were exposed to VZV in their early life and the virus remains dormant in their ganglia, T cell responses to the first dose of the RZV vaccine offer an opportunity to examine responses of quiescent VZV-specific memory and naïve T cells upon antigen re-exposure. The current one-year longitudinal study used HLA class II tetramers and single cell RNA sequencing to follow phenotypic and transcriptomic changes of VZV gE-specific T cell responses pre- and post-RZV vaccination, at both the epitope-specific level and the TCR clonotypic level. For phenotypic analysis, markers being examined included both PD-1 and ICOS, which are activation markers but also play a role in regulating T cell proliferation and differentiation¹⁸. For transcriptional changes, we focus on T cells with identical TCR clonotype. Our observations show that a substantial proportion of memory T cells on day 365 had identical TCR clonotypes with the memory T cells on day 0. However, the study also found that the transcriptomic state of memory T cells 365 days post-vaccination was distinct from pre-vaccinated quiescent memory T cells on day 0.

The mechanisms by which RZV effectively prevents disease due to reactivation of a latent VZV infection are unknown. Because shingles is severe in persons with CD4⁺ T cell deficiency, and RZV strongly and durably increases immunogen-specific Th1 CD4⁺ T cells in blood, we perform detailed epitope- and clonotype-level longitudinal studies in humans. We find that the transcriptional profile of specific memory CD4⁺ T cells is durably altered compared to pre-vaccine baseline, with changes that persist long after receipt of the final dose of vaccine. These results imply that administration of recombinant protein with the adjuvants used in the licensed RZV product programs a long-

lasting phenotypic state in circulating CD4⁺ T cells that may be causally associated with desirable clinical outcomes.

Results

Frequencies and phenotype of VZV gE epitope-specific T cell before and after vaccination

Tetramer-guided epitope mapping (TGEM)¹⁹ was used to identify CD4⁺ T cell epitopes for VZV gE in healthy participants. Table 1 shows the 14 epitopes identified with 3 DRB1*01:01 (DR0101) restricted, 5 DRB1*04:01 (DR0401) restricted, 3 DRB1*04:04 (DR0404) restricted and 3 DRB1*07:01 (DR0701) restricted. Three of these epitopes have been previously reported by Voic et al.²⁰. To track VZV gE epitope-specific CD4⁺ T cells responses pre- and post-vaccination longitudinally, we recruited 7 healthy participants that had HLA corresponding to DR0101, DR0401, DR0404 and DR0701. Blood was collected at days 0 (baseline, immediately prior to first vaccine dose), 14, 60 (immediately prior to second dose), 74 (14 days post-second dose) and 365 (Fig. 2A). Table S1 lists the HLA and demographics for this cohort; all were between 53 and 67 years of age with known history of varicella during childhood.

A tetramer activation assay was used to track the frequency and phenotype of VZV gE-specific cells over time. For this assay, PBMCs were incubated with tetramers at 37 °C for 5 h before analysis. Only cells that were positive for both tetramer and the activation marker CD154 were considered specific. This assay was chosen as it provided higher quality single cell RNA sequencing (scRNA-seq) data than direct ex vivo staining without activation. The short stimulation also mimics the response of T cells after encountering the cognate antigen. Four tetramer reagents: DR0101/gE₁₉₀₋₂₀₉, DR0401/gE₂₈₀₋₂₉₉, DR0404/gE₂₈₀₋₂₉₉ and DR0701/gE₁₉₀₋₂₀₉ were chosen for these assays because their corresponding peptides elicited the highest T cell responses in the TGEM assays (Table 1). Figure S1 shows the gating strategy to identify VZV gE-specific CD4⁺ T cells, and Fig. 1 shows the staining results for a representative participant (Donor 1).

The frequencies and surface phenotype of VZV gE-specific CD4⁺ T cells for all 7 participants over the length of the study are shown in Fig. 2; note, donor 1 and donor 7 are represented twice showing results from DR0101/gE₁₉₀₋₂₀₉, DR0401/gE₂₈₀₋₂₉₉ and DR0701/gE₁₉₀₋₂₀₉ tetramers, as donor 1 is DR0401/DR0701 and donor 7 is DR0101/DR0401. At baseline, the frequencies of gE-specific CD4⁺ T cells ranged from 1.2 to 15.7/10⁶ (median 4.8/10⁶) CD4⁺ T cells. The percentage of CCR7⁺CD45RA⁺ naïve gE-specific T cells at baseline ranged from 0% to 16%, with a median of 9.2% (Fig. S1D). The kinetics of the T cell responses post-vaccination was as expected, with an increase in frequency that peaked on day 14 after the first dose of vaccine, followed by decline until day 60 and then another increase in frequency after the second dose (Fig. 2A, B). In our cohort, the frequency of VZV gE-specific CD4⁺ T cells was on average 15-fold higher than baseline after the first dose of vaccine, 81-fold higher than the pre-vaccination baseline 14 days after the second dose and 16-fold higher than baseline one year after the first vaccination (Fig. 2B). In addition, the frequency of VZV gE-specific CD4⁺ T cells was 33-fold higher after the second dose compared to immediately before this dose (Fig. 2B).

There were also changes in the phenotype of VZV gE-specific CD4⁺ T cell over time (Fig. 2C). Compared to baseline, there was a significant increase in the percentage of CXCR5⁺ cells at day 14, and a significant decrease in the percentage of CCR4⁺ cells at day 365 (Fig. 2C). There was also a decrease in the percentage of CCR7⁺ cells at day 74 (14 days post-second dose) compared to day 60 (Fig. 2C). A trend of an increase of CCR7⁺ cells at day 60 compared to day 0 was also observed. Up-regulation of CD38, a marker for recently activated CD4⁺ T cells, was observed 14 days after each dose of vaccine (Fig. 2C). The same pattern was observed for the activation markers PD-1 and ICOS with significant increases 14 days after each vaccine dose (Fig. 2C). Furthermore, although there was a significant decline in the percentage of ICOS⁺

Table 1 | VZV gE-specific CD4⁺ T cell epitopes

HLA	Epitope Position	AA seq	% Tetramer Positive
DRB1*01:01	gE ₁₉₀₋₂₀₉	RIYGVRYTETWSFLPSLTCT	8.2 (2) ^a
	gE ₂₈₀₋₂₉₉	EIEPGVLKVLRTKQYLGVI	2.4 (1)
	gE ₃₀₇₋₃₂₆	DGTSTYATFLVTWKGDEKTR	1.6 (1)
DRB1*04:01	gE ₄₆₋₆₅	DTNSVYEPYHSDHAESSWV	1.1 (2)
	gE ₁₂₇₋₁₄₆	GDDTGIHVIPTLNGDDRHKI	0.6 (1)
	gE ₁₅₄₋₁₇₃	YGDVFKGDLNPKPQGRRLIE	2.3 (2)
	gE ₂₈₀₋₂₉₉	EIEPGVLKVLRTKQYLGVI	5.3 (2)
	gE ₅₆₈₋₅₈₇	AYRVDKSPYNQSMYYAGLPV	0.6 (1)
DRB1*04:04	gE ₉₁₋₁₁₀	FLENAHEHHGVYNQGRGIDS	0.4 (1)
	gE ₁₂₇₋₁₄₆	GDDTGIHVIPTLNGDDRHKI	1.7 (2)
	gE ₂₈₀₋₂₉₉	EIEPGVLKVLRTKQYLGVI	18.0 (2)
DRB1*07:01	gE ₁₉₀₋₂₀₉	RIYGVRYTETWSFLPSLTCT	10.9 (2)
	gE ₂₈₉₋₃₀₈	LRTEKQYLGVIWNNMRGSDG	0.8 (1)
	gE ₄₄₂₋₄₆₁	CLGISHMEPSFGLILHDGGT	0.6 (1)

^aNumber in parentheses indicates the number of participants tested. % Tetramer positive shows % tetramer positive T cells in TGEM assay. If there was more than one participant being examined, the average is shown.

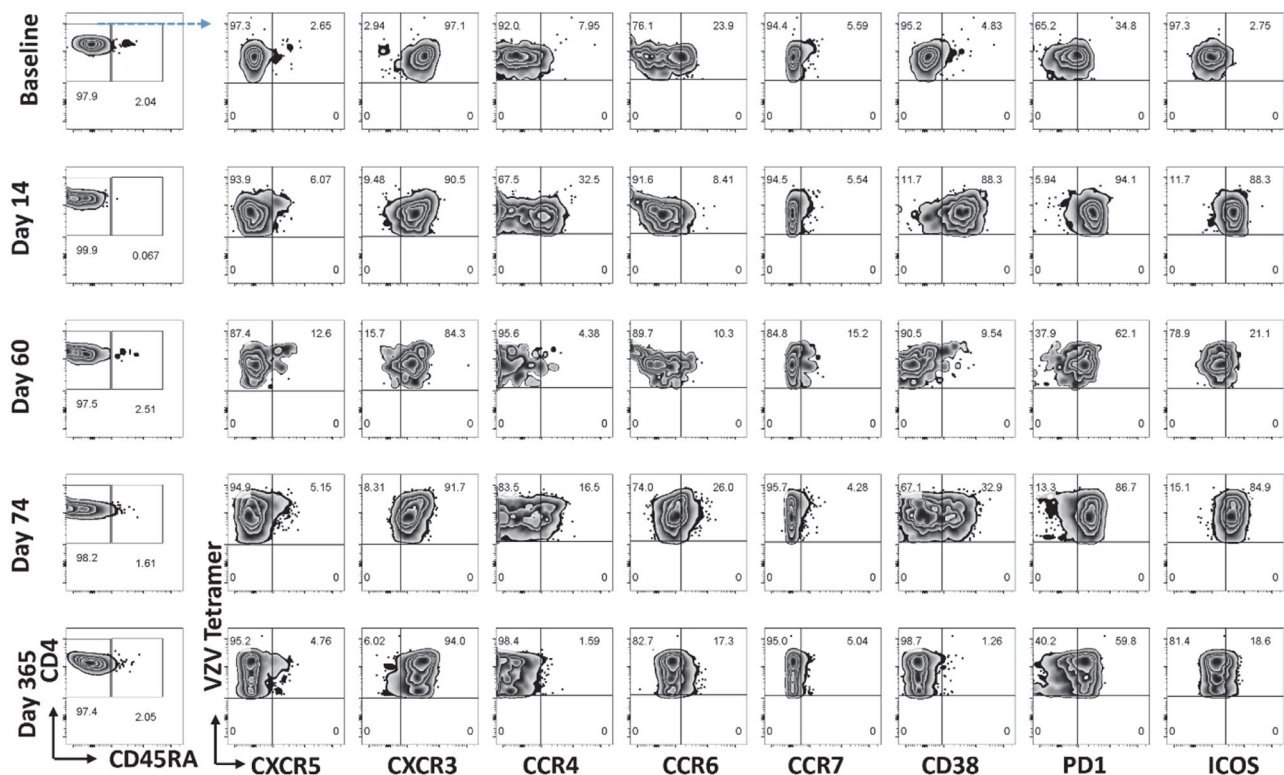


Fig. 1 | Phenotypic change of VZV gE-specific CD4⁺ T cells during RZV vaccination. Representative FACS analysis on VZV gE-specific CD4⁺ T cells by surface staining of CD45RA, CXCR5, CXCR3, CCR4, CCR6, CCR7, CD38, PD-1 and ICOS on CD4⁺CD154⁺Tetramer (DR0401/gE₂₈₀₋₂₉₉)⁺ cells obtained from donor 1 at days 0

(baseline), 14, 60, 74 and 365 post-first dose of Shingrix (second dose was administered after sample collection at day 60). The CD154(PE)⁺Tetramer(PE-Cy7 or PE-Dazzle 594)⁺ cells were enriched by anti-PE MicroBeads and MS column then stained with other surface antibodies described above before flow cytometry analyses and sorting.

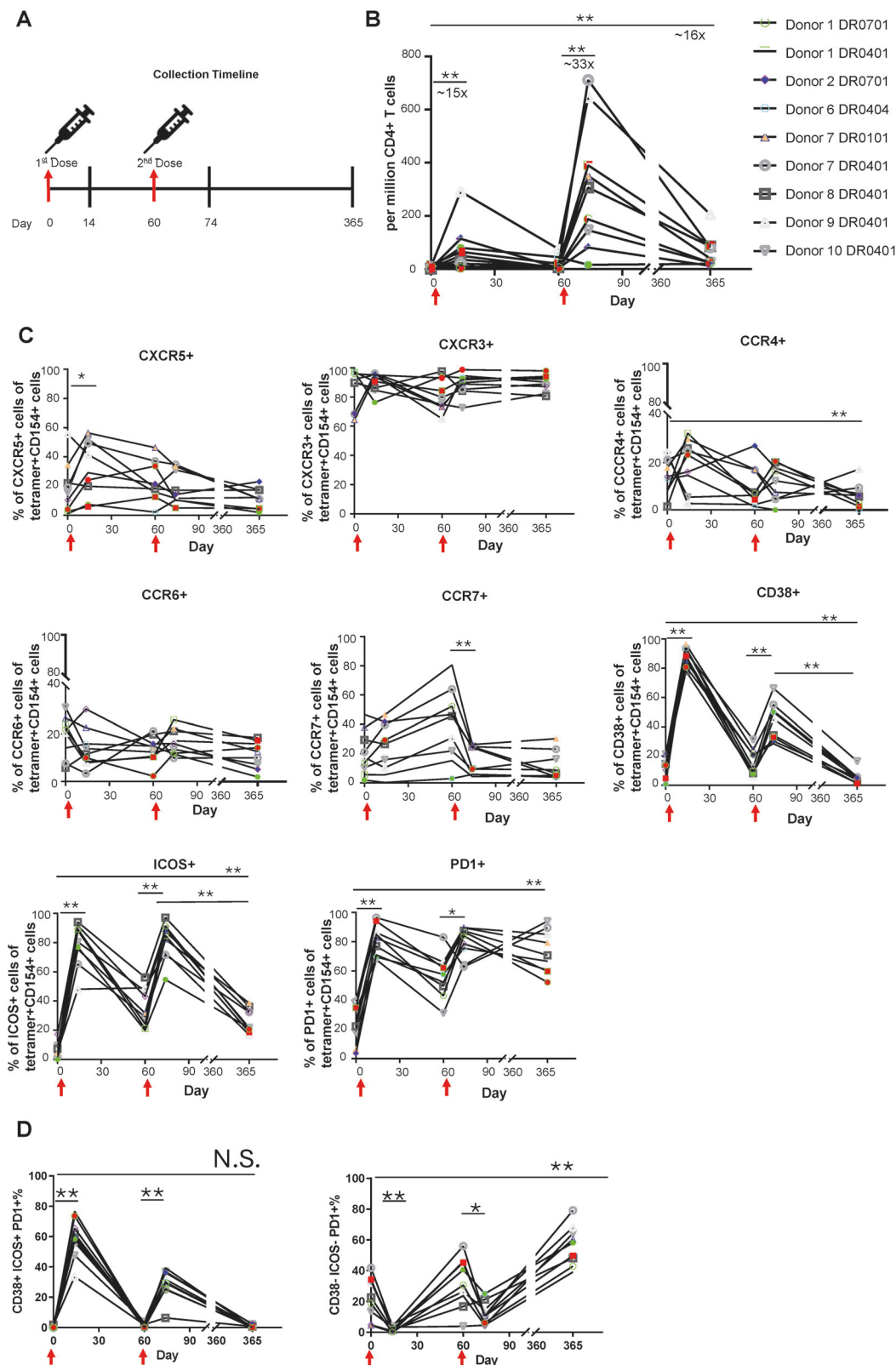
cells at day 365 compared to day 74, this activation marker was significantly up-regulated compared to baseline. In contrast, the percentage of PD-1⁺ cells at day 365 remained elevated and was not significantly different than observed for day 74 (Fig. 2C). We also noted up-regulation and co-expression of CD38, ICOS and PD-1 at day 14 and 74 (Fig. 2D). Up-regulation and co-expression of CD38, ICOS and PD-1 have also been observed at day 14 post-vaccination in a yellow fever virus vaccine study²¹. The kinetics of the appearance of these triple positive cells was directly opposite to that of the CD38⁺ICOS⁺PD-1⁺ population (Fig. 2D). This likely reflects the transient up-regulation of CD38 and ICOS in contrast to the more persistent up-regulation of PD-1 after activation. Correlation between expression of PD-1 and chemokine receptors, CCR4, CXCR3, CCR6 and CXCR5, was not observed (Fig. S2). The significant increase in frequency of T cells emphasized the robustness of VZV gE-specific CD4⁺ T cell responses post-vaccination, enabling a strong protective response from the vaccine. The data also demonstrated that VZV gE-specific CD4⁺ T cells had long-term memory of the previous activation, as PD-1 and ICOS were still up-regulated compared to baseline one year post-vaccination.

TCR usages and TCR diversity of gE-specific T cells pre- and post-vaccination

Single cell RNA sequencing was performed to characterize the TCRs and transcriptome of individual VZV gE-specific CD4⁺ T cells. Cells were available from every time point for donors 1, 2, 6, 7 and 8; and only from baseline, day 14 and day 365 for donors 9 and 10. A total of 2231 tetramer⁺CD154⁺ cells were index sorted for TCR analyses: 1665 cells (74.6% of cells sequenced) provided data for both TRAV and TRBV chains, 155 cells (6.9%) only TRAV chains and 336 cells (15.1%) only TRBV chains. Neither TRAV nor TRBV chain sequence was recovered from 75 cells (number of cells sorted and number of TCRαβ pairs recovered for each individual at each time point is shown in Table S2).

A total of 730 and 671 unique TCRαβ pairs based on nucleotide and amino acid sequence, respectively, were identified. Based on nucleotide sequence, 237 clonotypes were expanded (present in 2 or more cells), while 493 clonotypes were singletons. The number of unique TCR clonotypes based on nucleotide sequence for DR0101/gE₁₉₀₋₂₀₉, DR0401/gE₂₈₀₋₂₉₉, DR0404/gE₂₈₀₋₂₉₉ and DR0701/gE₁₉₀₋₂₀₉ was 59, 443, 100 and 128, respectively.

Within an individual, sharing of TCRαβ clonotypes (both VDJβ and VJα with CDR3 with identical nucleotide sequence) at different time points was common (Figs. 3A and S3). Sharing of clonotypes at the amino acid level (but not at the nucleotide level) between donors was also observed (Fig. 3B), including 2 public TCRαβ clonotypes that were shared by 4 out of 5 DRB1*04:01 donors in our cohort (Fig. 4A). There was also sharing of either Vα or Vβ usage (at the amino acid level) amongst 5 out of 5 DRB1*04:01 donors (Fig. S4). A strong preference for usage of TRAV29DV5/AJ36 and TRBV28/BJ1-1 was observed for DR0401/gE₂₈₀₋₂₉₉-specific T cells. Amongst the 347 unique DRB1*04:01 restricted TCR clonotypes that include exactly one alpha and one beta chain (as according to amino acid sequences), 32 utilized TRAV29DV5/AJ36 and TRBV28/BJ1-1 (Fig. 4B, Table S3). TCR sequence similarity analysis using TCRdist shows sequence similarity amongst gE-specific T cells using a strict neighbor distance of 24 units (Fig. 4C, D). 37 % (166/449) of these TCRs were grouped into 33 clusters with two or more clonotypes and 4 additional public TCR clusters with a single TCR clonotype (clusters 27, 31, 32 and 33). 24 of these 37 clusters have clonotypes from two or more individuals (Fig. 4D and Table S3). For example, cluster 0 has 18 different clonotypes from 6 different individuals. TCRs within each of these clusters can be considered meta-clonotypes, i.e., a group of biochemically similar TCR that recognize an identical T cell epitope^{20,22}. The CDR3 motifs of the 3 largest, sequence-related clusters are shown (Fig. 4E²²). These clusters, as shown in Fig. 4C, were fairly restricted by the HLA restriction element. Most of



the clusters consist of TCRs restricted by a single HLA-DR allele. Of note, both DRB1*04:01 and DRB1*04:04 restricted clonotypes could be observed within some clusters. In this context, it should be noted that DRB1*04:01 and DRB1*04:04 only differ at two amino acids in their first exon, and in our studies, present an identical VZV T cell epitope.

Whether T cells of identical epitope specificity with high and low TCR clonotype abundance expanded to a similar degree after

vaccination was also investigated. There was a weak direct correlation between frequencies of T cells of identical clonotype pre- and post-vaccination (day 0 versus day 14) and this was irrespective of clonotype abundance (Fig. 5A). A similar trend was observed for T cell frequency at day 74 versus day 60 (Fig. 5B). This observation suggests that T cells with dominant clonotypes are capable of expanding, but the slope of the curve suggests that T cells with lower clonotype

Fig. 2 | Analysis of VZV gE-specific CD4⁺ T cells during RZV vaccination.

A Timeline of vaccination and PBMC collection. **B** The frequency of VZV gE-specific CD4⁺ T cells driven by RZV vaccination. VZV gE-specific CD4⁺ T cells were identified as CD4⁺CD154⁺Tetramer⁺ cells by tetramer activation assay. As shown is their frequency change for each participant at baseline, day 14, 60, 74 and 365 post-first dose of RZV (second dose was administered after sample collection at day 60). $p = 0.0039$: day 0 vs day 14; $p = 0.0078$: day 60 vs day 74; $p = 0.0039$: day 0 vs day 365. **C, D** The phenotype changes of VZV gE-specific CD4⁺ T cells driven by RZV vaccination. The percentages of CD4⁺CD154⁺Tetramer⁺CD45RA⁺ memory cells expressing surface markers including CXCR5, CXCR3, CCR4, CCR6, CCR7, CD38, PD-1 and ICOS are shown for each participant during their visits. CXCR5%:

$p = 0.0391$ for day 0 vs day 14; CCR4%: $p = 0.0078$ for day 0 vs day 365; CCR7%: $p = 0.0078$ for day 60 vs day 74; CD38%: $p = 0.0039$ for day 0 vs day 14, $p = 0.0039$ for day 60 vs day 74, $p = 0.0039$ for day 74 vs day 365, $p = 0.0039$ for day 0 vs day 14, $p = 0.0039$ for day 60 vs day 74, $p = 0.0039$ for day 74 vs day 365, $p = 0.0039$ for day 0 vs day 365; PD1%: $p = 0.0039$ for day 0 vs day 14, $p = 0.0117$ for day 60 vs day 74, $p = 0.0039$ for day 0 vs day 365; CD38⁺ICOS⁺PD1%: $p = 0.0039$ for day 0 vs day 14, $p = 0.0039$ for day 60 vs day 74; CD38⁺ICOS⁺PD1%: $p = 0.0039$ for day 0 vs day 14, $p = 0.0195$ for day 60 vs day 74, $p = 0.0039$ for day 0 vs day 365. Wilcoxon matched-pairs signed rank test (two-tailed) from GraphPad Prism 9 was used ($n = 9$). * $p < 0.05$; ** $p < 0.01$. Source data are provided as a Source data file for (B–D).

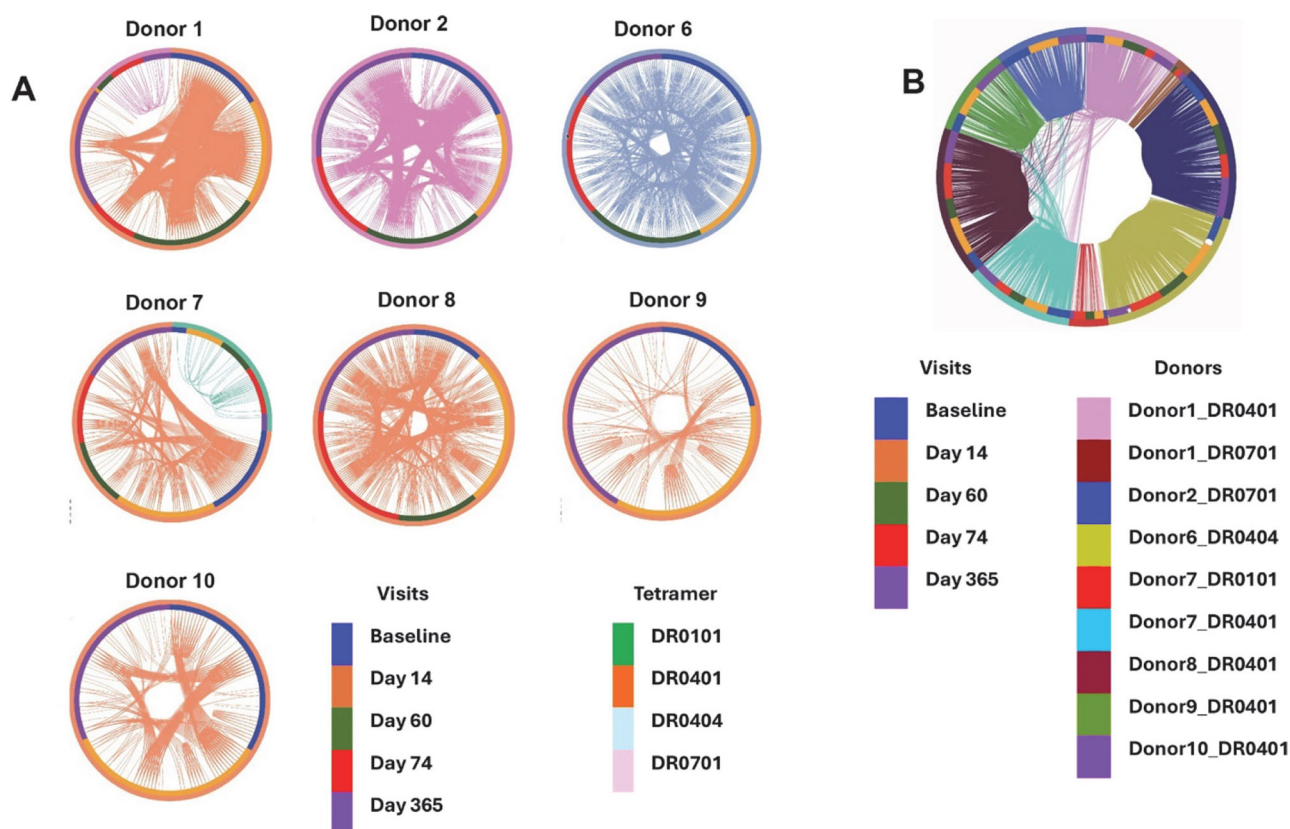


Fig. 3 | Intra-individual TCR sharing at different time points. A Circos style plot demonstrating clonal sharing within individual donors, across timepoints. Clone groups were identified by the sharing of identical β - and α -chain T cell receptor CDR3 junction sequences at the nucleotide level. Outer circle and arcs connecting cells are colored by tetramer grouping. Inner circle indicates the timepoint.

B Circos style plot demonstrating clonal sharing within and across donors at all timepoints. Clone groups were identified by the sharing of identical β - and α -chain T cell receptor CDR3 junction sequences at the amino acid level. Outer circle and arcs connecting cells are colored by donor/tetramer grouping. Inner circle indicates the timepoint.

frequency expand to a greater extent than T cells with higher clonotype frequency. Nevertheless, T cells with the most dominant clonotype at baseline were still present at day 365 for all participants examined, and the most abundant clonotype at day 0 was still the most abundant clonotype at day 365 for 4 out of the 7 donors examined (Table 2). Overall, 45% of unique TCR clonotypes from memory cells at day 0 and 38% from day 60 were also detected at subsequent time points. For the 7 participants examined, 10%–48% (with a median of 23%) of unique TCR clonotypes at day 365 matched those at day 0 and 14%–64% (with a median of 41%) of T cells at day 365 with clonotypes matched those at day 0 (Table 2). In contrast, TCR clonotypes present in CD45RA⁺CCR7⁺ naïve T cells (by flow) pre-vaccination at either day 0 or day 60 could not be detected at all subsequent time points.

Despite the persistence of memory T cells of identical clonotypes, an increase in diversity of gE-specific TCR usage at day 365 compared

to those at day 0, as evaluated by Shannon diversity index, was still observed (Fig. 5C). This increase in gE-specific TCR diversity also implied a preferential expansion of memory T cells of low clonotype frequency compared to those of high clonotype frequency.

Collectively, these results indicate that T cells of identical clonotype persist within an individual after activation by vaccination. These findings emphasize the significance of the memory TCR repertoires rather than the naïve TCR repertoires in recall responses.

Differential gene expression pre- and post-vaccination

Using the same cells sorted for the TCR clonotype analysis, we characterized the transcriptomic response of VZV gE-specific CD4⁺ T cells to vaccination. To track changes in identical T cells over time, comparisons were limited to epitope-specific T cells with identical TCR $\alpha\beta$ clonotypes at the nucleotide level at the time points of interest.

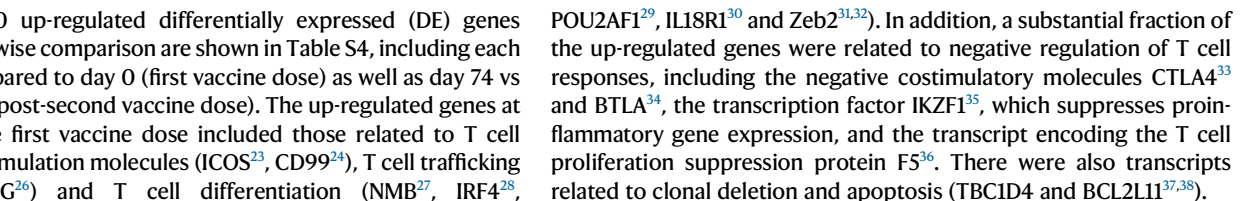


Fig. 4 | gE-specific TCR usage in DR0401 individuals. **A** Public TCR usage. TCR α and β amino acid CDR3 sequences are defined as a pair if they are detected in the same cell, and TCR pairs that are detected in cells across multiple donors are considered “shared.” 14 unique pairs were shared across 5 DR0401 donors, with 2 pairs shared by 4 donors, another 2 pairs shared by 3 donors and 10 pairs shared by 2 donors, as indicated. The number of cells containing each pair in each donor is depicted by the size of the dot, as shown. **B** V and J gene segment usage of gE-specific CD4⁺ T cells in DRB1*04:01 restricted donors. Different colors of the alluvia flowing between the strata (gene segments) indicate distinct TCR clonotypes, as determined by amino acid sequence, while width of the alluvia color indicates the number of libraries in each clonotype. TCR gene segments used in more than

twenty libraries are labeled in the strata. **C, D** Paired T cell receptor (TCR) sequence similarity graph of VZV specific T cells identified by tetramer sort with tetramers listed in figure legend (see Methods). Network graph shows 166 unique clonotypes (defined by amino acid sequence) which form 37 clusters defined by close sequence-similarity (nearest neighbor distance 24), indicated by grey lines. Size of circle reflects the number of copies of each sequence observed. Cluster members are shaded by tetramer (**C**) or person (**D**) in which the clonotype was identified. **E** Logoplots demonstrating the amino acid sequences of the CDR3 sequences in clonotypes contributing to clusters 0, 1 and 2. Flanking Sankey diagrams demonstrate the V and J chain usage in each cluster. Amino acid residues are shaded by chemical characteristics.

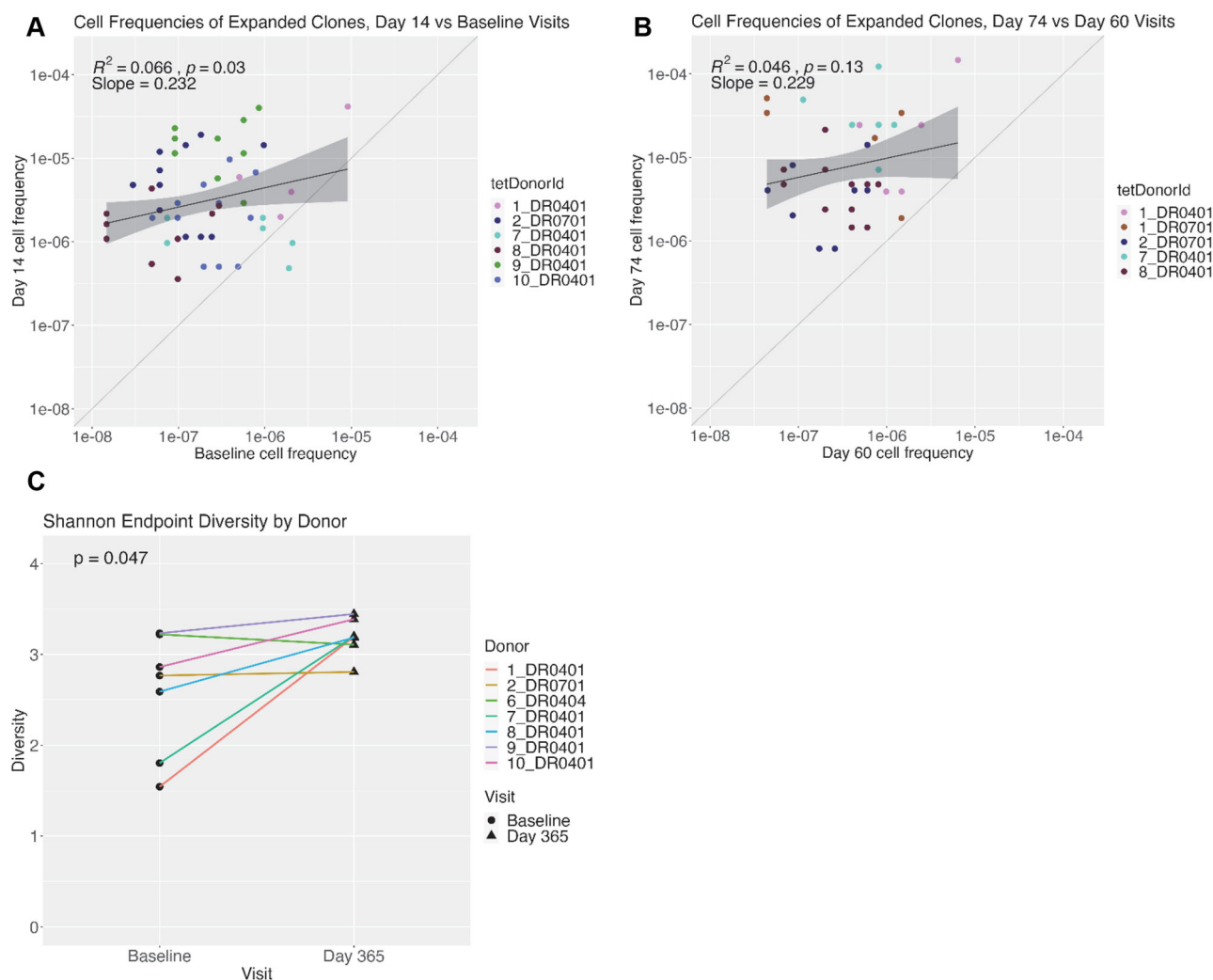


Fig. 5 | Frequency and diversity of TCR clonotype pre- and post-vaccination. **A** Comparison of tetramer positive, expanded T cells frequency per million CD4⁺ T cells, contrasting day 14 to baseline. Each point represents a single clonal group of T cells pre- (baseline) and post-vaccination (Day 14) colored by donor/tetramer combination. **B** Similarly, contrasting day 74 to day 60 (booster shot). Frequency of epitope-specific T cells of a specific clonotype were determined by the equation: Frequency of epitope-specific T cells \times % of cells with specific clonotype of interest of all the cell analyzed at that particular time point. For statistical evaluation, a linear model was fitted to the data, and Pearson's correlation coefficient and p -

value, using a two-sided alternative, were calculated (R package ggpubr, stat_cor() function, not adjusted for multiple comparisons). The grey shaded area indicates the 95% confidence interval at each fitted ordinate location from the linear model. **C** Simpson's diversity was calculated for TCR clonotypes between the baseline and day 365 visits using seven independent donor plus tetramer combinations, down-sampled to match the smaller time point sample size per donor. A two-sided Wilcoxon test for paired samples was used to provide a significant p -value when comparing the two timepoint diversity value distributions (not adjusted for multiple comparisons). Source data are provided as a Source Data file.

Genes that are related to T cell activation, trafficking, differentiation and negative regulation of T cell responses were also observed in all other pairwise comparisons with day 0. For the comparison between day 74 and day 0, DE genes related to T cell activation and trafficking are ICOS²³, ELMO1²⁵, GMFG²⁶, TNFRSF4 and TNFRSF9³⁹;

genes related to T cell differentiation are NMB²⁷, IRF4²⁸, JARID2⁴⁰, RBPJ⁴¹, ID2⁴², TET2⁴³ and ZEB2³¹; genes related to negative regulation of T cell responses and exhaustion are F5³⁶, CTLA4³³, FURIN⁴⁴, IQGAP1⁴⁵, CREM⁴⁶, MAF⁴⁷, SNX9⁴⁸ and BCL2L1³⁸ (Table S4). Note, these gene sets are similar to the DE gene set between day 14 and 0 but have more

Table 2 | Abundant and persistent TCR clonotypes

	% of the most dominant TCRαβ Clonotype at baseline	% of the identical TCRαβ Clonotype at day 14	% of the identical TCRαβ Clonotype at day 365	% of unique TCR clonotype in day 365 identical to those at day 0	% of T cells in day 365 with clonotype identical to those at day 0
Donor 1	58%	64% ^a	8%	10%	14%
Donor 2	33%	12%	28% ^a	25%	48%
Donor 6	10%	8%	7%	47%	64%
Donor 7	28%	5%	15% ^a	11%	21%
Donor 8	20%	8%	14% ^a	23%	47%
Donor 9	9%	14% ^a	2%	29%	40%
Donor 10	16%	13%	10% ^a	19%	41%

The 2nd column shows the % of T cells with the most dominant TCRαβ pair at baseline. The 3rd and 4th columns show the % of T cells at day 14 and day 365 that retained the most abundant clonotype at baseline

^aIn column 3 and 4 denoted those abundant TCR clonotypes at baseline that were also the most abundant clonotypes at later time points.

genes in each of the three categories discussed above. For comparisons between day 60 vs day 0 and day 365 vs day 0, genes related to negative T cell signaling, such as *F5*³⁶, *MAF*⁴⁷, *TBCID4*³⁷ (day 60 only); *IQGAP1*⁴⁵, *DOK2*⁴⁹, *CDKN3*⁵⁰ (day 365 only); and *CTLA4*³³ and *BCL2L1*³⁸ (day 60 and day 365) were at a significantly higher level. *TNFRSF9*, an activation marker, was also significantly higher at day 365 compared to baseline (Table S4). Transcripts of other activation-related genes, including *ICOS* and *TNFRSF4*, were also at higher levels for the day 60 vs day 0 comparisons but were not in the top 30 significant transcripts.

DE gene analyses between day 74 (14 days post-second dose) vs day 60 showed less up-regulation of genes related to activation, trafficking and differentiation in the top 30 DE gene sets. However, those related to negative regulation of T cells, including *FURIN*, *IQGAP1*, *CTLA4*, *CREM*, *SNX9*, *DOK2* and *CD47*³¹, were present. *TNFSF1B*, which is associated with suppression and pathogenic/dysfunctional T cells, was also present⁵² (Table S4). GO term analyses of the top 30 DE genes between day 60 vs 0, day 74 vs day 0 and day 74 vs day 60 also confirmed that these DE genes are related to T cell activation, differentiation and regulation of T cell activation (Table S5). These analyses also show DE genes related to cytoskeletal organization, which likely reflect both involvement of the cytoskeleton in cellular activation and post-vaccinated T cells being more active in trafficking than quiescent cells. Identical analyses of the top 30 DE genes between day 14 vs day 0 and day 365 vs day 0 did not show enrichment of genes related to the activation, differentiation and regulation. The vast number of DE genes for the day 14 vs day 0 comparison and the more subtle difference between day 365 vs day 0 comparison might account for our inability to show enrichment of these genes with GO term analysis.

The set of data for the four time points post-vaccination as compared to the pre-vaccination time point at baseline is summarized as a heatmap in Fig. 6A. Genes related to negative regulation of T cell expansion were mostly up-regulated at all time points and were a major component of the top-regulated genes. The up-regulation of genes related to the inhibition of T cell responses likely reflects a self-regulatory mechanism to limit the exuberant expansion of T cells after vaccination. Although the analyses of DE genes between different time points were focused on expanded T cells with identical TCRs, a very similar set of up-regulated genes was also observed when all the sorted epitope-specific cells were analyzed. For example, in the day 14 vs day 0 comparison with total sorted epitope-specific cells irrespective of their TCR clonotypes, 21 of the top 30 genes were the same as those for T cells that have identical TCR. In the day 365 vs day 0 comparison, 18 of the top 30 genes were identical (Table S6), including genes related to negative T cell signaling, activation and T cell differentiation. Thus, both global and focused analyses of sorted cells show that T cells of identical epitope specificity have a distinct transcriptomic profile one year post-vaccination.

Since the presence of transcripts related to negative T cell signaling in the DE gene analyses implied prior T cell activation, we

investigated whether post-vaccinated T cells expressed a higher level of activation transcripts in comparison to pre-vaccinated T cells. We generated an activation index using a T cell activation gene set signature defined by the following 8 genes: *CD137*, *CD134*, *PD-1*, *ICOS*, *CTLA4*, *CD38*, *CD82*⁵³ and *CD83*⁵⁴. The log expression of these genes was averaged across persistent cells at each time point for every donor. As shown in Fig. 6B, the activation index increases after both the first and second vaccine doses and is still elevated at day 365 compared to day 0. Interestingly, in previous studies, some of these activation markers, such as *CD137*, *PD-1*, *CTLA4* and *CD83*, were also associated with regulatory functions, implicating self-regulatory features of activated T cells to prevent exuberant activation.

For further examination of changes in gene expression of T cells with identical TCR clonotype over time, we also examined gene expression of the hallmark gene sets in the molecular signature database (MSigDB) at different time points⁵⁵ (Fig. 6C). We found hallmark gene sets that are related to mitotic spindle, IL2 and IL6 signaling, interferon gamma response and *KRAS* signaling were highly up-regulated at day 14 and 74 and down-regulated at day 60. These data suggest cells at day 14 and 74 were highly activated and in a proliferative state.

Up-regulation of gene expression at days 14 and 74 should be the outcomes of a combination of both vaccination and a 5 hrs in vitro tetramer activation. Down-regulation of hallmark gene sets related to cell activation at day 60 compared to all time points suggests that recently activated T cells are more resistant to a 5 hrs in vitro stimulation compared to those at day 0 and day 360.

These data also show that T cells at day 360 were distinct from those at day 0. Hallmark genes, including those related to glycolysis, fatty acid metabolism, oxidative phosphorylation, notch signaling and hedgehog signaling are differentially expressed at these 2 time points. This suggests that cells at day 360 and those at day 0 were in different metabolic, differentiation and activation states.

Pathway of memory VZV gE-specific CD4⁺ T cells post-vaccination

To further characterize the phenotype of the VZV gE-specific CD4⁺ T cells pre- and post- vaccination, we analyzed the scRNA-seq transcript profiles from all five time points using Monocle 3⁵⁶. The VZV gE-specific CD4⁺ T cells were segregated into five clusters as shown in the Uniform Manifold Approximation and Projection (UMAP) plot (Fig. 7A). We next examined the distribution of VZV gE-specific CD4⁺ T cells by time point and found that the majority cells from pre-vaccination (day 0) were in cluster 1, the cells from days 14, 60 and 74 were distributed across clusters 2, 3 and 4 with a slightly higher proportion of day 74 cells in cluster 4, and the majority of cells from day 365 in cluster 5 (Figs. 7B and S5). These data indicate that VZV gE-specific CD4⁺ T cells have distinct transcriptomic profiles at days 0 (baseline, immediately before the first vaccine dose), 60 (immediately

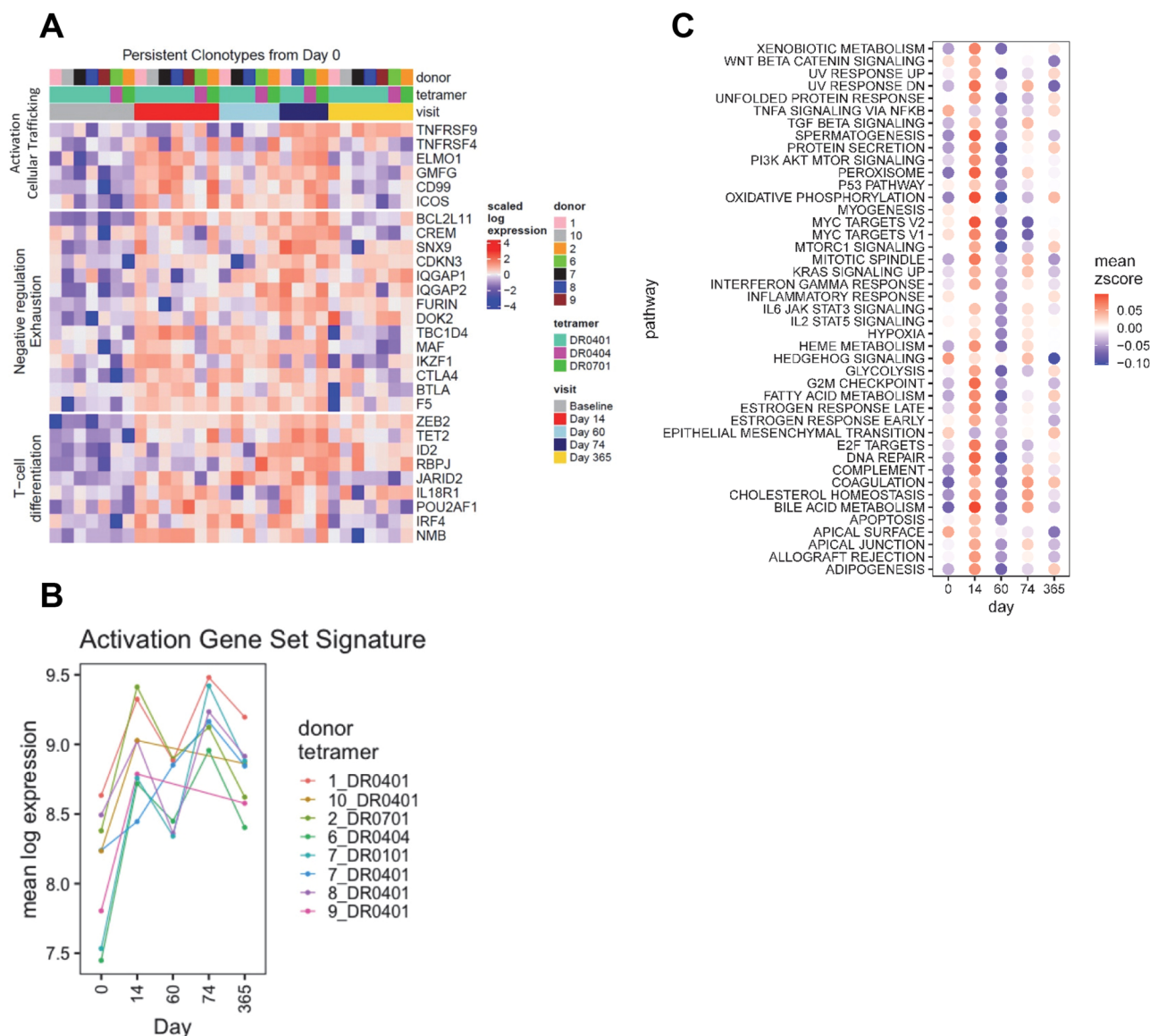


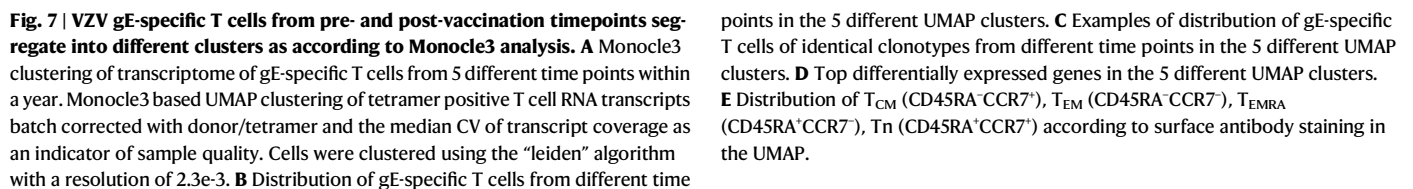
Fig. 6 | Top differentially expressed genes from T cells with persistent clonotypes. **A** For each comparison between two time points, cells containing TCR α/β nucleotide sequence pairs that persisted across both time points were identified, and differential expression analysis was performed between cells from the two time points. For each comparison, a selection of genes was shown from those differentially expressed at 5% FDR that are known to be related to activation, regulation and differentiation. **B** The log expression of 8 genes associated with activation

(TNFRSF9, TNFRSF4, PDCD1, ICOS, CTLA4, CD38, CD82, CD83) was averaged across persistent cells at each time point and displayed. **C** Gene set enrichment analysis (GSEA) identified 43/50 hallmark gene sets that are enriched in at least one cluster at 5% FDR. The dot plot displays via color the mean z-score of log-transformed expression for the genes in each gene set, averaged across the cells in each cluster. Source data are provided as a Source Data file for Fig. 6B.

before the second vaccine dose) and 365 (Fig. 7B), while cells at day 14 and 74 are more alike and resemble those at day 60. We also examined the distribution of two identical TCR clonotypes with one DR0401 restricted and the other DR0701 restricted (Fig. 7C). Both were detected in all 5 clusters with the clonotypes at day 0 primarily in cluster 1 and those at day 365 in cluster 5 (Fig. 7C). This highlights the changes in transcriptomic profiles of an individual T cell clone at different time points pre- and post-vaccination.

To further define the monocle clusters, we determined the top DE genes for each of the clusters (Table S7 and Fig. 7D). Cluster 1 DE genes are associated with cell quiescence (ZFP36L1⁵⁷), cell survival (IL2, IL7R⁵⁸ and BIRC3⁵⁹) and immediate early genes (NR4A2 and NR4A3⁶⁰). Cluster 2 DE genes are associated with newly activated T cells, such as those associated with one-carbon metabolism (SFXN1⁶¹), purine de novo synthesis (PAICS and GART⁶²) and cell division (SERBP1⁶³). F5, a T cell

proliferation inhibitory gene, is in cluster 2 with CCR7 and SELL, which are associated with central memory T cells (T_{CM}). Cluster 3 DE genes are related to T cell activation/co-stimulation (ICOS), differentiation (IRF4²⁸, JARID2⁴⁰, STK4⁶⁴ and ZEB2³¹), apoptosis (BCL2L1) and T cell dysfunction (TNFRSF1B). Cluster 4 DE genes are related to negative TCR signaling (AKAP13⁶⁵, CBLB⁶⁶ and SAMSNI⁶⁷) and non-apoptotic cell death (LAPTM5⁶⁸). There is also substantial overlap between the DE genes in Cluster 4 with those in Cluster 3, including those associated with activation and T cell dysfunction. Cluster 5 DE genes are associated with T cell homeostasis and polyfunctionality (PDIA6⁶⁹ and PRMT1⁷⁰), T cell activation (CD82 and TNFRSF9) and oxidative phosphorylation (MT-ND5⁷¹ and MRPL3⁷²). Signature DE genes, as described above in these different clusters, are in general agreement with the quiescent or activation state of cells from different time points within these clusters. Pre-vaccinated and day 365 cells were mostly located in



clusters 1 and 5, respectively, while recently activated T cells at day 14, 60 and 70 were located within clusters 2, 3 and 4.

As these cells in the monocle clusters were index sorted, we also examined the monocle clusters by surface markers CD45RA and CCR7 (Fig. 7E). This showed that the majority of CD45RA⁺CCR7⁺ (T_{CM}) and CD45RA⁺CCR7⁺ (naïve) T cells are primarily located in Cluster 2, whereas most of the CD45RA⁺CCR7⁺ T cells were in the other four clusters (Fig. 7E). We observed that the CCR7 antibody staining data corresponded well with the CCR7 transcriptomic data (Fig. S6A). A high percentage of CCR7⁺ T cells also have CD27 and CD28 transcripts (Fig. S6B), implying that these CCR7⁺ cells are not terminally differentiated T_{EM} (CD45RA⁺CCR7⁺CD27⁺CD28⁺) and resembled transitional memory T cells (T_{TM}, CCR7⁺CD45RA⁺CD27⁺CD28⁺)⁷³. The absence of Zeb2 transcript in cluster 1 also confirmed that cells in this cluster are mostly pre-vaccinated T cells (Fig. S6C). We noted that cells of identical clonotype at day 60 had a higher level of CCR7 transcript compared to those at day 0 (Fig. S7). The transcriptomic data aligns with the flow data, in which a higher percentage of T cells at day 60 had higher surface expression of CCR7 compared to day 0 (Fig. 2C). The main trajectory of the pseudotime pathway analysis shows that cells transited from cluster 1, through clusters 2 and 3 into clusters 4 or 5 (Fig. S8). This trajectory data is supported by the chronological data where cells in cluster 1 were mostly pre-vaccinated T cells, and cells in cluster 5 were mostly T cells at the one-year time point. As cells in cluster 1 were CCR7^{lo}, these data show CCR7^{lo} T_{TM/EM} (cluster 1) can convert to CCR7^{hi} T_{CM} (cluster 2) during activation and support the idea that antigen-specific T_{TM} cells play an important role in perpetuating the immune response upon re-encountering cognate antigens.

Discussion

Responses of VZV gE-epitope specific CD4⁺ T cells at the single cell level after RZV vaccination were examined in a TCR clonotypic fashion in this 1-year longitudinal study. We observed that memory T cells with abundant clonotypes at the pre-vaccination time point were fully capable of expanding, and the most dominant clonotype was still the dominant clonotype one year later after two doses of the vaccine in the majority of individuals. A relatively high percentage of the T cells at the one-year time point had TCR clonotypes that matched those identified from memory gE-specific T cell populations at the day 0 time point. In contrast, TCR clonotypes derived from naïve gE-specific T cells at day 0 or day 60 were not detected at any subsequent time points. In examining T cells with identical TCR clonotype, up-regulation of transcripts related to T cell activation and negative regulation of T cell responses were observed post-vaccination through the remaining length of the study. These data show that a substantial portion of memory T cells one year post-vaccination were derived from the pre-vaccinated memory T cells and have a distinct transcriptomic profile compared to the pre-vaccinated memory T cells.

TCR clonotype analyses in this study also led to identification of public TCR meta-clonotypes that are specific to the DRB1*04:01-DRB1*0404/gE₂₈₀₋₂₉₉ epitope. For confirmation of the publicity of these meta-clonotypes, we evaluated data from a prior RZV study⁷. In that study, CD4⁺ T cells that proliferated in response to gE antigen after 5 days of in vitro culture were bulk-sequenced at the *TRB* locus using a genomic DNA platform that yields quantitative information concerning cell abundance. We observed exact amino acid matching between TRB sequences from single cells in the current report and sequences from each of 5 HLA-DRB1*04:01 participants in the prior study. In contrast, none of 4 participants lacking HLA-DRB1*04:01 or HLA-DRB1*04:04 alleles had TRB matches (Table S8). These data highlight the structural constraint of the TCR in its interaction with peptide-MHC and the potential usage of meta-clonotypes in interrogating the TCR repertoires for specific immune responses.

Previous studies with YFV vaccine and zoster vaccine live (ZVL) examined whether T cells with abundant and infrequent clonotype

expanded to the same extent upon vaccination^{5,74}. In the YFV study, CD4⁺ T cells with dominant clonotypes expanded to the same extent as those with infrequent clonotypes. In the ZVL study, VZV T cells with low clonotype frequency expanded to a greater extent after a single dose of ZVL⁵. Note, both studies only used TCRβ to track T cell expansion, and the T cells were of unknown epitope specificity. In the current study, TCRαβ was used to track epitope-specific T cells. Although T cells with high clonotype frequency were capable of expanding, statistical analysis showed that cells with high clonotype frequency expanded less than those with lower clonotype frequency. Nevertheless, epitope-specific T cells with high TCR clonotype frequency at baseline remained at relatively high clonotype frequency one year post-vaccination in most participants.

Based on TCR sequencing, a recent report suggested that RZV is effective at recruiting T cells from the naïve repertoire⁷. One assumption was that post-vaccination TCRs that could not be detected in deep sequencing of pre-vaccinated CD4⁺ T cells were derived from naïve T cells. Data from the current study emphasized the persistence of memory T cells post-vaccination. Our data show less than 10% of gE-specific T cells in our cohort had the naïve phenotype at baseline, while >90% had the memory phenotype. The high proportion of memory gE-specific T cells at baseline and the ability of memory T cells with high clonotype frequency to expand are factors that favor the persistence of memory cells post-vaccination. In the current study, 47% of the TCR clonotype detected at day 365 could not be detected at any of the previous time points. It was unclear whether these “unseen” TCRs were derived from the naïve or memory population. The low number of cells analyzed in the current study precludes us from concluding the role of naïve gE-specific T cells in response to RZV. The relatively high percentage of T cells (median of 41%) at day 365 with TCR that matched those from memory cells at day 0 emphasized the central role of pre-existing memory T cells in maintaining long-term memory post-vaccination. This data also aligns with a recently published COVID mRNA vaccine study, in which a substantial portion of T cells one year post-vaccination were derived from memory T cells⁴.

Following activation through vaccination, T cells up-regulated a large set of genes related to the negative regulation of T cell responses. This is observed in the pairwise comparisons of T cells of identical clonotypes between days 14, 60 and 74 vs day 0. Up-regulation of a subset of these genes was also observed in the day 365 vs day 0 comparison. These genes could represent a self-regulatory mechanism that prevents excessive responses after activation. Although this should be an expected outcome within the first few weeks of vaccination, the up-regulation of some of these genes 365 days post-vaccination was unexpected.

The fact that post-vaccinated T cells at the one-year time point were distinct from the pre-vaccinated resting memory T cells was also reflected in the flow cytometric data. The data show that a higher percentage of gE-specific cells express the activation markers PD-1 and ICOS one year post-vaccination. Up-regulation of PD-1 expression in antigen-specific CD4⁺ T cell memory populations has been described in previous publications in relation to memory T cells in reaction to influenza viruses⁷⁵, Hepatitis B & C⁷⁶ and herpesvirus infections⁷⁷. These prior studies' findings of up-regulated PD-1 expression align with our results, suggesting that antigen-specific CD4⁺ T cells play an important role in providing long lasting protection against viral infections. Monocle analyses also show that pre-vaccinated T cells and T cells one year post-vaccination were distinct.

ZEB2 and CTLA4 are two genes in the top 5 up-regulated DE genes for all DE comparisons between the post-vaccination time points and day 0, including the day 365 vs day 0 comparison. CTLA-4 is a negative co-stimulation molecule which transmits a negative signal upon its interaction with its ligand B71/B72. Up-regulation of this marker aligns with our general observation of up-regulation of genes related to negative signaling. Zeb2 is an

E-box binding transcription factor³¹. Previous studies showed that Zeb2 was up-regulated in CD8⁺ T cells upon repeated re-stimulation⁷⁸ and Zeb2 expression in CD8⁺ T cells leads to terminal differentiation^{32,79}. In the current study, 19% and 66% of VZV gE-specific CD4⁺ T cells expressed Zeb2 at days 0 and 365, respectively. We also observed that 60% of CCR7⁺ T cells (a marker for T_{CM}) expressed Zeb2, and 57% of Zeb2⁺ T cells expressed CCR7, all at the transcript level. A study of CD4⁺ T cells responding to SARS-CoV-2 in patients early in the COVID19 pandemic also found CD4⁺ T cell subsets that were enriched for CCR7 and ZEB2 following exposure to the virus⁸⁰. Although it is possible that these CCR7⁺Zeb2⁺ T cells will differentiate into terminal CCR7⁺Zeb2⁺ T cells, the possibility that Zeb2⁺CD4⁺ T cells could turn off the expression of Zeb2 and become CCR7⁺Zeb2⁺ T_{CM} (during their life span) upon further homeostatic division cannot be ruled out.

Both flow and transcriptomic data in the current study also show that a large percentage of pre-vaccination quiescent gE-specific T cells have CCR7^{lo} T_{TM}/T_{EM}-like phenotype, while T cells at day 60 post-vaccination have a CCR7^{hi} T_{CM}-like phenotype. This finding highlights the plasticity of the T_{CM} and T_{TM}/T_{EM} subsets and the role of gE-specific T_{TM} cells in maintaining long-term memory.

The distinct T cell transcriptomics among memory cells of identical clonotype one year apart implies that there may also be epigenetic changes within these cells after vaccination. We speculate secondary vaccination and vaccine boosters not only increase the number of specific T cells but also likely change the transcriptomic landscape of these cells, leading them to respond differently compared to long-resting, quiescent memory T cells. These results also suggest that the time intervals between vaccinations will influence the T cell responses to the later dose.

There are limitations to this study. The tetramer activation assay with a 5 hrs stimulation may not detect low avidity antigen-specific T cells and may be less effective at detecting antigen-specific naïve T cells. The low number of T cells as detected by tetramer activation and assayed in this study can also lead to sampling error and our ability to examine the TCR of VZV naïve T cells. The lack of deep TCR sequencing data of total PBMCs precluded us from examining our data with a metric to evaluate repertoire coverage. It is also important to note that 6 out of 7 of the participants in this study were female. Therefore, if there are differences between the sexes in the immune response to RZV, the results reported in this study will be skewed due to this imbalance in the participants.

RZV, which contains MPL/QS adjuvant, has proven to be a highly effective vaccine. The robust CD4⁺ T cell responses elicited by the RZV are likely one of the major factors in its high efficacy. As more recombinant vaccines with very similar adjuvants are being used, further studies of antigen-specific CD4⁺ T cell responses to these vaccines, including transcriptomic and epigenetic changes in T cells, will be helpful in understanding the roles of CD4⁺ T cells in providing long lasting protective responses and T cell differentiation in general.

Material and Methods

Human participants

Healthy individuals with specific HLA class II haplotypes of interest were recruited at Benaroya Research Institute at Virginia Mason under a study approved by the Benaroya Research Institute Institutional Review Board (IRB Protocol 07109-563). All participants provided a written informed consent prior to study participation, in accordance with the Declaration of Helsinki and the IRB-approved protocols. A total of 15 participants with HLA-DRB1*01:01, DRB1*04:01, DRB1*04:04 or DRB1*07:01 were recruited. Eight individuals were recruited for epitope mapping studies, and an additional seven were recruited for

the longitudinal study. All seven longitudinal study participants were between 53 and 67 years of age with known history of varicella-zoster in their childhood (Table S1).

Tetramer guided epitope mapping

Tetramer-guided epitope mapping (TGEM)¹⁹ was performed to identify MHC class II-restricted peptide epitopes for glycoprotein E (gE) of VZV. For these experiments, a total of 68 peptides (20 amino acids in length with a 12 amino acid overlap) spanning the entire length of the VZV ORF68 encoding glycoprotein E were synthesized (Mimotopes, Mulgrave, Australia). Peptides were divided up into pools of 5 peptides each. Each pool of peptides at a final concentration of 2 µg/ml for each individual peptide was used to stimulate 4–5 × 10⁶ PBMCs in a well of a 24-well plate. IL-2 at a final concentration of 10 U/ml was added every other day starting at day 5. Cells were cultured for 14 days before staining with pooled tetramers and then with individual tetramers¹⁹. Tetramers were generated by loading pooled peptides or individual peptides (4 mg/ml) onto the specified biotinylated HLA class II proteins (0.5 mg/ml) by incubation for 48 h at 37 °C in 100 mM sodium phosphate pH 6.0 and 0.2% n-Octyl-Beta-D-Glucopyranoside and then conjugated with streptavidin-phycoerythrin (PE) at a molar ratio of 8:1 (monomer:streptavidin-PE) to generate tetramers^{21,81}. All four different biotinylated DR monomers were produced by the Benaroya Research Institute Tetramer Core.

Tetramer activation assay

Peripheral blood mononuclear cells (PBMCs) were isolated from 50 to 120 ml whole blood samples by Ficoll gradient method. Autologous plasma was also saved to make T cell medium (TCM; consisting of RPMI with 10% autologous human plasma and 1% sodium pyruvate, glutamine, and penicillin/streptomycin; Invitrogen) for each blood donor. Ten to 15 million PBMCs in 2 ml of TCM were plated into each well of a 12-well plate and were activated with peptide-MHC-II tetramers (5 µg/ml) in the presence of anti-CD40 (clone HB-14, Miltenyi Biotec) at 37 °C, 5% CO₂ in a humidified incubator for 5 h. Specific tetramers conjugated to either PE-Cy7 or PE Dazzle 594 were used according to the HLA class II haplotype of the individual. Cells were washed and then stained with anti-CD154-PE (clone 5C8, Miltenyi Biotec). A 1/10th fraction of the cells was saved for the purpose of calculating the total CD4⁺ T cells used, and the rest of cells were processed with anti-PE MicroBeads and MS column (Miltenyi Biotec) to magnetically enrich CD154⁺PE⁺/Tetramer-PE-Cy7⁺ and/or Tetramer-PE-Dazzle 594⁺ cells. The enriched cells were further stained with other surface antibodies including anti-CD45RA-AF700 (clone HI100), anti-CD4-BUV737 (clone SK3, BD Biosciences), anti-CXCR5-allophycocyanin (APC)-Cy7 (clone J252D4), anti-CCR4-PerCP-Cy5.5 (clone L291H4), anti-CXCR3-BUV395 (clone 1C6/CXCR3, BD Biosciences), anti-CCR6-BV650 (clone G034E3), anti-CCR7-BV421 (clone G043H7), anti-CD38-APC (clone HB-7), anti-PD1-BV605 (clone EH12.2H7), anti-ICOS-BV785 (clone C398.4A) and a dump channel consisting of anti-CD14-BV510 (clone M5E2), anti-CD19-BV510 (clone SJ25C1), anti-CD56-BV510 (clone HCD56) and Fixable Viability Stain 510 (BD Biosciences). All anti-human antibodies mentioned above were purchased from BioLegend, except as otherwise noted. The gating strategy for identifying CD4⁺CD154⁺Tet⁺ VZV-specific CD4⁺ T cells is shown in Fig. S1. Briefly, FSC, SSC, dump channel (consisting of anti-CD14, anti-CD19, anti-CD56 antibodies and viability stain 510) and anti-CD4 antibody were used to identify single CD4⁺ T cells. Tetramer reagents and anti-CD154 antibodies were then used to identify CD154⁺Tet⁺ cells. Gated CD154⁺Tet⁺ cells were analyzed separately for expression of other surface markers. FACS Diva software was used during sorting, and Flowjo 10.7.1 was used for data analysis. The frequency of VZV gE-specific CD4⁺ T cells was calculated by using the formula $F = n/N$, where n designates the number of CD154⁺Tetramer⁺ cells after enrichment and N is the total number of CD4⁺ T cells used in the assay²¹.

Single cell sorting and RNA sequencing

For all experiments in this longitudinal study, Rainbow calibration, 8 peaks beads (BD 559123) were used to achieve the same (+/−5%) MFI value of the peak 7 in each channel of the BD FACSARIA Fusion sorter. Then CD4⁺CD154⁺Tetramer⁺ cells were index sorted at single-cell purity directly into individual wells of a 96-well PCR plate with 5 µl of reaction buffer from the SMART-Seq v4 Ultra Low Input RNA Kit for Sequencing (Takara). Index sorting was performed, allowing the phenotype of surface staining to be traced and confirmed for each individual cell. Reverse transcription was performed, followed by PCR amplification to generate full-length amplified cDNA. Sequencing libraries were constructed using the NexteraXT DNA sample preparation kit (Illumina) to generate Illumina-compatible barcoded libraries. Libraries were pooled and quantified using a Qubit® Fluorometer (Life Technologies). Dual-index, single-read sequencing of pooled libraries was carried out on a HiSeq2500 sequencer (Illumina) with 58-base reads, using HiSeq v4 Cluster and SBS kits (Illumina) with a target depth of 1 million reads per sample. Base calls were processed to FASTQs on BaseSpace (Illumina), and a base call quality-trimming step was applied to remove low-confidence base calls from the ends of reads. The FASTQs were aligned to the human GRCh38 release 91 reference genome, using STAR v.2.4.2a and gene counts were generated using htseq-count (v0.4.1). Quality control and metrics analysis were performed using the Picard family of tools (v1.134). To identify TCR chains, the Trinity assembler was used to generate contigs. TCR sequences were identified and annotated from these contigs using MiXCR (v2.1.3). Cells with high-quality sequencing were identified as having at least 500,000 reads, 70% read alignment, and a median covariance of coverage of 1.

Differential gene expression analysis

To analyze differential gene expression between persistent clones at different time points, persistent clones were defined as cells with TCR alpha and beta full nucleotide sequence pairs that were detected within the same donor at both the baseline time point and a subsequent time point. Differential gene expression in these persistent clones between the baseline and subsequent time points was computed separately for each subsequent time point, using a MAST⁸² model with formula “expression ~ cdr+clonotype+visit”, where cdr is the “cellular detection rate”, or the number of genes detected in each cell. Gene expression was normalized from raw counts using the deconvolution method in the R package scran⁸³ on genes expressed in at least 5% of cells.

To find differential expression between clusters, cells were “pseudobulked” into groups according to cluster and sample⁸⁴. Gene expression in the groups of cells was defined as TMM normalized (using the EdgeR package)⁸⁵ summed counts across all cells within each group. Differential expression between clusters was computed from a LIMMA⁸⁶ model with the formula expression ~ 0+cluster+tetDonorID, where tetDonorID indicates the tetramer and donor of the sample. Genes differentially expressed between each cluster N and the rest were computed using the contrast cluster N − mean (other clusters).

For gene set enrichment analysis on all cells rather than just the persistent clonotypes, pseudobulk differential expression analysis was similarly performed between clusters using LIMMA⁸⁶, where each pseudobulk sample consisted of the TMM-normalized sum of the gene counts of all cells in a particular cluster from each sample. Comparisons were performed for each cluster vs the rest of the clusters. The roast() function within the LIMMA was then used on the 50 Hallmark gene sets⁵⁵ to identify enriched gene sets for each cluster comparison.

T cell transcriptome dimension reduction, clustering and trajectory analysis

To evaluate the T cell transcriptome results, Monocle3⁵⁶ was selected to perform normalization, batch correction, UMAP presentation⁸⁷, clustering⁸⁸ and trajectory analysis⁸⁹. Tetramer-positive T cell RNA

transcripts were pre-processed using 25 dimensions of PCA information and log normalization. The resulting transcript data was then batch corrected using donor/tetramer groupings and the Median CV of transcript coverage to control for sample quality. Cells were clustered using the “Leiden” algorithm⁸⁸ with a resolution of 2.3e-3.

T cell receptor sequence similarity and clustering analysis

To evaluate the structure of the tetramer-identified T cell repertoires, paired sequences were clustered into sequence-similar groups using the TCR sequence clustering algorithm TCRdist^{22,90}. A distance limit of 24 was used, a highly stringent limit which would allow up to 8 synonymous amino acid exchanges or 2 chemically dissimilar amino acid exchanges (including deletions or insertions) in the TRA and TRB CDR3 regions. Sequences were then clustered into efficient sequence-similar groupings (numbered clusters) using the Louvain communities package in Networkx⁹¹. For the purposes of TCR clustering, only sequence pairs with a single TRB chain and either one or two TRA chains were considered. In the instances of a TRB paired with two TRA sequences, two unique TRA/TRB pairs were generated from the TRB with each TRA to account for each potential expression pair.

TCR diversity analysis

Diversity analysis of the TCR clonotypes was performed using the diversity function by the vegan R package, using the provided “shannon” diversity index option. Visit data was randomly down-sampled as necessary to provide the same number of libraries at the beginning and ending visit for each donor and tetramer combination.

Statistics and reproducibility

Statistical tests in clonotype analysis were performed using the R programming language. The stats R package provided the Wilcoxon rank-sum test. Differential gene expression analysis was performed either at the single-cell level using the MAST R package or using pseudo-bulked cells aggregated according to various experimental variables (time point, cluster, clonal expansion) using the LIMMA R package, as noted in the figure legends. A Benjamini-Hochberg FDR-adjusted *p*-value of less than 0.05 was used to define differential gene expression unless otherwise noted in figure legends. Pearson’s product moment correlation coefficient tests were performed to assess correlations of clonotype frequencies at different time points in R.

No statistical method was used to determine sample size. No data were excluded from the analysis. The investigators were not blinded to allocation during experiments and outcome assessments.

Inclusion and ethics statement

This research is committed to inclusivity and actively recruited participants from diverse backgrounds to ensure the findings are representative of the broader population.

Benaroya Research Institute and University of Washington pursue excellence and high standards of performance, professionalism and ethical conduct. Both organizations strictly prohibit any form of discrimination against individuals on the basis of gender, race, age, religion, sexual orientation, veteran status or disability status.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

Source data are provided with this paper. All RNA-seq data has been submitted to GEO, and the series has been assigned the accession number GSE249632. The code used for TCR clustering analysis, comparison to Laing et al.⁷, and to generate TCR clustering figures is available at https://github.com/esford3/wen_2024. Source data are provided with this paper.

References

- Hand, T. W. & Kaech, S. M. Intrinsic and extrinsic control of effector T cell survival and memory T cell development. *Immunol. Res.* **45**, 46–61 (2009).
- DeWitt, W. S. et al. Dynamics of the cytotoxic T cell response to a model of acute viral infection. *J. Virol.* **89**, 4517–4526 (2015).
- Pogorelyy, M. V. et al. Precise tracking of vaccine-responding T cell clones reveals convergent and personalized response in identical twins. *Proc. Natl Acad. Sci. USA* **115**, 12704–12709 (2018).
- Ford, E. S. et al. Repeated mRNA vaccination sequentially boosts SARS-CoV-2-specific CD8(+) T cells in persons with previous COVID-19. *Nat. Immunol.* **25**, 166–177 (2024).
- Qi, Q. et al. Diversification of the antigen-specific T cell receptor repertoire after varicella zoster vaccination. *Sci. Transl. Med.* **8**, 332ra346 (2016).
- Minervina, A. A. et al. SARS-CoV-2 antigen exposure history shapes phenotypes and specificity of memory CD8(+) T cells. *Nat. Immunol.* **23**, 781–790 (2022).
- Laing, K. J. et al. Recruitment of naive CD4+ T cells by the recombinant zoster vaccine correlates with persistent immunity. *J. Clin. Invest.* **133**, e172634 (2023).
- Sureshchandra, S. et al. Single-cell profiling of T and B cell repertoires following SARS-CoV-2 mRNA vaccine. *JCI Insight* **6**, e153201 (2021).
- Pan, Y. G. et al. Vaccination reshapes the virus-specific T cell repertoire in unexposed adults. *Immunity* **54**, 1245–1256 e1245 (2021).
- Saggau, C. et al. The pre-exposure SARS-CoV-2-specific T cell repertoire determines the quality of the immune response to vaccination. *Immunity* **55**, 1924–1939 e1925 (2022).
- Waickman, A. T. et al. Dissecting the heterogeneity of DENV vaccine-elicited cellular immunity using single-cell RNA sequencing and metabolic profiling. *Nat. Commun.* **10**, 3666 (2019).
- Youngblood, B. et al. Effector CD8 T cells dedifferentiate into long-lived memory cells. *Nature* **552**, 404–409 (2017).
- Akondy, R. S. et al. Origin and differentiation of human memory CD8 T cells after vaccination. *Nature* **552**, 362–367 (2017).
- Hill, D. L. et al. Impaired HA-specific T follicular helper cell and antibody responses to influenza vaccination are linked to inflammation in humans. *Elife* **10**, e70554 (2021).
- Mahnke, Y. D., Brodie, T. M., Sallusto, F., Roederer, M. & Lugli, E. The who's who of T-cell differentiation: human memory T-cell subsets. *Eur. J. Immunol.* **43**, 2797–2809, (2013).
- Heineman, T. C., Cunningham, A. & Levin, M. Understanding the immunology of Shingrix, a recombinant glycoprotein E adjuvanted herpes zoster vaccine. *Curr. Opin. Immunol.* **59**, 42–48 (2019).
- Lal, H. et al. Efficacy of an adjuvanted herpes zoster subunit vaccine in older adults. *N. Engl. J. Med.* **372**, 2087–2096 (2015).
- Chen, L. & Flies, D. B. Molecular mechanisms of T cell co-stimulation and co-inhibition. *Nat. Rev. Immunol.* **13**, 227–242 (2013).
- Archila, L. L. & Kwok, W. W. Tetramer-guided epitope mapping: a rapid approach to identify HLA-restricted T-cell epitopes from composite allergens. *Methods Mol. Biol.* **1592**, 199–209 (2017).
- Voic, H. et al. Identification and characterization of CD4(+) T cell epitopes after shingrix vaccination. *J. Virol.* **94**, e01641-20 (2020).
- DeGottardi, Q. et al. Ontogeny of different subsets of yellow fever virus-specific circulatory CXCR5(+) CD4(+) T cells after yellow fever vaccination. *Sci. Rep.* **10**, 15686 (2020).
- Mayer-Blackwell, K. et al. TCR meta-clonotypes for biomarker discovery with tcrdist3 enabled identification of public, HLA-restricted clusters of SARS-CoV-2 TCRs. *Elife* **10**, e68605 (2021).
- Simpson, T. R., Quezada, S. A. & Allison, J. P. Regulation of CD4 T cell activation and effector function by inducible costimulator (ICOS). *Curr. Opin. Immunol.* **22**, 326–332, (2010).
- Takheaw, N., Earwong, P., Laopajon, W., Pata, S. & Kasinrerker, W. Interaction of CD99 and its ligand upregulates IL-6 and TNF-alpha upon T cell activation. *PLoS One* **14**, e0217393 (2019).
- Makino, Y. et al. Tyr724 phosphorylation of ELMO1 by Src is involved in cell spreading and migration via Rac1 activation. *Cell Commun. Signal* **13**, 35 (2015).
- Lippert, D. N. & Wilkins, J. A. Glia maturation factor gamma regulates the migration and adherence of human T lymphocytes. *BMC Immunol.* **13**, 21 (2012).
- Inclan-Rico, J. M. et al. Basophils prime group 2 innate lymphoid cells for neuropeptide-mediated inhibition. *Nat. Immunol.* **21**, 1181–1193 (2020).
- Huber, M. & Lohoff, M. IRF4 at the crossroads of effector T-cell fate decision. *Eur. J. Immunol.* **44**, 1886–1895 (2014).
- Betzler, A. C. et al. T cell specific BOB.1/OBF.1 expression promotes germinal center response and T helper cell differentiation. *Front. Immunol.* **13**, 889564 (2022).
- Good, S. R. et al. Temporal induction pattern of STAT4 target genes defines potential for Th1 lineage-specific programming. *J. Immunol.* **183**, 3839–3847 (2009).
- Scott, C. L. & Omilusik, K. D. ZEBs: novel players in immune cell development and function. *Trends Immunol.* **40**, 431–446 (2019).
- Dominguez, C. X. et al. The transcription factors ZEB2 and T-bet cooperate to program cytotoxic T cell terminal differentiation in response to LCMV viral infection. *J. Exp. Med.* **212**, 2041–2056 (2015).
- Tai, X. et al. Basis of CTLA-4 function in regulatory and conventional CD4(+) T cells. *Blood* **119**, 5155–5163 (2012).
- Liu, X. et al. Cutting edge: a critical role of B and T lymphocyte attenuator in peripheral T cell tolerance induction. *J. Immunol.* **182**, 4516–4520 (2009).
- Lyon de Ana, C., Arakcheeva, K., Agnihotri, P., Derosia, N. & Winandy, S. Lack of Ikaros deregulates inflammatory gene programs in T cells. *J. Immunol.* **202**, 1112–1123 (2019).
- Wang, J. et al. Coagulation factor V is a T-cell inhibitor expressed by leukocytes in COVID-19. *iScience* **25**, 103971 (2022).
- Parish, I. A. et al. The molecular signature of CD8+ T cells undergoing deletional tolerance. *Blood* **113**, 4575–4585 (2009).
- Puthalakath, H. et al. ER stress triggers apoptosis by activating BH3-only protein Bim. *Cell* **129**, 1337–1349 (2007).
- Ward-Kavanagh, L. K., Lin, W. W., Sedy, J. R. & Ware, C. F. The TNF receptor superfamily in co-stimulating and co-inhibitory responses. *Immunity* **44**, 1005–1019, (2016).
- Zhang, L. et al. MiR-22 regulated T cell differentiation and hepatocellular carcinoma growth by directly targeting Jarid2. *Am. J. Cancer Res.* **11**, 2159–2173 (2021).
- Brandstadter, J. D. & Maillard, I. Notch signalling in T cell homeostasis and differentiation. *Open Biol.* **9**, 190187 (2019).
- Shaw, L. A. et al. Id2 reinforces TH1 differentiation and inhibits E2A to repress TFH differentiation. *Nat. Immunol.* **17**, 834–843 (2016).
- Li, J. et al. Role of Tet2 in Regulating Adaptive and Innate Immunity. *Front. Cell Dev. Biol.* **9**, 665897 (2021).
- Pesu, M. et al. T-cell-expressed proprotein convertase furin is essential for maintenance of peripheral immune tolerance. *Nature* **455**, 246–250 (2008).
- Gorman, J. A. et al. The cytoskeletal adaptor protein IQGAP1 regulates TCR-mediated signaling and filamentous actin dynamics. *J. Immunol.* **188**, 6135–6144 (2012).
- Maine, C. J., Teijaro, J. R., Marquardt, K. & Sherman, L. A. PTPN22 contributes to exhaustion of T lymphocytes during chronic viral infection. *Proc. Natl Acad. Sci. USA* **113**, E7231–E7239 (2016).
- Verdeil, G. MAF drives CD8(+) T-cell exhaustion. *Oncoimmunology* **5**, e1082707 (2016).
- Li, H. et al. Dysfunctional CD8 T cells form a proliferative, dynamically regulated compartment within human melanoma. *Cell* **176**, 775–789 e718 (2019).

49. Yasuda, T. et al. Dok-1 and Dok-2 are negative regulators of T cell receptor signaling. *Int. Immunol.* **19**, 487–495 (2007).
50. Chen, C. F. et al. Regulation of T cell proliferation by JMJD6 and PDGF-BB during chronic hepatitis B infection. *Sci. Rep.* **4**, 6359 (2014).
51. Bouguermouh, S. et al. CD47 expression on T cell is a self-control negative regulator of type 1 immune response. *J. Immunol.* **180**, 8073–8082 (2008).
52. Chen, X. et al. Expression of costimulatory TNFR2 induces resistance of CD4⁺FoxP3⁺ conventional T cells to suppression by CD4⁺FoxP3⁺ regulatory T cells. *J. Immunol.* **185**, 174–182 (2010).
53. Shibagaki, N. et al. Functional analysis of CD82 in the early phase of T cell activation: roles in cell adhesion and signal transduction. *Eur. J. Immunol.* **28**, 1125–1133 (1998).
54. Su, L. L., Iwai, H., Lin, J. T. & Fathman, C. G. The transmembrane E3 ligase GRAIL ubiquitinates and degrades CD83 on CD4 T cells. *J. Immunol.* **183**, 438–444, (2009).
55. Liberzon, A. et al. The Molecular Signatures Database (MSigDB) hallmark gene set collection. *Cell Syst.* **1**, 417–425 (2015).
56. Cao, J. et al. The single-cell transcriptional landscape of mammalian organogenesis. *Nature* **566**, 496–502 (2019).
57. Galloway, A. et al. RNA-binding proteins ZFP36L1 and ZFP36L2 promote cell quiescence. *Science* **352**, 453–459 (2016).
58. Barata, J. T., Durum, S. K. & Seddon, B. Flip the coin: IL-7 and IL-7R in health and disease. *Nat. Immunol.* **20**, 1584–1593 (2019).
59. de Almagro, M. C. & Vucic, D. The inhibitor of apoptosis (IAP) proteins are critical regulators of signaling pathways and targets for anti-cancer therapy. *Exp. Oncol.* **34**, 200–211, (2012).
60. Chen, J. et al. NR4A transcription factors limit CAR T cell function in solid tumours. *Nature* **567**, 530–534 (2019).
61. Kory, N. et al. SFXN1 is a mitochondrial serine transporter required for one-carbon metabolism. *Science* **362**, eaat9528 (2018).
62. Li, S. X. et al. Octameric structure of the human bifunctional enzyme PAICS in purine biosynthesis. *J. Mol. Biol.* **366**, 1603–1614 (2007).
63. Martini, S. et al. A genetically-encoded crosslinker screen identifies SERBP1 as a PKCepsilon substrate influencing translation and cell division. *Nat. Commun.* **12**, 6934 (2021).
64. Ueda, Y., Kondo, N. & Kinashi, T. MST1/2 balance immune activation and tolerance by orchestrating adhesion, transcription, and organelle dynamics in lymphocytes. *Front. Immunol.* **11**, 733 (2020).
65. Zhang, S. et al. AKAP13 couples GPCR signaling to mTORC1 inhibition. *PLoS Genet.* **17**, e1009832 (2021).
66. Bachmaier, K. et al. Negative regulation of lymphocyte activation and autoimmunity by the molecular adaptor Cbl-b. *Nature* **403**, 211–216 (2000).
67. Wang, D. et al. Enhanced adaptive immunity in mice lacking the immunoinhibitory adaptor Hacs1. *FASEB J.* **24**, 947–956 (2010).
68. Mrschik, M. & Ryan, K. M. Lysosomal proteins in cell death and autophagy. *FEBS J.* **282**, 1858–1870, (2015).
69. Choi, J. H. et al. Essential cell-extrinsic requirement for PDIA6 in lymphoid and myeloid development. *J. Exp. Med.* **217**, e20190006 (2020).
70. Sung, B. Y. et al. Wnt activation promotes memory T cell poly-functionality via epigenetic regulator PRMT1. *J. Clin. Investig.* **132**, e140508 (2022).
71. Bourges, I. et al. Structural organization of mitochondrial human complex I: role of the ND4 and ND5 mitochondria-encoded subunits and interaction with prohibitin. *Biochem J.* **383**, 491–499 (2004).
72. Sylvester, J. E., Fischel-Ghodsian, N., Mougey, E. B. & O'Brien, T. W. Mitochondrial ribosomal proteins: candidate genes for mitochondrial disease. *Genet Med* **6**, 73–80 (2004).
73. Kwon, K. J. et al. Different human resting memory CD4⁺ T cell subsets show similar low inducibility of latent HIV-1 proviruses. *Sci. Transl. Med.* **12**, eaax6795 (2020).
74. Minervina, A. A. et al. Primary and secondary anti-viral response captured by the dynamics and phenotype of individual T cell clones. *Elife* **9**, e53704 (2020).
75. Pauken, K. E. et al. The PD-1 pathway regulates development and function of memory CD8⁺ T cells following respiratory viral infection. *Cell Rep.* **31**, 107827 (2020).
76. Jubel, J. M., Barbati, Z. R., Burger, C., Wirtz, D. C. & Schildberg, F. A. The role of PD-1 in acute and chronic infection. *Front. Immunol.* **11**, 487 (2020).
77. Parry, H. M. et al. PD-1 is imprinted on cytomegalovirus-specific CD4⁺ T cells and attenuates Th1 cytokine production whilst maintaining cytotoxicity. *PLoS Pathog.* **17**, e1009349 (2021).
78. Wirth, T. C. et al. Repetitive antigen stimulation induces stepwise transcriptome diversification but preserves a core signature of memory CD8⁺ T cell differentiation. *Immunity* **33**, 128–140 (2010).
79. Omilusik, K. D. et al. Transcriptional repressor ZEB2 promotes terminal differentiation of CD8⁺ effector and memory T cell populations during infection. *J. Exp. Med.* **212**, 2027–2039 (2015).
80. Meckiff, B. J. et al. Imbalance of regulatory and cytotoxic SARS-CoV-2-reactive CD4⁺ T cCells in COVID-19. *Cell* **183**, 1340–1353 e1316 (2020).
81. Wen, X. et al. Increased islet antigen-specific regulatory and effector CD4⁺ T cells in healthy individuals with the type 1 diabetes-protective haplotype. *Sci. Immunol.* **5**, eaax8767 (2020).
82. Finak, G. et al. MAST: a flexible statistical framework for assessing transcriptional changes and characterizing heterogeneity in single-cell RNA sequencing data. *Genome Biol.* **16**, 278 (2015).
83. Lun, A. T., McCarthy, D. J. & Marioni, J. C. A step-by-step workflow for low-level analysis of single-cell RNA-seq data with Bioconductor. *F1000Res* **5**, 2122 (2016).
84. Squair, J. W. et al. Confronting false discoveries in single-cell differential expression. *Nat. Commun.* **12**, 5692 (2021).
85. Robinson, M. D. & Oshlack, A. A scaling normalization method for differential expression analysis of RNA-seq data. *Genome Biol.* **11**, R25 (2010).
86. Ritchie, M. E. et al. limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.* **43**, e47 (2015).
87. Becht, E. et al. Dimensionality reduction for visualizing single-cell data using UMAP. *Nat. Biotechnol.* **37**, 38–44 (2018).
88. Traag, V. A., Waltman, L. & van Eck, N. J. From Louvain to Leiden: guaranteeing well-connected communities. *Sci. Rep.* **9**, 5233 (2019).
89. Liu, S. X., Gustafson, H. H., Jackson, D. L., Pun, S. H. & Trapnell, C. Trajectory analysis quantifies transcriptional plasticity during macrophage polarization. *Sci. Rep.* **10**, 12273 (2020).
90. Dash, P. et al. Quantifiable predictive features define epitope-specific T cell receptor repertoires. *Nature* **547**, 89–93 (2017).
91. Hagberg AA, Schult D. A., Swart P. J. in *Proceedings of the 7th Python in Science Conference*. (ed Vaught T Varoquaux G, Millman J) 11–15.

Acknowledgements

This study was supported by NIH contract HHSN272201400049C. The authors would like to thank the Genomics Core and Bioinformatics Core at Benaroya Research Institute for scRNA-seq and data processing, the Translational Core at Benaroya Research Institute for patient recruitment, Olivia Doyle and Anne Hocking for critical reading of the manuscript and Cynthia Cousens-Jacobs for manuscript preparation.

Author contributions

D.M.K. and W.W.K. designed the study. X.W. performed experiments. E.S.F., X.W., A.K.H., S.R.P. and W.W.K. analyzed data and generated figures. X.W., A.K.H., S.R.P. and W.W.K. wrote the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41467-025-57562-7>.

Correspondence and requests for materials should be addressed to William W. Kwok.

Peer review information *Nature Communications* thanks Alexandra Sharland and the other anonymous reviewers for their contribution to the peer review of this work. A peer review file is available.

Reprints and permissions information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2025