

Precursors of exhausted T cells are pre-emptively formed in acute infection

<https://doi.org/10.1038/s41586-024-08451-4>

Received: 10 February 2024

Accepted: 25 November 2024

Published online: 08 January 2025

Open access

 Check for updates

Talyn Chu^{1,9,10}, Ming Wu^{1,10}✉, Barbara Hoellbacher^{2,3,10}, Gustavo P. de Almeida^{1,4,10}, Christine Wurmser^{1,4}, Jacqueline Berner^{1,4}, Lara V. Donhauser^{1,4}, Ann-Katrin Gerullis^{1,4}, Siran Lin^{1,4}, J. Diego Cepeda-Mayorga^{1,4}, Iman I. Kilb^{1,4,5}, Lukas Bongers¹, Fabio Toppeta¹, Philipp Strobl¹, Ben Youngblood⁶, Anna M. Schulz^{1,4}, Alfred Zippelius⁷, Percy A. Knolle^{4,5}, Matthias Heinig^{2,8}✉, Carl-Philipp Hackstein^{1,4,5}✉ & Dietmar Zehn^{1,4}✉

T cell exhaustion limits effector T cell function in chronic infection and tumours^{1,2}. The development of these hypofunctional T cells and of their precursors was considered to require stimulatory conditions that are met only after persistent exposure to antigen and inflammation. Here we show, however, that similar T cell populations exist in the early phase of acute infections^{1,2}. At that stage, the early developing TCF1⁺ precursor population exhibits an unexpected diversity; it includes precursors of normal memory T cells, but also cells with phenotypic, gene-expression and epigenetic profiles that resemble those of precursors of exhausted T cells found in chronic infections. We show that high ligand affinity promotes and PD-1 signalling restricts the development of these precursors. Although the exhausted precursors are at first found frequently, they decline without being completely lost in infections that the immune system resolves. We therefore conclude that precursor T cells with at least two distinct phenotypes are pre-emptively generated irrespective of the outcome of an infection.

T cell exhaustion refers to the formation of hypofunctional T cells with a reduced effector function and reduced potency in eliminating target cells in the setting of prolonged antigen exposure in chronic infections and tumours. Characterized by a reduced cytokine-secretion capacity and upregulated expression of inhibitory receptors (PD-1, LAG3, 4-1BB or OX40)^{1–4}, exhausted CD8 T cells also express a unique pattern of transcription factors, which includes high levels of TOX, NR4A1, NR4A2 and EOMES, and lower levels of T-bet (refs. 5–12) and of hypoxia-inducible factor 1 α (HIF-1 α)¹³. Moreover, the hypofunctional phenotype is epigenetically enforced and ‘exhaustion’ imprints are retained even when cells are transferred from chronic into acute infections¹⁴, and after their reactivation through checkpoint inhibitors^{2,15–18}. Although this exhausted state is crucial for preventing immunopathology in the context of chronic antigen stimulation, the low effector capacity of exhausted T cells is a major obstacle for effective immunotherapy against chronic infections and tumours^{19–21}. Understanding the kinetics and mechanisms that drive the induction and maintenance of exhausted T cells is therefore essential for devising advanced immunotherapeutic solutions that aim to prevent or overcome their development.

The vast majority of exhausted T cells are terminally differentiated effector cells, which lack the capacity to undergo further expansion. Nonetheless, the long-term maintenance of exhausted T cell populations, which can persist for years in people with tumours or chronic

hepatitis C infections, requires a population of precursor cells, with a stem-cell-like function. These cells have been shown to be produced in the early phase of chronic infections²⁴, but so far, they have been considered to be formed exclusively in chronic infections and tumours, or in infections that the immune system struggles to eliminate—such as severe COVID-19 infections²⁵. In sharp contrast to this view, here, by systematically studying early differentiation events in acute infection, we observed substantial heterogeneity in the cell population expressing TCF1, which is thought to give rise to memory precursor cells. We discovered that this population contains a subset that exhibits key features of precursors of exhausted T (T_{pe}) cells¹. These precursors express TOX and PD-1, and have epigenetic and transcriptional profiles similar to those of exhausted T cells. We further show that these T cells are driven by strong TCR stimulation and that their development is promoted by ligand affinity and antagonized by PD-1 signalling. Accordingly, our observations challenge the concept that precursors of exhausted T cells are exclusively generated in chronic infections or tumours. Instead, we show that these cells are pre-emptively or constitutively formed in infections that the immune system successfully eliminates.

Exhausted T cells exist in acute infections

Precursors of exhausted effector T cells have been shown to be detectable in infections that become chronic as early as five days after

¹Division of Animal Physiology and Immunology, School of Life Sciences Weihenstephan, Technical University of Munich, Freising, Germany. ²Institute of Computational Biology, Helmholtz Zentrum München Deutsches Forschungszentrum für Gesundheit und Umwelt, Neuherberg, Germany. ³Department of Informatics, Technical University of Munich, Garching, Germany. ⁴Center for Infection Prevention (ZIP), School of Life Sciences Weihenstephan, Technical University of Munich, Freising, Germany. ⁵Institute of Molecular Immunology, School of Medicine and Health, Technical University of Munich, Munich, Germany. ⁶Department of Immunology, St. Jude Children’s Research Hospital, Memphis, TN, USA. ⁷Cancer Immunology, Department of Biomedicine, University Hospital Basel, Basel, Switzerland. ⁸Department of Computer Science, TUM School of Computation, Information and Technology, Technical University of Munich, Garching, Germany. ⁹Present address: Cancer Immunology, Department of Biomedicine, University Hospital Basel, Basel, Switzerland. ¹⁰These authors contributed equally: Talyn Chu, Ming Wu, Barbara Hoellbacher, Gustavo P. de Almeida. ✉e-mail: ming.wu@tum.de; matthias.heinig@helmholtz-muenchen.de; carl-philipp.hackstein@tum.de; dietmar.zehn@tum.de

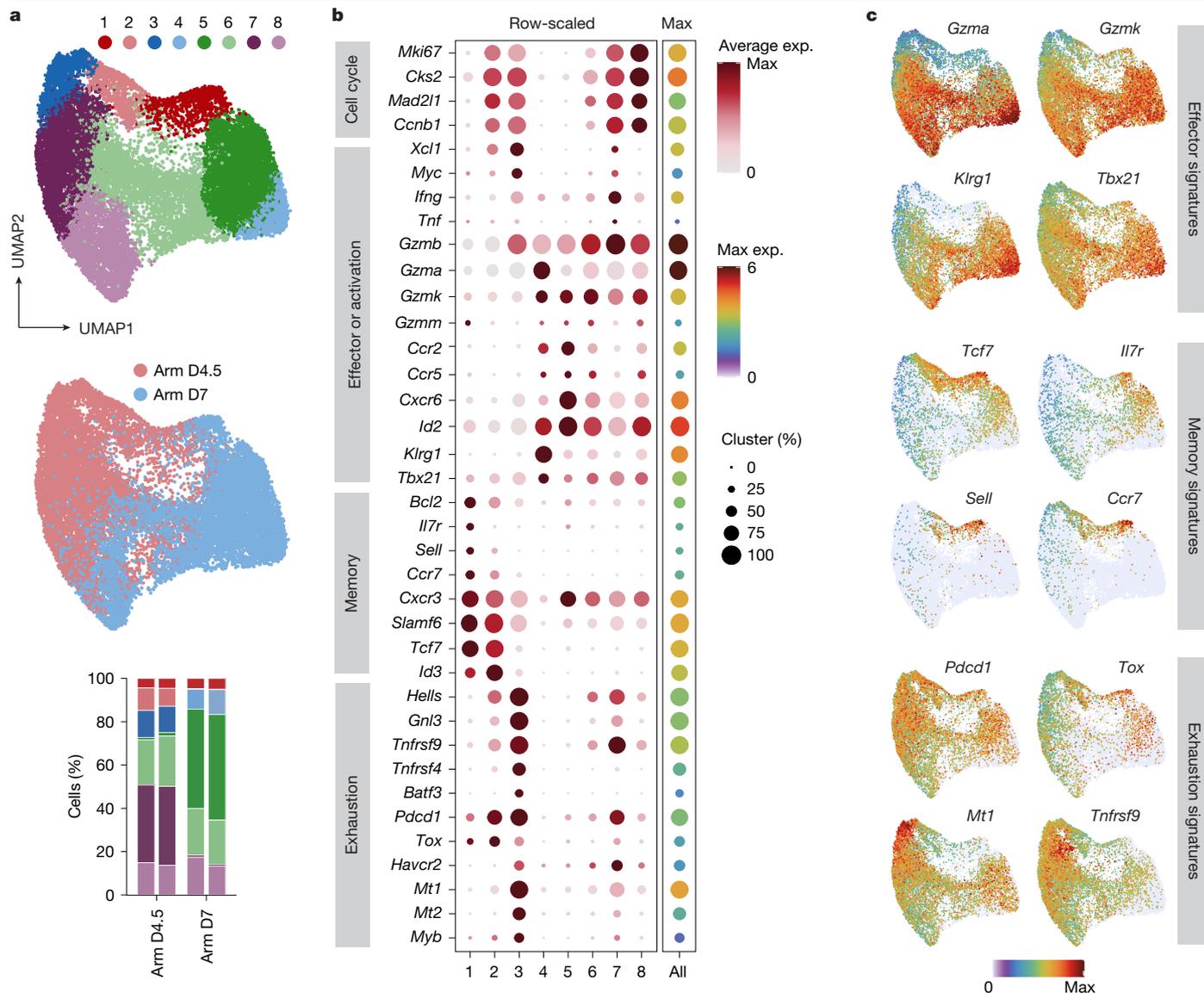


Fig. 1 | Acute LCMV-specific T cell populations contain cells with transcriptional similarity to exhausted T cells. **a–c**, A published dataset²⁶ was used in which LCMV-specific, TCR-transgenic P14 T cells were transferred into host mice that were subsequently challenged with acute LCMV Armstrong (Arm) infection. P14 cells were re-isolated on days (D) 4.5 and 7 after infection. **a**, Uniform manifold approximation and projection (UMAP) shows eight identified clusters (top), the distribution of the two time points projected

infection^{1,24} and thus long before the infection turns into a chronic phase. This early commitment raised our interest in investigating whether similar cells might also arise in the early phase of infections that eventually resolve. To interrogate cell heterogeneity in an unbiased manner, we used the public single-cell RNA sequencing (scRNA-seq) dataset GSE119943 (ref. 26). These data stem from experiments in which lymphocytic choriomeningitis virus (LCMV)-specific, TCR-transgenic P14 T cells were transferred into host mice, followed by an acute LCMV infection. Single-cell-resolved levels of gene expression were determined after isolating the P14 T cells from the spleen on days 4.5 or 7 after infection. By reanalysing the datasets, we identified eight clusters with distinct gene-expression signatures (Fig. 1a,b). Clusters 4–8 seemed to be effector cells typically associated with acute infection, expressing high levels of *Id2* and varying levels of *Tbx21* and effector molecules (*Gzmb*, *Gzma*, *Gzmk*, *Ifng* and *Tnf*). Cluster 4 predominantly included day-7 T cells expressing a transcriptional signature similar to that of

over the Louvain clusters (middle) and an illustration of the relative size of the cluster in each sample (bar graph at the bottom). **b**, Signature dot plot for the clusters. Colours encode average normalized expression values within the clusters, and circle size indicates the percentage of the cluster expressing the gene of interest. **c**, Feature plots illustrating selected effector, memory and exhaustion markers. Colours encode normalized expression values. Data are derived from a single previously published dataset.

terminal effector T cells, with enhanced levels of genes such as *Klrg1* and those encoding granzymes (*Gzma* and *Gzmk*; Fig. 1b,c). Clusters 5 (day 7) and 6 (days 4.5 and 7) showed an early effector phenotype, with expression of granzymes but with lower expression of *Klrg1* compared with the terminal effectors seen on day 7. Clusters 7 (day 4.5) and 8 (days 4.5 and 7) contained high levels of cycling effector cells, with upregulated cell-cycle signatures such as *Mki67*, *Cks2*, *Mad2l1* and *Ccnb1* (Fig. 1b,c), and cluster 7 also expressed the highest levels of *Ifng*, *Tnfa* and *Havcr2*. By contrast, cluster 1, which contained T cells from both day-4.5 and day-7 datasets (Fig. 1b,c) exhibited a transcriptional signature known for memory precursor cells (T_{mpc}), which included expression of *Tcf7*, *Il7r*, *Sell* and *Id3* and downregulation of effector molecules such as *Klrg1*, *Gzma*, *Gzmk* and *Tbx21* (Fig. 1b,c). Cluster 2 (day 4.5) contained a combination of markers of progenitor cells (*Tcf7*, *Slamf6* and *Id3*) and markers of T cell exhaustion (*Xcl1*, *Pdcd1* and *Tox*). T cells in the day-4.5 restricted cluster 3 (Fig. 1b) expressed

high levels of early activation markers, including *Xcl1* and *Myc*, high levels of well-known genes associated with exhaustion, such as *Pdcd1*, *Tox*, *Tnfrsf9* (4-1BB), *Myb* and *Havcr2* (TIM3), low levels of traditional effector molecules (*Id2*, *Tbx21*, *Cxcr6* and *Ccr5*) and reduced levels of granzymes (*Gzma*, *Gzmk* and *Gzmm*) (Fig. 1b,c). Thus, the phenotype of cells found in clusters 2 and 3 suggests that cells with transcriptional similarity to that seen in T cell exhaustion are detectable in the early phase of acute infection. This encouraged us to examine the origin and fate of these cells in more detail.

T_{ex} precursors in acute infections

To investigate the clusters showing signs of exhaustion in more detail, we specifically reanalysed the *Tcf7*-overexpressing precursor clusters 1 and 2 and the exhausted cluster 3. (Fig. 2a). Differential expression analysis comparing the two precursor clusters identified 1,482 genes that were upregulated in cluster 2 compared with cluster 1. Because *Tox*, *Lag3*, *Pdcd1*, *Ikzf2*, *Tnfrsf9* and *Myb* were detected in cluster 2, we termed these cells precursors of exhausted T (T_{pex}) cells (Fig. 2a–c). By contrast, among the 151 upregulated genes in cluster 1, we identified some that are typically associated with conventional memory cells. We therefore refer to these cells as memory precursor T (T_{mpc}) cells (Fig. 2b). Of note, T_{pex} cells shared several features with the day 4.5 exhausted T (T_{ex}) cells in cluster 3 that we had included in our analysis as a reference population (Fig. 2c,d). To further examine differentially expressed signatures identified in two types of precursors, we extracted acute and exhausted signatures generated from two previously published studies^{5,27}. Gene set enrichment analysis (GSEA) revealed that the cells from cluster 2 showed an enrichment of exhaustion gene signatures, whereas those in cluster 1 showed an enrichment of the corresponding gene signatures derived from cells from an acute infection (Fig. 2e). We also performed a pseudotime and velocity analysis on the selected clusters (Fig. 2f). The lowest pseudotime values were observed at the border between the T_{pex} and the T_{mpc} clusters and increased from there into the far ends of both clusters, suggesting that unique differentiation avenues result in the existence of these two types of precursor cells. Pseudotime further increased from the T_{pex} to the T_{ex} cluster, and plateaued in the latter, positioning the T_{ex} population downstream of the T_{pex} one. Because these results pointed towards the existence of a common precursor of T_{mpc} and T_{pex}, we screened for such a cell before day 4.5. scRNA-seq of P14 T cells that were re-isolated at day 3.25 after infection revealed two different *Tcf7*-expressing precursor clusters (Extended Data Fig. 1a,b). Notably, these populations showed only marginal differences in the expression of genes such as *Pdcd1*, *Tnfrsf9*, *Ccr7* and *Il7r*, and no differences in the expression of *Tox* (Extended Data Fig. 1b,c). This shows that distinguishable T_{pex} and T_{mpc} populations arise after day 3.25 and are visible at day 4.5 after infection. Altogether, our analysis reveals that the early precursor population contains not only cells that are related to classical memory precursor cells, but also a subpopulation that resembles the T_{pex} cells found in chronic infections.

Exhaustion markers in early T_{pex} cells

Next, we sought to confirm the presence of T_{pex} cells through flow-cytometry-based analysis in various acute infections. We transferred P14 or OT-1 TCR-transgenic T cells into CD45.1 congenic host mice, which were subsequently infected either with the acute LCMV Armstrong strain (Fig. 3a) or with a recombinant vesicular stomatitis virus that expresses ovalbumin (VSV-N4Ova) (Extended Data Fig. 2a). Transgenic T cells were re-isolated either five or seven days after infection and analysed by flow cytometry. We excluded KLRG1-positive terminally differentiated cells and selected from the remaining cells the population with the highest or lowest PD-1 expression (Fig. 3a and Extended Data Fig. 2a). To identify TCF1- and TOX-positive cells, we established a gating strategy (shown in Extended Data Fig. 3). We used

total endogenous CD8 T cells as reference. We gated on CD44^{hi} cells, and selected KLRG1⁻ cells. The resulting cells revealed clear TCF1 populations that showed robust differences in TOX expression. Moreover, endogenous CD44^{hi} CD62L⁺ naive cells acted as an internal negative control for TOX expression (Extended Data Fig. 3). Applying these gates to the P14 and OT-1 T cells (Fig. 3a and Extended Data Fig. 2a), we found that PD-1^{hi} and PD-1^{lo} cells differed significantly in their expression of TOX, which was high in the PD-1^{hi} group (Fig. 3a and Extended Data Fig. 2a). Moreover, both PD-1^{hi} and PD-1^{lo} cells contained a notable population of TCF1⁺ precursor cells expressing TOX in the PD-1^{hi} but not in the PD-1^{lo} group (Fig. 3a and Extended Data Fig. 2a). This indicates that cells expressing signs of exhaustion and markers of progenitor cells can also be found using flow-cytometry analysis. Of note, upregulation of NUR77 in both PD-1^{hi} and PD-1^{lo} P14 T cells suggests that both subsets were activated and received recent TCR signalling (Fig. 3b). NUR77 levels were slightly higher in the PD-1^{hi} P14 T cells, which indicates that the PD-1^{hi} phenotype is specifically induced in T cells receiving higher levels of TCR signalling.

To further examine the diversity between the two types of precursors, we used a previously generated P14 *Tcf7*^{gfp(bright)mCherry} reporter mouse that shows very bright GFP fluorescence, compared with a previously generated mouse¹⁷, and enables a better selection of early TCF1-expressing precursors. For technical reasons, the P14 TCF1 reporter cells were adoptively transferred into Vβ5 TCRβ-chain-only transgenic hosts to prevent the rejection of transferred mCherry-expressing reporter cells during LCMV infections. These mice contain a polyclonal but restricted TCR repertoire. They have endogenously rearranged TCRα chains but they express the TCRβ chain that is also contained in OT-1 TCR-transgenic mice. This fixed TCRβ chain biases the TCR repertoire and prevents rejection of the *Tcf7*^{gfp(bright)mCherry} cells. Host mice were then infected with LCMV seven days after infection. TCF1-expressing P14 precursors were isolated, sorted into PD-1^{lo} or PD-1^{hi} precursors (Extended Data Fig. 2b) and submitted for RNA-seq. Principal component analysis (PCA) showed that the component capturing most of the variance in the data corresponds to the PD-1 status of the samples (Extended Data Fig. 2c). We performed differential gene-expression analysis comparing the PD-1^{hi} with the PD-1^{lo} precursors and found a total of 2,550 differentially expressed genes. Of these, 1,236 genes were more highly expressed in the PD-1^{hi} population and 1,314 in the PD-1^{lo} population ($P_{\text{adj}} < 0.05$ and $\log_2\text{FC} > \log_2(1.5)$) (Fig. 3c and Extended Data Fig. 2d). Next, we used a different public Gene Expression Omnibus (GEO) dataset (GSE142687) to compare the transcriptional differences between our PD-1^{hi} and PD-1^{lo} precursors with the differences between precursors taken from chronically infected LCMV Docile and acutely infected LCMV Armstrong at day 21 after infection²⁴. Transcriptional changes of the two comparisons correlated with $r = 0.517$ (Fig. 3d). GSEA using lists of genes significantly upregulated in PD-1^{hi} or PD-1^{lo} cells showed an enrichment in T_{pex} or T_{mpc}, respectively (Extended Data Fig. 2e,f). All of these observations show that PD-1^{hi} sorted precursors share a strong exhaustion signature, as seen in the T_{pex} cells described in chronic infections^{1,24}. In addition, we found that PD-1^{hi} precursors transcribed increased levels of receptors associated with exhaustion (*Pdcd1*, *Havcr2* and *Tnfrsf9*) as well as other genes that are significantly upregulated in exhaustion, such as *Tox*, *Tox2* and *Ikzf2* (Fig. 3e). By contrast, PD-1^{lo} precursors shared characteristics with conventional memory cells, with high expression of *Il7r*, *Eomes*, *Ccr7* and *Id3* (Extended Data Fig. 2g). Among the genes that are differentially expressed between the two types of precursors, we noted that PD-1^{hi} precursors transcribed lower levels of *Id2* and of effector cytokines (*Ifng* and *Tnf*) (Fig. 3e). To confirm the reduced expression of effector cytokines at the protein level, we isolated total splenocytes on day 4.5 after infection and briefly re-stimulated ex vivo using LCMV-derived gp33 and np396 peptides. Again, we found that T_{pex}-like cells showed an impaired capacity to co-produce TNF and IFNγ on in vitro re-stimulation, compared with their T_{mpc} counterparts (Fig. 3f). Together, the presented data led us

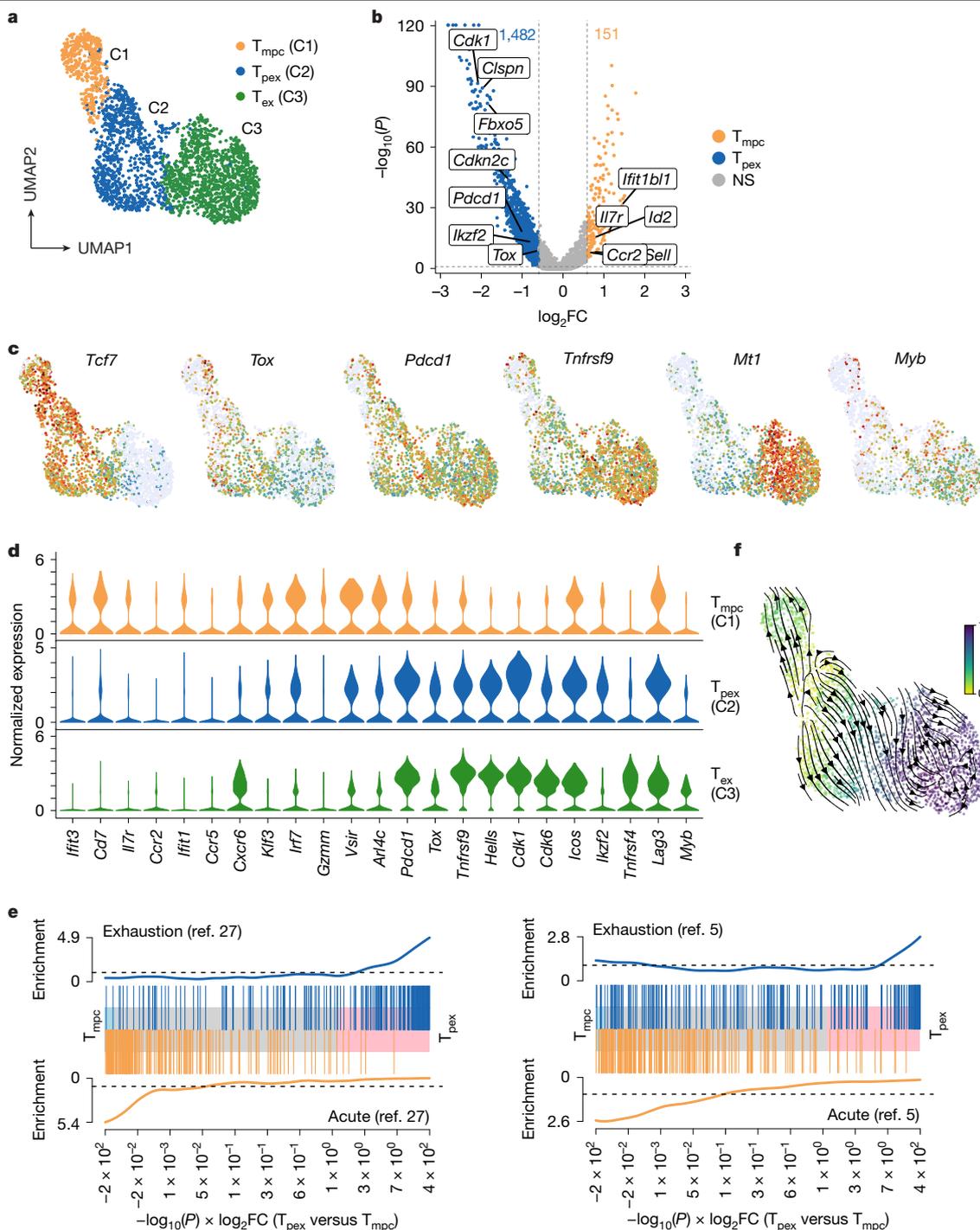


Fig. 2 | Cells similar to memory precursor T cells and to exhausted precursors exist early in acute infection. Day-4.5 cells from Fig. 1 located in progenitor clusters 1 and 2 and T_{ex} cluster 3 were selected, re-clustered and analysed. **a**, T_{mpc} , T_{pex} and T_{ex} clusters after re-clustering. **b**, Volcano plot of differential gene expression between the T_{mpc} cluster and the T_{pex} cluster (as shown in **a**). Genes with $P_{adj} < 0.05$ and $\log_2FC > \log_2(1.5)$ were considered differentially expressed. Significance was calculated with Wilcoxon rank-sum test and adjusted for multiple testing with Benjamini-Hochberg correction. **c**, Feature plots illustrating the representative markers. **d**, Violin plots showing expression levels of selected signature genes by cells assigned to the new clusters depicted in **a**. **e**, GSEA for exhausted (top) and acute (bottom) signatures from two distinct sources (left, ref. 27; right, ref. 5) on T_{pex} (cluster 2) versus T_{mpc} (cluster 1). The background genes were sorted by differential expression levels between these two clusters, weighted by the significance obtained with a Wilcoxon rank-sum test. **f**, Combined velocity and pseudotime analysis using the subclusters shown in **a**. Data are derived from a single previously published dataset.

Colours encode normalized expression values. **d**, Violin plots showing expression levels of selected signature genes by cells assigned to the new clusters depicted in **a**. **e**, GSEA for exhausted (top) and acute (bottom) signatures from two distinct sources (left, ref. 27; right, ref. 5) on T_{pex} (cluster 2) versus T_{mpc} (cluster 1). The background genes were sorted by differential expression levels between these two clusters, weighted by the significance obtained with a Wilcoxon rank-sum test. **f**, Combined velocity and pseudotime analysis using the subclusters shown in **a**. Data are derived from a single previously published dataset.

to conclude that precursor cells resembling chronic-infection-associated T_{pex} cells are formed in the early phase of acute infections. Furthermore, the data show that a diverse set of precursor T cells is generated in the early phase of infections independently of whether or not the infection eventually becomes chronic or resolved.

Epigenetics of T_{pex} cells in acute infections

A key element of T_{pex} commitment is the epigenetic imprinting of the exhausted phenotype^{14–18}. Accordingly, after identifying T_{pex} cells through flow analysis and transcriptional profiling, we were wondering

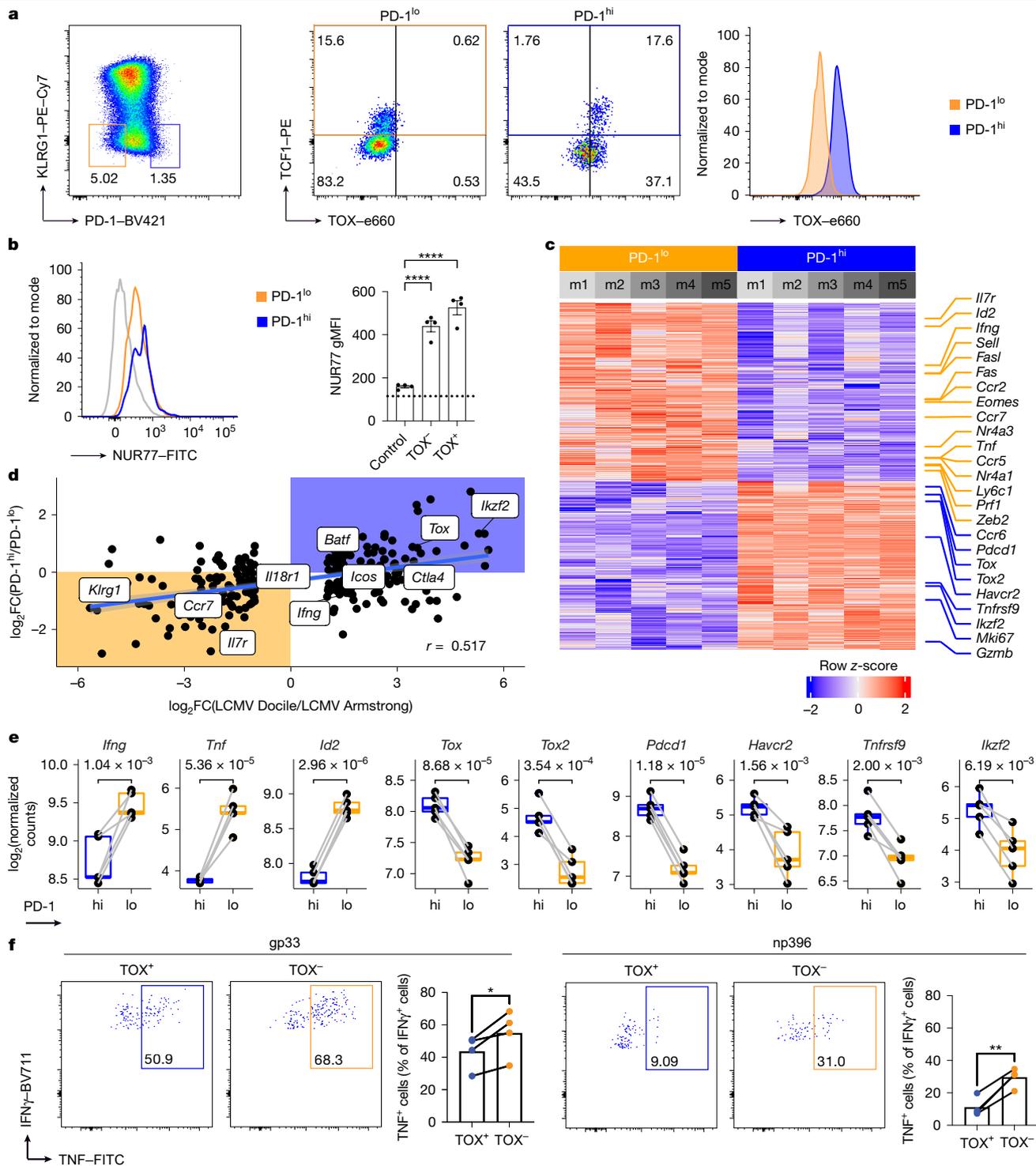


Fig. 3 | Transcriptional changes between PD-1^{hi} and PD-1^{lo} precursors reveal an exhausted phenotype. **a**, CD45.1 congenic, TCR-transgenic P14 (10³) T cells were transferred into C57BL/6 hosts, which were then infected with LCMV Armstrong. CD44⁺ KLRG1⁺ PD-1^{hi} and PD-1^{lo} populations were selected and analysed for TCF1 and TOX expression seven days after infection. **b**, NUR77 expression in PD-1^{hi} and PD-1^{lo} subsets of transferred P14 T cells on day 4.5 after infection with LCMV Armstrong. *n* = 4; mean ± s.e.m.; *****P* < 0.0001 (ordinary one-way ANOVA with Dunnett's multiple comparisons test). gMFI, geometric mean fluorescence intensity. **c**, Heat map of differentially expressed genes (*P*_{adj} < 0.05 and |log₂FC| > 0.58) between PD-1^{hi} and PD-1^{lo} P14. **d**, Correlation of the log₂FC between PD-1^{hi} and PD-1^{lo} P14 with the log₂FC between acute LCMV Armstrong and chronic LCMV Docile infection at day 21 after infection. Line represents a linear model regression and the bands show a confidence interval

of 95%. **e**, Log₂-normalized counts of bulk RNA-seq with paired PD-1^{hi} and PD-1^{lo} P14 samples from the same donor mouse. Significance values were calculated using empirical Bayes moderated *t*-statistics and were adjusted with Benjamini-Hochberg correction. *n* = 5 for all groups; lines represent mean and symbols represent individual mice. Box plots show median, first and third quartiles (hinges); lines show the smallest or largest observation within a distance from the nearest hinge of 1.5 times the size of the box. Observations outside this range are shown as outliers. **f**, Comparison of TNF co-production by IFN γ ⁺ cells after brief ex vivo stimulation with gp33 or np396 peptide. Spleen cells were isolated on day 4.5 after infection with LCMV Armstrong and pre-gated as KLRG1⁺ TCF1⁺ TOX⁺ (TOX⁺) or KLRG1⁺ TCF1⁺ TOX⁻ (TOX⁻) T cells. *n* = 4; **P* = 0.0417, ***P* = 0.0046 (two-tailed paired *t*-tests). Data are representative of three (**a**) or two (**b, f**) independent experiments or derived from a single experiment (**c-e**).

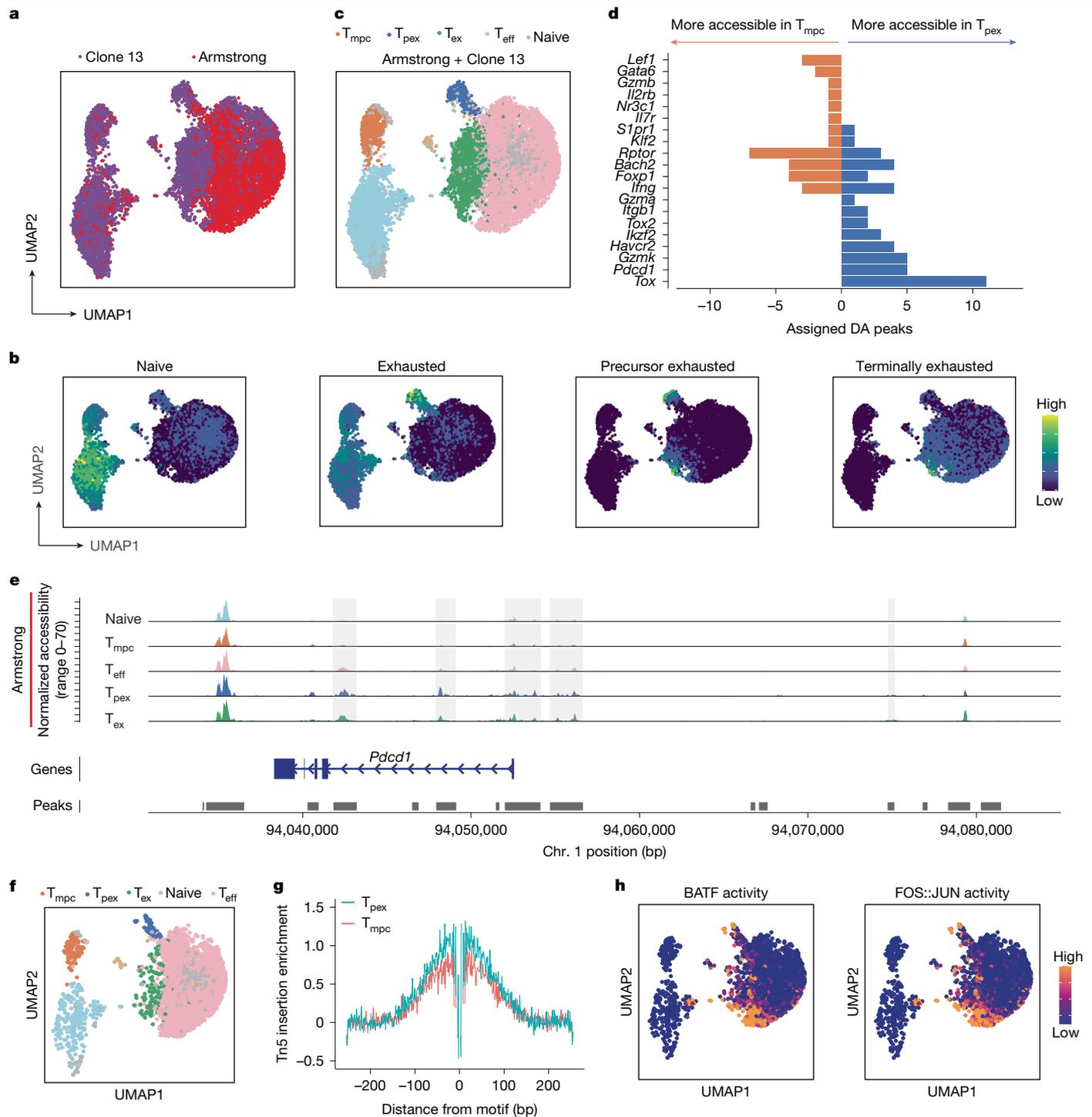


Fig. 4 | T_{pex} cells from acute infections epigenetically resemble their counterparts from chronic infections. Public scATAC-seq data²⁸ for total splenic CD8 T cells analysed seven days after mice were infected with either LCMV Armstrong or LCMV clone 13. **a**, UMAP embedding of cell similarity. Cells are coloured by infection type. **b**, Signatures for naive, exhausted, precursor exhausted (T_{ex}) and terminally exhausted cells derived from bulk ATAC-seq data are used to compute signature scores for each cell. Colours represent the mean expression level of the signature within the cell shown on the UMAP embedding. **c**, Cell-type annotations were added on the basis of reference

signatures derived from the bulk ATAC-seq experiments in **b**. T_{eff} effector T cells. **d**, Number of differentially accessible (DA) regions between T_{pex} and T_{mpc} annotated to genes of interest. **e**, *Pdccl1* locus accessibility signal of cells from acute infection. Marker regions of the T_{pex} cluster are highlighted in grey. **f**, UMAP embedding of cell similarity subset on cells from Armstrong infection only. Colours represent cell assignments to Leiden clusters. **g**, Accessibility signal at FOS::JUN motif locations. **h**, Transcription-factor activity for the FOS::JUN motif. Data are derived from a single previously published dataset.

whether the T_{pex} cells found in early acute infection show the epigenetic landscape of chronic-infection-associated T_{pex} cells. To this end, we used the public dataset GSE164978. In this study, C57BL/6j mice were infected with either LCMV Armstrong or LCMV clone 13.

Seven days later, total CD8 T cells were isolated from spleen and analysed by single-cell-resolved assay for transposase-accessible chromatin with sequencing (ATAC-seq)²⁸. As a first step in our analysis of the published datasets, we jointly clustered cells from both types of

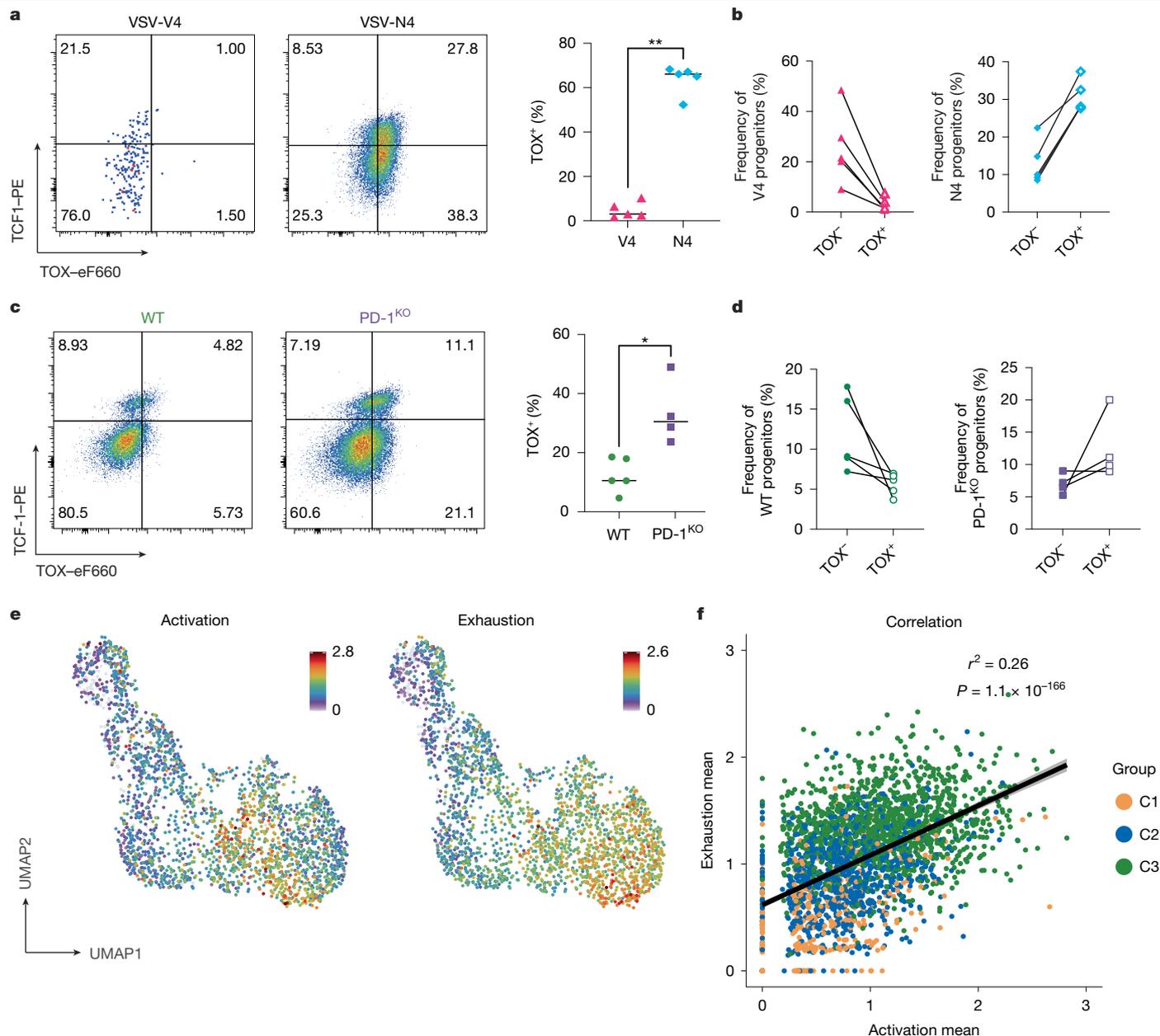


Fig. 5 | Strong TCR stimulation and the absence of PD-1 signalling promote T_{pex} formation in acute infection. a–d, Wild-type (WT) OT-1 (a,b) and wild-type or PD-1^{KO} P14 transgenic (c,d) T cells were adoptively transferred into CD45.2 congenic hosts. Mice were subsequently infected with VSV-V4 (low-affinity stimulation) or VSV-N4 (high-affinity stimulation) (a,b) or with LCMV Armstrong (c,d). Splenocytes were collected seven days after infection. $n = 5$ (OT-1 and wild-type P14 recipients); $n = 4$ (PD-1^{KO} recipients). a,c, Flow plots show TCF1 versus TOX expression by CD44⁺KLRG1⁻ gated OT-1 (a) or P14 (c) T cells, and graphs show the frequency of TOX⁺ OT-1 (a) or P14 (c) precursor T cells. * $P < 0.05$, ** $P < 0.01$ (two-tailed Mann–Whitney test). b,d, Frequency of TOX⁻ and TOX⁺

TCF1⁺KLRG1⁻ OT-1 (b) and P14 (d) T cells. e, Average normalized expression of genes from activation and exhaustion signatures (Extended Data Table 1) projected onto the UMAP embedding shown in Fig. 2. f, Pearson correlation between activation and exhaustion scores for individual cells. Colours encode the subcluster assignment as shown in Fig. 2. Line represents a linear model regression and the bands show a confidence interval of 95%. Each symbol in a–d represents an individual mouse. Data in a–d are representative of at least three independent experiments. Lines in a,b represent the mean. Lines in b,d mark cells originating from the same donor. Data in e,f are derived from a single previously published dataset.

infection (Fig. 4a). In the original publication, the authors had mined numerous ATAC-seq datasets obtained by sequencing bulk populations of sorted exhausted T cells (T_{ex}), precursors of exhausted T cells (T_{pex}) and terminally exhausted T cells, but also naive cells and memory precursor cells (T_{mpc})²⁸.

We then superimposed these reference bulk signatures onto our clustered single-cell-resolved ATAC-seq analysis (Fig. 4b). We saw a significant overlap between cells showing high signal intensities for the naive and for the T_{mpc} signature. This is in line with the fact that

both populations are known to share a large fraction of their epigenetic profiles and transcriptional mechanisms. Similarly, the population that had the strongest T_{pex} signature scored highly in the T_{mpc} signature, again reflecting partially shared epigenetic profiles. On the basis of the results shown in Fig. 4b, we then assigned clusters as naive cells, T_{mpc} , T_{pex} , T_{eff} and T_{ex} (Fig. 4c). As a final step, we performed differential accessibility analysis between the T_{pex} and T_{mpc} clusters and aggregated differentially accessible regions on the basis of the gene they annotate to (Fig. 4d). As expected, we found that T_{pex} cells are more accessible

at several regions in proximity to *Tox*, *Tox2* and *Pdcd1*, compared with T_{mpc} (Fig. 4d). Notably, we observed that the T_{pex} cluster contained cells from both acute and chronic infections (Fig. 4a). Moreover, when directly comparing the accessibility of T_{pex} marker regions in the *Pdcd1* locus of cells isolated from Armstrong or clone 13 infected mice, we did not find significant differences (Extended Data Fig. 4a). This underlines that T_{pex} cells from both acute and chronic infection share similar epigenetic profiles. As a next step, we selected the cells infected with LCMV Armstrong only and plotted the coverage at the *Pdcd1* locus split by cluster assignment (Fig. 4e). The plots highlight in grey the regions identified on the basis of differential accessibility between the T_{pex} and the T_{mpc} cluster. We observed similar accessibility between T_{ex} and T_{pex} and clear differences between T_{pex} and T_{mpc} , indicating that differences between T_{pex} and T_{mpc} can be reproduced with cells originating solely from acute infection (Fig. 4f). We continued to use the subset of cells obtained from Armstrong infection and integrated data on known transcription-factor-binding motifs to find putative regulators that drive the differences between T_{pex} and T_{mpc} . Among the marker regions associated with higher accessibility in T_{pex} than in T_{mpc} , a large portion were part of the AP-1 family of transcription factors (Extended Data Fig. 4b). We see that regions around FOS::JUN sites are more accessible in T_{pex} than in T_{mpc} (Fig. 4g), and infer higher transcription-factor activity (Fig. 4h). Altogether, these observations show that epigenetic marks that are typically associated with T_{pex} cells can also be detected in the T_{pex} population that is found in the early phase of acute infection.

TCR affinity and PD-1 affect T_{pex} formation

Because increased TCR signal strength drives exhaustion in chronic infection^{17,29–31} we sought to determine how ligand affinity and the reception of inhibitory signals affect T_{pex} activation in acute infection. To do so, we used a previously established system in which OT-1 T cells are exposed either to their native ligand SIINFELK (referred to as N4) or to a lower-affinity altered peptide ligand (SIIVFEKL; V4) during an infection with recombinant vesicular stomatitis virus expressing N4 (VSV-N4) or V4 (VSV-V4)¹. Mice infected with high-affinity VSV-N4 had a large shift in the frequency of TOX⁺ precursors and effectors at seven days after infection, whereas mice infected with VSV-V4 had little to no TOX⁺ expression in either the precursors or the effector compartment (Fig. 5a,b), indicating that the formation of T_{pex} in acute infections depends on receiving strong TCR signals.

A key feature of T_{pex} cells is the upregulation of PD-1. Being an inhibitory receptor, PD-1 is thought to modulate signals that promote activation through the TCR and CD28 by engaging its ligands³². Therefore, we hypothesized that PD-1 expression might limit the extent of T_{pex} formation. To test this, we transferred wild-type or PD-1-knockout (PD-1^{KO}) P14 transgenic CD8 (CD45.1) T cells into B6 hosts (CD45.2) and infected these mice with LCMV Armstrong. Samples were analysed by flow cytometry seven days after infection. When comparing PD-1^{KO} P14 to wild-type P14, we noted that PD-1^{KO} cells had a higher frequency of total TOX⁺ cells in both the TCF1⁺ precursor and the TCF1⁻ effector population (Fig. 5c). Moreover, we noted a shift in the precursor population from a dominance of TOX⁻ precursors in wild-type cells towards a higher fraction of TOX⁺ precursors in PD-1^{KO} P14 T cells (Fig. 5d). These data were further corroborated by the fact that the single-cell transcriptomic data for day-4.5 Armstrong-infection-derived T cells (presented in Figs. 1 and 2) also show that highly activated cells exhibit a stronger exhausted signal, which confirms the positive correlation between activation strength and the acquisition of exhaustion signatures (Fig. 5e,f and Extended Data Table 1). We therefore concluded that although strong TCR signalling promotes T_{pex} formation in early acute infection, the expression of PD-1 restricts or limits the generation of T_{pex} cells in acute infection.

Plasticity of early T_{pex} progenitors

To determine the stability of T_{pex} cells formed in the early phase of acute infection, we transferred TCF1-reporter-positive P14 PD-1^{hi} or PD-1^{lo} cells (as shown in Fig. 2a) from day-4.5 Armstrong-infected mice into naive hosts. Host mice were then infected with LCMV Armstrong on the same day of transfer (Fig. 6a). scRNA-seq of transferred day-4.5 P14 cells re-isolated at day 15 after infection revealed six distinct clusters (Fig. 6b), spanning the spectrum from cells exhibiting features of progenitors (*Tcf7*, *Il7r* and *Slamf6*; clusters 1, 2, 3 and 5) to effector cells that lack progenitor markers and express higher levels of effector function (*Ifng*, *Gzmb*, *Klrg1* and *Cx3cr1*; clusters 4 and 6) (Extended Data Fig. 5a). Of note, the progeny of PD-1^{hi} and PD-1^{lo} cells produced no unique clusters, with both input populations producing cells that were found in all clusters (Fig. 6b). Therefore, to determine whether we could still identify T_{pex} , we excluded the effector clusters 4 and 6 and focused our analysis on the progenitor clusters (1, 2, 3 and 5) (Fig. 6c). Progenitor clusters 1 and 2 expressed higher levels of the stemness marker *Sell*, along with increased expression of *Ccr7*. Progenitor cluster 3 expressed higher levels of effector-memory-associated genes, maintaining the highest levels of *Zeb2*, *Gzma*, *Gzmb*, *Cx3cr1* and *Ifng*. Notably, cells found in progenitor cluster 5 were defined by the expression of genes that regulate key features associated with T cell exhaustion and maintenance, including *Tox*, *Pdcd1*, *Ikzf2*, *Tcf1* and *Cd200r1* (Fig. 6c and Extended Data Fig. 5b). To further confirm that the cells in cluster 5 were T_{pex} , we analysed the expression of gene signatures associated with either acute or exhausted CD8 T cells; progenitor cluster 5 showed higher expression of exhaustion-associated genes, but lacked genes associated with an acute phenotype (Fig. 6d). Together, the data suggest that early day-4.5 PD-1^{hi} and PD-1^{lo} precursors show considerable plasticity towards their differentiation into various effector and progenitor subsets. Nonetheless, although the progeny of both PD-1^{hi} and PD-1^{lo} P14 cells were found in all clusters (Fig. 6b) and showed comparable gene expression (Fig. 6c), our data also revealed a minor fate bias. We observed, for instance, that in the largest cluster (cluster 1), most cells originated from PD-1^{lo} cells, whereas most cells in T_{pex} cluster 5 were derived from PD-1^{hi} cells (Fig. 6b,e). Furthermore, PD-1^{hi}- and PD-1^{lo}-derived cells showed some differences in the expression of important exhaustion-associated genes; for example, *Tox*, *Pdcd1* and *Ikzf2* were expressed more highly in PD-1^{hi}-derived cells, whereas key effector genes such as *Ifng* and *Tnf* were slightly more highly expressed in the progeny of PD-1^{lo} cells (Fig. 6c). This indicates that features of T cell exhaustion persisted after the transfer. Moreover, we also transferred TCF1-reporter-positive P14 cells with high PD-1 and low PD-1 expression from day-7 Armstrong-infected mice into naive hosts that were then infected with LCMV Armstrong on the day of transfer. Although both types of cells underwent robust secondary expansion, the progeny of the transferred cells still mostly differed in their TOX expression levels (Extended Data Fig. 5c,d). These data show that although there is substantial plasticity at early stages, the T_{pex} phenotype seems to stabilize at later time points.

T_{pex} stability in acute infections

Next, we sought to determine the longevity and stability of T_{pex} cells. We thought at first that these cells might be only temporarily detectable and that their persistence might require continuous antigen exposure, because we could detect T_{pex} among the sequenced day-4.5 P14 T cells (in Fig. 1) but had difficulties in identifying them in significant numbers among the day-7 cells (Fig. 1a). However, this outcome needs to be viewed with caution as the relatively low sample size of scRNA-seq datasets poses particular challenges in detecting rare subpopulations. Of note, day-7 LCMV-specific T cells contain only around 10% precursor cells, of which only a small fraction might be the descendants of the day-4.5 T_{pex} . Thus, the day-7 P14 T_{pex} numbers might be too low to

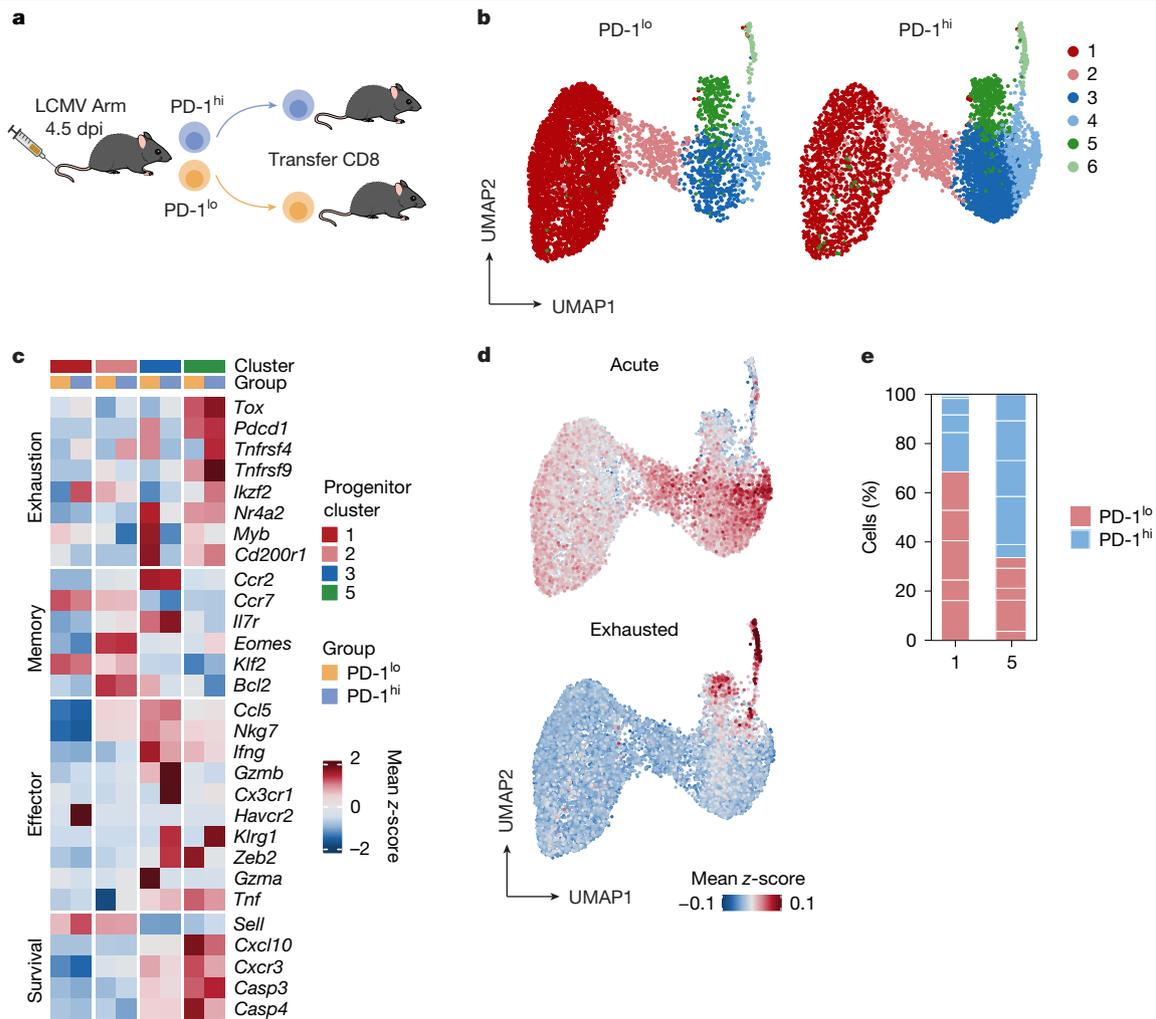


Fig. 6 | Plasticity of T_{pex} cells formed in the early infection phase. **a**, CD45.1 congenic P14 T cells were transferred into C57BL/6 host mice that were infected with LCMV Armstrong. PD-1^{hi} and PD-1^{lo} P14 cells (gating strategy shown in Extended Data Fig. 3b) were collected at 4.5 days post-infection (dpi) and subsequently retransferred into naive C57BL/6 host mice that were infected with LCMV Armstrong on the same day of P14 transfer. Retransferred P14 cells from PD-1^{hi} versus PD-1^{lo} recipients were collected at 15 days after infection and

sent for single-cell sequencing. Graphics were created with Biorender.com. **b**, UMAP representation of Louvain clusters from PD-1^{hi} and PD-1^{lo} recipients. **c**, Heat map showing z-scored gene expression segregated by PD-1^{hi} and PD-1^{lo} samples. **d**, Signature score of genes associated with either acute or exhausted CD8 T cells projected onto the UMAP. **e**, Fractions of PD-1^{hi} and PD-1^{lo} recipients in clusters 1 and 5. Data are derived from a single experiment.

detect them in the scRNA-seq dataset. In fact, rare but detectable numbers of day-7 P14 T_{pex} can be found when large sample sizes and thus larger precursor numbers are analysed by flow cytometry, as shown in Figs. 3a and 5c,d.

Moreover, we noted a considerable number of TCF1- and TOX-co-expressing CD8 T cells of unknown specificity 15 and 35 days after an LCMV infection among endogenous cells (Extended Data Fig. 5e). Cells expressing increased levels of PD-1 and TOX can also be found among CD8⁺ T cells specific to the LCMV-derived np396 epitope at days 15 and 30 after infection (Fig. 7a,b). This indicates that T_{pex} cells that are formed in the early infection period can indeed survive long term and can be found among the memory T cell population. Notably, these cells are detectable only among np396- and not among gp33-multimer-positive populations (Fig. 7a,b). The np396 epitope was previously shown to provide a stronger level of TCR stimulation to T cells than the gp33 peptide³³. That T_{pex} cells persist among these cells further corroborates our idea that strong TCR stimulation is needed to form and maintain these cells. To characterize this in more depth, we isolated TCF1-expressing memory gp33⁺ or np396⁺ T cells from week-4 LCMV-infected Tcf7^{gfp(bright)}mCherry reporter mice and performed scRNA-seq together with paired α and β TCR repertoire analysis. Clustering identified five populations

(Extended Data Fig. 6a), and np396- and gp33-specific cells contributed to all of these in an approximately similar manner (Extended Data Fig. 6b). Although we did not see any clusters that were mostly dominated by either gp33- or np396-specific T cells, we nevertheless noted differences in the expression of key genes associated with T cell exhaustion. In particular, *Tox* and *Pdcd1* were more highly expressed in np396-specific CD8 T cells (Extended Data Fig. 6c). We further noted that *Tox* and *Pdcd1* expression was particularly strong in certain clusters, especially cluster 2, suggesting the existence of a somewhat heterogeneous transcriptional landscape. To analyse this in more depth, we decided to include TCR sequences into our analysis, and aggregated cells into clonotype modules³⁴. These are defined as assemblies of TCR nucleotide sequences that are predicted to share similar binding properties to specific antigens and similar TCR signal strength (Fig. 7c). Clonotype modules are grouped on the basis of TCRs that share amino-acid sequences that lead to similar secondary and tertiary structures. Hence, TCR sequences that contain substitutions with functionally equivalent amino acids would be grouped into the same clonotype module, whereas sequences that contain substitutions with amino acids with different biochemical properties would be assigned to another clonotype module. To simplify our findings,

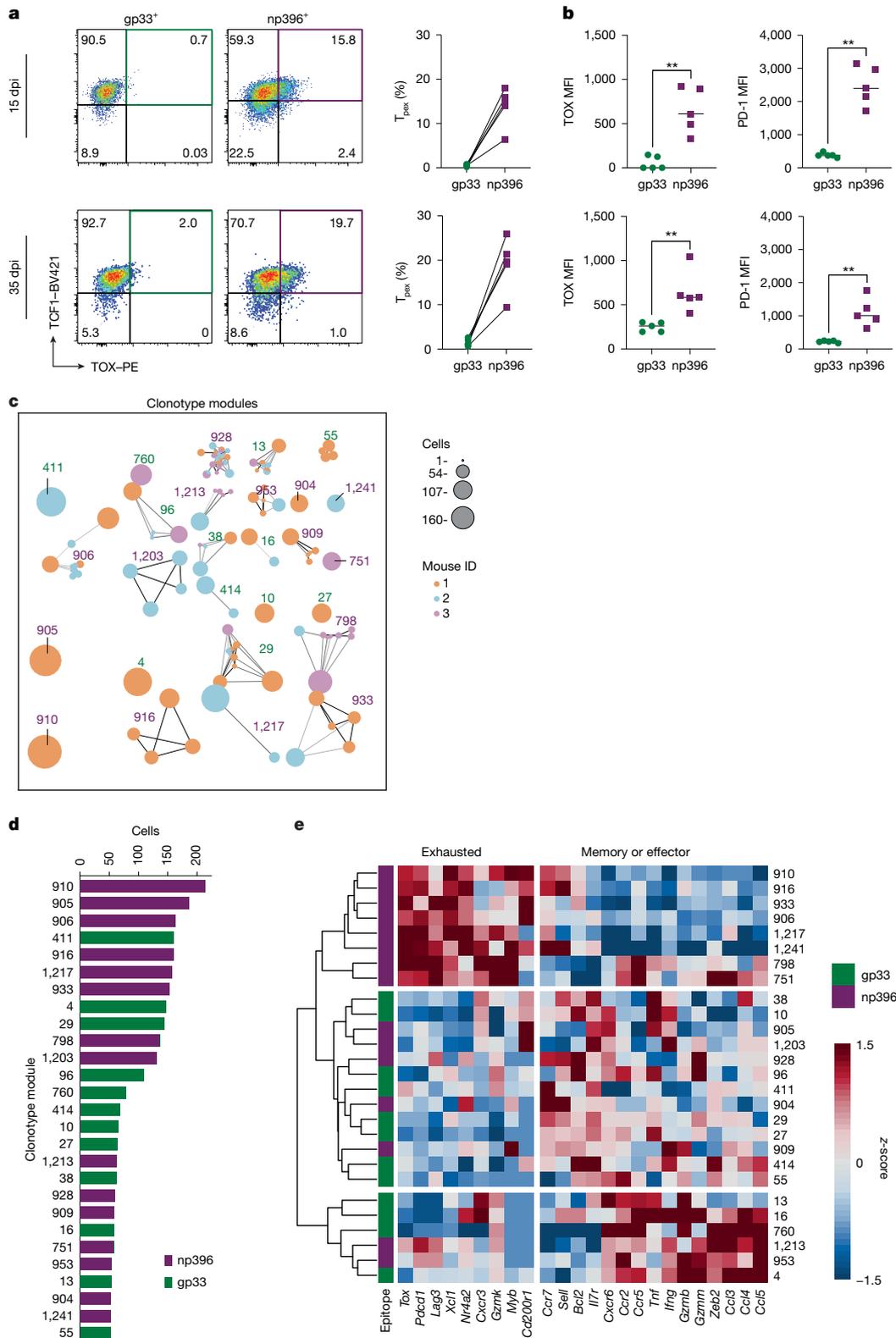


Fig. 7 | T_{pex} cells are detectable in the memory phase after stronger TCR stimulation. **a**, TCF1 and TOX expression in CD8 T cells stained with gp33⁺ and np396⁺ multimers at 15 or 30 days after infection. Each symbol represents an individual mouse ($n = 5$ mice per group). **b**, Scatter plots showing the mean intensity of expression of TOX (left) or PD-1 (right) by multimer-positive T cells. Each symbol represents an individual mouse ($n = 5$ mice per group); lines represent the mean; * $P < 0.05$, ** $P < 0.01$ (two-tailed Mann–Whitney test). **c**, gp33-specific (green) and np396-specific (purple) TCF1⁺ precursors were purified from TCF1 reporter mice ($n = 3$) four weeks after LCMV Armstrong

infection. Purified cells were subjected to scRNA-seq with paired TCR repertoire analysis. Clonotype pairs were aligned using their amino-acid sequence. Shown are clonotype networks of all modules containing at least 50 cells, with similar clonotypes connected by an edge. Node size depicts clonotype size, the arbitrary clone number, and colours represent mouse ID. **d**, Number of cells per clonotype module. The dendrogram was generated using Euclidean distance and complete linkage. Data are representative of three independent experiments (**a, b**) or derived from a single experiment.

we included into our subsequent analysis any module that contained at least 50 cells (Fig. 7c,d). To validate this method, we confirmed that there was no significant sharing of modules between gp33+ and np396+, emphasizing the unique binding properties between gp33+ and np396+ clonotypes (Fig. 7c,d). Of note, we can see module sharing between different mice within certain gp33+ or np396+ clonotypes—confirming that different mice develop unique but also shared TCR amino-acid sequences to the same antigen (Fig. 7c). Aggregating T cells on the basis of clonotype modules showed a clear pattern of modules that express an exhausted signature, including *Tox*, *Pdcd1*, *Lag3*, *Myb* and *Cd200r1* (Fig. 7e). Exhausted modules were specific to np396+, whereas the expression of genes associated with effector and acute phenotypes was seen in gp33+ and some np396+ clonotype modules. (Fig. 7e). Together, these findings are in line with our previous flow cytometric data (Fig. 7a), in which we saw the expression of TOX in some np396+ but in very few gp33+ CD8 T cells. Taking into consideration our observations that TCR signal strength is crucial for T_{pe} formation (Figs. 3b, 5a,b and 7e) and also that np396 is considered a higher-affinity epitope than gp33 (ref. 33), our data strongly suggest that strong TCR stimulation is required for the formation of a long-lived T_{pe} population.

Using a combination of comprehensive sequencing-data analysis and flow-cytometry-based verification, we have shown that T_{pe} cells are generated in acute infections. We show that they are formed in the early phase of infection, and thus long before the infection becomes resolved or chronic. Notably, these T_{pe} cells are found at first in almost as high numbers as typical memory precursor T cells, but they decline substantially thereafter. We take this to mean that the consecutive formation allows the environment to subsequently shape the precursor repertoire on the basis of the infection outcome (Supplementary Fig. 1). Such dynamics have considerable advantages for the infected host, in that they proactively prepare the host for the possibility that the infection becomes chronic, by generating the type of cells that are needed in such infections. This interpretation takes a more recent view of the phenomenon of T cell exhaustion into consideration, which is that forming exhausted T cells can be beneficial for the host^{5,20}. Although their reduced effector capacity limits the magnitude by which these cells can protect the host from the pathogen, exhausted T cells are at the same time causing much less pathology than are normal effector T cells²⁰. Thus, their presence might ensure that chronic infections are controlled to a necessary level, without causing detrimental organ damage and massive immunopathology.

Mechanistically, we show that T_{pe} formation in acute infections requires strong TCR signals, which are mediated, for instance, by high-affinity TCR ligation. Furthermore, inhibiting T cell activation through PD-1 signalling restricts the formation of T_{pe} cells. The latter is particularly noteworthy, because PD-1 is often thought to induce exhausted T cells. Instead, our data suggest that PD-1 signalling actually prevents exhaustion and limits the formation of T_{pe} cells. Altogether, our observations support a major conceptual change to the commonly held view that exhausted T cells are only formed in the context of long-term antigen persistence. By contrast, we show that the formation of exhausted T cells occurs independently of the outcome of the infection, and that chronic infections propagate and expand these pre-emptively formed T_{pe} cells.

Online content

Any methods, additional references, Nature Portfolio reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at <https://doi.org/10.1038/s41586-024-08451-4>.

- Zehn, D., Thimme, R., Lugli, E., de Almeida, G. P. & Oxenius, A. 'Stem-like' precursors are the fount to sustain persistent CD8⁺ T cell responses. *Nat. Immunol.* **23**, 836–847 (2022).
- Blank, C. U. et al. Defining 'T cell exhaustion'. *Nat. Rev. Immunol.* **19**, 665–674 (2019).
- Barber, D. L. et al. Restoring function in exhausted CD8 T cells during chronic viral infection. *Nature* **439**, 682–687 (2006).
- Blackburn, S. D. et al. Coregulation of CD8⁺ T cell exhaustion by multiple inhibitory receptors during chronic viral infection. *Nat. Immunol.* **10**, 29–37 (2009).
- Alfei, F. et al. TOX reinforces the phenotype and longevity of exhausted T cells in chronic viral infection. *Nature* **571**, 265–269 (2019).
- Khan, O. et al. TOX transcriptionally and epigenetically programs CD8⁺ T cell exhaustion. *Nature* **571**, 211–218 (2019).
- Scott, A. C. et al. TOX is a critical regulator of tumour-specific T cell differentiation. *Nature* **571**, 270–274 (2019).
- Seo, H. et al. TOX and TOX2 transcription factors cooperate with NR4A transcription factors to impose CD8⁺ T cell exhaustion. *Proc. Natl. Acad. Sci. USA* **116**, 12410–12415 (2019).
- Chen, J. et al. NR4A transcription factors limit CAR T cell function in solid tumours. *Nature* **567**, 530–534 (2019).
- Buggert, M. et al. T-bet and Eomes are differentially linked to the exhausted phenotype of CD8⁺ T cells in HIV infection. *PLoS Pathog.* **10**, e1004251 (2014).
- McLane, L. M. et al. Role of nuclear localization in the regulation and function of T-bet and Eomes in exhausted CD8 T cells. *Cell Rep.* **35**, 109120 (2021).
- Wherry, E. J. et al. Molecular signature of CD8⁺ T cell exhaustion during chronic viral infection. *Immunity* **27**, 670–684 (2007).
- Doedens, A. L. et al. Hypoxia-inducible factors enhance the effector responses of CD8⁺ T cells to persistent antigen. *Nat. Immunol.* **14**, 1173–1182 (2013).
- Utzschneider, D. T. et al. T cells maintain an exhausted phenotype after antigen withdrawal and population reexpansion. *Nat. Immunol.* **14**, 603–610 (2013).
- Ghoneim, H. E. et al. De novo epigenetic programs inhibit PD-1 blockade-mediated T cell rejuvenation. *Cell* **170**, 142–157 (2017).
- Pauken, K. E. et al. Epigenetic stability of exhausted T cells limits durability of reinvigoration by PD-1 blockade. *Science* **354**, 1160–1165 (2016).
- Philip, M. et al. Chromatin states define tumour-specific T cell dysfunction and reprogramming. *Nature* **545**, 452–456 (2017).
- Sen, D. R. et al. The epigenetic landscape of T cell exhaustion. *Science* **354**, 1165–1169 (2016).
- Tsui, C. et al. MYB orchestrates T cell exhaustion and response to checkpoint inhibition. *Nature* **609**, 354–360 (2022).
- Speiser, D. E. et al. T cell differentiation in chronic infection and cancer: functional adaptation or exhaustion? *Nat. Rev. Immunol.* **14**, 768–774 (2014).
- Wherry, E. J. T cell exhaustion. *Nat. Immunol.* **12**, 492–499 (2011).
- Utzschneider, D. T. et al. T cell factor 1-expressing memory-like CD8⁺ T cells sustain the immune response to chronic viral infections. *Immunity* **45**, 415–427 (2016).
- Im, S. J. et al. Defining CD8⁺ T cells that provide the proliferative burst after PD-1 therapy. *Nature* **537**, 417–421 (2016).
- Utzschneider, D. T. et al. Early precursor T cells establish and propagate T cell exhaustion in chronic infection. *Nat. Immunol.* **21**, 1256–1266 (2020).
- Kusnadi, A. et al. Severely ill COVID-19 patients display impaired exhaustion features in SARS-CoV-2-reactive CD8⁺ T cells. *Sci. Immunol.* **6**, eabe4782 (2021).
- Yao, C. et al. Single-cell RNA-seq reveals TOX as a key regulator of CD8⁺ T cell persistence in chronic infection. *Nat. Immunol.* **20**, 890–901 (2019).
- West, E. E. et al. Tight regulation of memory CD8⁺ T cells limits their effectiveness during sustained high viral load. *Immunity* **35**, 285–298 (2011).
- Pritykin, Y. et al. A unified atlas of CD8 T cell dysfunctional states in cancer and infection. *Mol. Cell* **81**, 2477–2493 (2021).
- Schietinger, A. et al. Tumor-specific T cell dysfunction is a dynamic antigen-driven differentiation program initiated early during tumorigenesis. *Immunity* **45**, 389–401 (2016).
- Shakiba, M. et al. TCR signal strength defines distinct mechanisms of T cell dysfunction and cancer evasion. *J. Exp. Med.* **219**, e20201966 (2022).
- Utzschneider, D. T. et al. High antigen levels induce an exhausted phenotype in a chronic infection without impairing T cell expansion and survival. *J. Exp. Med.* **213**, 1819–1834 (2016).
- Sharpe, A. H. & Pauken, K. E. The diverse functions of the PD1 inhibitory pathway. *Nat. Rev. Immunol.* **18**, 153–167 (2018).
- Wherry, E. J., Blattman, J. N., Murali-Krishna, K., van der Most, R. & Ahmed, R. Viral persistence alters CD8 T-cell immunodominance and tissue distribution and results in distinct stages of functional impairment. *J. Virol.* **77**, 4911–4927 (2003).
- Drost, F. et al. Multi-modal generative modeling for joint analysis of single-cell T cell receptor and gene expression data. *Nat. Commun.* **15**, 5577 (2024).

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 4.0 International License, which permits use, sharing, adaptation, 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 changes were made. 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/4.0/>.

© The Author(s) 2024

Methods

Mice

Mice were bred and housed in specific-pathogen-free facilities at the Technical University of Munich in Germany. P14 TCR $\alpha\beta$ -transgenic mice³⁵ were provided by A. Oxenius and V β 5 TCR β -only transgenic mice³⁶ by P. Fink. OT-1 mice were purchased from the Jackson Laboratory. PD-1^{KO} P14 mice were provided by P.-C. Ho. TCF1 reporter mice were generated by us and described previously¹⁹. Mice were housed under the following conditions: light from 07:00 to 19:00; temperature 22–26 °C; humidity 30–70 g m⁻³. Experiments were performed on male and female mice aged 6–13 weeks in compliance with the institutional and governmental regulations of Germany. They were approved by the responsible veterinarian authorities of the Regierung von Oberbayern in Germany. Statistical methods were not used to predetermine sample size. We chose sample sizes on the basis of previous experience and with the goal of producing statistically robust data while respecting animal welfare regulations. Mice of the appropriate genotype and age were randomly assigned to the experimental groups. We did not perform any readouts involving subjective evaluations, such as histological analysis or clinical scoring. Owing to biosafety regulations, we were instructed not to perform blinding in experiments involving biohazardous substances.

Viral infections

LCMV virus strains were originally provided by M. J. Bevan. LCMV was expanded in BHK-21 cells and titrated on Vero cells (originally provided by M. J. Bevan) using a focus-forming assay³⁷. Recombinant ovalbumin expressing VSV-N4 containing the original SIINFEKL epitope³⁸ or the low-affinity altered peptide ligand SIIVFEKL (VSV-V4, originally provided by L. Lefrancois) were expanded and titrated on BHK-21 cells (University of Washington) using a lytic plaque assay³⁹. Frozen stocks of LCMV Armstrong, LCMV clone 13 and VSV-N4 or VSV-V4 were diluted in phosphate-buffered saline (PBS). Mice were infected by intraperitoneal injection of 2×10^5 plaque-forming units (PFU) of LCMV Armstrong or intravenous injection of 2×10^6 PFU of recombinant VSV-N4 expressing the SIINFEKL³⁸ or VSV-V4 OVA strain displaying the SIIVFEKL epitope. Before experiments, we ensured that the cell lines used showed the expected results and were in good condition. Mycoplasma tests were routinely performed and were negative.

Cell preparations, T cell purification and adoptive T cell transfers

Single-cell splenocyte suspensions were prepared from naive or infected mice by mashing spleen through a 100- μ m cell strainer followed by hypotonic ACK lysis. Total splenocytes were used for flow-cytometry analysis and sorting. P14 TCR $\alpha\beta$ transgenic wild-type, P14 TCR-transgenic PD-1^{KO} or OT-1 mice were used as donors for adoptive transfers. CD8⁺ T cells were isolated from splenocyte suspension using the negative selection CD8⁺ T cell enrichment kit II (Miltenyi Biotec) in accordance with the manufacturer's protocol. For LCMV Armstrong experiments, 1,000 P14 T cells were transferred intravenously by tail-vein injection. For VSV infections, 5,000 OT-1 cells were transferred intravenously by tail-vein injection.

Flow-cytometry analysis and cell sorting

After isolation, cells were incubated with Zombie NIR dye (BioLegend, 423106) and Fc-blocking reagent 2.4G2 in PBS (Thermo Fisher Scientific) for 15 min at room temperature. Next, cells were washed and resuspended in freshly prepared master mix containing fluorescent staining antibodies in fluorescence-activated cell sorting (FACS) buffer (PBS, 2% fetal calf serum (FCS) and 0.1% sodium azide). When the master mix contained two or more brilliant violet polymer dyes, super bright staining buffer (Invitrogen, SB-4401-42) was used. For tetramer or CXCR5 staining, cells were stained for one hour at room

temperature. FoxP3/Transcription factor staining buffer (Thermo Fisher Scientific, 00-5532-00) was used to perform intranuclear staining following manufacturer recommended protocols. Samples were acquired on BD FACS Fortessa or Beckman Coulter CytoFlex LX instruments. For sorting: after live–dead stain and Fc block, cells were incubated with freshly prepared master mix for 20 min at room temperature, followed by washing in magnetic-activated cell sorting (MACS) buffer (PBS, 1% FCS and 2 mM EDTA). Cells were immediately run on a BD FACS Aria Fusion sorter. Single-stained controls were prepared using Ultracomp eBeads (Invitrogen 01-2222-42) for every experiment. Beads were treated like the cell-containing samples (including fixation protocol). A detailed list of staining antibodies is provided in Supplementary Table 1. FACS data were analysed using FlowJo software (BD, v.10.09) and exported as Excel (Microsoft, v.16.89.1) tables for subsequent statistical analysis.

Statistical analysis

GraphPad Prism software was used for statistical analysis (except for the gene-expression data). Student's *t*-tests (two-tailed) for parametric and Mann–Whitney tests for non-parametric data were used to compare between two independent conditions. Experiments with $n < 5$ were considered non-parametric. When comparing samples originating from the same mouse, a paired Student's *t*-test was used for parametric data and a Wilcoxon matched pairs signed-rank test for non-parametric data. All measurements were taken from distinct samples.

De novo 10X Genomics scRNA-seq and data analysis

TCF1 reporter mice were infected directly with LCMV Armstrong. Splenocytes were collected four weeks after infection and stained for np396 and gp33. Tetramer and TCF1-reporter-positive CD8⁺ T cells were purified by flow-cytometry-based sorting. Gene expression and T cell receptor V(D)J libraries were prepared by using the Chromium Next GEM Single Cell 5' Reagent Kit v2 (PN-1000265, 10X Genomics), Chromium Single Cell Mouse TCR Amplification Kit (PN-1000254, 10X Genomics) and Chromium Next GEM Chip K Single Cell Kit (PN-1000287, 10X Genomics) following the manufacturer's protocol (CG000331 Rev E). The Dual Index Kit TT Set A (PN-1000215, 10X Genomics) was used for multiplexing (i7 and i5 index read, 10 bp). The samples were sequenced in a paired-end run (read 1: 26 bp; read 2: 90 bp) on a NovaSeq6000 platform using S1 v.1.5 (100 cycles) sequencing kits (20028319, Illumina). Bcl2fastq software (v.2.20.0.422) was used for demultiplexing and generating .fastq files.

De novo BD Rhapsody scRNA-seq and data analysis

Spleens were collected from LCMV-Armstrong-infected mice at various days after infection (see below) and cut into small pieces. These pieces were then exposed to a medium containing digestive enzymes (DNase and liberase) for one hour at 37 °C shaking at 240 rpm. The digested tissue was then macerated through a 100- μ m Nytex filter and washed with Dulbecco's modified Eagle's medium (DMEM), and erythrocyte lysis was performed using a hypotonic ACK medium. Lysis was stopped by adding DMEM containing 10% FCS. Cells were counted using a Neubauer counting chamber and a CD90.2 MACS enrichment (Miltenyi) was performed according to the manufacturer's instructions. Finally, cells were stained with fluorescently labelled antibodies against CD8a, CD45.1 and CD44 as well as a fixable live–dead dye. For the experiment on day 3.25 cells, single, live CD8a⁺CD45.1⁺ cells were sorted and subsequently used to prepare scRNA-seq libraries. For the retransfer experiment, single, live CD8a⁺CD45.1⁺ cells from day 4.5 after infection were sorted into PD-1^{hi} and PD-1^{low} and adoptively transferred into new host mice, which were subsequently infected with LCMV Armstrong as well. On day 15 after infection, spleens from these secondary hosts were isolated, CD8a⁺CD45.1⁺ were sorted and subsequently used to prepare scRNA-seq libraries using the BD Rhapsody system (BD Biosciences). First, sample tags were added to individual cell samples (BD Ms Single

Cell Sample Multiplexing Kit, protocol 23-21340(02)). Next, single cells of the pooled samples were captured using the BD Rhapsody HT Xpress system for each experiment separately, followed by reverse transcription and cDNA amplification. Sequencing libraries were constructed using the BD Rhapsody whole transcriptome analysis approach. Quality control of libraries was done with Bioanalyzer 2100 using the High Sensitivity DNA Kit (Agilent) and Qubit 2.0 fluorometer with the Qubit 1X dsDNA HS Assay Kit (Invitrogen). Libraries were sequenced on a NovaSeq X Plus (Illumina) with paired-end 2 × 150-bp reads. Detailed information on the analysis pipelines is provided in the Supplementary Methods (scRNA-seq analysis).

ScRNA-seq analysis of previously published data

Read counts from a previously published dataset²⁶ were retrieved from the GEO database under the accession number GSE119943. Data integration was performed on the two published replicates. Subsequently, PCA, identification of nearest neighbours and Louvain⁴⁰ clustering were performed, followed by UMAP dimensional reduction. Detailed information on these steps and further downstream analyses is provided in the Supplementary Methods (scRNA-seq analysis).

De novo bulk RNA-seq

Total RNA was isolated using the ReliaPrep RNA Cell Miniprep System (Promega). The quality and quantity of isolated RNA were analysed with the Bioanalyzer RNA Pico Chip (Agilent). cDNA synthesis and library preparation were performed with the SMART-Seq v4 PLUS Ultra Low Input RNA kit (Takara) following the manufacturer's protocol. Libraries were sequenced on a single-end run (1 × 100 bp, dual-index) on a NovaSeq6000 (Illumina) using SP100 v.1.5 chemistry (Illumina).

De novo bulk RNA-seq data analysis

Reads were processed with an adapted version of the nf-core⁴¹ pipeline for RNA-seq, using Nextflow⁴² (v22.04). Adapters were trimmed with TrimGalore (v.0.6.7) and trimmed reads were aligned to GRCh38 with STAR⁴³ (v.2.6.1d) and quantified with Salmon⁴⁴ (v1.5.2). Sorting and indexing of the bam files was done with SAMtools⁴⁵ (v.1.14). PCA was performed on library-size normalized log values computed with edgeR's (v.3.36.0) cpm function. Mean-variance trend was estimated with limma's (v.3.50.3) voom function. Model fitting was performed with limma lmFit followed by empirical Bayes moderation, and changes with $P_{\text{adj}} < 0.05$ and $|\log_2\text{FC}| > 0.58$ were considered significant. Differential expression between PD-1^{hi} and PD-1^{lo} precursors for all genes from the core exhaustion signature described previously²⁴ was correlated with differential expression between precursor T cells after LCMV Docile and LCMV Armstrong infections at 21 days after infection (GSE142687)²⁴.

scATAC-seq analysis of previously published data

Barcodes, peaks and matrix files from the scATAC-seq data of CD8 T cells seven days after Armstrong or clone 13 infection were downloaded from the GEO database under the accession number GSE164978. Reads were mapped to the prebuilt mm10 reference from 10X (<https://cf.10xgenomics.com/supp/cell-atac/refdata-cellranger-arc-mm10-2020-A-2.0.0.tar.gz>) using cellranger-atac (v.2.0.0). Data were analysed in R (v.4.1.0) using Seurat (v.4.0.3)⁴⁶ and Signac (v.1.3.0)⁴⁷. Features detected in a minimum of 10 cells and cells with at least 200 features were kept. Most frequently observed features were identified using FindTopFeatures with min.cutoff 10. Data were normalized using term-frequency inverse-document-frequency normalization by running RunTFIDF function before reducing dimensionality with RunSVD. Because the first LSI dimension had a strong negative correlation with sequencing depth, only dimensions 2 to 30 were considered for finding neighbours and UMAP reduction. Clusters were identified with the Leiden algorithm (Python v.3.8.10) and a resolution of 0.25. Clusters were annotated on the basis of expression scores calculated

using AddModuleScore for region-based bulk ATAC-seq signatures of naive, exhausted–dysfunctional, terminally exhausted–dysfunctional and precursor exhausted–dysfunctional cells reported previously²⁸. Differentially accessible features between T_{pex} and the other clusters were identified with FindMarker, restricting the search to features detected in at least 10% of either of the populations and using the logistic regression framework as a test. Differences with $P_{\text{adj}} < 0.05$ and average $\log_2\text{FC} > 0.3$ were considered significant. Regions were annotated to genes using the R package EnsDb.Mmusculus.v79 (v.2.99.0) and Signac's function ClosestFeature. Features differentially accessible between T_{pex} and T_{mpe} were considered significant when $P_{\text{adj}} < 0.05$ and average $\log_2\text{FC} > 0.15$. Known transcription-factor-binding sites were retrieved using getMatrixSet from the package TFBSTools (v.1.32.0) and added to the object with Signac's AddMotifs function using the genomic reference from the package BSgenome.Mmusculus.UCSC.mm10 (v.1.4.3). Transcription factors associated with the differential accessibility between T_{pex} and T_{mpe} were identified with the Signac wrapper for chromVAR (v.1.16.0) and FindMarkers applied to the chromvar assay. Tn5 insertion frequency at peaks containing FOS::JUN motifs was visualized with Signac's Footprint function.

Reporting summary

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

Data availability

All RNA-seq datasets have been deposited in the NCBI GEO public database under accession number GSE278807. The following published datasets available in the GEO database were reanalysed: GSE119943 (scRNA-seq), GSE142687 (bulk RNA-seq), GSE164978 (scATAC-seq). All other data supporting this study are available in the main Article and its Supplementary Information.

Code availability

Scripts allowing the reproduction of the data will be shared by D.Z. upon reasonable request.

35. Pircher, H., Burki, K., Lang, R., Hengartner, H. & Zinkernagel, R. M. Tolerance induction in double specific T-cell receptor transgenic mice varies with antigen. *Nature* **342**, 559–561 (1989).
36. Dillon, S. R., Jameson, S. C. & Fink, P. J. V beta 5 + T cell receptors skew toward OVA+H-2Kb recognition. *J. Immunol.* **152**, 1790–1801 (1994).
37. Bategay, M. et al. Quantification of lymphocytic choriomeningitis virus with an immunological focus assay in 24- or 96-well plates. *J. Virol. Methods* **33**, 191–198 (1991).
38. Kim, S. K. et al. Generation of mucosal cytotoxic T cells against soluble protein by tissue-specific environmental and costimulatory signals. *Proc. Natl Acad. Sci. USA* **95**, 10814–10819 (1998).
39. Zehn, D., Lee, S. Y. & Bevan, M. J. Complete but curtailed T-cell response to very low-affinity antigen. *Nature* **458**, 211–214 (2009).
40. Traag, V. A., Waltman, L. & van Eck, N. J. From Louvain to Leiden: guaranteeing well-connected communities. *Sci. Rep.* **9**, 5233 (2019).
41. Ewels, P. A. et al. The nf-core framework for community-curated bioinformatics pipelines. *Nat. Biotechnol.* **38**, 276–278 (2020).
42. Di Tommaso, P. et al. Nextflow enables reproducible computational workflows. *Nat. Biotechnol.* **35**, 316–319 (2017).
43. Dobin, A. et al. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–21 (2013).
44. Patro, R., Duggal, G., Love, M. I., Irizarry, R. A. & Kingsford, C. Salmon provides fast and bias-aware quantification of transcript expression. *Nat. Methods* **14**, 417–419 (2017).
45. Li, H. et al. The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**, 2078–2079 (2009).
46. Hao, Y. et al. Integrated analysis of multimodal single-cell data. *Cell* **184**, 3573–3587 (2021).
47. Stuart, T., Srivastava, A., Madad, S., Lareau, C. A. & Satija, R. Single-cell chromatin state analysis with Signac. *Nat. Methods* **18**, 1333–1341 (2021).

Acknowledgements We thank H. McNamara for insights; T. Walzthöni and X. Pastor Hostench for bioinformatics support; B. Dötterböck, I. Celler, A. Ramirez and W. Schmid for technical assistance; and C. Lechner for animal husbandry. B.H. is supported by the Helmholtz Association under the joint research school Munich School for Data Science (MUDS). Additional funding was provided by the San Salvatore Foundation. Work in the D.Z. laboratory

Article

was supported by a European Research Council starting grant (ProtecTC) and subsequently by a European Research Council consolidator grant (ToCCaTa); grants from the German Research Foundation (SFB1054 and SFB1371); and a grant from the German Israeli Foundation (GIF) (1440).

Author contributions Initial observations that lead to the study: M.W. Conceptualization: M.W., B.H., T.C. and D.Z. Methodology: M.W., B.H., G.P.d.A., C.W., J.B., L.V.D., L.B., F.T., P.S., T.C., B.Y., D.Z. and A.M.S. Investigation: M.W., B.H., G.P.d.A., C.W., J.B., L.V.D., A.-K.G., J.D.C.-M., I.I.K., S.L., L.B., F.T., P.S., T.C., A.M.S. and C.-P.H. Visualization: M.W., B.H., G.P.d.A., M.H., T.C. and D.Z. Funding acquisition: M.H., A.Z., P.A.K. and D.Z. Project administration: M.H. and D.Z. Supervision: M.H., T.C., C.-P.H. and D.Z. Writing (original draft): M.W., B.H., T.C. and D.Z. Writing (review and editing): T.C., M.W., B.H., M.H., C.-P.H. and D.Z.

Funding Open access funding provided by Technische Universität München.

Competing interests The authors declare no competing interests.

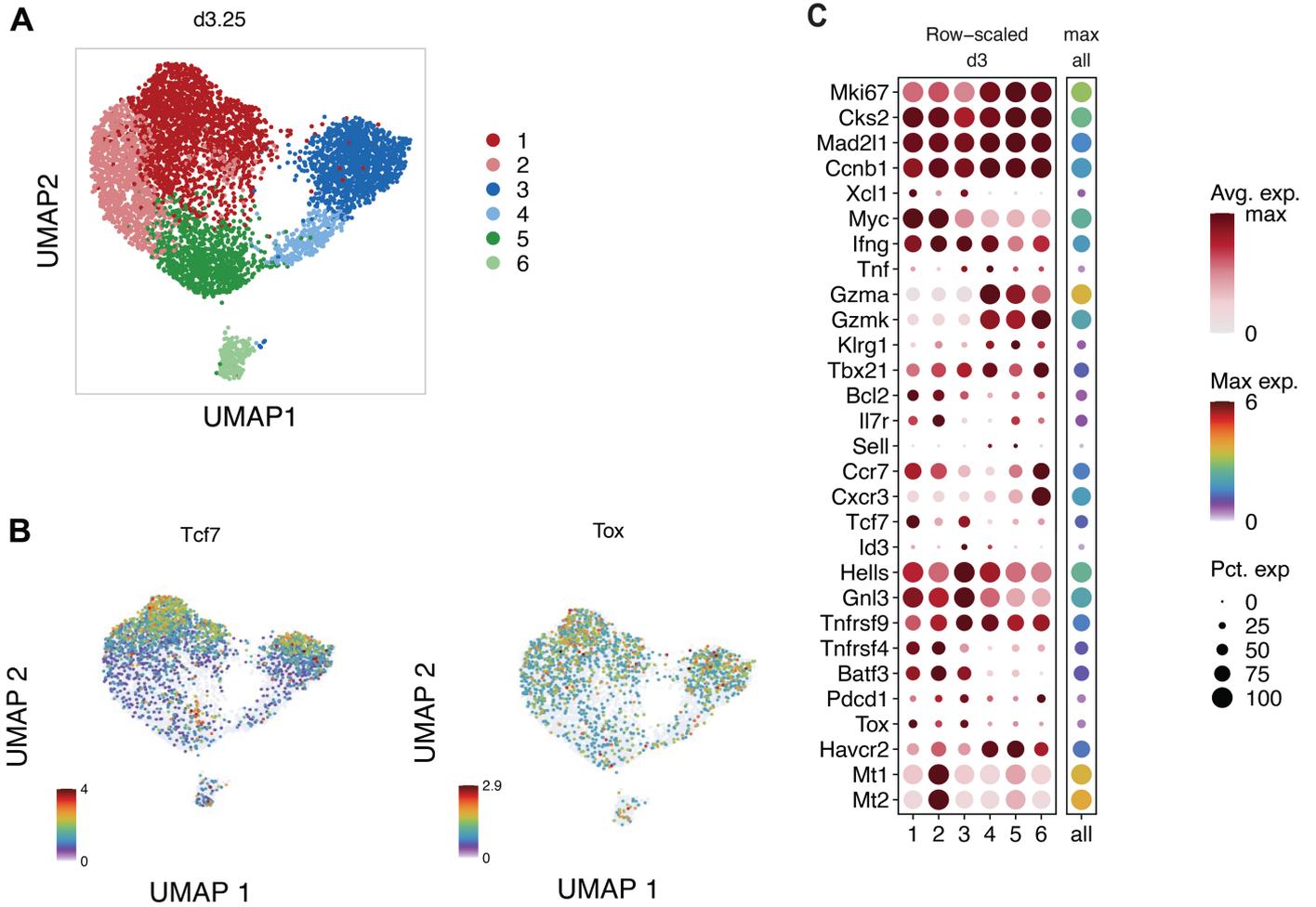
Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41586-024-08451-4>.

Correspondence and requests for materials should be addressed to Ming Wu, Matthias Heinig, Carl-Philipp Hackstein or Dietmar Zehn.

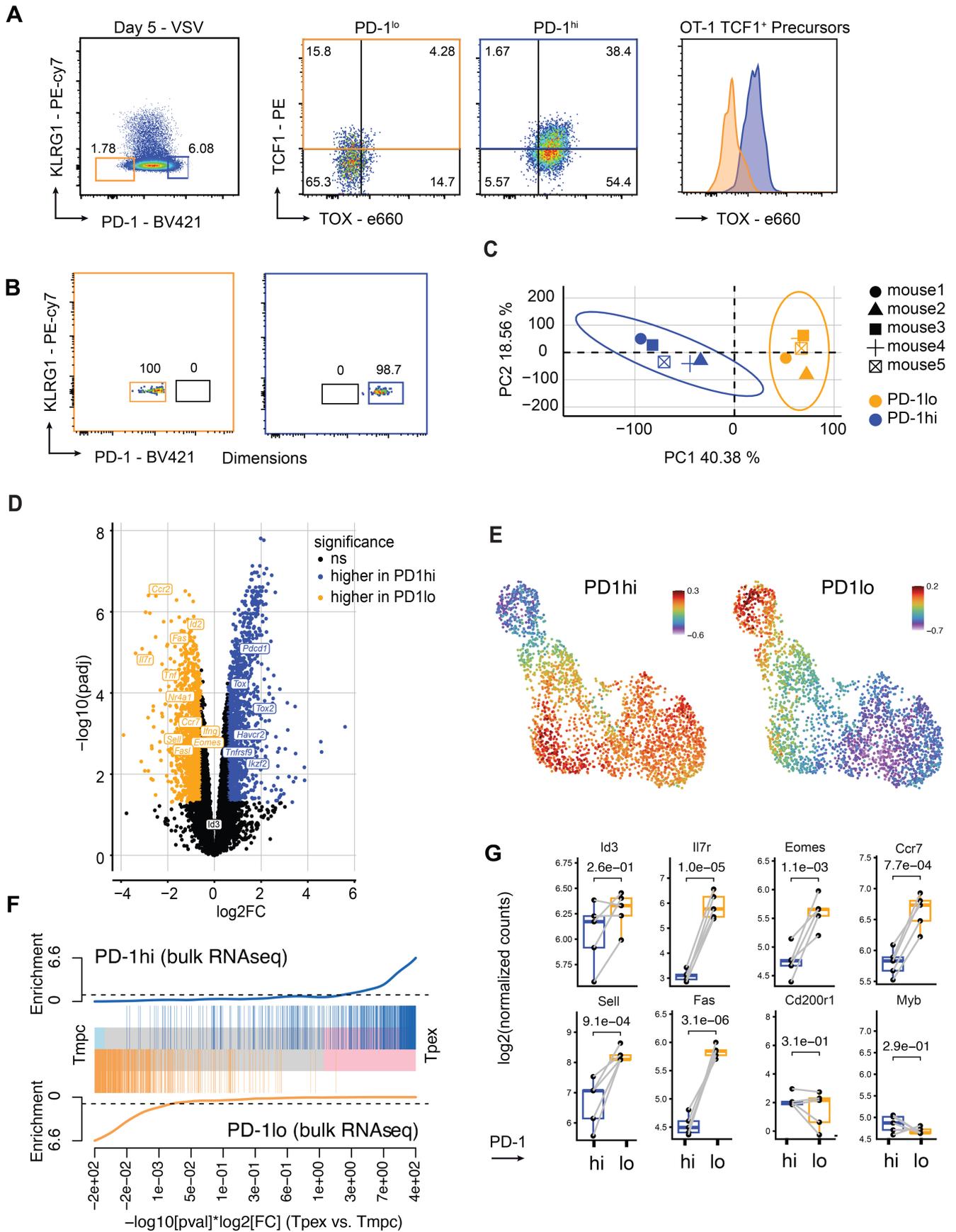
Peer review information *Nature* thanks Stephen Jameson and Tuoqi Wu for their contribution to the peer review of this work.

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



Extended Data Fig. 1 | Precursor populations present at 3.25 days after infection show only minimal differences in the expression of T_{pe} -associated genes. **a**, TCR-transgenic P14 T cells were re-isolated from LCMV ARM infected host mice at day 3.25 after infection and subjected to scRNA-seq. Depicted are

the six Louvain clusters identified at this time point. **b**, Feature plots showing the expression of *Tcf7* and *Tox* in the clusters as in **a**. **c**, Average cluster expression of genes associated with CD8 T cell differentiation, effector function and memory. Data were derived from a single experiment.

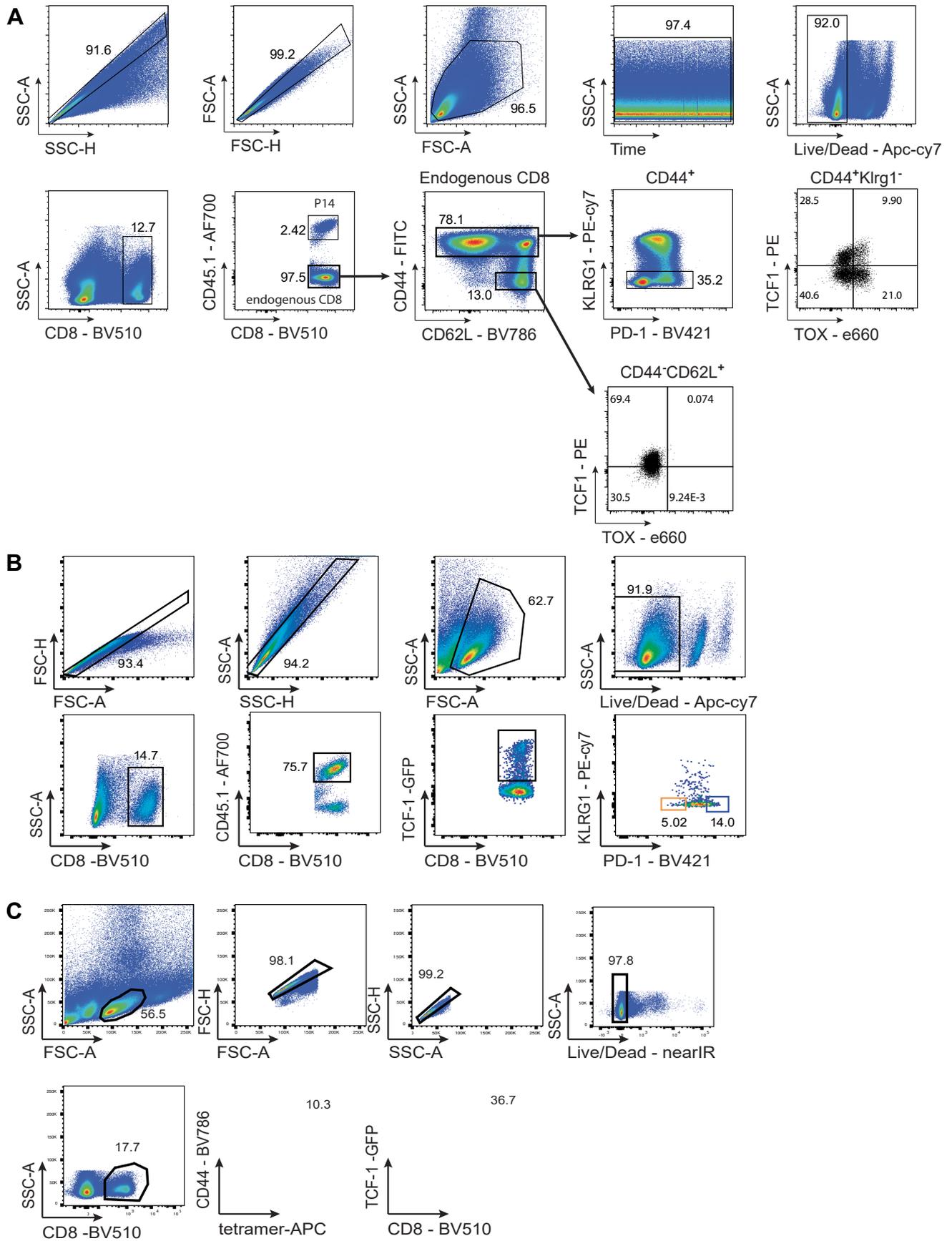


Extended Data Fig. 2 | See next page for caption.

Extended Data Fig. 2 | Transcriptional changes between PD-1^{hi} and PD-1^{lo} precursors are strong and match differences seen in comparisons of LCMV-specific CD8 T cells isolated from chronic versus acute infections. CD45.1 congenic OT-1 (5×10^3) T cells were transferred into C57BL/6 hosts, which were then infected with recombinant ovalbumin expressing vesicular stomatitis virus (VSV). **a**, Expression of TCF1 and TOX were analysed in PD-1^{hi} (blue) or PD-1^{lo} (yellow) KLRG1⁻ OT-I cells from VSV-infected mice. **b**, Adoptively transferred Tcf7^{gfp(bright)mCherry} reporter transgenic P14 T cells were re-isolated from LCMV-infected mice on day 7 after infection and separated into GFP⁺ CD44⁺ KLRG1⁻ PD-1^{lo} vs PD-1^{hi} cells and submitted for sequencing. **c**, Samples of PD-1^{hi} and PD-1^{lo} precursors segregate along principal component 1. **d**, Volcano plot comparing PD-1^{hi} and PD-1^{lo} samples. Significance values were calculated using empirical Bayes moderated t-statistics and were adjusted with

Benjamini–Hochberg correction. **e**, Signature score of genes associated either with PD-1^{hi} or PD-1^{lo} CD8 T cells projected onto the UMAP from Fig. 2a. **f**, Gene set enrichment analysis of signatures from PD-1^{hi} and PD-1^{lo} samples on T_{pex} (cluster 2) vs. T_{mpc} (cluster 1) from Fig. 2a. The background genes were sorted by differential expression levels between these two clusters, weighted by their significance. *P* values were calculated using Wilcoxon rank-sum test. **g**, Log₂ normalized counts of bulk RNA-seq with paired PD-1^{hi} and PD-1^{lo} P14 samples from the same donor mouse. Significance values were calculated as in **d**. *n* = 5 for all groups, lines in **g** represent the mean and symbols identify individual mice. Box plots show median, first and third quartiles (hinges), the lines show the smallest or largest observation within a distance from the nearest hinge of 1.5 times the size of the box. Observations outside this range are shown separately as outliers. Data were derived from three independent experiments (**a, b**) or a single experiment (**c-g**).

Article



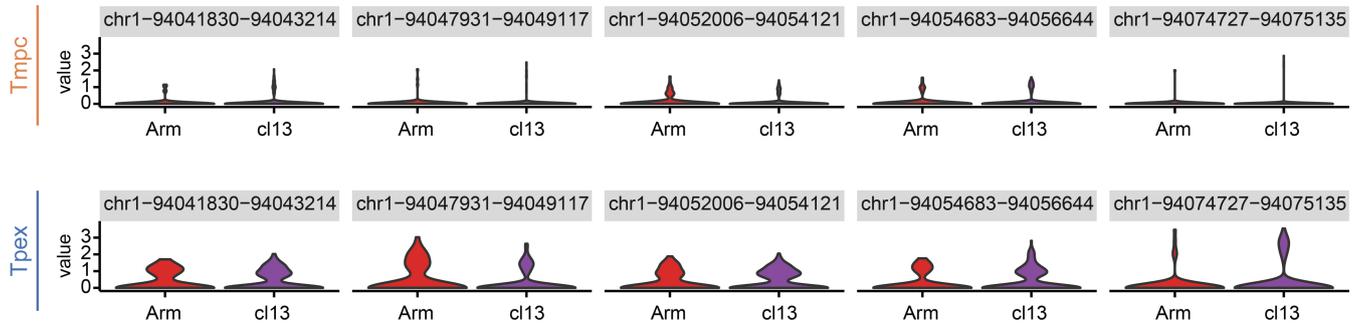
Extended Data Fig. 3 | See next page for caption.

Extended Data Fig. 3 | Gating strategies for FACS analyses and cell sorting.

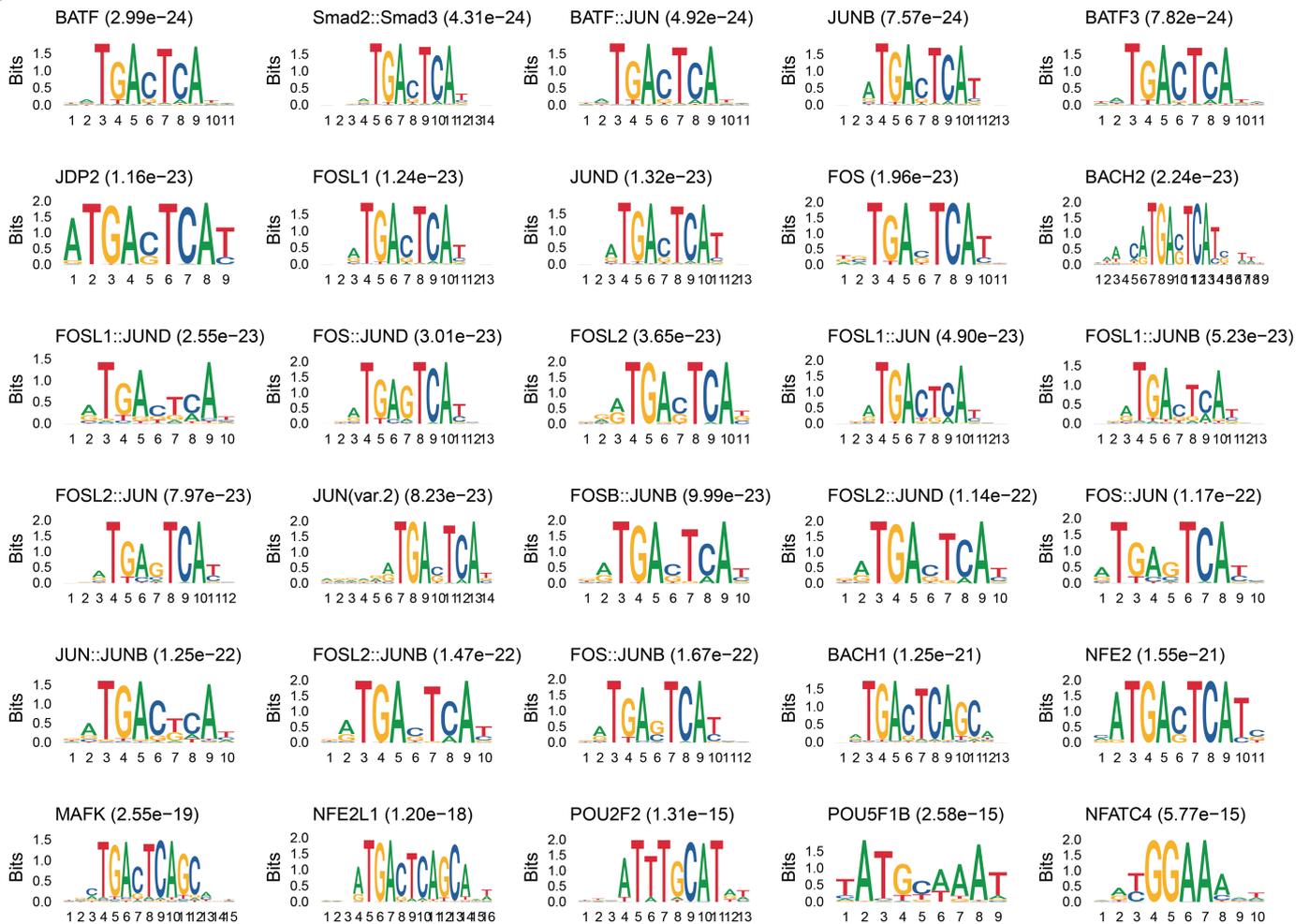
a, Gating strategy to analyse expression of TCF1 and TOX (Fig. 3a,b, Fig. 5a–d and Extended Data Fig. 2a) or the detection of IFN γ and TNF within the TOX+ and TOX- TCF1+ cells (Fig. 3f). Cells were gated for singlets/fsc-ssc/time/live/CD8⁺/CD45.1⁻ or CD45.1⁺/CD44⁺/KLRG⁻ followed by analysis of TCF1 vs TOX. TCF1 and TOX gates were also applied on the endogenous CD44⁻CD62L⁺ naive population to confirm proper placement of TOX gates. **b**, Gating strategy for the sort of precursor populations (Fig. 3c–e and Extended Data Fig. 2b–d,g,h) Cells were purified by MACS column (Miltenyi) through positive selection of

CD45.1⁺ cells and then stained and sorted for singlets/fsc vs ssc/live/CD8⁺/CD45.1⁺/TCF1⁻GFP⁺ and for PD-1^{hi} (blue) vs PD-1^{lo} (orange). For the sort of total P14 cells (Extended Data Fig. 1) the same strategy was used. Here, within the singlets/fsc-ssc/live/CD8⁺ population the CD45.1⁺ cells were sorted. **c**, Gating strategy for the sort of endogenous precursor populations (Fig. 7c–e and Extended Data Fig. 6). Cells were gated for fsc-ssc/singlets/live/CD8⁺/CD44⁺tetramer⁺/TCF1-GFP⁺. The same gating strategy without the last gate was also used for the analysis of TOX and TCF1 in endogenous tetramer⁺ populations (Fig. 7a,b).

A



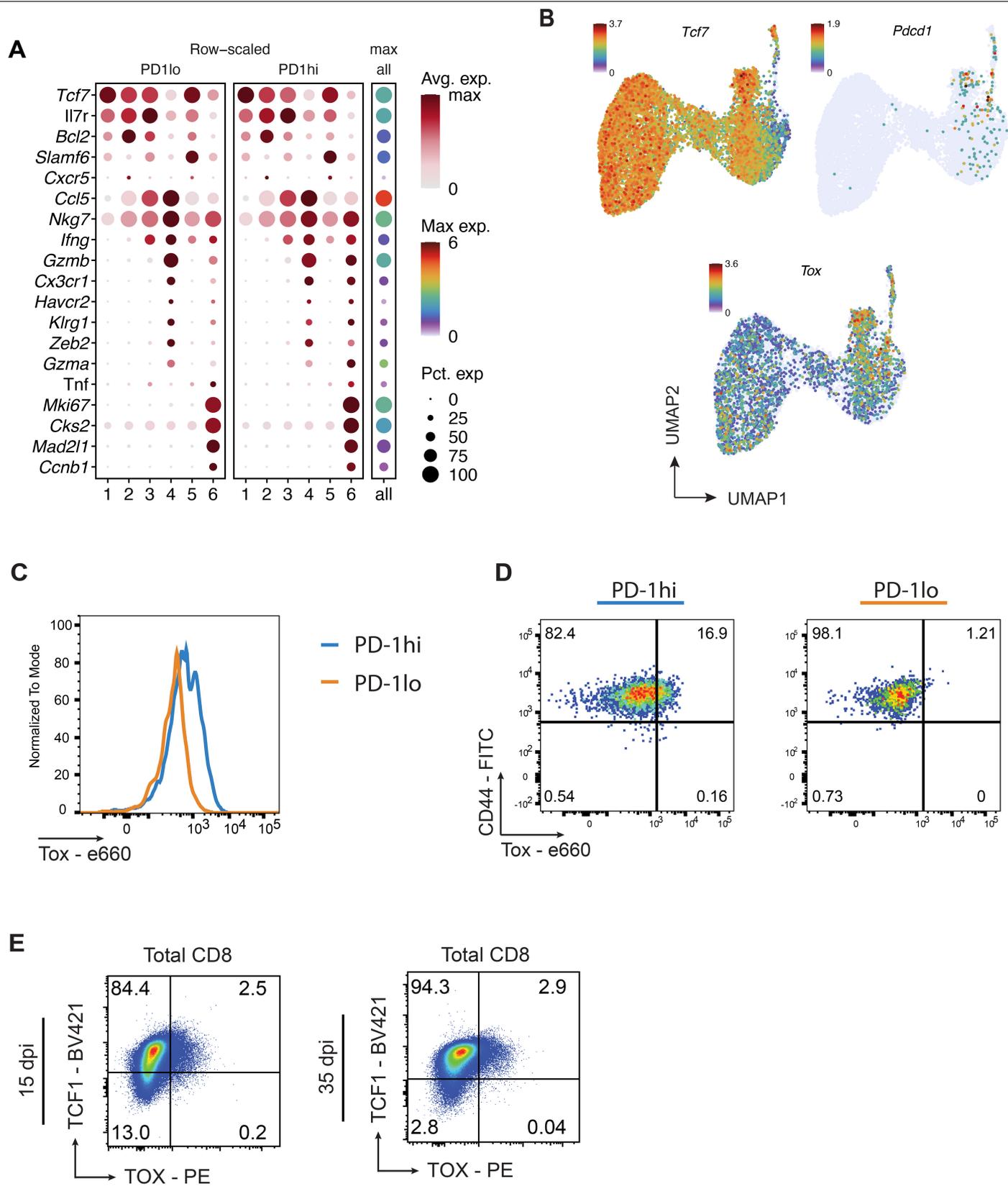
B



Extended Data Fig. 4 | Accessibility at *Pcd1* locus and motif analysis.

a. Accessibility levels of T_{pex} marker regions at *Pcd1* locus (highlighted regions in Fig. 4d) directly comparing T_{pex} and T_{mpc} clusters of Armstrong and clone

13 infections. **b.** Top 30 most differentially active transcription factors comparing T_{mpc} and T_{pex} clusters of Armstrong infection. Data were derived from a single previously published experiment.

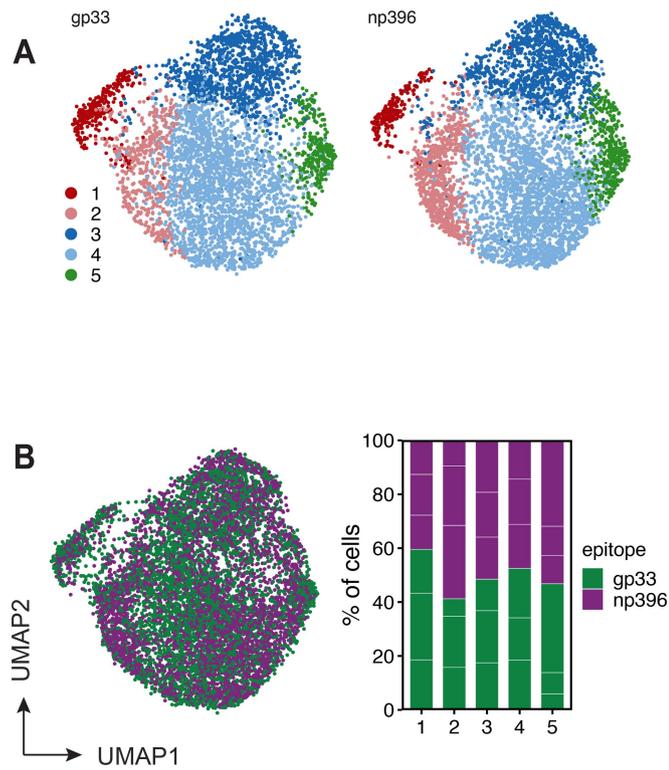


Extended Data Fig. 5 | See next page for caption.

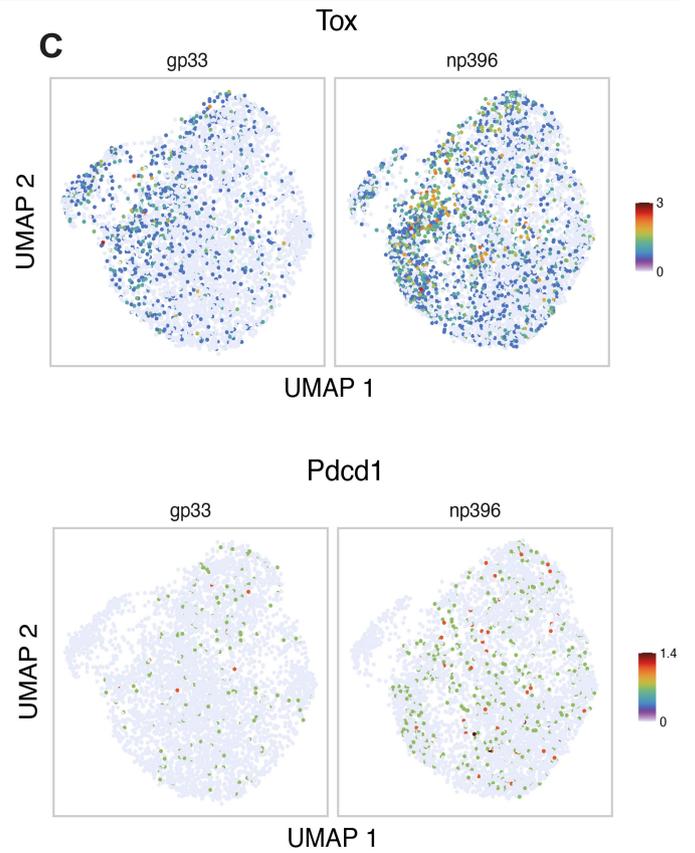
Article

Extended Data Fig. 5 | Early but not late T_{pe} cells show developmental plasticity. **a,b**, CD45.1 congenic P14 TCR-transgenic T cells were transferred into C57BL/6 host mice as indicated in Fig. 6a. PD-1^{hi} and PD-1^{lo} P14s were re-isolated on day 4.5 after an acute LCMV infection and transferred into a new C57BL/6 host that was then also infected with acute LCMV. 15 days later, the progeny of the PD-1^{hi} and PD-1^{lo} P14 cells were again isolated and analysed by scRNA-seq. **a**, Average cluster expression segregated on PD-1^{hi} and PD-1^{lo}. Circle size indicates the percentage of cells expressing the gene of interest. **b**, Feature plots with colours encoding normalized expression values. **c,d**, CD45.1 congenic PD-1^{hi} and PD-1^{lo}

P14 TCR-transgenic T cells were isolated at 7 days after infection and transferred into naive C57BL/6 hosts. 7 days after infecting the secondary hosts with LCMV Armstrong, the progeny of the transferred PD-1^{hi} and PD-1^{lo} P14 T cells was collected and analysed for TOX and CD44 expression. **c**, TOX expression in PD-1^{hi} and PD-1^{lo} cells. **d**, Co-expression of TOX with CD44 in the indicated cell populations. **e**, C57BL/6 mice were infected with LCMV Armstrong and total splenocytes were isolated 15 and 35 days after infection. Shown are TCF1 and TOX expression by total endogenous CD8 T cells. Data were derived from a single experiment (**a-d**) or three independent experiments (**e**).



Extended Data Fig. 6 | Signs of exhaustion can be found in subsets of np396- but not gp33-specific CD8 T cells. **a**, gp33 and np396-specific TCF1⁺ precursors were purified from TCF1 reporter mice (n = 3) 4 weeks after infecting the mice with LCMV Armstrong. Purified cells were subjected



to scRNA-seq. **a**, UMAP representation of Louvain clusters. **b**, Epitope fractions per cluster. **c**, Feature plots depicting the expression of *Tox* or *Pdcd1* in either gp33- or np396-specific CD8 T cells. Data were derived from a single experiment.

Article

Extended Data Table 1 | Activation and exhausted gene signatures

Signature type	Signatures
Activation signatures	MYC
Activation signatures	XCL1
Activation signatures	TNFRSF4
Activation signatures	IL2RA
Activation signatures	ID3
Activation signatures	IRF4
Activation signatures	TNF
Activation signatures	IFNG
Exhaustion signatures	PDCD1
Exhaustion signatures	TOX
Exhaustion signatures	TNFRSF9
Exhaustion signatures	HELLS
Exhaustion signatures	IKZF2
Exhaustion signatures	LAG3
Exhaustion signatures	NR4A1
Exhaustion signatures	NR4A2
Exhaustion signatures	PRF1
Exhaustion signatures	RGS16
Exhaustion signatures	DUSP4
Exhaustion signatures	MT1
Exhaustion signatures	MT2

Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our [Editorial Policies](#) and the [Editorial Policy Checklist](#).

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

- | n/a | Confirmed |
|-------------------------------------|--|
| <input type="checkbox"/> | <input checked="" type="checkbox"/> The exact sample size (n) for each experimental group/condition, given as a discrete number and unit of measurement |
| <input type="checkbox"/> | <input checked="" type="checkbox"/> A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly |
| <input type="checkbox"/> | <input checked="" type="checkbox"/> The statistical test(s) used AND whether they are one- or two-sided
<i>Only common tests should be described solely by name; describe more complex techniques in the Methods section.</i> |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> A description of all covariates tested |
| <input type="checkbox"/> | <input checked="" type="checkbox"/> A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons |
| <input type="checkbox"/> | <input checked="" type="checkbox"/> A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals) |
| <input type="checkbox"/> | <input checked="" type="checkbox"/> For null hypothesis testing, the test statistic (e.g. F , t , r) with confidence intervals, effect sizes, degrees of freedom and P value noted
<i>Give P values as exact values whenever suitable.</i> |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes |
| <input type="checkbox"/> | <input checked="" type="checkbox"/> Estimates of effect sizes (e.g. Cohen's d , Pearson's r), indicating how they were calculated |

Our web collection on [statistics for biologists](#) contains articles on many of the points above.

Software and code

Policy information about [availability of computer code](#)

Data collection Facs Fortessa (BD Biosciences), Facs Aria III (BD Biosciences), Cytoflex LX (Beckman Coulter), 10X genomics, Illumina NovaSeq 6000

Data analysis FACS data – BD Bioscience FlowJo v10.09, Graph Pad Prism v.8, Microsoft Excel v16.89.1

scRNA-Seq: R v4.0.4, v4.1.0. and v4.3, python v3.8.12- a detailed overview describing the bioinformatic pipelines is included in the Supplementary Information.

RNA-seq: bcl2fastq (v2.20.0.422), R (v4.0.4, v4.1.0, v4.1.1 , v4.3.0), python (v3.8.10, v3.8.12, v3.9.6), jupyter notebook (v6.4.3), cellranger multi (v7.1.0), cellranger-atac (v2.0.0), BD Rhapsody™ Sequence Analysis Pipeline (r16), soupX (v1.6.1), scanpy (v1.8.1) , scDbfFinder (v1.8.0), sctransform (v0.3.2, v0.3.3, v0.3.5), rpy2 (v3.4.5), anndata2ri (v1.1), SingleCellExperiment (v1.16.0), Seurat (v4.0.1, v4.0.3, v4.1.1, v4.3.0), SeuratDisk (v0.0.0.9021), Signac (v1.3.0), EnsDb.Mmusculus.v79 (v2.99.0), mvTCR (v0.1.1), diffxpy (v0.7.4), limma (v3.50.3, v3.56.1), velocity (v0.17), scvelo (v0.2.4), AnnData (v0.7.6), scirpy (v0.12.2) , nextflow (v22.04), nf-core (v3.9), TFBSTools (v1.32.0), BSgenome.Mmusculus.UCSC.mm10 (v1.4.3), chromVAR (v1.16.0), TrimGalore (v0.6.7), and Cutadapt (v3.4), STAR (v2.6.1d), Salmon (v1.5.2), SAMtools45 (v1.14), edgeR (v3.36.0).

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio [guidelines for submitting code & software](#) for further information.

Data

Policy information about [availability of data](#)

All manuscripts must include a [data availability statement](#). This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our [policy](#)

All RNA-sequencing datasets have been deposited in the NCBI GEO public database under accession number GSE278807.

The following published data sets available in GEO database were reanalyzed: GSE119943 (scRNAseq), GSE142687 (bulk RNA-seq), GSE164978 (scATAC-seq).

The following genome references and annotation were used: refdata-cellranger-arc-mm10-2020-A-2.0.0, refdata-cellranger-vdj-GRCm38-alt-ensembl-7.0.0 (GRCm38 - gencode M23) and RhapRef_Mouse_WTA_2023-02 (GRCm38, gencode vM24).

Source data are provided with this paper. All other data supporting this study are available in the main article and Supplementary Information.

Research involving human participants, their data, or biological material

Policy information about studies with [human participants or human data](#). See also policy information about [sex, gender \(identity/presentation\), and sexual orientation](#) and [race, ethnicity and racism](#).

Reporting on sex and gender	<input type="text" value="n/a"/>
Reporting on race, ethnicity, or other socially relevant groupings	<input type="text" value="n/a"/>
Population characteristics	<input type="text" value="n/a"/>
Recruitment	<input type="text" value="n/a"/>
Ethics oversight	<input type="text" value="n/a"/>

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

Life sciences Behavioural & social sciences Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see [nature.com/documents/nr-reporting-summary-flat.pdf](https://www.nature.com/documents/nr-reporting-summary-flat.pdf)

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	Statistical methods were not used to predetermine sample size. We chose sample sizes based on previous experience (Alfei et al., Nature, 2019, von Hoesslin et al. Science Immunology, 2022, Scherer et al. Nature Immunology 2023) and with the goal of producing statistically robust data while respecting animal welfare regulations.
Data exclusions	For TCR-scRNAseq – one sample was excluded during the FACS sorting because we were unable to obtain sufficient cell numbers owed to an unsuccessful infection, all other samples were included for sequencing and used for the subsequent analysis. We did not exclude any samples except for rare events where the adoptive cell transfer or the infection were unsuccessful (2+1 cases in the entire manuscript).
Replication	Experiments were repeated to ensure reproducibility. Number of repeats and sample size are provided in each figure legend.
Randomization	Mice of the appropriate genotype and age were randomly assigned to the experimental groups. No experiments using material from any other species were performed in this study.
Blinding	We did not perform any readouts involving subjective evaluations, such as histological analysis or clinical scoring. Due to bio-safety regulations, we were instructed not to perform blinding in experiments involving biohazardous substances.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
<input type="checkbox"/>	<input checked="" type="checkbox"/> Antibodies
<input type="checkbox"/>	<input checked="" type="checkbox"/> Eukaryotic cell lines
<input checked="" type="checkbox"/>	<input type="checkbox"/> Palaeontology and archaeology
<input type="checkbox"/>	<input checked="" type="checkbox"/> Animals and other organisms
<input checked="" type="checkbox"/>	<input type="checkbox"/> Clinical data
<input checked="" type="checkbox"/>	<input type="checkbox"/> Dual use research of concern
<input checked="" type="checkbox"/>	<input type="checkbox"/> Plants

Methods

n/a	Involved in the study
<input checked="" type="checkbox"/>	<input type="checkbox"/> ChIP-seq
<input type="checkbox"/>	<input checked="" type="checkbox"/> Flow cytometry
<input checked="" type="checkbox"/>	<input type="checkbox"/> MRI-based neuroimaging

Antibodies

Antibodies used	All antibodies used, their clone and catalogue numbers as well as dilutions used are listed in Table 2.
Validation	Antibodies were all validated by the manufacturer or previous experiments and studies (Alfei, F. et al. Nature, 2019.)

Eukaryotic cell lines

Policy information about [cell lines and Sex and Gender in Research](#)

Cell line source(s)	Vero cells, originally provided by M. J. Bevan, University of Washington BHK-21 cells, originally provided by M. J. Bevan, University of Washington
Authentication	Since we used only cell-lines that we have been using in the lab for several years, we did not specifically authenticate them. However, we made sure that all used cell-lines showed the expected results and were in good condition prior to using them.
Mycoplasma contamination	Mycoplasma tests were routinely performed and were negative.
Commonly misidentified lines (See ICLAC register)	To the best of our knowledge, no commonly misidentified cell lines were used in the study.

Animals and other research organisms

Policy information about [studies involving animals; ARRIVE guidelines](#) recommended for reporting animal research, and [Sex and Gender in Research](#)

Laboratory animals	C57BL/6 mice were obtained from Charles River and C57BL/6.SJL from Jackson Laboratory and both lines were maintained by intercrossing. P14 TCR transgenic (provided by A. Oxenius, ETH, Zürich, Switzerland), Vb5 TCR transgenic (provided by P. Fink, University of Washington, Seattle, USA), OT-I TCR transgenic mice (purchased from Jackson Laboratories) and PD-1ko P14 TCR transgenic (provided by Ping-Chih Ho (University of Lausanne, Switzerland) mice were obtained on a C57BL/6 background and maintained by crossing them to C57BL/6.SJL mice. P14 TCR transgenic Tcf7-GFP mice were generated by the Zehn lab in the past and have been published previously (Tsui et al., Nature, 2022) 6-13 week old male and female mice from these were used for experiments and were housed under the following conditions: - light cycle is 7:00 AM on, 7:00 PM off - Temperature at 22-26°C - Humidity is between 30/70 g/m3
Wild animals	No samples from wild animals were used in this manuscript.
Reporting on sex	There was no discrimination of sex for these experiments. All experiments used a combination of male and female recipients and CD45.1 P14 or OT-I donors.
Field-collected samples	No field-collected samples were used in this manuscript.
Ethics oversight	Experiments performed with mice were approved by the veterinarian authorities of the 'Regierung von Oberbayern' in Germany.

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Plants

Seed stocks	n/a
Novel plant genotypes	n/a
Authentication	n/a

Flow Cytometry

Plots

Confirm that:

- The axis labels state the marker and fluorochrome used (e.g. CD4-FITC).
- The axis scales are clearly visible. Include numbers along axes only for bottom left plot of group (a 'group' is an analysis of identical markers).
- All plots are contour plots with outliers or pseudocolor plots.
- A numerical value for number of cells or percentage (with statistics) is provided.

Methodology

Sample preparation	Single cell splenocyte suspensions were prepared from naive or infected mice by mashing spleen through a 100µm cell strainer followed by hypotonic ACK lysis. Total splenocytes were used for flow cytometry analysis and sorting. Following isolation, cells were incubated with Zombie NIR dye (Biolegend, 423106) and Fc-blocking reagent 2.4G2 in PBS (ThermoFisher) for 15 minutes at room temperature (RT). Next, cells were washed and resuspended in freshly prepared master mix containing fluorescent staining antibodies in fluorescence-activated cell sorting (FACS) buffer (PBS, 2% FCS, 0.1% sodium azide). When master mix contained 2 or more brilliant violet polymer dyes, Super bright staining buffer (Invitrogen SB-4401-42) was used. For tetramer or CXCR5 staining, cells were stained for one hour at room temperature. FoxP3/Transcription factor staining buffer (Thermo Fisher 00-5532-00) was used to perform intranuclear staining following manufacturer recommended protocols. Samples were acquired on BD FACS Fortessa or Beckman Coulter CytoFlex LX instruments. For sorting: following live/dead stain and Fc block, cells were incubated with freshly prepared master mix for 20 minutes at RT, followed by wash in Magnetic-activated cell sorting buffer (PBS, 1% FCS, 2 mM EDTA). Cells were immediately run on BD FACSAria Fusion sorter. Single-stained controls were prepared using Ultracomp eBeads (Invitrogen 01-2222-42) for every experiment. Beads were treated like the cell containing samples (including fixation protocol).
Instrument	FACS data collection took place on a FACS Fortessa (BD Biosciences) or a Cytoflex LX (Beckman Coulter). Cell sorting was executed on FACS Aria Fusion (BD Biosciences).
Software	FACSDiva version, FlowJo version 10.09 (BD Bioscience).
Cell population abundance	Purity of sorted cell populations was >95%.
Gating strategy	Cells from mice lacking the Tcf7-reporter construct were gated as single, lymphocyte (SSC-A vs FSC-A), time (consistent), live CD8a+, CD45.1+ or CD45.1-, CD44+ or CD44- CD62L+ CD44- . Within the CD44+ populations cell were further gated as KLRG1-. Subsequently expression of transcription factors, or within TOX+ and TOX- populations of the TCF1+ subset, the expression of Nur77 or effector molecules (IFNγ and TNF) was analysed. Cells expressing the Tcf7-GFP reporter were gated as single, lymphocyte (SSC-A vs FSC-A), live, CD8a+, CD45.1+ Tcf7-GFP+, KLRG1-, PD-1low or PD-1hi . Gating strategies are displayed in Extended Data Figure 3.

- Tick this box to confirm that a figure exemplifying the gating strategy is provided in the Supplementary Information.