1 **Divergent clonal differentiation trajectories of T cell exhaustion**

2

3 Bence Daniel[1,2,9], Kathryn E. Yost[2,9], Katalin Sandor[1], Yu Xia[4], Yanyan Qi[1], Kamir J.
4 Hiam-Galvez[1], Stefanie L. Meier[1,5], Julia A. Belk[1], Josephine R. Giles[6,7,8], E. John
5 Wherry[6,7,8], Howard Y. Chang[2,3*], Takeshi Egawa[4*], Ansuman T. Satpathy[1,5*#]

6

7 [1]Department of Pathology, Stanford University, Stanford, CA, USA.
8 [2]Center for Personal Dynamic Regulomes, Stanford University, Stanford, CA, USA.
9 [3]Howard Hughes Medical Institute, Stanford University, Stanford, CA, USA.
10 [4]Department of Pathology and Immunology, Washington University School of Medicine,
11 MO, St. Louis.
12 [5]Parker Institute for Cancer Immunotherapy, San Francisco, CA 94129, USA.
13 [6]Department of Systems Pharmacology and Translational Therapeutics, University of
14 Pennsylvania, Philadelphia, PA, USA.
15 [7]Institute for Immunology, Perelman School of Medicine, University of Pennsylvania,
16 Philadelphia, PA, USA.
17 [8]Parker Institute for Cancer Immunotherapy at University of Pennsylvania, Philadelphia,
18 PA, USA.
19 [9]*These authors contributed equally to this work.*
20 *Co-senior authors*
21 [#]*Corresponding author*

22

23 *To whom correspondence should be addressed. E-mail: satpathy@stanford.edu*

24

25 Key words: T cell exhaustion, single-cell genomics, clonal T cell differentiation, chronic
26 viral infection

27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46

SUMMARY

T cells activated by chronic antigen exposure in the setting of viral infections or cancer can adopt an exhausted T cell (Tex) state, characterized by reduced effector function and proliferative capacity, and the upregulation of inhibitory receptors. However, whether all antigen-specific T cell clones follow the same molecular and cellular Tex differentiation trajectory remains unclear. Here, we generate a single-cell multi-omic atlas of T cell exhaustion that redefines the phenotypic diversity and molecular regulation of Tex phenotypes. Longitudinal analysis during chronic viral infection identifies an early effector phenotype that is epigenetically primed for Tex differentiation and two late-stage Tex cell states with either a terminal exhaustion or a killer cell lectin-like receptor (KLR)-expressing cytotoxic gene signature. We define clonal trajectories of antigen-specific T cells using paired single-cell RNA and T cell receptor sequencing and reveal distinct differentiation trajectories resulting in terminal Tex-biased, KLR Tex-biased, or divergent clones that differentiate into both phenotypes. Comparison of Tex phenotypes among shared T cell clones that traffic to multiple organs reveals that clonal differentiation trajectories are maintained across tissues. Finally, we show that differences in clonal differentiation trajectory are driven by TCR signal strength, whereby high-affinity T cell clones preferentially adopt a terminal Tex fate, while low-affinity clones adopt an effector-like KLR Tex fate that is detectable long-term but depleted in high antigen settings. These findings reveal clonal heterogeneity in the T cell response to chronic antigen and genomic programs that underlie Tex fates and persistence.

Highlights

- A single-cell atlas of T cell exhaustion identifies novel early effector and KLR Tex states.
- Clonal T cell analysis defines divergent differentiation trajectories during chronic viral infection leading to terminal and KLR Tex fates.
- The heterogeneity of the Tex pool arises from three primary differentiation patterns and are differentially persistent in the setting of high antigen.
- Clonal Tex differentiation patterns are conserved across organ sites and driven by TCR signal strength.

93

Introduction

95   Chronic antigen exposure during chronic viral infections and cancer leads to
96 impaired CD8[+] T cell responses, termed T cell exhaustion [1, 2]. Exhausted T cells
97 (Tex) are characterized by reduced effector function, diminished proliferative capacity,
98 and high expression of inhibitory receptors, including PD-1, LAG-3, and TIM3. However,
99 Tex are able to maintain some effector functions and persist long-term, suggesting that
100 T cell exhaustion may represent a mechanism to control pathogen burden while
101 maintaining immune homeostasis [3, 4]. Recent studies have identified heterogeneity in
102 Tex phenotypes, which are characterized by distinct surface receptors, functionality,
103 proliferative capacity, and tissue localization during chronic viral infections and cancer
104 [5-12]. Some of these studies have proposed a linear differentiation model, whereby
105 progenitor Tex (Tex[prog]; marked by expression of TCF1 and CXCR5) self-renew and
106 maintain downstream Tex subsets, including CX3CR1[+]PD-1[+] intermediate Tex (Tex[int])
107 with proliferative, cytolytic and memory potential, and PD-1[+] TIM3[+] terminal Tex
108 (Tex[term]; marked by high expression of inhibitory receptors, and limited effector or
109 proliferative potential) [5-9, 13, 14, 60]. These subpopulations exhibit distinct epigenetic
110 states and rely on distinct transcription factors (TFs). TCF1 (encoded by *Tcf7*) and
111 BACH2 are indispensable for the formation of the Tex[prog] phenotype, while the high
112 mobility group TF, TOX, orchestrates the establishment and maintenance of the
113 molecular program of exhaustion in all Tex states, including Tex[term], and is required for
114 their survival [4, 7, 15-19]. Finally, these Tex populations are further distinguished by
115 their ability to respond to immune checkpoint blockade (ICB); Tex[term] possess a stable
116 epigenetic program and cannot be efficiently reinvigorated by ICB, while Tex[prog] can
117 proliferate in response to ICB and may be important for the therapeutic responses [3, 6,
118 20]. Despite these advances, we lack a comprehensive view of Tex states, their clonal
119 relationships, and the molecular programming underlying their differentiation,
120 particularly in polyclonal T cell responses.

121   Here, we generate a comprehensive atlas of Tex differentiation using single-cell
122 chromatin accessibility, transcriptome, and T cell receptor (TCR) sequencing of antigen-
123 specific CD8[+] T cells in the setting of chronic lymphocytic choriomeningitis virus (LCMV)
124 infection. We discover previously unappreciated Tex subsets, including an early effector
125 exhausted subset (Tex[eeff]) in the early phase of infection that initiates the molecular
126 program of exhaustion, and a killer cell lectin-like receptor-expressing Tex subset
127 (Tex[KLR]) as a late-stage phenotype concurrent with terminal Tex, which suggests a
128 divergent developmental path during Tex differentiation. T cell clone tracing based on
129 paired scRNA/TCR-seq nominates unexpected diversity in Tex differentiation
130 trajectories; namely, chronic antigen-specific T cell clones can adopt Tex[term]-biased,
131 Tex[KLR]-biased, or divergent fates, comprising both cell types. Multi-organ clonal analysis
132 reveals that Tex clones traffic to multiple organ sites and that their differentiation
133 trajectories are conserved across tissues; however, Tex[KLR]-biased clones are depleted
134 in the liver, suggesting that Tex[term] cells may be phenotypically adapted for high-antigen
135 tissue microenvironments. Finally, we show that clone behaviors are programmed by
136 TCR affinity to cognate antigen; high-affinity TCR clones are biased towards a Tex[term]
137 differentiation trajectory, while low-affinity TCR clones are biased towards a Tex[KLR]
138 trajectory. Overall, these results provide an in-depth view of the gene regulatory

139 programs and clonal dynamics of Tex states during chronic infection and suggest that a
140 polyclonal T cell response to chronic antigen may balance T cell states that perform
141 effector and memory functions.
142 **Results**
143
144 **A multi-omic single-cell atlas of CD8$^+$ T cell differentiation during acute and**
145 **chronic viral infection.**
146   To profile CD8$^+$ T cell differentiation during T cell exhaustion, we used mouse
147 models of acute (LCMV Armstrong – Arm) or chronic (LCMV Clone 13 – Cl13) viral
148 infection. These two viral strains only differ by two amino acids, and the
149 immunodominant epitopes are identical, enabling direct comparisons of antigen-specific
150 T cell responses in both models [21, 22]. We generated paired single-cell RNA- and T
151 cell receptor (TCR)-sequencing (scRNA/TCR-seq) and single-cell assay for transposase
152 accessible chromatin with sequencing (scATAC-seq) data from LCMV glycoprotein 33-
153 41 tetramer positive (gp33$^+$) and tetramer negative (gp33$^-$) splenic CD8$^+$ T cells at two
154 timepoints (Day 8 and Day 21 post-infection) for both infection models (**Figure 1A-C**).
155 At Day 21 (D21) of Cl13 infection, we also generated scRNA/TCR-seq of gp33$^+$ and
156 gp33$^-$ populations from two additional organs with known differences in viral antigen
157 load (lung and liver; **Figure 1B and Figure S1A**) [23, 24]. Finally, we sorted D21 Cl13
158 splenic T cells using previously defined surface markers that identify Tex^{prog} (SLAMF6$^+$),
159 Tex^{int} (CX3CR1$^+$), and Tex^{term} (PD-1$^+$, SLAMF6$^-$ and CX3CR1$^-$) phenotypes and
160 performed scRNA/TCR- and scATAC-seq (**Figure S1B**) [7-9, 13]. In total, we obtained
161 96,750 scRNA-seq profiles that passed quality control filters based on the detected
162 gene count (>200 genes/cell), mitochondrial content (<5% mitochondrial RNA
163 content/cell), and predicted doublets (**Figure 1D and S1C, Methods**). Of scRNA-seq
164 profiles passing quality control filters, we detected TCR alpha and beta sequences in
165 88,696 T cells (91.7%) and 5,197 expanded T cell clones (clones >1 cell; **Figure 1D**).
166 We obtained 62,731 scATAC-seq profiles that passed quality control filters based on the
167 unique ATAC-seq fragment count (>1,000 fragments/cell), median read enrichment at
168 transcription start sites (>4 TSS score), and predicted doublets (**Figure 1E, S1D and**
169 **S1E, Methods**).
170   After scRNA-seq quality control filtering, we performed uniform manifold
171 approximation and projection (UMAP), followed by dimensionality reduction and
172 identified 11 scRNA-seq clusters, which were annotated based on differentially
173 expressed genes (DEGs; log$_2$ FC >0.25, Bonferroni adjusted p-value <0.01). In Arm
174 infection, we observed expected T cell phenotypes, including naïve T cells (T^{naïve}; *Ccr7*,
175 *Sell,* and *Lef1,* 248 DEGs), effector T cells (T^{eff}; *Klrg1* and *Ly6c2*, 143 DEGs), effector
176 memory T cells (T^{em}; *Klrb1c*, *Klrd1* and *S1pr1*, 193 DEGs), and memory T cells (T^{mem};
177 *Il7r*, *Arl4c* and *Il18r1*, 35 DEGs; **Figure 1F, Table S1**). In Cl13 infection, we observed
178 Tex^{prog} (*Tcf7*, *Slamf6* and *Id3*, 117 DEGs), Tex^{int} (*Lgals3*, *S100a4* and *Mki67*, 113
179 DEGs), and Tex^{term} (*Gzma*, *Bcl2, Cd101* and *Entpd1*; n=138), as expected (**Figure 1D**
180 **and 1F**). In addition, we also observed early effector exhausted cells (Tex^{eeff}; *Xcl1*,
181 *Top2a* and *Mif*, 1,059 DEGs; a predominant population on D8 of C13 infection), killer
182 cell lectin-like receptor (KLR)-expressing exhausted cells (Tex^{KLR}; *S1pr5*, *Cx3cr1*, *Klrc1*
183 and *Zeb2*, 260 DEGs; emerging specifically late in C13), lung terminal exhausted cells
184 primarily detected in the lung (Tex^{lung}; *Lag3*, *Ifng*, *Ccl3* and *Ccl4*, 247 DEGs), and

4

185 interferon signature gene (ISG) exhausted T cells (Tex^{ISG}; *Isg15*, *Ifit1*, *Ifit3* and *Isg20*,
186 273 DEGs; **Figure 1D and F**).
187   We observed 8 analogous T cell populations in the scATAC-seq data and
188 annotated each cluster based on differential chromatin accessibility at marker gene loci
189 identified in scRNA-seq clusters (i.e., Gene Score, log$_2$ FC > 0.5, FDR < 0.01; **Figure**
190 **1G**) and integration with scRNA-seq data (**Figure S1F, Methods**). Since our primary
191 goal was to analyze Tex differentiation, we did not perform scATAC-seq at D21 in Arm
192 infection, or in lung or liver T cells in Cl13 infection; thus, scATAC-seq clusters did not
193 include T^{em}, T^{mem}, or Tex^{lung} subsets. However, scATAC-seq clusters did reveal
194 additional heterogeneity in the early effector subsets, including three early
195 activated/effector populations from the D8 time point in the two infection models. These
196 subsets did not co-cluster with D21 Tex populations, and included two effector
197 populations (T^{eff} and T^{eff2}) - mainly derived from the Arm condition - and an early
198 effector exhausted population from the Cl13 condition (Tex^{eeff}; **Figure 1E and G, Table**
199 **S2**).
200   scATAC-seq profiles were analyzed at the level of: (1) chromatin accessibility of
201 *cis*-elements (open chromatin regions; OCRs), (2) gene activity scores, computed from
202 the accessibility of enhancers linked to a single gene promoter based on proximity and
203 co-accessibility, and (3) transcription factor (TF) activity, computed from the enrichment
204 of TF binding sites in OCRs or the accessibility of TF binding sites genome-wide in each
205 single cell [25, 26]. Analysis of *cis*-elements identified cell type-specific OCRs (T^{naive} -
206 12,444; T^{eff} - 331; T^{eff2} - 4,070; and Tex^{eeff} - 1,463; Tex^{prog} - 4,532; Tex^{int} - 448; Tex^{term} -
207 2,264; Tex^{KLR} - 668; **Figure 1H, Table S2**), and accessibility was correlated with gene
208 expression at marker gene loci that define Tex subsets, including *Tcf7* (T^{naive}, Tex^{prog}),
209 *Pdcd1* (Tex^{prog}, Tex^{int}, and Tex^{term}), and *Tox* (Tex^{prog}, Tex^{int}, and Tex^{term}; **Figure 1I**,
210 **Methods**). TF motif enrichment analysis at cell type-specific OCRs identified TF motifs
211 enriched in specific T cell subsets. As expected, T^{naive}-specific OCRs were enriched for
212 the TCF/LEF motifs, which were also enriched in Tex^{prog}, along with other known Tex^{prog}
213 TFs (e.g., BATF, AP-1 and BACH), and two with undescribed functions (HIVEP and
214 NFKB) [14, 17, 27, 28]. Tex^{eeff} showed NFAT motif enrichment, while KLF motifs were
215 specifically enriched in the Tex^{int}, T^{eff}, and Tex^{KLR} populations. Finally, Tex^{term}-specific
216 OCRs exhibited strong enrichment for NR4A, RUNX and NFAT TF motifs (**Figure S1G,**
217 **Table S3**) [29-32]. Together, these datasets describe the landscape of transcriptional
218 and epigenetic CD8$^+$ T cell states, including previously unidentified Tex populations,
219 that emerge in response to chronic LCMV infection.
220
221 **CX3CR1$^+$ exhausted T cells comprise three distinct Tex subsets.**
222   We first examined heterogeneity within CX3CR1$^+$ Tex cells, since these cells
223 have recently been described as a highly proliferative and multi-functional intermediate
224 cell state between Tex^{prog} and Tex^{term} [8, 9, 13]. scRNA-seq of sorted CX3CR1$^+$ T cells
225 from D21 of Cl13 infection revealed substantial heterogeneity that primarily spanned
226 three distinct phenotypes (Tex^{eeff}, abundant at D8, and Tex^{int} and Tex^{KLR}, abundant at
227 D21; **Figure 2A, Figure S2A**). To better understand the temporal gene expression
228 programs of Tex^{eeff} and Tex^{int}, we performed DEG analysis (log$_2$ FC > 0.25, Bonferroni
229 adjusted p-value < 0.01) and identified 382 genes with significantly higher expression in
230 Tex^{eeff} (e.g., *Rplp0*, *Rpsa*, *Gapdh* and *Cenpa*) and 286 genes with significantly higher

231 expression in Tex^{int} (e.g., *Ccl5*, *Cd3e*, *Lcp2* and *Nfatc1*; **Figure S2A, Table S4**).
232 Ingenuity pathway analysis linked protein translation (EIF2 Signaling), cell cycle
233 (Kinetochore Metaphase Signaling Pathway), and glycolysis (Glycolysis I.) to the D8
234 population, indicative of a highly proliferative, activated T cell subset. In contrast, the
235 transcriptional program of Tex^{int} at D21 was related to TCR stimulation and downstream
236 signaling (**Figure S2A**). These results suggest that the Tex^{eeff} subset possesses higher
237 cycling and glycolytic activities, while the Tex^{int} subset is more differentiated and
238 expresses genes related to TCR signaling, which seeds downstream Tex populations.
239   To similarly determine the transcriptional programs that distinguish Tex^{int} cells
240 from Tex^{KLR} cells, we performed DEG analysis and found 97 Tex^{KLR}-biased genes and
241 340 Tex^{int}-biased genes (**Figure 2B, Table S5**). Pathway analysis of Tex^{KLR} and Tex^{int}
242 genes demonstrated the enrichment of cell cycle- and T cell exhaustion-related
243 biological terms in the Tex^{int} population, while linking T cell activation signaling and T
244 cell motility-related functions to the Tex^{KLR} subset **(Figure 2B)**. Notably, many markers
245 of terminal effector and effector memory T cells, such as the killer cell lectin-like
246 receptor (KLR) family members (e.g., *Klrd1*, *Klrk1*, *Klrc1*, *Klre1* and *Klrg1*), the TF,
247 *Zeb2*, and its target gene, *S1pr5* (a marker of tissue emigrating antigen-experienced T
248 cells), showed a highly specific expression pattern in the Tex^{KLR} subset [33, 34]. In
249 contrast, Tex^{int} cells expressed canonical exhaustion markers, such as *Tox*, *Tox2*,
250 *Ctla4*, *Pdcd1*, and *Lag3*, along with cell cycle genes (e.g., *Cdk6*), and TCR signaling
251 genes (e.g., *Coro1a*, *Tapbpl* and *Sh2d2a*; **Figure 2B**) [35-37]. We focused on the
252 Tex^{KLR} subset and assessed the expression of the gene signature of terminal effector
253 memory T cells (T-T^{em}), a recently described subset of T^{em} identified during acute LCMV
254 infection, which express effector T cell markers, including KLRs (e.g., *Cx3cr1*, *Zeb2*,
255 *S1pr5*, and *Klre1*) [38]. The T-T^{em} gene signature was highly expressed in the Tex^{KLR}
256 cells and was also observed at the single cell level by scoring the cells based on the
257 expression of this gene panel; thus, cells we term 'KLR-expressing Tex' (Tex^{KLR}) may
258 represent a parallel differentiation path to T-T^{em} with strong effector function and the
259 potential for memory formation (**Figure 2C-E)** [13]. In summary, the $CX3CR1^{+}$ T cell
260 pool contains additional T cell subsets with distinct functionalities and dynamics during
261 the course of chronic infection, which may explain the multitude of effector- and
262 exhaustion-related functions that have previously been linked to this population [8, 9,
263 13].
264
265 **Tex acquire organ-specific terminal exhaustion signatures.**
266   Next, we asked whether chronic viral infection leads to similar T cell states in
267 different tissues. We re-clustered scRNA-seq profiles from animal-matched $gp33^{+}$ and
268 $gp33^{-}$ $CD8^{+}$ T cell fractions from the spleen, lung, and liver at D21 of Cl13 infection. We
269 annotated $CD8^{+}$ T cell subsets based on the previously-defined markers and examined
270 their distribution across organs (**Figure 2F**). Relative to splenic T cells, cells in the lung
271 exhibited an alternative terminal exhausted phenotype (Tex^{lung}) and a reduced Tex^{prog}
272 population, while the proportions of Tex^{KLR} and Tex^{int} populations were similar to the
273 spleen (**Figure 2F and G**). Strikingly, T cells in the liver almost exclusively adopted the
274 Tex^{term} phenotype, with dramatically reduced numbers of other Tex phenotypes, as
275 previously described (6.1% of the total**; Figure 2F and G**) [60]. We further examined
276 tissue-specific differences in the exhaustion signature by pairwise differential gene

277 expression analyses (log$_2$ FC > 0.25, Bonferroni adjusted p-value < 0.01). Compared to
278 splenic Tex^{term} cells, liver-derived Tex^{term} cells possessed a strong tissue-resident
279 memory T cell signature, including the expression of *Cd69*, *Cxcr6*, *Ccl3* and *Ccl4*, and
280 heightened mTOR and glycolytic activity (**Figure 2H, Table S6**). Similarly, lung-derived
281 Tex^{term} cells also exhibited typical markers of lung-resident memory T cells, including
282 *Cxcr6*, *Cd44* and several integrin genes (*Itga4*, *Itgad*, *Itgab7* and *Itgab1*; **Figure 2H**,
283 **Table S7**). Furthermore, both liver- and lung-derived Tex^{term} cells expressed higher
284 levels of pro-survival genes than splenic Tex^{term} cells, including, *Bcl2*, *Bcl2a1b,* and
285 *Bcl2a1d* (**Figure 2H, Table S8**). These results suggest that Tex^{term} cells can obtain
286 tissue residency signatures and persist in tissues in the setting of chronic antigen.
287　Despite tissue specific differences in gene expression of Tex^{term}, we observed a
288 common Tex^{term} gene signature across all organs. This signature (n=35 genes)
289 contained previously described exhaustion-related genes, such as immune checkpoint
290 inhibitory receptors, *Pdcd1*, *Lag3*, and *Tigit*, and the key TF, *Tox*, which imprints the
291 transcriptional and epigenetic signature of T cell exhaustion (**Figure 2I**). Finally, we
292 constructed an exhaustion gene signature based on previously defined CXCR5$^+$ and
293 CXCR5$^-$ T cells subsets and scored the severity of exhaustion among Tex^{term} cells from
294 each organ [6]. We observed that liver-derived Tex^{term} cells scored the highest for the
295 exhaustion signature, followed by splenic and lung-derived Tex^{term} cells (**Figure 2J**). We
296 also scored Tex^{term} and Tex^{int} cells based on cell cycle activity, which ranked liver-
297 derived cells as the least proliferative, followed by the lung and spleen, inversely
298 correlating with the severity of exhaustion (**Figure S2B**). These results demonstrate that
299 T cell exhaustion develops across multiple organs with a common gene expression
300 signature but microenvironment-specific effects; namely, exhaustion is most
301 pronounced in the liver niche, which is perhaps driven by higher antigen burden or
302 anatomical differences [39].
303
Regulatory programs underlying Tex subsets and early fate commitment to the
Tex lineage.
306　The chromatin state of Tex subsets is dynamically regulated and represents a
307 major point of epigenetic imprinting [20] [40]. Two open questions are: (1) the earliest
308 cell cell stage of Tex epigenetic priming, and (2) the temporal regulation of the Tex
309 epigenetic program. To address these questions, we focused our analysis on scATAC-
310 seq data from Cl13 infection and analyzed gp33$^+$ and gp33$^-$ T cells from two time points
311 (D8 and D21) that encompassed previously-defined and our newly-defined Tex subsets
312 (**Figure 3A and B**). Next, we defined differential OCRs for each Tex subset, including
313 Tex^{eeff} (3,567 OCRs), Tex^{prog} (4,818 OCRs), Tex^{int} (235 OCRs), Tex^{KLR} (1,223 OCRs),
314 and Tex^{term} (1,594 OCRs; **Figure 3C**). Cl13-focused scATAC-seq analysis identified a
315 second, more effector-like Tex^{eeff2} population at D8 that exhibited higher accessibility at
316 *Klrc1* and *Gzmm* genes (2,296 OCRs). In addition to the previously described Tex
317 subset-specific motif enrichments (**Figure S1G**), the Cl13-focused analysis allowed us
318 to observe potential relatedness of the subsets based on their chromatin features and
319 enriched TF motifs **(Figure 3C)**. Namely, we observed that: (1) the open chromatin
320 landscape of Tex^{prog} and Tex^{eeff} partially overlap, indicative of developmental
321 relatedness, and (2) the Tex^{int} subset exhibits an intermediate chromatin state between
322 the Tex^{term} and Tex^{KLR} subsets, with very few unique OCRs, suggesting that this cell

7

323 state is an intermediate cell stage and a potential bifurcation point of Tex differentiation,
324 supported by our observation of Tex^{int} at D8 and the emergence of Tex^{term} and Tex^{KLR}
325 by D21 (**Figure 1B**).

326   Next, we focused on the early programming of exhaustion by comparing D8
327 scATAC-seq phenotypes in Arm and Cl13 infection **(Figure S3A)**. As previously
328 described, memory precursor cells (T^{mp}) are present at D8 in Arm infection and cluster
329 with an early Tex^{prog} population present at D8 in Cl13 infection that expresses *Tox* and
330 *Tcf7* (herein referred to as precursor exhausted - Tex^{prec}), but these subsets were
331 relatively infrequent compared to the effector populations in both infection models (1.4%
332 of D8 cells in Arm infection, 3.3% of cells at D8 in Cl13 infection) [14, 41]. We first
333 compared the gene expression and chromatin state programs of T^{mp} and Tex^{prec}
334 subsets, which revealed strong exhaustion- and interferon-induced programs in Tex^{prec},
335 as expected (**Figure S3B, S3C and S3D, Table S9**). Second, we analyzed DEGs of
336 effector cells in Arm and Cl13 infection, which revealed a strong Tex signature in the
337 Tex^{eeff} subset compared to T^{eff}; T^{eff} showed a *bona fide* effector program (e.g., *Gzma*,
338 *Klrd1*, *Ccr2*, and *Cx3cr1*, 371 DEGs), while Tex^{eeff} expressed high levels of exhaustion
339 marker genes (e.g., *Tox*, *Lag3*, *Pdcd1*, *Havcr2*, *Ctla4*, and *Tigit,* 618 DEGs; **Figure 3D**).
340 These observations were also supported by the chromatin state programs of these
341 subsets (T^{eff} - 7,066 OCRs vs. Tex^{eeff} - 5,211 OCRs) that were associated with T^{eff}-
342 specific (ETS and RUNX) and Tex^{eeff}-specific (NFAT and BATF) TF motifs (**Figure 3D,**
343 **Table S10**). Altogether, these results support recent studies demonstrating the
344 formation of Tex^{prec} early during chronic infection that exhibit molecular signatures of
345 exhaustion, distinct from T^{mp} [14, 41]. However, we find that the exhaustion program,
346 including *Tox* expression, is present in an earlier Tex^{eeff} stage and driven by NFAT and
347 BATF, which may prime the chromatin state of TCF1$^-$ cells for exhaustion, supporting a
348 model in which an initial wave of effector cells undergoes contraction and gives rise to
349 Tex^{prog} cells that seed additional Tex subsets.

350   Although Tex differentiation downstream of Tex^{prog} is currently thought to follow a
351 linear path, our identification of a Tex^{KLR} subset, which emerges late in infection
352 alongside Tex^{term}, suggests that the Tex^{int} population may represent a potential
353 bifurcation point between Tex^{KLR} and Tex^{term} phenotypes (**Figure 1D, 2A and 3A**) [3].
354 We compared scATAC-seq profiles of Tex^{KLR} and Tex^{term} subsets to Tex^{int} cells and
355 identified 405 Tex^{term}-specific, 364 Tex^{KLR}-specific OCRs, and only 4 common OCRs,
356 suggesting that these two cell states are epigenetically divergent (**Figure 3E**).
357 Accordingly, TF motif enrichment analysis showed increased accessibility of NFAT, IRF,
358 STAT, and NR4A TF motifs in Tex^{term} and increased accessibility of RUNX, MGA, KLF,
359 TBET/EOMES and ETS TF motifs in Tex^{KLR} (**Figure 3F**). Differential gene expression
360 analysis (log$_2$ FC > 0.25, Bonferroni adjusted p-value < 0.01) identified 97 Tex^{KLR}-
361 biased genes (e.g., *Klrg1*, *Arl4c* and *Zeb2*) and 340 Tex^{term}-biased genes (e.g., *Tox*,
362 *Tox2*, *Lag3* and *Pdcd1*; **Figure S3E, Table S11**). These results indicate that Tex^{KLR} and
363 Tex^{term} cells exhibit distinct chromatin and gene expression programs, supporting the
364 idea that these phenotypes represent late stages of a divergent differentiation trajectory
365 of exhaustion that bifurcates at the Tex^{int} stage.

366   Finally, we analyzed 15,809 variable OCRs for TF motif enrichments across
367 three differentiation trajectories nominated by longitudinal timepoint data and/or
368 chromatin state similarities: (1) Tex^{prog} trajectory (Tex^{eeff} → Tex^{prec} → Tex^{prog}), (2)

369   Tex^{term} trajectory (Tex^{prog} → Tex^{int} → Tex^{term}), and (3) Tex^{KLR} trajectory (Tex^{prog} → Tex^{int}
370   → Tex^{KLR}; **Figure 3G**). The Tex^{prog} trajectory showed a gradual loss of HOMEOBOX TF
371   motifs and enrichment of BATF, AP-1, BACH, NFKB, TCF and CTCF motifs. In contrast,
372   in both Tex^{term} and Tex^{KLR} trajectories, we observed a gradual loss of Tex^{prog} specific TF
373   motifs (e.g., TCF, BACH and BATF) upon entry to the Tex^{int} cell state. Differentiation
374   trajectories that bifurcated from the Tex^{int} state showed the enrichment of specific TF
375   motifs that might bind TFs which can guide the differentiation program of Tex^{KLR} (e.g.,
376   ZEB, ID, IRF, KLF, ETS, RFX, HIVEP and RUNX) and Tex^{term} (e.g., RUNX and NR4A;
377   **Figure 3G**). Finally, we studied the accessibility of the *Tox* locus, encoding the TF that
378   is critical for Tex differentiation [4, 15, 16, 18, 19]. Accessibility of the *Tox* locus
379   gradually increased as cells transitioned from Tex^{eeff} to Tex^{prog}, while it gradually
380   decreased as they transitioned to Tex^{KLR}. The Tex^{term} trajectory demonstrated a
381   decrease in *Tox* accessibility during the Tex^{prog} to Tex^{int} transition and a subsequent
382   increase in the Tex^{term} state (**Figure 3G**). We annotated differentially accessible OCRs
383   (compared to T^{naive} cells) in a +/- 250kb window around the transcription start site of *Tox*
384   and identified 88 OCRs. Of these OCRs, 16 and 8 were differentially accessible in
385   Tex^{eeff} or Tex^{prog}, respectively, which was also supported by high *Tox* expression in
386   these subsets (relative to T^{naive}), indicating that TOX executes the molecular
387   programming of Tex differentiation in these subsets (**Figure S3F, Table S12**). These
388   results identify the Tex^{eeff} population as a novel point of the molecular programming of
389   exhaustion and nominate the Tex^{int} population as a potential bifurcation point of Tex cell
390   differentiation states.
391
392   **Clone tracing reveals divergent Tex differentiation trajectories during chronic**
393   **viral infection.**
394   We next leveraged paired scRNA/TCR-seq data to analyze clonal trajectories of
395   T cells in Arm and Cl13 (D8 and D21) infection (**Figure 4A**). We identified 212 and 280
396   expanded T cell clones (> 1 cell) at D8 and D21 of Arm infection, respectively, and 134
397   and 338 expanded clones at D8 and D21 of Cl13 infection, respectively. As expected, at
398   D8 of Arm infection, clonally expanded T cells were largely restricted to the T^{eff} pool,
399   while at D21, clonally expanded T cells exhibited a balanced distribution between T^{em}
400   and T^{mem} phenotypes (**Figure 4B and Figure S4A**). In contrast, clonal expansion at D8
401   in Cl13 infection occurred almost exclusively in Tex^{eeff} **(Figure 4C)**. Importantly, the
402   Tex^{prog} population did not show strong clonal expansion at this early time point, further
403   supporting our prior observation that cells in this population are infrequent at D8 and
404   subsequently expand by transition from TEX-eeff or self-renewal from cells not
405   significantly present at D8 **(Figure 3)**. At D21 in Cl13 infection, we observed expanded
406   clones across multiple Tex phenotypes, including Tex^{prog}, Tex^{int}, Tex^{term}, and Tex^{KLR}
407   (**Figure 4C and Figure S4B**).
408   To further investigate Tex clonal differentiation trajectories, we visualized the
409   distribution of cellular phenotypes for the top 10 expanded clones at D8 and D21 in
410   each infection. At D8, cells of the top expanded clones from the Arm condition almost
411   exclusively acquired the T^{eff} phenotype, with clone sizes ranging from 77-321 cells
412   (mean 153 cells, 3.8% of 4,030 total cells). At D21, top expanded clones acquired both
413   T^{em} and T^{mem} phenotypes, with clone sizes ranging from 145-366 (mean 246 cells, 3.5%
414   of 7,033 total cells; **Figure 4D and Figure S4A**). In contrast, the top expanded clones in

9

415 Cl13 infection acquired the Tex^{eeff} phenotype at D8, with clone sizes ranging from 26-95
416 cells (mean 49 cells, 3.5% of 1,414 total cells). Notably, expanded clones only
417 contained small numbers of cells with the Tex^{prog} phenotype at this time point (**Figure**
418 **4D**). Analysis of D21 of Cl13 infection identified substantially larger clone sizes, ranging
419 from 146-2,026 cells (mean 525 cells, 7.0% of 7,489 total cells; **Figure 4D**). Strikingly,
420 these large clones contained cells with multiple Tex phenotypes (Tex^{prog}, Tex^{int}, Tex^{KLR},
421 and Tex^{term}), although the frequency of each phenotype varied considerably between
422 individual clones. Namely, individual clones either preferentially acquired the Tex^{term} or
423 the Tex^{KLR} phenotypes, or developed into both phenotypes (**Figure 4D and Figure**
424 **S4B**). This observation prompted us to perform a more detailed analysis of the
425 phenotypic distribution of all large clones (> 3 cells detected) of the top 7 most dominant
426 clonal phenotype combinations, which revealed three main clonal differentiation
427 patterns (referred to as clone behaviors): 1) Tex^{term}-biased clones, consisting of cells
428 that predominantly acquired the Tex^{term} and not Tex^{KLR} phenotype (45% of clones), 2)
429 Tex^{KLR}-biased clones, consisting of T cells that predominantly acquired the Tex^{KLR}
430 phenotype (18% of clones), and 3) divergent clones, consisting of cells that acquired
431 Tex^{term} and Tex^{KLR} phenotypes (37% of clones; **Figure 4E**). Divergent clones were the
432 most clonally expanded and ranged from 7-2,026 cells (mean 197 cells) per clone, while
433 Tex^{term}-biased clones ranged from 4-111 cells (mean 19 cells) per clone. Interestingly,
434 Tex^{KLR}-biased clones were relatively small and ranged from 4-21 cells (mean 8 cells)
435 per clone (**Figure 4E-G**). We also noted several larger clones (4-233 cells, mean 53
436 cells) that skewed heavily to the Tex^{KLR} phenotype (>50% of cells acquire the Tex^{KLR}
437 phenotype), but had a small percentage of Tex^{term} cells (**Figure S4C**). To account for
438 sampling biases where not all relevant phenotypes may be observed for small clones,
439 we randomized T cell phenotype and TCR clone assignment to generate a null
440 distribution of clone patterns if each clone randomly acquired all observed phenotypes
441 (**Methods**). This analysis revealed a striking enrichment of Tex^{KLR}- and Tex^{term}-biased
442 clone behavior over random chance, whereas the divergent clonal differentiation pattern
443 was twice as likely to be detected by random chance than observed in our data,
444 suggesting that the observed biases in clone behavior are not simply the result of
445 sampling bias (**Figure S4D**). Altogether, these results reveal novel clonal Tex
446 differentiation trajectories during chronic infection **(Figure 4H)**.

447

448 **Antigen-specific expanded Tex clones and phenotypes are shared across tissues.**
449 Next, we asked if clonal differentiation patterns are intrinsically programmed,
450 perhaps by the TCR, or stochastic. We first determined whether expanded Tex clones
451 could be found across different tissues by analyzing antigen-specific $gp33^+$ and $gp33^-$
452 $CD8^+$ T cells across organs (animal-matched) in Cl13 at D21 (**Figure 5A**). In spleen-,
453 liver-, and lung-derived scRNA/TCR-seq datasets, we detected expanded T cell clones
454 across all three tissues, and as expected, the $gp33^+$ and $gp33^-$ fractions showed
455 minimal TCR overlap, validating our sorting strategy (**Figure 5B and Figure S5A**).
456 Importantly, there was significant TCR sharing across the different organs for both
457 $gp33^+$ and $gp33^-$ fractions (**Figure S5A**). We identified expanded organ-shared T cell
458 clones that had at least 5 T cells, which consisted of at least 1 cell from each organ.
459 This analysis identified 100 shared T cell clones among all organs, 37 clones shared
460 between the lung and spleen, and 22 clones specific to the spleen (**Figure 5C and D**).

461 We examined the degree of expansion of TCR clones that were detected across organs
462 and observed a strong correlation in clone frequency in each pairwise organ
463 comparison (spleen:liver - R=0.66; spleen:lung - R=0.65; liver:lung - R=0.76, **Figure**
464 **5C**).
465 Next, we examined the distribution of phenotypes for clones shared across
466 organs with different differentiation trajectories defined by their trajectory in the spleen
467 (**Figure 5E, F and Figure S5B, C**). First, we focused on comparisons between the
468 spleen and lung, since they exhibited similar heterogeneity in Tex phenotypes.
469 Strikingly, differentiation trajectories were highly conserved between the two tissues.
470 Divergent clones in the spleen also maintained Tex^{term} and Tex^{KLR} phenotypes in the
471 lung (although instead exhibiting the aforementioned terminal Tex^{lung} phenotype; 35/48
472 divergent clones detected in both organs, **Figure 5E, F, Figure S5B, C**). Similarly, the
473 majority of splenic Tex^{KLR}-biased clones remained Tex^{KLR}-biased in the lung (4/7 clones
474 detected in both organs) and the majority of splenic Tex^{term}-biased clones remained
475 Tex^{term}-biased in the lung (11/14 clones detected in both organs; **Figure 5E and Figure**
476 **S5C**). In particular, we did not observe appreciable interconversion between Tex^{KLR}-
477 and Tex^{term}-biased clones between these two organs (0/19 shared clones). Accordingly,
478 quantification of the Tex^{KLR} and Tex^{term} frequencies within individual clones showed a
479 high concordance across organs (Tex^{KLR} spleen:lung - R=0.96, Tex^{term} spleen:lung -
480 R=0.75; **Figure 5G**). Altogether, these results demonstrate that clonally expanded Tex
481 clones are shared across organs and that clonal differentiation behavior is primarily an
482 intrinsically programmed, rather than stochastic, process.
483
484 **Depletion of Tex^{KLR} clones in the liver microenvironment.**
485 We next analyzed clonal behavior in the liver, which showed an enrichment of
486 Tex^{term} compared to other organs (94% Tex^{term}), perhaps driven by high antigen burden
487 [39]. Thus, in contrast to the lung, we expected an enrichment in clonal Tex^{term}
488 frequency; however, this could either be driven by: (1) depletion of Tex^{KLR} in the liver
489 microenvironment, or (2) interconversion of Tex^{KLR}-biased clones to Tex^{term}-biased
490 clones. To distinguish between these two possibilities, we first analyzed the Tex^{KLR}-
491 biased clones from the spleen and found that only one of these clones was present in
492 the liver (1/7 shared clones), suggesting that Tex^{KLR}-biased clones are depleted in the
493 liver niche. Similarly, although divergent clones were largely detectable in the liver
494 (52/58 clones shared between spleen and liver), we again observed a depletion of
495 Tex^{KLR} cells, resulting in Tex^{term}-biased behavior in the majority of the cases (32/52
496 shared clones). In contrast, the majority of Tex^{term}-biased clones remained Tex^{term}-
497 biased in the liver, although they were heavily skewed towards the Tex^{term} phenotype,
498 with relative loss of Tex^{prog} and Tex^{int} phenotypes (9/9 clones, **Figure 5E, F and Figure**
499 **S5C**). Importantly, Tex^{term}-biased clones did not adopt a Tex^{KLR} phenotype.
500 Quantification of frequencies of Tex^{KLR} and Tex^{term} phenotypes of shared clones in the
501 spleen and liver confirmed the depletion of Tex^{KLR} cells in the liver and a skewing of
502 Tex^{term}-biased clones to the Tex^{term} fate **(Figure 5G and H)**. Altogether, these results
503 demonstrate that Tex clones entering the liver exhibit changes in clonal behavior due to
504 the loss of Tex^{KLR}, suggesting that Tex^{KLR} are not able to persist in high antigen
505 environments, perhaps due to activation-induced cell death.
506

TCR affinity can program Tex clone behavior and phenotypic fate commitment.

The difference in expansion levels between Tex^{term}-biased clones and Tex^{KLR}-biased clones led us to examine whether Tex differentiation trajectories were driven by differences in TCR affinity. We used tetramer staining as a proxy for TCR affinity against the immunodominant LCMV epitope, gp33, and sorted gp33$^-$ (n=8,914), gp33-intermediate (gp33^{int}; n=5,875), and gp33-high (gp33^{high}; n=8,194) CD8$^+$ T cells from the spleen of Cl13-infected mice at D21 and performed scRNA/TCR-seq (**Figure 6A and Figure S6A**). Analysis of TCR sequences identified 313 TCRs in gp33^{high} cells, 1,576 TCRs in gp33^{int} cells, and 3,803 TCRs in gp33$^-$ cells (**Figure S6B**). The TCR repertoire showed a relatively small overlap between gp33^{high} and gp33$^-$ cells (13 shared TCRs), compared to the overlap between gp33^{high} and gp33^{int} cells (158 shared TCRs), or gp33^{int} and gp33$^-$ cells (306 shared TCRs), and quantification of TCR repertoire similarity using the Morisita overlap index demonstrated that gp33^{int} sorting captured a distinct TCR repertoire compared to gp33^{high} and gp33$^-$ fractions (**Figure 6B and Figure S6B**).

Next, we evaluated the clone size distribution of the sorted populations, which revealed an increase in the percentage of large clones (clones with 5-200 or >200 cells) as a function of higher tetramer fluorescence, with an accompanying decrease in clonal diversity (**Figure 6C**). To link unique TCRs to each gp33-tetramer fraction, we compared the overlap of clones between gp33 fractions and identified 592 unique gp33$^-$ clones, 114 unique gp33^{int} clones, and 88 unique gp33^{high} clones (**Figure 6D**). Importantly, the distribution of cellular phenotypes for these unique clones showed considerable phenotypic skewing (**Figure 6E-G**). Namely, gp33^{high} cells contained ~3.3 times more cells with Tex^{term} and Tex^{int} phenotypes, compared to either gp33$^-$ or gp33^{int} cells (39% Tex^{term} and 19% Tex^{int} in gp33^{high}; 11% Tex^{term} and 6.7% Tex^{int} in gp33^{int}; 11% Tex^{term} and 7.0% Tex^{int} in gp33$^-$), indicating a pronounced phenotypic skewing towards terminal exhaustion. In contrast, gp33^{int} cells exhibited phenotypic skewing towards the Tex^{KLR} phenotype in the population, compared to the gp33^{high} and gp33$^-$ populations (27% Tex^{KLR} in gp33^{int}; 7.9% Tex^{KLR} in gp33^{high}; 13% Tex^{KLR} in gp33$^-$; **Figure 6F and G**).

To further analyze differentiation trajectories at a clonal level, we visualized the top 10 unique expanded clones in each gp33-tetramer fraction and assessed their phenotypic composition. We found that the top clones in the gp33$^-$ and gp33^{high} fractions were biased towards Tex^{term} or divergent phenotypes (10/10 gp33$^-$ clones and 10/10 gp33^{high} clones), while in contrast, the largest clones in the gp33^{int} pool exhibited phenotypic skewing towards the Tex^{KLR} phenotype (5/10 gp33^{int} clones; **Figure 6H**). Finally, we analyzed the clone behaviors of the unique clones of the three gp33 fractions in the top 7 most dominant phenotypic patterns that define clone behaviors (**Figure 6I**). Clones of the gp33$^-$ fraction exhibited two major clone behaviors, Tex^{term}-biased and Tex^{KLR}-biased. Interestingly, clones from the gp33^{int} fraction were heavily enriched for Tex^{KLR}-biased differentiation. Finally, expanded clones unique to the gp33^{high} fraction were biased towards Tex^{term}-biased and divergent clone behaviors, and no Tex^{KLR}-biased clones were identified **(Figure 6I)**. Surprisingly, divergent clone behaviors were much more common in the unique T cell clones of the gp33^{int} and gp33^{high} fractions compared to gp33$^-$ clones, suggesting that this differentiation path may be more common among T cell clones that recognize this immunodominant

553 epitope. These results establish that T cell clones distinguished by their affinity for the
554 immunodominant LCMV epitope have divergent differentiation paths, with lower affinity
555 TCR clones favoring the development of Tex^{KLR} and higher affinity TCR clones biasing
556 toward Tex^{term} and divergent behavior.
557
558 **Discussion**
559 Here we report a single-cell multi-omic atlas of T cell exhaustion during chronic
560 viral infection, which reveals novel Tex subsets, identifies multiple differentiation
561 trajectories of Tex clones, and nominates TCR signal strength as a key driver of clonal
562 behavior. We define an early effector Tex differentiation state (Tex^{eeff}), where the
563 molecular program of exhaustion is initiated, and identify a bifurcation point of Tex
564 differentiation (Tex^{int}), which can give rise to two alternative late-stage Tex phenotypes
565 (Tex^{KLR} and Tex^{term}) with the potential to balance effector function, immunological
566 memory, and persistence in high antigen environments. Using the TCR sequence as an
567 endogenous molecular barcode, we track the fate of individual T cell clones and
568 establish three main clonal developmental trajectories that give rise to the
569 heterogeneous Tex pool. Surprisingly, we find that clonal differentiation patterns are
570 shaped by TCR affinity and affect the resulting phenotype and clonal expansion in
571 different tissue microenvironments. These findings highlight the importance of studying
572 the polyclonal T cell repertoire at single cell resolution to fully uncover the diversity and
573 function of T cell states in the immune response.
574 Prior studies have described multiple Tex subsets with distinct phenotypic and
575 functional traits, primarily within the spleen microenvironment during chronic viral
576 infection [3,10, 11, 42, 43]. In addition to Tex^{term} and Tex^{prog} subsets, transitory
577 exhausted cells have more recently been characterized as a multi-functional $CX3CR1^{+}$
578 population with high cytolytic activity, proliferative capacity, and the ability to contribute
579 to the memory T cell pool [8, 9, 13]. Here we show that this $CX3CR1^{+}$ population
580 encompasses three T cell subsets with distinct functionalities: 1) an early effector
581 exhausted subset (Tex^{eeff}) with high proliferative capacity early in infection that is largely
582 absent at later stages; 2) intermediate exhausted T cells (Tex^{int}), which maintain a high
583 proliferation signature and upregulate signaling downstream of TCR stimulation, and 3)
584 a Tex^{KLR} subset with a strong cytolytic gene expression program, and a terminal effector
585 memory cell-like signature that has been described in acute infection [38].
586 Given the distinct, stable epigenetic state of Tex, which persists after antigen
587 clearance [20, 40, 44-46], a key question is the stage at which Tex epigenetic imprinting
588 occurs. Previous studies have shown that early $TCF1^{+}$ Tex^{prec} cells possess the
589 epigenetic signature of Tex and can seed additional Tex subsets [14, 41]. Here, we find
590 that the Tex program is initiated at an earlier stage in $TCF1^{-}$ Tex^{eeff}. scATAC-seq
591 analysis suggests that this fate decision is initially driven by NFAT and BATF, which
592 may prime the chromatin state of $TCF1^{-}$ cells to develop into Tex^{prec}, which
593 subsequently activate BACH2 and TCF-1 to give rise to Tex^{prog} [14, 17]. This finding
594 supports a model in which the Tex^{prec} pool, and eventually the Tex^{prog} pool originates
595 from Tex^{eeff}, analogous to memory differentiation from memory precursors or short-lived
596 effector cells during acute infection.
597 Downstream of the Tex^{prog} population, the differentiation trajectory of Tex has
598 largely been shown to follow a linear cellular path [3]. However, our data suggests that

13

599 there are two late-stage cell types that result from a divergent differentiation path
600 (Tex^{KLR} and Tex^{term}), and that individual clones can follow three differentiation
601 trajectories resulting in Tex^{term}-biased, Tex^{KLR}-biased, or divergent fates, comprising
602 both cell types. Furthermore, we find that the differentiation trajectory of Tex clones is
603 intrinsically programmed by TCR affinity and conserved across specific tissue
604 microenvironments; high-affinity TCR clones are biased towards divergent and Tex^{term}
605 differentiation trajectories, while low-affinity TCR clones are biased towards a Tex^{KLR}
606 trajectory. However, the presence of clones with divergent behavior suggests that there
607 may be additional paths to induce TCR signal strength variation – perhaps via inhibitory
608 receptor signaling, access to antigen, antigen-presenting cell type, or other factors – to
609 generate Tex^{KLR}. Importantly, Tex^{KLR}-biased clones were dramatically depleted in the
610 liver microenvironment, suggesting that these clones and this phenotype are sensitive
611 to the antigen-rich environment of the liver and are unable to persist. Given the high
612 viral load and inflammatory microenvironment of the liver during infections, these results
613 suggest that the Tex^{term} phenotype precludes activation-induced cell death, improves
614 Tex persistence, and preserves anti-viral effector function in the organ system [47].
615   Finally, these findings may have several implications for cancer, where T cell
616 exhaustion can limit the T cell response and efficacy of immunotherapies. First, several
617 ongoing therapeutic strategies aim to reverse exhaustion; however, our results suggest
618 that Tex^{term} may be specifically adapted to survive in high antigen niches, and that
619 inhibiting Tex^{term} differentiation may be deleterious, rather than beneficial, to the T cell
620 response [4, 15, 16, 18, 19, 48]. Whether the pro-survival aspects of T cell exhaustion
621 can be specifically maintained, while still reinvigorating other aspects of effector function
622 will require further study. Second, our findings reinforce the notion that TCR signal
623 strength directs the phenotypic fate of T cells, in addition to mediating recognition of
624 specific antigens [49, 50]. Thus, the generation of TCR-based cellular therapies should
625 incorporate the assessment of phenotypic outcomes of TCR activation, in addition to
626 peptide-MHC binding properties. Finally, the observation that a polyclonal T cell
627 response to chronic antigens balances persistence, effector, and potential memory
628 functions via the development of two Tex states suggests that future cellular therapies
629 should also aim to establish divergent phenotypes, encompassing Tex^{term} and Tex^{KLR}.
630 Future studies should investigate whether Tex^{KLR} develop during tumor-specific T cell
631 responses. A recent study identified a natural killer (NK) cell-like signature in chronic
632 antigen-induced exhausted human chimeric antigen receptor (CAR)-T cells, which
633 resembles the Tex^{KLR} signature described here, suggesting that this cell type may be
634 present in adoptive cell therapy settings as well [51]. Manipulation of these Tex states
635 and their underlying gene regulatory programs and differentiation pathways may provide
636 avenues to improve T cell-based immunotherapies in the future.
637
638 **Acknowledgements:**

655

656 **Author contributions:**
657 B.D., K.E.Y. and A.T.S. conceptualized the study. B.D., K.E.Y. and A.T.S. wrote and
658 edited the manuscript and all authors reviewed and provided comments on the
659 manuscript. B.D., K.S., K.E.Y., K.J.H.G., X.Y. and Y.Q. performed experiments. K.E.Y.,
660 S.L.M. and J.A.B. analyzed data. B.D., K.E.Y., J.R.G., E.J.W., H.Y.C., T.E. and A.T.S.
661 guided experiments and data analysis.
662

663 **Declaration of interests:**
664 A.T.S. is a founder of Immunai and Cartography Biosciences and receives research
665 funding from Allogene Therapeutics, Merck Research Laboratories, and 10x Genomics.
666 H.Y.C. is a co-founder of Accent Therapeutics, Boundless Bio and Cartography
667 Biosciences, and an advisor to 10x Genomics, Arsenal Biosciences, and Spring
668 Discovery. K.E.Y. is a consultant for Cartography Biosciences. J.A.B. is a consultant for
669 Immunai.
670

671 **Data availability**
672 Reviewer access for sequencing data is available under GEO accession: GSE188670.
673

674 **Methods**
675

676 **Mice and infection**

677 Male C57BL/6N mice were purchased from Charles River Laboratories. All mice were
678 housed in a specific pathogen-free facility at Washington University in St. Louis and
679 were used for infection at 8–12 week of age. LCMV infection was performed essentially
680 as described previously [52]. All experiments were performed according to a protocol
681 approved by Washington University's Institutional Animal Care and Use Committee.

682 **Tissue preparation**
683 Single cell suspension of the different organs was prepared by manual dissociation.
684 Organs were minced and gently pushed through a 40-micron strainer. Spleen single cell
685 suspensions were spun, and red blood cells were lysed with ACK-lysis buffer by
686 resuspending the cell pellet followed by 2 minutes incubation. Cells were then washed
687 with ice-cold PBS and stained for sorting in FACS buffer (PBS, 0.1% BSA, 2mM EDTA,
688 5% FBS). For the lung and liver single-cell suspension, organs were cut into small
689 pieces and gently pushed through a 40-micron diameter strainer. Single-cell

690 suspensions were then layered on top of Ficoll-Paque Plus (Cytiva) and centrifuged
691 according to the manufacturer's recommendations. The lymphocyte fraction was
692 collected and washed with ice-cold PBS, and then stained for sorting.
693
694 **Staining T cells for sorting**
695 Single cell suspensions were stained with the following antibodies: CD8b (PerCP-
696 Cy5.5), PD-1 (PE-Cy7), CX3CR1 (APC), SLAMF6 (BV605) and the class I tetramer, H-
697 2Db LCMV gp33-41 (KAVYNFATC) (PE). Cells were stained with the tetramer for 20
698 minutes at 4C followed by staining with the combination of the other antibodies for 20
699 minutes. Cells were washed in FACS buffer and stained with LIVE/DEAD Fixable Aqua
700 dead cell stain for 20 minutes in PBS.
701
702 **scATAC-seq sample and library generation**
703 Single cell ATAC-seq experiments were performed on the 10x Chromium platform as
704 described earlier [53]. Briefly, after sorting, T cells were washed with PBS + 0.04% BSA
705 and then subjected to nuclei isolation according to the protocol of the manufacturer.
706 Nuclei were counted and on average ~10,000 nuclei were submitted for tagmentation.
707 After tagmentation, nuclei were loaded for capture using the 10x Chromium controller.
708 After Gel emulsion generation, linear amplification was performed, followed by DNA
709 purification according to the manufacturer's protocol. The resulting DNA was used for
710 library construction as described on the website of the manufacturer. Libraries were
711 quantified by Agilent Bioanalyzer and were sequenced on an Illumina NovaSeq S4
712 sequencer, using the following setup: 50bp read 1N, 8bp i7 index, 16bp i5 index and
713 50bp read 2N. In this reaction, 1N and 2N refers to the DNA insert sequencing, while i5
714 and i7 sequencing identifies the individual barcodes of single cells.
715
716 **Single-cell RNA-seq library preparation**
717 Single-cell RNA-seq libraries were prepared using the 10X 5' Single Cell Immune
718 Profiling Solution Kit (v1.1 Chemistry), according to the manufacturer's instructions.
719 Briefly, FACS sorted cells were washed once with PBS + 0.04% BSA and on average
720 10,000 cells were submitted for capture using the 10x Chromium controller. Following
721 reverse transcription and cell barcoding in droplets, emulsions were broken, and cDNA
722 was purified using Dynabeads MyOne SILANE followed by PCR amplification (98°C for
723 45 sec; 14 cycles of 98°C for 20 sec, 67°C for 30 sec, 72°C for 1 min; 72°C for 1 min).
724 For gene expression library construction, 50 ng of amplified cDNA was fragmented,
725 end-repaired, and double-sided size selected with SPRIselect beads. Purified DNA was
726 subjected to PCR amplification with sample indexing primers (98°C for 45 sec; 14 cycles
727 of 98°C for 20 sec, 54°C for 30 sec, 72°C for 20 sec; 72°C for 1 min). Amplified DNA
728 was double-sided size selected with SPRIselect beads and were quantified using
729 Agilent Bioanalyzer. Single-cell RNA-seq libraries were sequenced on an Illumina
730 NovaSeq S4 sequencer using the following read configuration 26bp Read1, 8bp i7
731 Index, 91bp Read2.
732
733 **Single-cell TCR library generation**
734 Single-cell TCR libraries were prepared with the 10x Chromium Single Cell V(D)J
735 Enrichment Kit for mouse T cells (v1.1 Chemistry) following the manufacturer's protocol.

16

736 Briefly, after cDNA amplification and clean up, 2ul of cDNA was used for target
737 enrichment. First, target enrichment 1 was performed by specific primers followed by a
738 SPRIselect bead clean-up. Second, target enrichment 2 was performed with specific
739 primers followed by double-sided size selection with SPRIselect beads. After the two
740 target enrichment steps, the quality of the product was assessed with Agilent
741 Bioanalyzer. Amplified product was then subjected for fragmentation, followed by end
742 repair and A-tailing. End repaired product was then subjected to adaptor ligation
743 followed by SPRIselect bead purification. Product was amplified and barcoded with
744 adaptor specific primers and the quality of the resulting libraries were determined by
745 Agilent Bioanalyzer. Single-cell TCR-seq libraries were sequenced on an Illumina
746 NovaSeq S4 sequencer using the following read configuration 26bp Read1, 8bp i7
747 Index, 91bp Read2.
748
749 **scATAC-seq data processing and analysis**
750 scATAC-seq datasets were processed as described previously [54]. Briefly, reads were
751 filtered, trimmed, and aligned to the mm10 reference genome using 10X Genomics'
752 cellranger-atac count pipeline (version 1.2.0).
753
754 Processed fragment files were loaded into ArchR (version 1.0.1) for additional
755 processing and analysis. All functions used default parameters unless otherwise
756 specified. Cells were filtered during Arrow file generation using ArchR's
757 createArrowFiles function to remove cells with an enrichment of Tn5 insertions in
758 transcription start sites (TSS enrichment) of less than 4 or less than 1000 unique
759 fragments. Doublets were identified using ArchR's addDoubletScores function and
760 predicted doublets removed using ArchR's filterDoublets function. Dimensionality
761 reduction was performed using Iterative Latent Semantic Indexing (LSI) using ArchR's
762 addIterativeLSI function. After initial clustering and UMAP projection, we excluded a
763 small cluster of non-T cells. Cell clustering was performed using ArchR's addClusters
764 function on IterativeLSI reduced dimensions 1:10 and a resolution of 0.4 (reducedDims
765 = "IterativeLSI", dimsToUse = 1:10, resolution = 0.4). The same dimensions were used
766 for single cell embedding by Uniform Manifold Approximation and Projection (UMAP)
767 using ArchR's addUMAP function using IterativeLSI reduced dimensions 1:10 and a
768 minimum distance of 0.1 (reducedDims = "IterativeLSI", dimsToUse = 1:10, minDist =
769 0.1). Cell clustering and UMAP projection for Chronic LCMV (D8 and D14, Figure 3) and
770 Day 8 (Chronic and Acute, Figure S3B) subsets were performed as described above
771 with the following modifications: dimsToUse = NULL, resolution = 0.2, and minDist =
772 0.4.
773
774 GeneScore matrices were computed by summing Tn5 insertions in the gene promoter
775 and gene body during Arrow file generation using ArchR's createArrowFiles function
776 [54]. Gene score imputation was performed with Magic using ArchR's
777 addImputeWeights function [55]. After clustering the cells, peaks were called by MACS2
778 on pseudoreplicates sampled from each cluster to obtain a reproducible peak set
779 retaining cell type specific peaks using ArchR's addReproduciblePeakSet function.
780 Peak co-accessibility and Peak2Gene linkages were computed using ArchR's
781 addCoAccessibility and addPeak2GeneLinks functions. Transcription factor (TF) motif

17

782 deviations were computed with chromVar using ArchR's addDeviationsMatrix function
783 [26]. Pseudo-bulk tracks for indicated groups of cells were plotted using ArchR's
784 plotBrowserTrack function with default normalization method based on reads in
785 transcription start sites ("ReadsInTSS"). Differential peak testing was performed using
786 ArchR's getMarkerFeatures function with testMethod = "wilcoxon" and bias =
787 c("TSSEnrichment", "log10(nFrags)"). TF motif enrichment in differential peavks was
788 performed using ArchR's peakAnnoEnrichment function. Trajectory analysis was
789 performed using ArchR's addTrajectory and plotTrajectory functions. Identification of
790 positive TF regulators was performed using ArchR's correlateMatrices function to
791 examine the correlation between chromVar deviation z-scores of TF motifs
792 ("MotifMatrix") and imputed gene expression ("GeneIntegrationMatrix") following cross-
793 platform linkage with scRNA-seq data using ArchR's addGeneIntegrationMatrix.
794

scRNA-, TCR-seq computational methods
795
796 scRNA-seq reads were aligned to the mm10 reference genome and quantified using
797 cellranger count (10x Genomics, version 3.1.0). Filtered gene-barcode matrices that
798 contained only barcodes with unique molecular identifier (UMI) counts that passed the
799 threshold for cell detection were used for further analysis. scTCR reads were aligned to
800 the mm10 reference genome and consensus TCR annotation was performed using
801 cellranger vdj (10x Genomics, version 3.1.0). TCR annotation was performed using the
802 10x cellranger vdj pipeline as described.
803

804 Additional analysis was performed in R (version 4.0.3) using Seurat (version 4.0.1)
805 using default function parameters unless otherwise noted [56]. Doublets were predicted
806 using DoubletFinder (version 2.0.3) [57]. Cell types were predicted using SingleR
807 (version 1.4.1) based on mouse bulk RNA-seq reference data (MouseRNAseqData)
808 from celldex (version 1.0.0) [58]. Cells with less than 200 genes detected, greater than
809 5% mitochondrial RNA content, predicted doublets from DoubletFinder, and cells
810 annotated as non-T and non-NK cells by SingleR were excluded from analysis. We
811 predicted cell cycle phase based on previously defined gene sets using the
812 CellCycleScoring function [59]. We then split cells by experimental batch and cell cycle
813 (non-cycling or G1 vs. cycling or G2M/S) into four datasets using Seurat's SplitObject
814 and performed batch correction using Seurat's reciprocal PCA workflow. First, we
815 normalized and identified variable features for each dataset independently using
816 Seurat's NormalizeData and FindVariableFeatures. Then we selected variable features
817 across datasets using Seurat's SelectIntegrationFeatures. We excluded variable TCR
818 (^Tr.v) genes, variable Ig (^Ig.v) genes, cell cycle genes (used for cell cycle scoring),
819 and mitochondrial genes (^mt-) from integration features used for downstream analysis.
820 We then scaled data and ran PCA on each dataset independently using these features
821 using Seurat's ScaleData and RunPCA. We identified integration anchors using
822 Seurat's FindIntegrationAnchors using non-cycling datasets as reference datasets and
823 rpca for dimensionality reduction. We integrated all datasets using Seurat's
824 IntegrateData using dims=1:50. Integrated data was used for data scaling with
825 ScaleData and PCA dimensionality reduction with RunPCA. After initial clustering we
826 noted three small clusters representing 7% of total cells which had low number of genes
827 detected and high mitochondrial RNA content which were excluded from further

828 analysis. Clusters were identified using shared nearest neighbor (SNN) based clustering
829 based on the first 15 PCs with resolution = 0.45. The same principal components were
830 used to generate the UMAP projections, which were generated with a minimum
831 distance of 0.1. Cell clustering and UMAP projection for Chronic Day 21 T cells (all
832 tissues, Figure 2 and Figure 5), spleen derived T cells (Chronic and Acute, Day 8 and
833 Day 21, Figure 4 and Figure 6), and Day 8 T cells (Spleen, Chronic and Acute,
834 Supplemental Figure 3) were performed as described above with the following
835 modifications:
836   Chronic Day 21 T cells: dims = 1:10, resolution = 0.25, min.dist = 0.1
837   Spleen derived T cells: dims = 1:8, k.param = 50, resolution = 0.45, min.dist = 0.1
838   Day 8 T cells: dims = 1:12, k.param = 40, resolution = 0.28, min.dist = 0.2
839
840 Expression of selected genes was plotted using log normalized gene expression values
841 based on original RNA count data prior to data integration. Marker genes were identified
842 using Seurat's FindAllMarkers using a cutoff of $p\_val\_adj < 0.01$. Differential gene
843 expression analysis was performed using Seurat's FindMarkers using a cutoff of
844 $p\_val\_adj < 0.05$ and $abs(avg\_log2FC) > 0.25$. Gene module scoring was performed
845 using Seurat's AddModuleScore. TCR clone behaviors were visualized using UpSetR
846 (version 1.4.0). Null distribution of TCR clone behaviors was determined by randomly
847 shuffling TCR clonotype and scRNA phenotype and generating a distribution of TCR
848 clone phenotype combinations (n=50 iterations). Morisita-Horn index for quantifying
849 TCR overlap was calculated using the mh function from the R package divo (version
850 1.0.1).
851

References

852 **References**
853 1.   Hashimoto, M., et al., *CD8 T Cell Exhaustion in Chronic Infection and Cancer:*
854   *Opportunities for Interventions.* Annu Rev Med, 2018. **69**: p. 301-318.
855 2.   McLane, L.M., M.S. Abdel-Hakeem, and E.J. Wherry, *CD8 T Cell Exhaustion*
856   *During Chronic Viral Infection and Cancer.* Annu Rev Immunol, 2019. **37**: p. 457-
857   495.
858 3.   Blank, C.U., et al., *Defining 'T cell exhaustion'.* Nat Rev Immunol, 2019. **19**(11):
859   p. 665-674.
860 4.   Scott, A.C., et al., *TOX is a critical regulator of tumour-specific T cell*
861   *differentiation.* Nature, 2019. **571**(7764): p. 270-274.
862 5.   Paley, M.A., et al., *Progenitor and terminal subsets of CD8+ T cells cooperate to*
863   *contain chronic viral infection.* Science, 2012. **338**(6111): p. 1220-5.
864 6.   Im, S.J., et al., *Defining CD8+ T cells that provide the proliferative burst after PD-*
865   *1 therapy.* Nature, 2016. **537**(7620): p. 417-421.
866 7.   Utzschneider, D.T., et al., *T Cell Factor 1-Expressing Memory-like CD8(+) T*
867   *Cells Sustain the Immune Response to Chronic Viral Infections.* Immunity, 2016.
868   **45**(2): p. 415-27.
869 8.   Hudson, W.H., et al., *Proliferating Transitory T Cells with an Effector-like*
870   *Transcriptional Signature Emerge from PD-1(+) Stem-like CD8(+) T Cells during*
871   *Chronic Infection.* Immunity, 2019. **51**(6): p. 1043-1058 e4.

872 9.   Zander, R., et al., *CD4(+) T Cell Help Is Required for the Formation of a Cytolytic*
873   *CD8(+) T Cell Subset that Protects against Chronic Infection and Cancer.*
874   Immunity, 2019. **51**(6): p. 1028-1042 e4.
875 10.   Beltra, J.C., et al., *Developmental Relationships of Four Exhausted CD8(+) T*
876   *Cell Subsets Reveals Underlying Transcriptional and Epigenetic Landscape*
877   *Control Mechanisms.* Immunity, 2020. **52**(5): p. 825-841 e8.
878 11.   He, R., et al., *Follicular CXCR5- expressing CD8(+) T cells curtail chronic viral*
879   *infection.* Nature, 2016. **537**(7620): p. 412-428.
880 12.   Wu, T., et al., *The TCF1-Bcl6 axis counteracts type I interferon to repress*
881   *exhaustion and maintain T cell stemness.* Sci Immunol, 2016. **1**(6).
882 13.   Raju, S., et al., *Identification of a T-bet(hi) Quiescent Exhausted CD8 T Cell*
883   *Subpopulation That Can Differentiate into TIM3(+)CX3CR1(+) Effectors and*
884   *Memory-like Cells.* J Immunol, 2021.
885 14.   Utzschneider, D.T., et al., *Early precursor T cells establish and propagate T cell*
886   *exhaustion in chronic infection.* Nat Immunol, 2020. **21**(10): p. 1256-1266.
887 15.   Khan, O., et al., *TOX transcriptionally and epigenetically programs CD8(+) T cell*
888   *exhaustion.* Nature, 2019. **571**(7764): p. 211-218.
889 16.   Alfei, F., et al., *TOX reinforces the phenotype and longevity of exhausted T cells*
890   *in chronic viral infection.* Nature, 2019. **571**(7764): p. 265-269.
891 17.   Yao, C., et al., *BACH2 enforces the transcriptional and epigenetic programs of*
892   *stem-like CD8(+) T cells.* Nat Immunol, 2021. **22**(3): p. 370-380.
893 18.   Yao, C., et al., *Single-cell RNA-seq reveals TOX as a key regulator of CD8(+) T*
894   *cell persistence in chronic infection.* Nat Immunol, 2019. **20**(7): p. 890-901.
895 19.   Seo, H., et al., *TOX and TOX2 transcription factors cooperate with NR4A*
896   *transcription factors to impose CD8(+) T cell exhaustion.* Proc Natl Acad Sci U S
897   A, 2019. **116**(25): p. 12410-12415.
898 20.   Pauken, K.E., et al., *Epigenetic stability of exhausted T cells limits durability of*
899   *reinvigoration by PD-1 blockade.* Science, 2016. **354**(6316): p. 1160-1165.
900 21.   Matloubian, M., et al., *Molecular determinants of macrophage tropism and viral*
901   *persistence: importance of single amino acid changes in the polymerase and*
902   *glycoprotein of lymphocytic choriomeningitis virus.* J Virol, 1993. **67**(12): p. 7340-
903   9.
904 22.   Matloubian, M., et al., *Genetic basis of viral persistence: single amino acid*
905   *change in the viral glycoprotein affects ability of lymphocytic choriomeningitis*
906   *virus to persist in adult mice.* J Exp Med, 1990. **172**(4): p. 1043-8.
907 23.   Stamm, A., et al., *An intermediate dose of LCMV clone 13 causes prolonged*
908   *morbidity that is maintained by CD4+ T cells.* Virology, 2012. **425**(2): p. 122-32.
909 24.   Wherry, E.J., et al., *Viral persistence alters CD8 T-cell immunodominance and*
910   *tissue distribution and results in distinct stages of functional impairment.* J Virol,
911   2003. **77**(8): p. 4911-27.
912 25.   Pliner, H.A., et al., *Cicero Predicts cis-Regulatory DNA Interactions from Single-*
913   *Cell Chromatin Accessibility Data.* Mol Cell, 2018. **71**(5): p. 858-871 e8.
914 26.   Schep, A.N., et al., *chromVAR: inferring transcription-factor-associated*
915   *accessibility from single-cell epigenomic data.* Nat Methods, 2017. **14**(10): p.
916   975-978.

27. Chen, Y., et al., *BATF regulates progenitor to cytolytic effector CD8(+) T cell transition during chronic viral infection.* Nat Immunol, 2021. **22**(8): p. 996-1007.

28. Stelekati, E., et al., *Long-Term Persistence of Exhausted CD8 T Cells in Chronic Infection Is Regulated by MicroRNA-155.* Cell Rep, 2018. **23**(7): p. 2142-2156.

29. Chen, J., et al., *NR4A transcription factors limit CAR T cell function in solid tumours.* Nature, 2019. **567**(7749): p. 530-534.

30. Liu, X., et al., *Genome-wide analysis identifies NR4A1 as a key mediator of T cell dysfunction.* Nature, 2019. **567**(7749): p. 525-529.

31. Shan, Q., et al., *The transcription factor Runx3 guards cytotoxic CD8(+) effector T cells against deviation towards follicular helper T cell lineage.* Nat Immunol, 2017. **18**(8): p. 931-939.

32. Martinez, G.J., et al., *The transcription factor NFAT promotes exhaustion of activated CD8(+) T cells.* Immunity, 2015. **42**(2): p. 265-278.

33. Chung, H.K., B. McDonald, and S.M. Kaech, *The architectural design of CD8+ T cell responses in acute and chronic infection: Parallel structures with divergent fates.* J Exp Med, 2021. **218**(4).

34. Evrard, M., et al., *Sphingosine 1-phosphate receptor 5 (S1PR5) regulates the peripheral retention of tissue-resident lymphocytes.* J Exp Med, 2022. **219**(1).

35. Mueller, P., et al., *Regulation of T cell survival through coronin-1-mediated generation of inositol-1,4,5-trisphosphate and calcium mobilization after T cell receptor triggering.* Nat Immunol, 2008. **9**(4): p. 424-31.

36. Lin, Y., et al., *Identification of TAPBPL as a novel negative regulator of T-cell function.* EMBO Mol Med, 2021. **13**(5): p. e13404.

37. Granum, S., et al., *Modulation of Lck function through multisite docking to T cell-specific adapter protein.* J Biol Chem, 2008. **283**(32): p. 21909-19.

38. Milner, J.J., et al., *Delineation of a molecularly distinct terminally differentiated memory CD8 T cell population.* Proc Natl Acad Sci U S A, 2020. **117**(41): p. 25667-25678.

39. Jenne, C.N. and P. Kubes, *Immune surveillance by the liver.* Nat Immunol, 2013. **14**(10): p. 996-1006.

40. Sen, D.R., et al., *The epigenetic landscape of T cell exhaustion.* Science, 2016. **354**(6316): p. 1165-1169.

41. Pritykin, Y., et al., *A unified atlas of CD8 T cell dysfunctional states in cancer and infection.* Mol Cell, 2021. **81**(11): p. 2477-2493 e10.

42. Leong, Y.A., et al., *CXCR5(+) follicular cytotoxic T cells control viral infection in B cell follicles.* Nat Immunol, 2016. **17**(10): p. 1187-96.

43. Miller, B.C., et al., *Subsets of exhausted CD8(+) T cells differentially mediate tumor control and respond to checkpoint blockade.* Nat Immunol, 2019. **20**(3): p. 326-336.

44. Philip, M., et al., *Chromatin states define tumour-specific T cell dysfunction and reprogramming.* Nature, 2017. **545**(7655): p. 452-456.

45. Abdel-Hakeem, M.S., et al., *Epigenetic scarring of exhausted T cells hinders memory differentiation upon eliminating chronic antigenic stimulation.* Nat Immunol, 2021. **22**(8): p. 1008-1019.

46. Yates, K.B., et al., *Epigenetic scars of CD8(+) T cell exhaustion persist after cure of chronic infection in humans.* Nat Immunol, 2021. **22**(8): p. 1020-1029.

963 47. Philip, M. and A. Schietinger, *Heterogeneity and fate choice: T cell exhaustion in*
964   *cancer and chronic infections.* Curr Opin Immunol, 2019. **58**: p. 98-103.
965 48. Stadtmauer, E.A., et al., *CRISPR-engineered T cells in patients with refractory*
966   *cancer.* Science, 2020. **367**(6481).
967 49. Chang, J.T., E.J. Wherry, and A.W. Goldrath, *Molecular regulation of effector and*
968   *memory T cell differentiation.* Nat Immunol, 2014. **15**(12): p. 1104-15.
969 50. Daniels, M.A. and E. Teixeiro, *TCR Signaling in T Cell Memory.* Front Immunol,
970   2015. **6**: p. 617.
971 51. Good, C.R., et al., *An NK-like CAR T cell transition in CAR T cell dysfunction.*
972   Cell, 2021.
973 52. Chou, C., et al., *The Transcription Factor AP4 Mediates Resolution of Chronic*
974   *Viral Infection through Amplification of Germinal Center B Cell Responses.*
975   Immunity, 2016. **45**(3): p. 570-582.
976 53. Satpathy, A.T., et al., *Massively parallel single-cell chromatin landscapes of*
977   *human immune cell development and intratumoral T cell exhaustion.* Nat
978   Biotechnol, 2019. **37**(8): p. 925-936.
979 54. Granja, J.M., et al., *ArchR is a scalable software package for integrative single-*
980   *cell chromatin accessibility analysis.* Nat Genet, 2021. **53**(3): p. 403-411.
981 55. van Dijk, D., et al., *Recovering Gene Interactions from Single-Cell Data Using*
982   *Data Diffusion.* Cell, 2018. **174**(3): p. 716-729 e27.
983 56. Hao, Y., et al., *Integrated analysis of multimodal single-cell data.* Cell, 2021.
984   **184**(13): p. 3573-3587 e29.
985 57. McGinnis, C.S., L.M. Murrow, and Z.J. Gartner, *DoubletFinder: Doublet Detection*
986   *in Single-Cell RNA Sequencing Data Using Artificial Nearest Neighbors.* Cell
987   Syst, 2019. **8**(4): p. 329-337 e4.
988 58. Aran, D., et al., *Reference-based analysis of lung single-cell sequencing reveals*
989   *a transitional profibrotic macrophage.* Nat Immunol, 2019. **20**(2): p. 163-172.
990 59. Tirosh, I., et al., *Dissecting the multicellular ecosystem of metastatic melanoma*
991   *by single-cell RNA-seq.* Science, 2016. **352**(6282): p. 189-96.
992 60. Sandu, I., et al., *Landscape of Exhausted Virus-Specific CD8 T Cells in Chronic*
993   *LCMV Infection.* Cell Rep, 2020. **32**(8): p. 108078.

Figure 1.
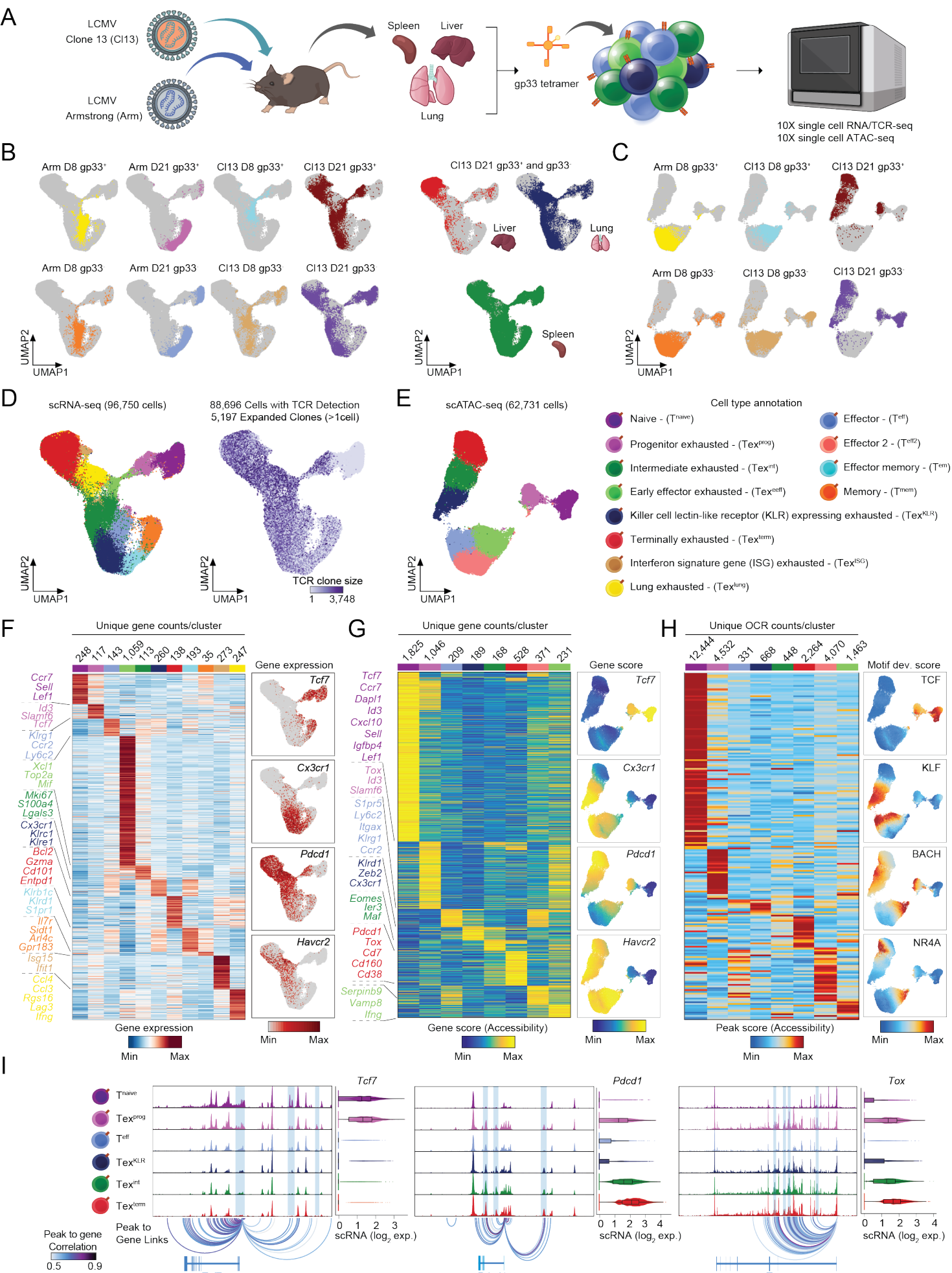
Figure 1. Single-cell genomic atlas of T cell exhaustion during LCMV infection.
(A) Schematics on the mouse model used, indicating the two viral strains, the tetramer sort and the single cell technologies applied. **(B)** UMAPs of scRNA-seq profiles colored by the samples (gp33$^+$ and gp33$^-$ fractions) sorted from the spleen of Arm- or Cl13-infected animals on the indicated days (D8 and D21) (left). UMAPs of scRNA-seq profiles colored by the samples originating from the different organs of Cl13-infected animals at D21 (right). **(C)** UMAPs of scATAC-seq profiles colored by the samples (gp33$^+$ and gp33$^-$ fractions) sorted from the spleen of Arm- or Cl13-infected animals on the indicated days. **(D)** UMAP of all scRNA-seq profiles, colored by the annotated T cell subsets (left). UMAP of scTCR-seq results colored by the size of the expanded clones from which individual T cells originate (right). **(E)** UMAP of all scATAC-seq profiles colored by the annotated T cell subsets. **(F)** Heat map of subset specific marker genes determined by scRNA-seq. Feature plots of specific gene markers that characterize T cell subsets. **(G)** Heat map of Gene score values (accessibility) determined by scATAC-seq. Feature plots of specific Gene score values that mark main T cell subsets. **(H)** Heat map of Peak score values at the unique open chromatin regions (OCRs) of the T cell subsets determined by scATAC-seq. Feature plots show the motifs that are accessible in specific T cell subsets (chromVAR deviation scores are depicted). **(I)** Genome browser snapshots on the indicated gene loci, showing the chromatin states of the different T cell subsets. Violin plots show the associated expression level of the indicated genes from the respective T cell subsets determined by scRNA-seq.

Figure 2.

A PD-1⁺ CX3CR1⁺ CD8⁺ sort

D21 Clone 13

Phenotypic distribution

Cell type annotation

Tnaive	Tem
Texeeff2	Tmem
Texeeff	Texint
Texprog	TexKLR
TexISG	Texlung
Texterm	

B

Texint vs. TexKLR
- Cell Cycle: G2/M Checkpoint Regulation
- Kinetochore Metaphase Signaling Pathway
- Cell Cycle Control of Chromosomal Replication
- Cyclins and Cell Cycle Regulation
- T Cell Exhaustion Signaling Pathway
- NUR77 Signaling in T Lymphocytes

TexKLR vs. Texint
- T Cell Activation Signaling
- PI3K/AKT Signaling
- FAK Signaling
- Integrin Signaling
- RAC Signaling
- Leukocyte Extravasation Signaling

-log₁₀(p-value)

C Expression of T-Tem gene signature from Milner et al.

D T-Tem gene signature score

E Gene expression

F Cell type annotation / Phenotype distribution

G Spleen / Lung / Liver

H

Spleen Texterm program (86 genes) — Texterm — Lung Texterm program (159 genes)

Spleen Texterm program (103 genes) — Texterm — Liver Texterm program (101 genes)

Enriched biological pathways

Texterm
1. EIF2 Signaling
2. T Cell Exhaustion Pathway
3. T Cell Receptor Signaling

1. Leukocyte Extravasation
2. Signaling by Rho GTPases
3. T Cell Activation Pathway

Enriched biological pathways

Texterm
1. EIF2 Signaling
2. T Cell Exhaustion Pathway
3. T Cell Activation Pathway

1. TNFR2 Signaling
2. Inflammatory Signaling
3. Glucocorticoid Receptor Signaling

I Common exhaustion signature

J Exhaustion score (Im et al.)

Figure 2. Identification of early effector, KLR-expressing, and organ-specific Tex subsets.

(A) UMAP of scRNA-seq results colored by the main T cell subsets of the sorted PD-1[+], CX3CR1[+] and CD8[+] T cells. Stacked bar plot shows the phenotypic distribution of the sorted population (right) **(B)** Volcano plot of differentially expressed genes between the Tex^{KLR} and Tex^{int} cell populations (left). Ingenuity pathway analyses on the differentially expressed genes show the enriched biological pathways in the two subsets. Top 6 hits are shown. **(C)** Heatmap of the expression of the marker genes of terminal effector memory (T-T^{em}) cells defined by Milner et al. in the indicated T cell subsets. **(D)** UMAP colored by the strength of the T-T^{em} gene signature (T-T^{em} module score) in the scRNA-seq dataset (left). Violin plot representation of the T-T^{em} score in the indicated T cell subsets. **(E)** UMAPs colored by the expression of the indicated marker genes of the T-T^{em} subset. **(F)** UMAP of scRNA-seq results from the three organs at D21 following Cl13 infection colored by the annotated T cell subsets (left). Stacked bar plot representation of the phenotypic distribution of the annotated T cell subsets in the three organs (right). **(G)** UMAPs colored by the cells from the three organs. **(H)** Volcano plots of differentially expressed genes comparing the Tex^{term} cell populations from the different organs. Ingenuity pathway analysis results on the differentially expressed gene groups (bottom). Top 3 hits are shown. **(I)** Heat map of the gene expression values of the common exhaustion gene signature among the organ specific Tex^{term} subsets. **(J)** Violin plot depicts the exhaustion scores of the three organs based on Im et al. 2016.

Figure 3.
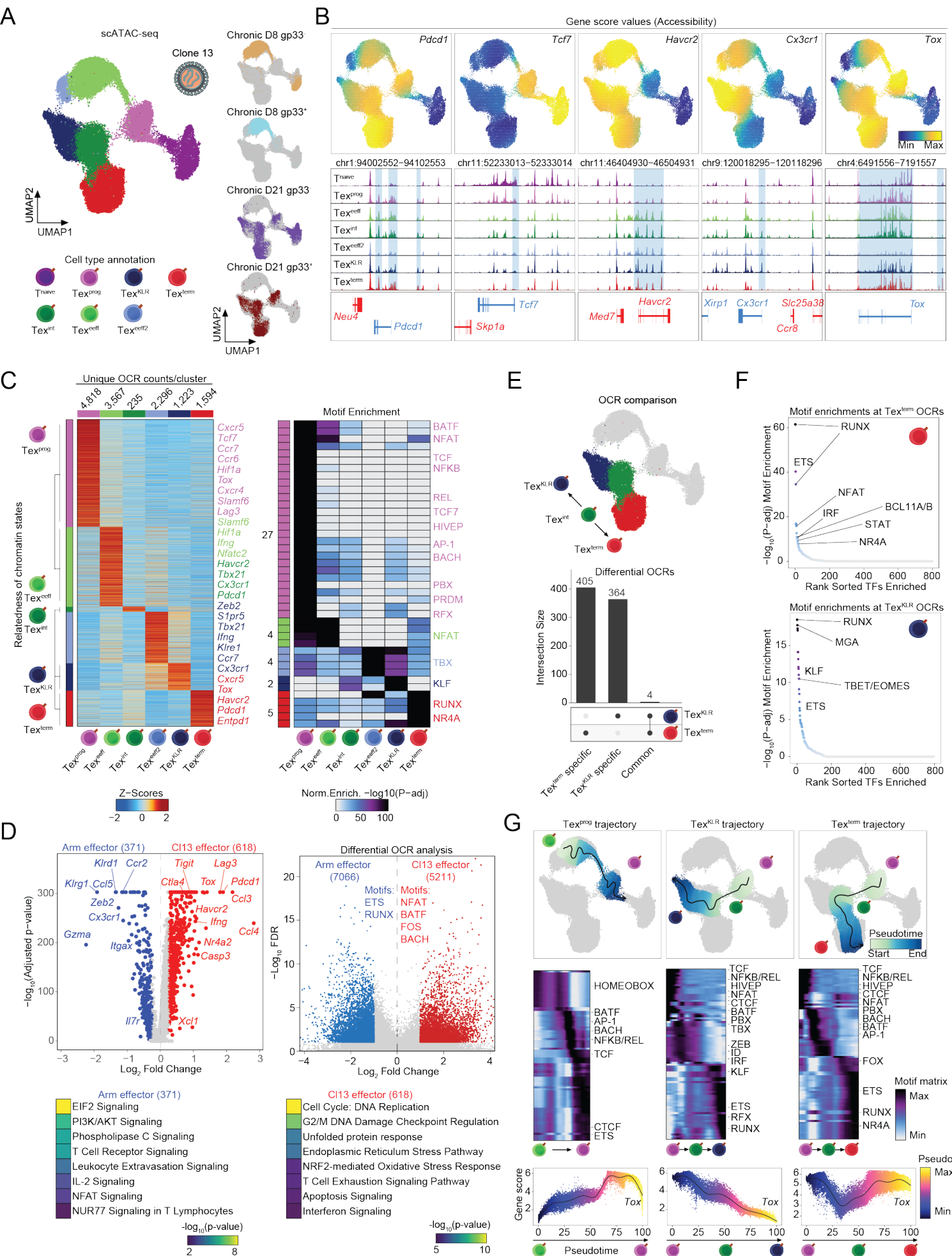
Figure 3. Tex^{int} represent a bifurcation point of exhausted T cell fate differentiation. **(A)** UMAP of scATAC-seq results of D8 and D21 gp33^{+} and gp33^{-} T cells from the Cl13 infection model. UMAP is colored by the annotated T cell subsets. Small UMAPs (right) show T cells that originate from the indicated gp33 fractions and timepoints. **(B)** Feature plots of the indicated Gene score values (accessibility) (top) and genome browser snapshots of the corresponding genomic loci (bottom). **(C)** Heat map of Peak score values at the unique open chromatin regions (OCRs) of the T cell subsets determined by scATAC-seq with a list of annotated putative target genes based on proximity (left). Heat map of motif enrichment results at the unique OCR sets of the annotated T cell subsets. **(D)** Volcano plot of differentially expressed genes between the Arm effector cells and Cl13 early effector cells (left). Ingenuity pathway analysis results show the top 8 enriched biological terms (bottom). Volcano plot depicts the differentially accessible OCRs between the Arm effector cells and Cl13 early effector cells (right). **(E)** UMAP depicts the populations used for differential OCR analysis (top). Upset plot of differentially accessible OCRs and their overlap among the Tex^{KLR} and Tex^{term} populations (bottom). **(F)** Hockey stick plots depict the enriched transcription factor motifs at the specific OCRs of the Tex^{term} and Tex^{KLR} subsets. **(G)** Pseudotime trajectory analyses of three potential Tex differentiation paths (top). Heat maps show transcription factor deviation scores that change over the pseudotime trajectories (middle). Gene score values of *Tox* on the three pseudotime trajectories (bottom).

Figure 4.
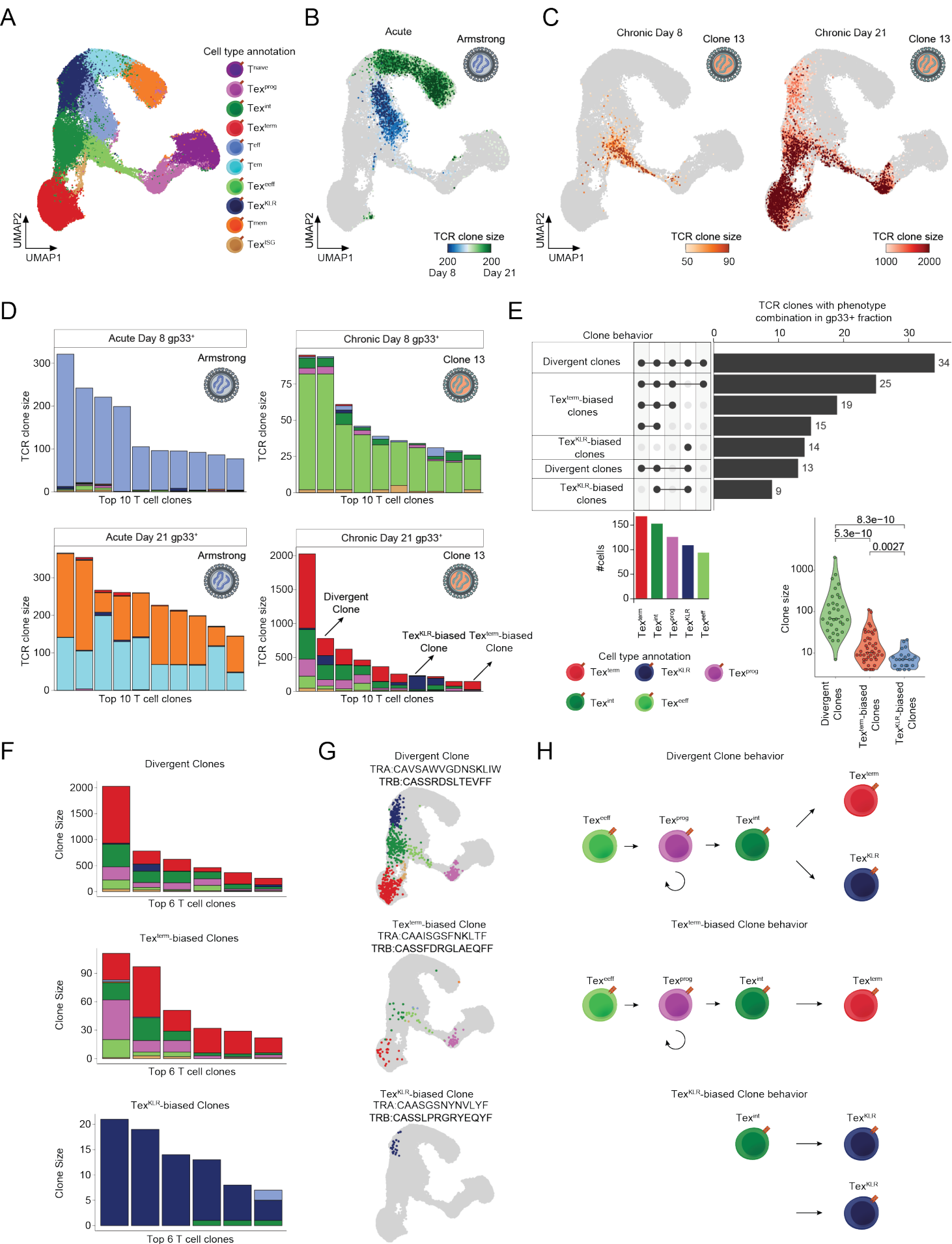
Figure 4. TCR-based lineage tracing reveals divergent Tex clonal trajectories.
(A) UMAP of scRNA-seq results from the gp33$^+$ and gp33$^-$ T cell factions of the Arm and Cl13 infection model from D8 and D21 following infection. UMAP is colored by the annotated T cell subsets. **(B)** UMAP of scRNA-seq results colored by the size of the detected TCR clones at D8 and D21 in the Arm infection model. **(C)** UMAP of scRNA-seq results colored by the size of the detected TCR clones at D8 in the Cl13 infection model (left). Same UMAP colored by the TCR clone size at the D21 time point in the Cl13 infection model (right). **(D)** Stacked bar plot of the phenotypic distribution of the top 10 expanded clones in the gp33$^+$ fraction of Arm D8 and D21 samples (left). Same stacked bar plots representing the top 10 expanded clones in the Cl13 infection model (right). **(E)** Upset plot depicting the expanded clones with specific phenotype combinations (clone behaviors). Barplot shows the number of cells with the indicated phenotypes that make up the expanded clones. Violin plot shows the clone size distribution of the detected clone behaviors. **(F)** Stacked bar plots show the top 6 expanded clones with the indicated clone behaviors. **(G)** UMAPs show representative examples for the detected clone behaviors. **(H)** Scheme on the phenotypic composition and the potential differentiation trajectories of the identified clone behaviors.

Figure 5.

A Day 21 — Clone 13

Cell type annotation:
- T^naive
- Tex^prog
- Tex^int
- Tex^ISG
- Tex^KLR
- Tex^term
- Tex^lung

UMAP2 / UMAP1

B
- Spleen TCR Clone Size — 36,272 Cells with TCR Detection — 2,365 Expanded Clones (>1 cell) — TCR clone size ($\log_{10}$) 1 4
- Lung TCR Clone Size — 13,650 Cells with TCR Detection — 922 Expanded Clones (>1 cell) — TCR clone size ($\log_{10}$) 1 3
- Liver TCR Clone Size — 12,549 Cells with TCR Detection — 571 Expanded Clones (>1 cell) — TCR clone size ($\log_{10}$) 1 4

C Clone frequencies across organs in gp33+ fraction

- $R = 0.66$ — Clone frequency in liver vs Clone frequency in spleen
  - 104 Shared clones
  - 6 Liver specific clones
  - 59 Spleen specific clones
- $R = 0.65$ — Clone frequency in lung vs Clone frequency in spleen
  - 137 Shared clones
  - 6 Lung specific clones
  - 26 Spleen specific clones
- $R = 0.76$ — Clone frequency in liver vs Clone frequency in lung
  - 106 Shared clones
  - 4 Liver specific clones
  - 37 Lung specific clones

D Overlap of expanded clones in the gp33+ fractions of organs

Liver 0, Lung 0, 6, 100, 4, 37, Spleen 22

E
- Divergent Clone 1
- Divergent Clone 2
- Tex^term-biased Clone 1
- Tex^term-biased Clone 2
- Tex^KLR-biased Clone 1
- Tex^KLR-biased Clone 2
- Tex^KLR-biased Clone 3
- Tex^KLR-biased Clone 4

Cell count in clone — Spleen / Lung / Liver — gp33+

Cell type annotation:
- T^naive
- Tex^prog
- Tex^int
- Tex^KLR
- Tex^term
- Tex^lung
- Tex^ISG

F
- Divergent Clone
- Tex^term-biased Clone
- Tex^KLR-biased Clone

Spleen / Lung / Liver

UMAP2 / UMAP1

G
- Tex^KLR Phenotype
  - $R=0.96$, $p < 2.2e-16$ — Lung / Spleen
  - $R=0.63$, $p < 0.00028$ — Liver / Spleen
- Tex^term Phenotype
  - $R=0.75$, $p < 3.5e-8$ — Lung / Spleen
  - $R=0.6$, $p < 0.0067$ — Liver / Spleen

Fraction of clone with phenotype (y) vs Fraction of clone with phenotype (x)

$\log_{10}$(Clone size): 1.0, 1.5, 2.0, 2.5, 3.0, 3.5

H
- Clone frequency in organs — Tex^KLR-biased Clones
  - 2.2e-13
  - 0.29
  - 1.6e-06
- Clone frequency in organs — Divergent and Tex^term-biased Clones

Figure 5. Conserved clonal T cell trajectories across organs and depletion of Tex^{KLR} in the liver microenvironment.

(A) UMAP of organ-derived T cells at D21in Cl13 infection colored by the annotated T cell subsets. **(B)** UMAPs colored by the detected TCR clone sizes in the different organs. **(C)** Scatterplots depicting the frequencies of expanded T cell clones from the indicated organ comparisons. The correlation coefficient, and specific and shared clone numbers are indicated for each comparison. **(D)** Venn diagram depicting the overlap of expanded T cell clones in the gp33$^+$ fraction of the indicated organs. **(E)** Stacked bar plot of the phenotypic composition of individual clones across organs. **(F)** UMAPs depict individual clones with specific clone behaviors among organs. **(G)** Scatter plots showing the fraction of the shared clones with Tex^{KLR} and Tex^{term} phenotypes between the indicated organs. **(H)** Violin plot depicts Tex^{KLR}-biased clone frequencies across the organs, which includes clones with >50% Tex^{KLR} phenotype (left). Violin plot of Tex^{term}-biased and divergent clone frequencies across the organs.

Figure 6.

A Gated on CD8⁺ T cells

CD8b - PerCP-Cy5.5 vs gp33 - PE
gp33int, gp33-, gp33high
Clone 13, Spleen

B Morisita index (TCR overlap)

Liver gp33⁺
Lung gp33⁺
Spleen gp33high
Spleen gp33int
Spleen gp33-
Lung gp33-
Liver gp33-

Similarity (0.2 – 1.0)

C Clone size distribution across tetramer sorted CD8⁺ populations

Spleen gp33- — 1786 expanded clones
>200 4.9%, 50–200 8.6%, 10–50 14%, 1–10 24%, =1 49%

Spleen gp33int — 567 expanded clones
>200 10%, =1 31%, 1–10 18%, 10–50 30%, 50–200 12%

Spleen gp33high — 248 expanded clones
=1 11%, 1–10 11%, 10–50 11%, 50–200 18%, >200 49%

D Expanded clones

gp33high: 88
gp33-: 592
gp33int: 114
gp33high ∩ gp33-: 2
gp33high ∩ gp33int: 147
gp33- ∩ gp33int: 295
all three: 11

E Unique clones across tetramer sorted CD8⁺ populations

Chronic Day 21 gp33- — Unique clones n=592, TCR clone size (log₁₀) 0.0–1.6
Chronic Day 21 gp33int — Unique clones n=114, TCR clone size (log₁₀) 0.0–1.0
Chronic Day 21 gp33high — Unique clones n=88, TCR clone size (log₁₀) 0.0–2.0

F Unique clones in gp33 fractions

Fraction of cells with phenotype — gp33-, gp33int, gp33high

Cell type annotation:
T^naive, Tex^prog, Tex^int, Tex^term, T^eff, T^em, Tex^eeff, Tex^KLR, Tex^ISG, T^mem

G Unique clones across tetramer sorted CD8⁺ populations

Chronic Day 21 gp33-
Chronic Day 21 gp33int
Chronic Day 21 gp33high

H

Chronic Day 21 gp33- — Clone Size vs Top 10 T cell clones, Clone 13
Chronic Day 21 gp33int — Clone Size vs Top 10 T cell clones, Clone 13
Chronic Day 21 gp33high — Clone Size vs Top 10 T cell clones, Clone 13

T^naive, Tex^prog, Tex^int, Tex^term, T^eff, T^em, Tex^eeff, Tex^KLR, Tex^ISG, T^mem

I

TCR clones with phenotype combination in gp33- fraction
Clone behavior:
Tex^KLR-biased 19
Tex^term-biased 10
Tex^term-biased 9
Tex^term-biased 8
Tex^term-biased 6
Tex^KLR-biased 5
Tex^term-biased 5

#cells: Tex^term, Tex^int, Tex^KLR, Tex^prog, Tex^eeff

Tex^eeff → Tex^prog → Tex^int → Tex^term
Dominant clone behavior → Tex^term-biased clone behavior

TCR clones with phenotype combination in gp33int fraction
Clone behavior:
Tex^KLR-biased 14
Tex^term-biased 4
Tex-biased 3
Tex^KLR-biased 2
Divergent 2
Tex^term-biased 1
Tex^term-biased 1

#cells: Tex^KLR, Tex^int, Tex^term, Tex^eeff, Tex^prog

Tex^int → Tex^KLR
Tex^KLR-biased clone behavior

TCR clones with phenotype combination in gp33high fraction
Clone behavior:
Tex^term-biased 20
Tex^term-biased 14
Divergent 11
Tex^term-biased 9
Tex^term-biased 8
Tex^term-biased 7
Divergent 6

#cells: Tex^term, Tex^int, Tex^prog, Tex^eeff, Tex^KLR

Tex^eeff → Tex^prog → Tex^int → Tex^KLR / Tex^term
Divergent clone behavior | Tex^int → Tex^term | Tex^term-biased clone behavior

Figure 6. Differences in TCR signal strength regulate clonal differentiation of Tex^{KLR} and Tex^{term}.

(A) Sorting strategy to obtain gp33$^-$, gp33^{int} and gp33^{high} T cell populations from the spleen of LCMV-Cl13 infected animals 21 days following infection. **(B)** Heat map depicting TCR repertoire overlap (Morisita index) among the different gp33 fractions from the indicated samples. **(C)** Pie chart representation of the fraction of the detected clone sizes in the three gp33 T cell fractions. **(D)** Venn diagram depicts the overlap of the expanded clones from the gp33 T cell fractions. **(E)** UMAPs colored by size of the unique expanded clones in the three gp33 T cell fractions. **(F)** Stacked bar plot of the phenotypic distribution of the unique expanded clones of the three gp33 T cell fractions. **(G)** UMAPs visualizing the unique expanded clones of the three gp33 T cell fractions colored by the annotated T cell subsets. **(H)** Stacked bar plot of the top 10 uniquely expanded T cell clones from the three gp33 T cell fractions colored by the annotated T cell phenotypes. **(I)** Upset plots depict the unique expanded clones with specific phenotype combinations (clonotype behavior) from the three gp33 T cell fractions. Barplots show the number of cells with the indicated phenotypes. Dominant clone behaviors are indicated at the bottom.

Figure S1.
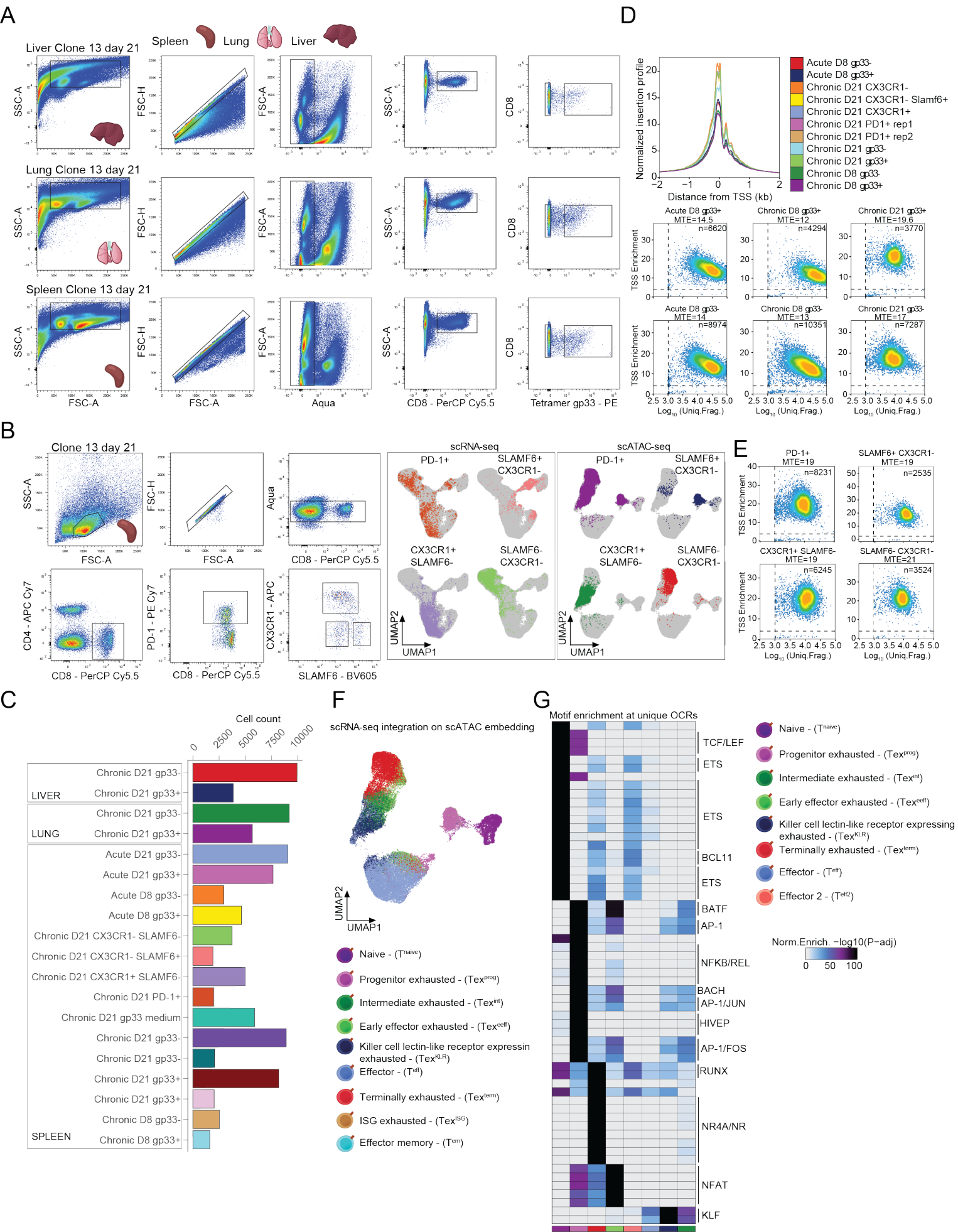
Figure S1. Sorting strategy and quality controls of scATAC-seq data. Related to Figure 1.

(A) Sorting strategy to obtain antigen specific gp33$^+$ and gp33$^-$ CD8$^+$ T cells from different organs. **(B)** Sorting strategy to obtain the main exhausted T cell subsets (left). UMAPs of scRNA-seq and scATAC-seq results, originating from the main, indicated exhausted T cell subsets. **(C)** Bar plot representation of cell counts from the scRNA-seq results. **(D)** Quality control of scATAC-seq data. Histogram shows normalized read enrichment on the transcription start sites (TSS) of genes from the indicated samples (top). Density plots depict the cells that passed the TSS enrichment and Log$_{10}$ unique fragment count threshold. Median TSS enrichment (MTE) is also indicated. **(E)** Density plots of scATAC-seq data from the main exhausted T cell populations depicting the same quality controls as on panel C. **(F)** UMAP of scATAC-seq data colored by the integrated scRNA-seq cluster labels. **(G)** Heat map of motif enrichments at the specific open chromatin regions (OCRs) of the annotated T cell populations.

Figure S2.

A
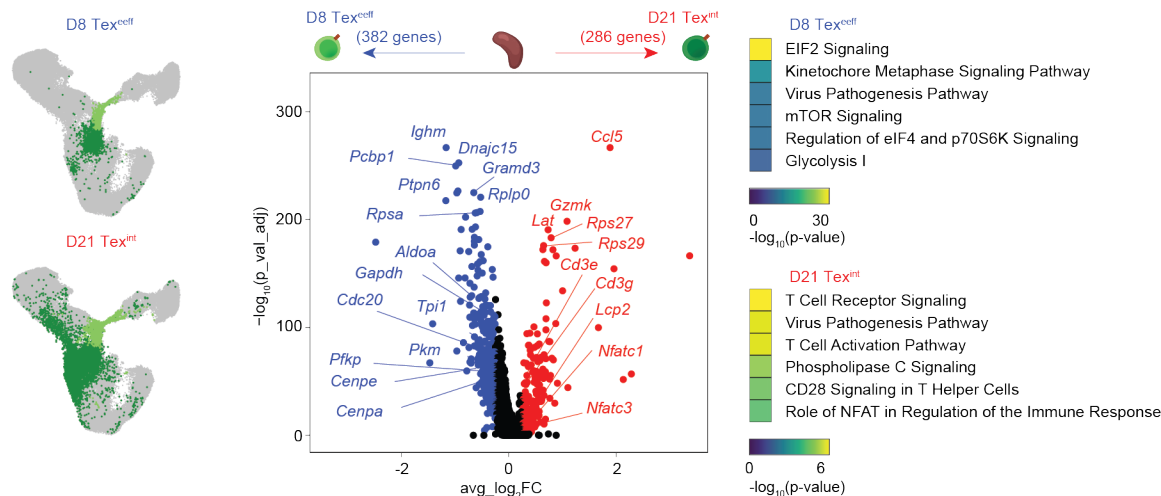
B
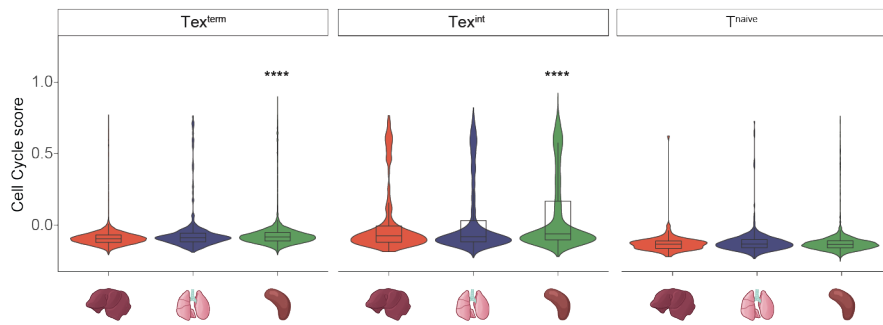
Figure S2. Early effector exhausted cells and intermediate exhausted T cells are phenotypically different populations with distinct temporal appearance. Related to Figure 2.

(A) UMAPs visualize the early effector exhausted population at D8 and the intermediate exhausted population at D21 (left). Volcano plot depicts the differentially expressed genes between the two populations (middle). Ingenuity pathway analysis results depict the top 6 enriched biological terms in the two populations. **(B)** Violin plots depict the Cell Cycle score of the indicated T cell populations across the indicated organs.

Figure S3.

A scATAC-seq Day 8   scRNA-seq Day 8   Cell type annotation

T^{naive} T^{mp}/Tex^{prog}
T^{eff2} Tex^{eff1}
T^{eff1} Tex^{ISG}
Tex^{eff2} ($Cx3cr1$ low)

Arm   Cl13   Arm   Cl13

B Arm T^{mp} program (93 genes)   Cl13 Tex^{prog} program (102 genes)

T^{mp} Tex^{prog}

$Ccl5$, $Klrd1$, $Ccr7$, $Il7r$, $Gzma$, $Klrc1$, $Cxcr3$
$Lag3$, $Pdcd1$, $Xcl1$, Tox, $Tigit$, $Cxcl10$, $Ifng$, $Icos$, $Ccl4$, $Ctla4$

x-axis: Avg. Log$_2$ Fold Change
y-axis: $-\log_{10}$(Adjusted p-value)

C Arm T^{mp} program (93 genes)

- EIF2 Signaling
- Th1 and Th2 Activation Pathway
- Th1 Pathway
- Granulocyte Adhesion and Diapedesis
- Virus Pathogenesis Pathway
- Agranulocyte Adhesion and Diapedesis
- Regulation of eIF4 and p70S6K Signaling
- Regulation of Actin-based Motility by Rho

$-\log_{10}$(P-adj)

Cl13 Tex^{prog} program (102 genes)

- Virus Pathogenesis Pathway
- EIF2 Signaling
- mTOR Signaling
- Role of Hypercytokinemia/hyperchemokinemia in the Pathogenesis of Influenza
- Regulation of eIF4 and p70S6K Signaling
- Activation of IRF Signaling Pathway
- Interferon Signaling
- Role of PKR in Interferon Induction and Antiviral Response

$-\log_{10}$(P-adj)

D Differential OCR analysis

Arm T^{mp} (4962)   Cl13 Tex^{prog} (5437)

T^{mp} Tex^{prog}

x-axis: Log$_2$ Fold Change
y-axis: $-\log_{10}$ FDR

Motif enrichment Arm T^{mp} Motif enrichment Cl13 Tex^{prog}

T^{mp}: ETS, RUNX, ETS
Tex^{prog}: NFAT, BATF, FOS, JUN

x-axis: Rank Sorted TFs Enriched
y-axis: $-\log_{10}$(P-adj) Motif Enrichment

Norm.Enrich. $-\log_{10}$(P-adj)
Min — Max

E Tex^{KLR} program (97 genes) Tex^{term} program (340 genes)

$Klrd1$, $Klrg1$, $Klrc1$, $Klre1$, $Klrk1$, $Zeb2$, $Hopx$, $Kcnj8$, $Arl4c$, $Klrb1c$, $Ifitm10$, $Klf3$, $Cd8b1$
$Sh2d2a$, $Gapdh$, $Trbc2$, $Ctla4$, $Tnfrsf9$, $Tox2$, Tox, $S100a6$, $Cxcr3$, $Cd27$, $Pdcd1$, $Lag3$, $Cdk6$, $Tapbpl$

x-axis: Avg. Log$_2$ Fold Change
y-axis: $-\log_{10}$(p_val_adj)

F

Cell type annotation
Tex^{prog} Tex^{int}
Tex^{eeff} Tex^{KLR}
Tex^{eeff2} Tex^{term}

Tox annotated OCRs
Upregulated OCRs (vs. Naive) n=88

Intersection Size: 17, 16, 8, 8, 8, 6, 4, 4, 3, 2, 2, 2, 2, 2, 1, 1, 1, 1

Peak set vs. Tnaive: 0
Tex^{eeff2}, Tex^{KLR}, Tex^{term}, Tex^{int}, Tex^{prog}, Tex^{eeff}

x-axis: Peak set/population

Tox

y-axis: scRNA (log$_2$ exp.)
x-axis: Tex^{prog}, Tex^{eeff2}, Tex^{eeff}, Tex^{int}, Tex^{KLR}, Tex^{term}

Figure S3. Early progenitor exhausted T cells possess the molecular program of exhaustion. Related to Figure 3.

(A) UMAPs depict scATAC-seq (left) and scRNA-seq (right) results from the D8 Arm and Cl13 infections. Cells on the small UMAPs are colored by their origin from the two infection models (bottom). **(B)** Volcano plot of differentially expressed genes between the memory precursor T cells (T^{mp}) of the Arm and the progenitor exhausted T cells (Tex^{prog}) of the Cl13 infection model. **(C)** Ingenuity pathway analyses of the T^{mp} and Tex^{prog} specific gene sets. Top 8 enriched biological terms are shown. **(D)** Volcano plot depicts the differential open chromatin regions (OCRs) of the T^{mp} and Tex^{prog} populations (left). Hockey stick plots show the enriched transcription factor motifs at the specific OCR sets of the T^{mp} and Tex^{prog} subsets. **(E)** Volcano plot of the differentially expressed genes between the Tex^{KLR} and Tex^{term} subsets. **(F)** Upset plot of differentially accessible OCRs annotated to the *Tox* gene relative to T^{naive} cells and their overlap among the different Tex cell subsets. Violin plot shows the gene expression level of *Tox* in the identified Tex subsets.

Figure S4.

A

Acute Day 8
Highly expanded clone 1

Acute Day 8
Highly expanded clone 2

Acute Day 21
Highly expanded clone 1

Acute Day 21
Highly expanded clone 2

Other
Other Acute Day 8
TRA:CAMREGLYGSSGNKLIF
TRB:CASGDGDWQDTQYF
321 cells

Other
Other Acute Day 8
TRA:CAANYGNEKITF
TRB:CASSDSGAGAEQFF
242 cells

Other
Other Acute Day 21
TRA:CAAFNTGNYKYVF
TRB:CASSDHSNTEVFF
354 cells

Other
Other Acute Day 21
TRA:CATDLNNAGAKLTF
TRB:CASSFDRGLTEVFF
267 cells

B

Chronic Day 8
Highly expanded clone 1

Chronic Day 8
Highly expanded clone 2

Chronic Day 21
Highly expanded clone 1

Chronic Day 21
Highly expanded clone 2

Other
Other Chronic Day 8
TRA:CAMRVVDSNYQLIW
TRB:CASSLWTNAEQFF
95 cells

Other
Other Chronic Day 8
TRA:CALSANTNKVVF
TRB:CASSLDDWGPSYEQYF
94 cells

Other
Other Chronic Day 21
TRA:CAFEVVGQLTF
TRB:CASSFRDSSYEQYF
2026 cells

Other
Other Chronic Day 21
TRA:CAVSAWVGDNSKLIW
TRB:CASSRDSLTEVFF
780 cells

C

Cell type annotation

T^{naive}
Tex^{prog}
Tex^{int}
Tex^{term}
T^{eff}
T^{em}
Tex^{eeff}
Tex^{KLR}
T^{mem}
Tex^{ISG}

Clone Size

Top 6 T cell clones

D

TCR clones with phenotype
combination in gp33+ fraction

Divergent clones — 34
Tex^{term}-biased clones — 25
— 19
— 15
Tex^{KLR}-biased clones — 14
Divergent clones — 13
Tex^{KLR}-biased clones — 9

observed
shuffled

Tex^{term} Tex^{int} Tex^{prog} Tex^{KLR} $Tex^{eff transit}$

Tex^{term} Tex^{KLR}
Tex^{int} Tex^{eeff}
Tex^{prog}

Figure S4. Highly expanded clones of the Arm and Cl13 infection model describe the dominant clone behaviors of exhausted T cell differentiation. Related to Figure 4.

(A) UMAPs depict highly expanded clones from the Arm infection model at the indicated time points. **(B)** UMAPs depict highly expanded clones of the Cl13 infection model at the indicated time points. **(C)** Stacked bar plot of the phenotypic composition of individual T cell clones with a bias towards the Tex^{KLR} fate, but also exhibiting the Tex^{term} phenotype. Top 6 clones are shown. **(D)** Upset plot of the phenotype combinations of the observed and shuffled TCR clones.

Figure S5.

A — TCR overlap

	Spleen gp33+	Lung gp33+	Liver gp33+	Spleen gp33-	Lung gp33-	Liver gp33-	
	NA	0.69	0.73	0.00	0.01	0.03	Spleen gp33+ (313)
	0.52	NA	0.76	0.00	0.01	0.03	Lung gp33+ (236)
	0.37	0.51	NA	0.00	0.01	0.02	Liver gp33+ (160)
	0.04	0.03	0.07	NA	0.27	0.38	Spleen gp33- (3803)
	0.06	0.08	0.12	0.18	NA	0.47	Lung gp33- (2469)
	0.09	0.12	0.14	0.11	0.20	NA	Liver gp33- (1052)

Correlation: 0 – 0.5

B — Divergent Clone (gp33+: Spleen, Lung, Liver)

Cell type annotation: Tnaive, Texprog, Texint, TexKLR, Texterm, Texlung, TexISG

C

Overlap of expanded clones in gp33+ fractions

Liver 0 — Lung 0 — 6 — 100 — 4 — 37 — Spleen 22
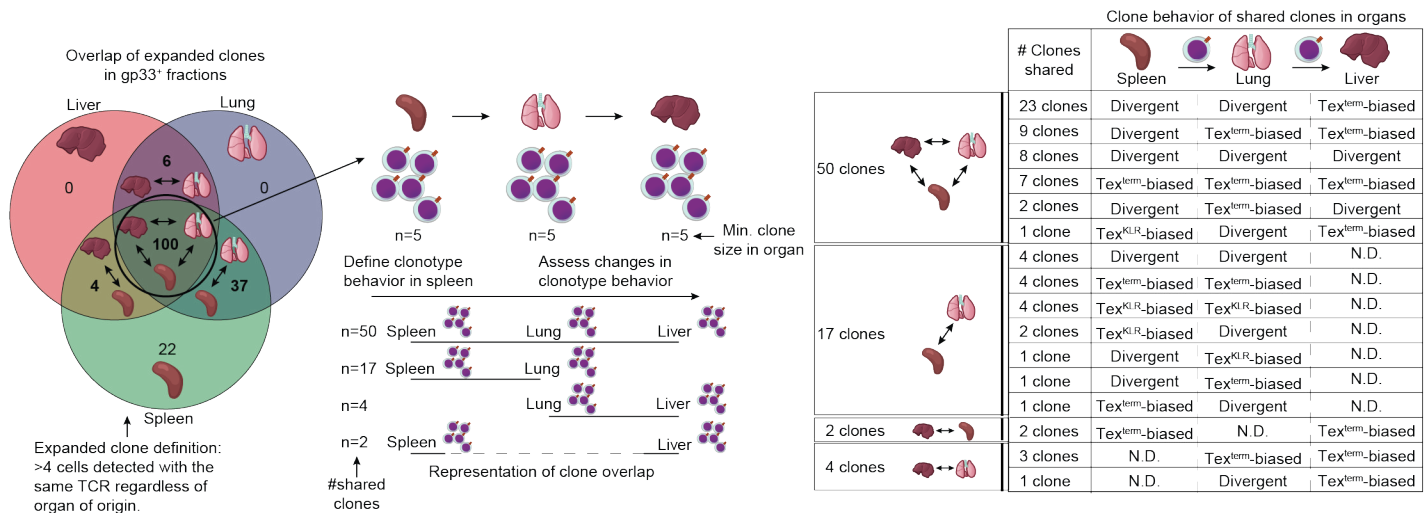
Expanded clone definition: >4 cells detected with the same TCR regardless of organ of origin.

Min. clone size in organ (n=5)

Define clonotype behavior in spleen — Assess changes in clonotype behavior

n=50 Spleen Lung Liver
n=17 Spleen Lung
n=4 Lung Liver
n=2 Spleen Liver

#shared clones — Representation of clone overlap

Clone behavior of shared clones in organs

# Clones shared	Spleen	Spleen → Lung	Lung → Liver
23 clones	Divergent	Divergent	Texterm-biased
9 clones	Divergent	Texterm-biased	Texterm-biased
8 clones	Divergent	Divergent	Divergent
7 clones	Texterm-biased	Texterm-biased	Texterm-biased
2 clones	Divergent	Texterm-biased	Divergent
1 clone	TexKLR-biased	Divergent	Texterm-biased
4 clones	Divergent	Divergent	N.D.
4 clones	Texterm-biased	Texterm-biased	N.D.
4 clones	TexKLR-biased	TexKLR-biased	N.D.
2 clones	TexKLR-biased	Texterm-biased	N.D.
1 clone	Divergent	TexKLR-biased	N.D.
1 clone	Divergent	Texterm-biased	N.D.
1 clone	Texterm-biased	Divergent	N.D.
2 clones	Texterm-biased	N.D.	Texterm-biased
3 clones	N.D.	Texterm-biased	Texterm-biased
1 clone	N.D.	Divergent	Texterm-biased

50 clones (Liver ↔ Lung ↔ Spleen); 17 clones (Liver ↔ Lung); 2 clones (Spleen ↔ Liver); 4 clones (Spleen ↔ Lung)

Figure S5. scRNA/TCR-seq reveals the clone behaviors of different organs. Related to Figure 5.

(A) Heat map representation of the correlation between the TCR repertoires of the indicated gp33$^+$ and gp33$^-$ CD8$^+$ T cell subsets from different organs. **(B)** Stacked bar plot of the phenotypic composition of individual clones across organs. **(C)** Schematics show the definition of an expanded, organ-shared T cell clone for clone behavior analysis. Only those clones were considered that had at least 5 T cells present in each organ. Shared clone numbers across the organs are indicated (left). Table depicting the number of expanded clones that are shared across tissues and their clone behaviors (right).
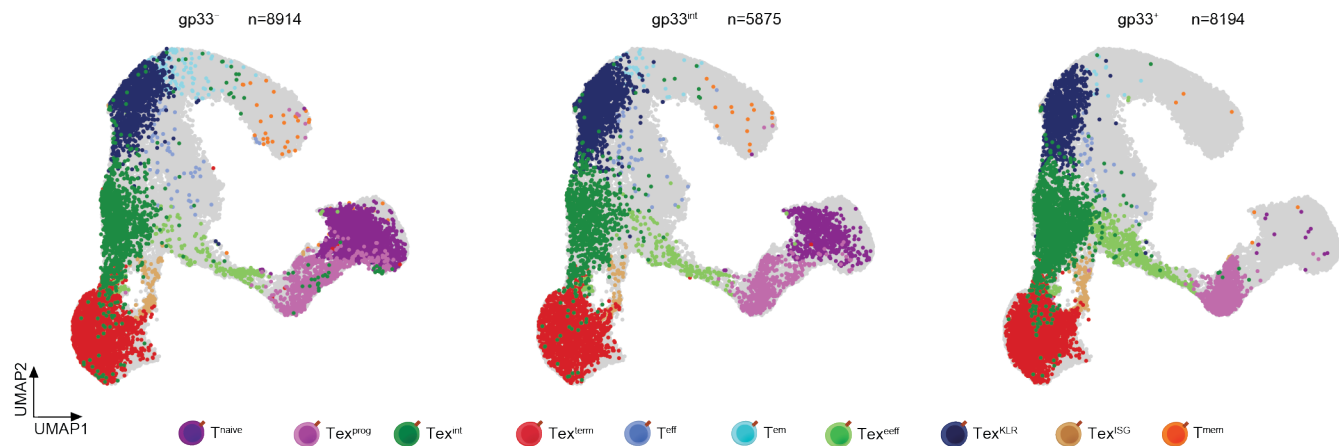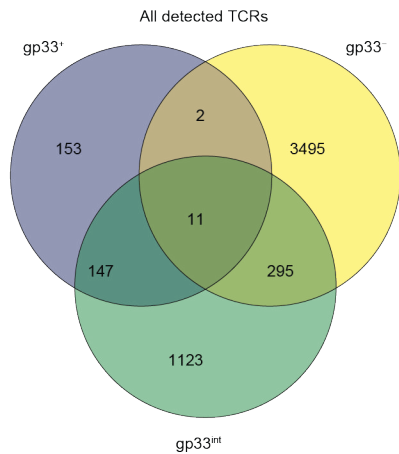
Figure S6.

A

gp33⁻   n=8914   gp33ⁱⁿᵗ   n=5875   gp33⁺   n=8194

UMAP2
UMAP1

● T^{naive} ● Tex^{prog} ● Tex^{int} ● Tex^{term} ● T^{eff} ● T^{em} ● Tex^{eeff} ● Tex^{KLR} ● Tex^{ISG} ● T^{mem}

B

All detected TCRs

gp33⁺

gp33⁻

153

2

3495

11

147

295

1123

gp33ⁱⁿᵗ

Figure S6. scRNA-seq reveals the phenotypic composition of T cell subsets with different affinities to recognize the immunodominant viral epitope. Related to Figure 6.

(A) UMAPs of scRNA-seq results colored by the phenotypic distribution of the three gp33 fractions of T cells. **(B)** Venn diagram shows the overlap of all detected TCR clones among the three gp33 T cell fractions.