Mapping the developmental trajectories of SAP-dependent innate-like gamma delta (γδ) T cells

Somen K Mistri[1], Brianna M. Hilton[1], Katherine J. Horrigan[1], Emma S. Andretta[1], Remi Savard[1], Oliver Dienz[1], Kenneth J Hampel[2], Diana L. Gerrard[2], Joshua T. Rose[2], Nikoletta Sidiropoulos[2], Devdoot Majumdar[1], and Jonathan E. Boyson[*1]

[1]Department of Surgery, Larner College of Medicine, University of Vermont, Burlington, Vermont 05405, USA

[2]Department of Pathology and Laboratory Medicine, Larner College of Medicine, University of Vermont Medical Center, Burlington, Vermont 05405, USA

*Corresponding Author

Abstract

During thymic development, most γδ T cells acquire innate-like characteristics that are critical for their function in tumor surveillance, infectious disease, and tissue repair. The mechanisms, however, that regulate γδ T cell developmental programming remain unclear. Recently, we demonstrated that the SLAM-SAP signaling pathway regulates the development and function of multiple innate-like γδ T cell subsets. Here, we utilized a single-cell proteogenomics approach coupled with γδ V(D)J profiling to map the transcriptional landscape, TCR repertoire, and developmental checkpoints of SAP-dependent mouse γδ T cells. We found that SAP mediates its effects on immature, uncommitted thymic γδ T cells and that it simultaneously promotes progression through the γδT17 developmental pathway while also inhibiting the diversion of γδ T cells into the αβ T cell pathway. The redirection of γδ T cell clonotypes into the αβ T cell pathway was associated with a decreased frequency of these mature clonotypes in neonatal thymus, and an altered γδT17 TCR repertoire in the periphery. Finally, we identify TRGV4/TRAV13-4(DV7)-expressing T cells as a novel, SAP-dependent γδT1 Vγ4 T cell subset. Altogether, these data suggest that SLAM/SAP signaling acts during the very early stages of γδ T cell development, where it regulates critical pathways involved in the development of specific γδ T effector subsets, and influences the establishment of a unique innate-like γδ TCR repertoire in both the thymus and the periphery.

32 **Introduction**

33 Like αβ T cells, γδ T cells can be broadly grouped into innate-like IFN-γ-producing γδT1,
34 IL-4-producing γδT2, or IL-17-producing γδT17 subsets that are programmed during thymic
35 development[1, 2, 3, 4], or inducible subsets that differentiate in the periphery[5]. The innate-like γδ T
36 cells are unconventional T cells that are highly enriched in mucosal tissues such as the skin, gut,
37 and lung and that play critical roles in the host response to pathogens, in tumor surveillance, and
38 in autoimmunity[6]. Yet, exactly why a developing γδ T cell acquires an innate-like phenotype
39 during thymic development, as opposed to being exported into the periphery in a naïve state
40 remains unclear.

41 A significant body of data supports a model in which weak or strong TCR signal strength
42 might contribute to the development of innate-like γδT17 (weak signal) and γδT1 (strong signal)
43 subsets[4, 7, 8, 9, 10], and a strong TCR signal has been suggested to be necessary for the development
44 of the IFN-γ and IL-4 producing γδNKT cell subset[11]. However, there is also evidence of a TCR-
45 independent model in which pre-existing hard-wired transcriptional programs regulate the
46 development of some innate-like γδT17 subsets[12], and accumulating data suggest that γδT17
47 programming is associated with the use of distinct signaling pathways[13, 14, 15].

48 SLAM family receptors comprise a family of nine receptors, SLAMF1 (SLAM; CD150),
49 SLAMF2 (CD48), SLAMF3 (Ly9; CD229), SLAMF4 (2B4; CD244), SLAMF5 (CD84),
50 SLAMF6 (Ly108; CD352), SLAMF7 (CRACC; CD319), SLAMF8 (BLAME), and SLAMF9
51 (SF2001; CD84H), that are expressed primarily on hematopoietic cells and which play diverse
52 roles in immune development and function[16]. Most SLAM family receptors interact in a
53 homophilic manner, the exception to this rule being SLAMF4 which serves as the ligand for
54 SLAMF2. SLAM family receptor signals can be both activating or inhibitory, depending on the

55 recruitment of the cytosolic signaling adapter proteins SAP and EAT-2, and inhibitory SHP-1,
56 SHP-2, or SHIP phosphatases to immunoreceptor tyrosine switch motifs in the SLAM family
57 receptor cytoplasmic domain[16]. Although SLAM/SAP signaling has long been known to play a
58 role in innate-like NKT[17] and γδNKT[11] cell development, recent findings from our group and
59 others demonstrating that this pathway is also involved in the development of innate-like γδT1,
60 γδT17[18] and innate-like MAIT cells[19, 20] indicates that SLAM/SAP signaling plays a fundamental
61 role in the development of a majority of innate-like T cell lineages.

62 In this study, we set out to define the stage of thymic development during which SAP
63 mediates its effect on γδ T cell developmental programming and to define SAP-dependent changes
64 in the γδ T cells transcriptional program during their development. Finally, since SAP-dependent
65 innate-like T cell subsets are often characterized by highly restricted TCRs (e.g., $V\gamma1^+/V\delta6.3^+$
66 γδNKT[21]), we wished to identify whether specific TCR clonotypes were also associated with SAP-
67 dependent γδT1 and γδT17 T cells. Therefore, we utilized a single-cell proteogenomics approach
68 with V(D)J profiling that allowed us to simultaneously compare cell surface protein expression,
69 gene expression, and TCR clonotypes between B6 and B6.SAP$^{-/-}$ γδ T cells from embryonic,
70 neonatal, and adult thymus, as well as from peripheral lung γδ T cells. Collectively, these data
71 provide a comprehensive map of the SAP-dependent γδ T cell subsets and their developmental
72 checkpoints in perinatal and adult thymus.

73 **Results**

74 *SLAM family receptor expression marks transcriptionally distinct developmental stages among*
75 *E17 γδ T cells*

76 γδT17 developmental programming is thought to occur primarily during
77 embryonic/neonatal thymic development[2, 12, 22], and is tightly linked to the expression of discrete
78 TCR gamma and delta pairings[23]. To identify the specific developmental stages of thymic
79 development during which SLAM/SAP signaling pathway exerts its effects, we employed a single-
80 cell RNAseq approach coupled with feature barcoding (i.e., CITEseq[24]) and V(D)J profiling to
81 compare embryonic day 17 (E17) B6 and B6.SAP$^{-/-}$ thymic γδ T cells **(Fig. 1A**). FACS-sorted
82 thymic γδ T cells (**Suppl. Fig. 1**) from individual embryos (n = 4 B6 and 4 B6.SAP$^{-/-}$ mice) were
83 identified using hashtags, and Vγ1^{+} and Vγ4^{+} γδ T cells were identified using barcoded antibodies
84 (ADTs). To help define the developmental trajectories of the E17 γδ T cells, we also stained cells
85 with barcoded antibodies specific for CD24, CD73, CD44, and CD45RB which define distinct
86 developmental pathways of γδT17 and γδT1 cells in the thymus[4, 25, 26], as well as for SLAMF1 and
87 SLAMF6. After data filtering and QC, we verified that clusters were evenly distributed among the
88 biological replicates (**Suppl. Fig. 1**). The resulting dataset represented a total of 5278 thymic γδ T
89 cells from B6 (n = 2597) and B6.SAP$^{-/-}$ (n = 2681) γδ T cells. Data were visualized using uniform
90 manifold approximation and projection (UMAP) for dimensionality reduction and a graph-based
91 clustering approach[27].

92 First, we annotated the B6 dataset using a combination of cell surface protein expression,
93 gene expression, and TCR clonotype sequence. This analysis revealed 11 B6 E17 γδ T cell clusters,
94 mostly distributed between two major branches (**Fig. 1B**). We observed the presence of
95 CD24^{low}73^{low}44^{high}45RB^{low}SLAMF1^{pos} (c8) and CD24^{low}73^{high}44^{high}45RB^{high}SLAMF6$^{low/pos}$ (c0,

96 c9) clusters, consistent with mature γδT17 phenotype and γδT1 phenotypes, respectively (**Fig. 1C,**
97 **Suppl. Fig. 1).** In support of this annotation, the c8 cluster was enriched in *Il17a*, *Il17f, Il23r,*
98 *Rorc*, and TRGV6/TRDV4 (IMGT nomenclature; encoding Vγ6/Vδ1) transcripts with limited
99 CDR3 diversity, while the c0/c9 cluster was enriched in *Ifng*, *Tbx21, Klrd1*, *Il2rb, Xcl1,* and
100 TRGV5/TRDV4 (encoding Vγ5/Vδ1) transcripts also with little to no diversity (**Figs. 1D – 1F,**
101 **Suppl. Fig. 2, Suppl. Tables 1 and 2**)[28, 29]. Consistent with previous observations[18], we noted that
102 the c0/c9 cluster was highly enriched in *Cd244a* (*Slamf4*), and was also enriched in transcripts
103 from another SLAM family member, *Slamf7* (**Figs. 1D and 1E**). Flow cytometric analysis
104 confirmed that SLAMF7 was expressed primarily on mature γδ T cells and that its expression was
105 mostly confined to γδ T cells expressing the canonical $CD44^{+}CD45RB^{+}$ markers of γδT1 cells
106 (**Suppl. Fig. 1**)[4, 30].

107 We identified a small $CD24^{low/int}73^{high}44^{high}45RB^{pos}$ c3 cluster characterized by enrichment
108 in *Cd200* (whose expression is upregulated upon TCR ligation[31]), *Il2rb, Nr4a1, Cd244a*, *Xcl1,* and
109 *Tnfrsf9* (encoding 4-1bb receptor). A high frequency of invariant TRGV5/TRDV4 clonotypes and
110 decreased TCR clonotype diversity (**Suppl. Fig. 2, Suppl. Table 2)** suggested that c3 represented
111 an immature γδT1 population. The $CD24^{low/pos}CD73^{low/pos}44^{low/pos}CD45RB^{low/pos}$ c4 cluster was
112 enriched in *Ms4a4b and Ms4a6b (*encoding CD20 homologues that modulate T cell activation[32,
113 33]), *S1pr1*, *Il6ra, Cd8a, Cd8b1,* and *Ccr9,* suggesting that the c4 cluster represents the previously
114 described emigrating thymic $Ccr9^{+}$ $S1pr1^{+}$ γδ T cells [34] (**Fig. 1C and 1E**). Consistent with the c4
115 cluster representing naïve emigrating thymic γδ T cells, the diversity of the c4 TCR clonotypes
116 was the highest among the $CD24^{low/int}$ (c0/c9, c8, c3) clusters (**Suppl. Fig. 2).**

117 We identified several immature $CD24^{high}CD73^{neg}$ ADT clusters (c1, c2, c7, c5, c6, and c10)
118 that also expressed SLAMF1 and/or SLAMF6 (**Fig. 1B and Suppl. Fig 1**). The c1 and c2 clusters

119 were enriched in γδT17-associated genes such as *Blk, Maf, Sox13, Ccr2, Tcf12, and Rorc*[26, 35, 36,
120 37, 38], as well as *Sh2d1a* (encoding SAP) and *Ccr9* [10] **(Fig. 1D, Suppl. Table 1)**. Within these two
121 clusters, the CD44$^{low/pos}$CD45RB^{low} c2 cluster was especially enriched in genes previously
122 implicated in γδ T cell development such as *Etv5, Tox, Sox4, Tcf7*, *Zbtb16* [38, 39, 40, 41, 42], *Igfbp4* and
123 *Igf1r* which were recently implicated in Th17 differentiation[43], *Tnfsf11* (RANKL) which regulates
124 Aire^{+} mTEC development [44], as well as *Themis* and *Tox2* **(Fig. 1D, Suppl. Fig. 1)**. A high level of
125 *Gzma* expression (**Fig. 1E**) identified c2 as the previously described *Gzma*$^{+}$ γδ T cell subset[34].
126 Interestingly, although enriched in genes associated with a γδT17 signature, the c2 cluster was also
127 enriched in γδT2-associated *Gata3* and *Il4* transcripts (**Fig. 1D, Suppl. Table 1**). The
128 CD44$^{low/pos}$CD45RB^{neg} c1 cluster, in contrast, exhibited significantly higher levels of *Il17a, Il17f,*
129 *Rorc*, *Maf* [36], as well as *Ifngr1,* consistent with the c1 cluster representing immature γδT17 cells
130 (**Fig. 1D – E, 1G, and Suppl. Fig. 1**).

131 Analysis of the c1 and c2 TCR transcripts revealed that the c2 cluster was notably enriched
132 in TRGV4-expressing cells and that 87% of Vγ4 T cells utilized just 3 TRDV chains (50%
133 TRDV5, 24% TRDV2-2, and 12% TRAV13-4/DV7) (**Fig. 1F, Suppl. Table 2**). Clonotype
134 analysis revealed a high frequency of semi-invariant TRGV4/TRDV5 and TRGV4/TRDV2-2
135 clonotypes previously associated with γδT17[45, 46, 47] (**Suppl. Fig. 2, Suppl. Table 2**). In contrast,
136 the c1 cluster possessed comparatively fewer of these TRGV4 clonotypes, and was enriched in
137 TRGV6, TRGV5, and TRGV1 paired with TRDV4 (**Suppl. Fig. 2, Suppl. Table 2)**. Indeed, we
138 noted that the Vγ1 repertoire, which is diverse in adult tissues, was dominated by TRGV1/TRDV4
139 clonotypes (35% of all TRGV1) in E17 thymus, and just two closely related semi-invariant
140 TRGV1/TRDV4 clonotypes comprised 17.4% of all TRGV1. The location of these clonotypes in

141 the immature γδT17-related clusters (**Suppl. Fig. 2)** suggested that they represent the dominant
142 IL-17-producing Vγ1 clonotypes at E17, which we verified using flow cytometry (**Suppl. Fig. 3**).
143 Since IL-4-producing Vγ1/Vδ6.3 γδNKT cell subset develop during fetal thymic
144 development[1], we also investigated whether the *Il4*-enriched γδ T cells in cluster 2 (**Fig. 1D**) were
145 immature Vγ1 γδNKT cells with previously identified canonical Vδ6.3 sequences[21]. Interestingly,
146 while we did identify a small number of Vγ1 T cells with these canonical sequences (encoded by
147 TRAV15-1/DV6-1 or TRAV15N-1), these were not present in the *Il4*-enriched population.
148 The $CD44^{pos}CD45RB^{pos}$ c5 cluster was enriched in *Il2ra*, *Scin, Bex6, Cd5, Slamf6, Cd28,*
149 *Hivep3*, and cell-cycle/metabolism genes, consistent with γδ T cells that have recently developed
150 from DN2/DN3[34, 48, 49]. Within the c5 and c7 clusters, we noted that while there was expression of
151 γδT17-associated genes such as *Blk, Maf, Sox13,* and *Etv5*, there was relatively little *Rorc*
152 expression. *Rorc* expression became apparent in the c2 cluster ($Rorc^{low}$) and was significantly
153 increased in the c1 cluster (**Suppl. Fig. 1)**. Flow cytometric analysis confirmed BLK and PLZF
154 expression in $CD25^{high}$ E17 Vγ4 T cells, while RORγt expression was primarily observed in the
155 $CD25^{neg}$ population, suggesting that BLK is expressed soon after γδ T cell selection, followed by
156 RORγt (**Suppl. Fig. 3)**. These data were consistent with MAF-directed expression of *Rorc*[34, 36]
157 and suggested a c5 – c7 – c2 – c10 / c1 – c8 γδT17 developmental trajectory, which was supported
158 by trajectory inference analysis **(Fig. 1H).**
159 Finally, analysis of the $CD24^{high}73^{neg}44^{neg}45RB^{int}SLAMF1^{pos}SLAMF6^{pos}$ c6 population
160 revealed that it was highly enriched in *Cd8a, Cd8b1, Cd4, Rag1, Rag2, Ptcra, Arpp21, and*
161 *Dgkeos,* consistent with cells developing along the αβ T cell lineage **(Figs. 1E, 1G, Suppl. Fig. 1,**
162 **and Suppl. Table 1**)[50]. Analysis of the Vγ1 and Vγ4 ADTs confirmed that c6 cells were not simply
163 contaminating αβ T cells as both Vγ1- and Vγ4-expressing cells were present in the c6 cluster

164 **(Suppl. Fig. 3**) and analysis of the c6 TCR clonotypes revealed the presence of TRGV1, GV4,
165 GV5, GV6, and GV7 clonotypes, none of which appeared to be specific to the c6 cluster (**Suppl.**
166 **Table 2)**.

167 Since trajectory inference analysis suggested a strong relationship between the *Il2ra*$^{+}$ c5
168 and c6 clusters (**Fig. 1H**), we conducted differential gene expression analysis between the c5 and
169 c6 from B6 mice. Interestingly, we found that the c6 cluster was highly enriched in *Rorc,*
170 suggestive of a γδT17 phenotype, but was mostly lacking in the expression of *Maf, Blk, Ccr2*, and
171 other canonical γδT17-associated transcripts (**Fig. 1E, 1G, and Suppl. Fig. 1**). Indeed, flow
172 cytometric analysis of E17 TCRβ^{neg}TCRδ^{pos} γδ T cells confirmed that CD4CD8 DP γδ T cells
173 exhibited an immature CD44^{low}CD25^{low}MAF^{neg}BLK^{neg}RORγt^{pos} phenotype (**Suppl. Fig. 3**), and
174 that RORγt expression in the absence of BLK or MAF is a characteristic of Lin^{neg} DN thymocytes
175 as they progress from CD44^{low}25^{high} DN3 to CD44^{low}25^{low} DN4 (**Suppl. Fig. 3**). Together, these
176 data suggested that the c6 cluster γδ T cells represented previously described CD4CD8 DP γδ T
177 cells that were developing along an alternative αβ T cell pathway[51, 52, 53].

178

179 *SAP regulates γδ T cell clonotype diversion into both the γδT17 and the αβ T cell developmental*
180 *pathways.*

181 To identify SAP-dependent developmental checkpoints, we next compared the hashtag-
182 separated B6 and B6.SAP$^{-/-}$ E17 γδ T cell datasets. A comparison of the cluster frequencies
183 between B6 and B6.SAP$^{-/-}$ γδ T cells revealed significantly decreased frequencies of the
184 CD24^{high}73^{neg}44^{int}45RB^{int}SLAMF1^{pos}SLAMF6^{pos} c2 and c1 clusters in B6.SAP$^{-/-}$ mice, and a
185 striking increase in the frequency of the CD24^{high}73^{neg}44^{neg}45RB^{neg}SLAMF1^{high}SLAMF6^{high} c6
186 cluster (**Fig. 2A)**. A comparison of differentially expressed genes among the immature (c1, c2, c7,

187 c5, c6, and c10) clusters using pseudobulk analysis revealed 234 differentially expressed genes
188 with a p_{adj} threshold < 0.005 and log_2 fold change greater than 0.5, with 148 genes exhibiting
189 decreased expression in B6.SAP$^{-/-}$ γδ T cells, and 86 genes exhibiting increased expression (**Fig.**
190 **2B, Suppl. Table 3**). Consistent with the decreased frequencies of the c1/c2 clusters in B6.SAP$^{-/-}$
191 mice, a significant fraction of the differentially expressed genes that were down-regulated in
192 B6.SAP$^{-/-}$ mice reflected γδT17-associated genes (e.g., *Maf, Blk, Ccr2, Sox13, Etv5, etc.*) that were
193 present in the c2 and c1 clusters, as well as a number of c1/c2-enriched genes (e.g., *Cpa3, Tox2,*
194 *Pdcd1, Tmem121, Bex6*, and *Scin*) whose importance in γδ T cell development is unclear (**Fig. 2B-**
195 **C, Suppl. Table 2**). This was not solely due to the decreased frequency of the c2 and c1 clusters
196 in B6.SAP$^{-/-}$ mice, as a closer analysis revealed a decreased number of transcripts for many
197 signature genes (e.g., *Blk*, *Etv5*, *Ccr2*) prevalent in the c2 cluster and to a lesser extent the c7
198 cluster **(Suppl. Fig. 4**). In contrast, a large number of genes whose expression was significantly
199 increased in B6.SAP$^{-/-}$ mice were those whose expression was mostly confined to the c6 cluster
200 (e.g., *Cd4, Cd8b1, Ptcra, Rag2*) or whose expression was highest in c6 (e.g., *Tcf7, Thy1*)(**Fig. 2B-**
201 **C, Suppl. Fig. 4)**. Together, these data identified the *Gzma*$^+$*Etv5*$^+$*Sox13*$^+$*Blk*$^+$ c2 and
202 *Cd4*$^+$*Cd8b1*$^+$*Ptcra*$^+$*Rag2*$^+$ c6 clusters as SAP-dependent developmental checkpoints.

203 Next, we asked whether these SAP-dependent alterations in γδ T cell development were
204 associated with any changes in the TCR repertoire. A comparison of all γδ TCR clonotypes
205 between E17 B6 and B6.SAP$^{-/-}$ γδ T cells revealed no significant changes in overall clonotype
206 frequency (**Fig. 2D**). However, a closer examination of TCR repertoire revealed significant
207 changes in the distribution of specific TCR clonotypes. For example, within the TRGV4 repertoire,
208 which represents the dominant TCR among immature clusters at E17, we observed a significant
209 decrease in TRGV4/TRDV5 clonotypes in the c2 cluster and a concomitant increase of these

210 clonotypes in the c6 cluster (**Fig. 2E**). Among the TRGV4 clonotypes, this appeared to be specific
211 to the TRDV5 clonotypes, as we observed no significant differences c2 and c6 frequencies among
212 TRGV4/TRDV2-2 and TRDV7 clonotypes (**Fig. 2E**). We did note, however, a significant increase
213 in the TRGV4/TRDV7 clonotypes in the *Il2ra*$^{+}$ c5 cluster (**Fig. 2E**). Within the TRGV1 repertoire,
214 we noted a significant increase in TRGV1/TRDV4 clonotypes within the c6 cluster (**Fig. 2E**).
215 Together, these data suggested that SAP operated at an early stage of γδ T cell development during
216 or soon after the down-regulation of *Il2ra,* and that it positively regulated entry into the γδT17
217 pathway at the c5/c7 to c2 transition, while at the same time inhibiting entry into the CD4^{+}CD8^{+}
218 αβ T cell-like c6 cluster. The net result of these changes appeared to be that specific TCR
219 clonotypes such as TRGV4/TRDV5 were diverted from the γδT17 pathway into the αβ T cell-like
220 pathway.
221
222 *SAP deficiency is associated with increased numbers of immature CD4^{+}CD8^{+}RORγt^{+} thymic γδ*
223 *T cells*
224 Next, we independently evaluated these observations using flow cytometry. This analysis
225 revealed that while SAP deficiency did not affect the overall number of E17 γδ T cells (**Suppl.**
226 **Fig. 5**), it did result in a significant decrease of immature CD24^{pos}BLK^{pos}MAF^{pos}RORγt^{pos} γδ T
227 cells in E17 B6.SAP$^{-/-}$ thymus (**Fig. 3A-B**). We found that this decrease was most pronounced in
228 the CD44^{pos} population after downregulation of CD25 (**Fig 3A, Suppl. Fig. 5**) and that this
229 population exhibited a high level of PLZF expression, suggesting it corresponded to the
230 CD44^{+}*Blk*$^{+}$*Maf*$^{+}$*Zbtb16*$^{+}$*Rorc*$^{+}$ c1 cluster (**Suppl. Fig. 5**). In addition, UMAP analysis of E17 Vγ4
231 T cells flow cytometric data revealed the presence of a SAP-dependent
232 CD24^{pos}BLK^{pos}PLZF^{pos}RORγt^{neg} population in B6.SAP$^{-/-}$ thymus **(Suppl. Fig. 5**) consistent with

233 the phenotype of the c2 cluster which expressed lower levels of *Rorc*. In contrast, we found no
234 evidence of a SAP-dependent decrease in the CD25^{pos} population frequency, consistent with the
235 notion that SAP exerts its effect during or after this stage of development (**Fig. 3A**). We also noted
236 that in addition to its effect on the numbers of immature thymic γδT17, SAP deficiency resulted
237 in the reduced expression levels of some of these markers, most notably BLK (**Fig. 3C**), a finding
238 that was consistent with the decreased number of *Blk* transcripts observed in Fig. 2. Interestingly,
239 as PLZF induction is often associated with SLAM/SAP signaling, we noted that SAP did not affect
240 overall PLZF expression levels in E17 thymic γδ T cells (**Suppl. Fig. 5**).

241 In contrast to the decreased numbers of CD24^{pos}BLK^{pos}MAF^{pos}PLZF^{pos} RORγt^{pos} γδ T
242 cells, we observed a significant increase in the frequency and number of
243 CD24^{pos}BLK^{neg}MAF^{neg}PLZF^{neg}RORγt^{pos} γδ T cells (**Figs. 3A-B**), whose phenotype was consistent
244 with the αβ T cell-like c6 cluster in Fig. 2. Consistent with our finding that the c6 cluster was
245 enriched in CD4 and CD8, we found that the frequency and number of CD4 and CD8-expressing
246 CD24^{pos} TCRδ^{pos}TCRβ^{neg} BLK^{neg}MAF^{neg}RORγt^{pos} thymic γδ T cells was significantly increased
247 in B6.SAP$^{-/-}$ mice (**Fig. 3D**). Together, these analyses confirmed our single-cell CITEseq analysis,
248 and suggested that SAP regulated both the throughput of immature γδ T cells in the γδT17 cell
249 pathway as well as the αβ T cell-like pathway. More specifically, the data suggested that SAP
250 promoted progression to the CD24^{pos} BLK^{pos}MAF^{pos}PLZF^{pos}RORγt^{low} c2 cluster that is enriched
251 in a number of genes (e.g., *Sox13, Etv5, Blk, Maf)* known to regulate γδT17 development, but that
252 it inhibited entry into the CD24^{pos}44^{neg}BLK^{neg}MAF^{neg}RORγt^{pos} αβ–T cell-like pathway (c6
253 cluster).

254 *Identification of SAP-dependent developmental checkpoints during neonatal and adult thymus γδ*
255 *T cell developmental programming.*

256 The neonatal period is a time of rapid expansion of numerous γδ T cell subsets in the
257 thymus including Vγ4 γδT17 and γδNKT. Therefore, we conducted a comparison between B6 and
258 B6.SAP$^{-/-}$ neonatal as well as adult thymic γδ T cells to both identify novel SAP-dependent
259 developmental checkpoints and to track the outcome of the SAP-dependent alterations in γδ T cell
260 development observed in E17 thymus. Neonate datasets revealed a similar branched clustering
261 pattern as in E17 with some notable differences. Similar to E17, we observed immature
262 CD24^{pos}73^{low}SLAMF1^{pos}SLAMF6^{pos} clusters (c0, c1, c2, c3, c9, c12, c14),
263 CD24^{pos}73$^{low/pos}$SLAMF1^{pos}SLAMF6^{pos} (c4, c5, c10, c15) *S1pr1*-enriched "naïve" clusters,
264 mature CD24^{neg}73^{low}SLAMF1^{pos}SLAMF6^{neg} (c6, c7) γδT17 clusters, and a mature
265 CD24^{neg}73^{pos}SLAMF1^{neg}SLAMF6$^{low/pos}$*Slamf7*^{high} (c8) γδT1 cluster (**Fig. 4A-B**, **Suppl. Fig. 6)**. In
266 general, we noted good correlation with cluster-specific gene expression between the E17 and D9
267 (Day 9) datasets. A notable exception was a striking decrease in *Zbtb16* expression among all
268 immature neonatal γδ T cells, which is consistent with previous reports[10, 34] (**Suppl. Fig. 6, Suppl.**
269 **Table 3**).

270 In contrast to the E17 dataset, the neonatal thymic γδ T cell dataset contained a
271 CD24^{neg}73^{pos}SLAMF1^{neg}SLAMF6^{high}*Slamf7*^{neg} (c13) γδT2 cluster enriched in *Il4, Il13, Zbtb16,*
272 *Cd40lig*, and *Nos1* (**Suppl. Fig. 6, Suppl. Table 1),** suggesting it represented IL-4-producing
273 γδNKT cells. Indeed, examination of the γδT2 (c13) and γδT1 (c8) cluster TCR repertoires
274 revealed a significant enrichment of canonical TRGV1 and TRAV15N-1 or TRAV15-1/DV6-1
275 sequences characteristic of γδNKT cells (**Fig. 4C, Fig. 5**)[21]. Analysis of the differentially
276 expressed genes in the γδT1 (c8) and γδT2 (c13) clusters revealed that *Zbtb16*, *Icos,* and *Slamf6*

277 were markers of γδT2 cells while *Tbx21*, *Klrb1c* (NK1.1), and *Slamf7* were markers of γδT1 cells
278 in neonatal thymus (**Suppl. Fig. 6**). Flow cytometric analysis confirmed that PLZF^{hi} γδT2
279 Vγ1Vδ6.3 cells co-express SLAMF6 and ICOS while Tbet^{hi} γδT1 Vγ1Vδ6.3 cells are NK1.1^{+} and
280 are either SLAMF6^{-}SLAMF7^{+} or SLAMF6^{+}SLAMF7^{+} **(Suppl. Fig. 7)**.

281 A comparison of B6 and B6.SAP$^{-/-}$ neonatal γδ T cells revealed multiple SAP-dependent
282 changes in clustering. Among the immature γδ T cells, we observed a decreased frequency of the
283 immature CD24^{high}73^{neg}SLAMF1^{pos}SLAMF6^{pos} c1 cluster (**Fig. 4D**) that was enriched in *Blk, Etv5,*
284 *Sox13, Rorc,* and *Maf* (**Suppl. Fig. 6**), suggesting it was analogous to the SAP-dependent c2
285 cluster in E17 γδ T cells. Similar to our observations in E17 γδ T cells, flow cytometric analysis
286 confirmed both a decrease in BLK^{pos}RORγt^{pos} and BLK^{pos}RORγt^{neg} γδ T cell populations, as well
287 as a decreased level of BLK expression (**Fig. 4E-F**). In addition, we observed an increased
288 frequency of immature CD24^{high}73^{neg}SLAMF1^{pos}SLAMF6^{pos} cluster with characteristics of αβ T
289 cells (c14) in B6.SAP$^{-/-}$ D9 thymic γδ T cells **(Fig. 4D, Suppl. Fig. 6**). While this increase did not
290 reach the level of significance, presumably due to the very low number of c14 cells in the dataset,
291 flow cytometric analysis revealed a significant increase in immature CD24^{pos}73^{neg}BLK^{neg}RORγt^{pos}
292 γδ T cells in D9 thymus (**Fig. 4E**). Last, we observed an increased frequency of immature
293 CD24^{pos}73$^{low/pos}$ SLAMF1^{pos}SLAMF6^{pos} *S1pr1*-enriched c4 and c5 clusters, but not the closely
294 related *Cd200*-enriched c10 cluster (**Fig. 4D).** While we did confirm an increased number
295 CD200^{neg}S1pr1^{pos} γδ T cells in neonatal thymus by flow cytometry, this increase was relatively
296 small and its significance is unclear **(Suppl. Fig. 7).** Independent confirmation of these SAP-
297 dependent changes in neonatal γδ T cell gene expression was obtained by conducting bulk
298 RNAseq analysis of FACSs-sorted CD24^{high} neonatal thymic γδ T cells from B6 and B6.SAP$^{-/-}$
299 mice (**Suppl. Table 4**).

300 Among the mature neonatal γδ T cells, we observed a decreased frequency of the
301 CD24^{low}73^{neg}SLAMF1^{pos}F6^{low} γδT17 c6 cluster which was consistent with our previous
302 observation of decreased IL-17 production in B6.SAP$^{-/-}$ neonatal γδ T cells[18]. Interestingly, we
303 observed a striking loss of the CD24^{low}73^{high}SLAMF1^{low}F6^{high}F7^{neg} γδT2 c13 cluster in B6.SAP$^{-/-}$
304 mice, as well as a decreased frequency of the mature CD24^{low}73^{high}SLAMF1^{low}F6^{low}F7^{high} γδT1
305 c8 cluster (**Fig. 4C**). These findings were consistent with a SAP-dependent loss of both the IL-4-
306 and IFN-γ-producing γδNKT as well as non-γδNKT γδT1 cells (discussed further below).
307 Unfortunately, while we were able to define mature SAP-dependent γδNKT subsets, we found no
308 evidence of immature γδNKT cells in our dataset, suggesting that they develop during a brief
309 developmental window between E17 and D9 after birth.

310 A comparison of B6 adult thymic γδ T cells between B6 and B6.SAP$^{-/-}$ mice recapitulated
311 many of our observations of the neonatal thymus, although to a somewhat lesser degree. For
312 example, we observed a small but significant decrease in the frequency of the CD24^{high}CD73^{low}
313 adult thymic c1 and c2 clusters in B6.SAP$^{-/-}$ mice that was coupled with an increased frequency of
314 the CD24^{pos}CD73$^{+/-}$ c0 cluster **(Suppl. Fig. 8)**. Like the SAP-dependent E17 c2 and D9 c1 clusters,
315 the adult thymus c1 and c2 clusters were enriched in γδT17-associated *Blk, Etv5, Sox13*, *Maf,* and
316 *Rorc* (mostly in c2) indicating that SAP regulates this γδ T cell thymic developmental checkpoint
317 at all ages. Likewise, the adult thymus c0 cluster that was increased in B6.SAP$^{-/-}$ mice was enriched
318 in *Ms4a4b, Ms4a6b, Dgka, Klf2,* and *S1pr1* similar to the neonatal c4 and c5 clusters whose
319 frequencies were increased in B6.SAP$^{-/-}$ γδ T cells. Just as we observed in neonatal thymic γδ T
320 cells, we noted a near complete loss of mature γδNKT cells (*Zbtb16, Icos*, and *Il4*-enriched γδT2
321 c8 cells and *Tbx21, Slamf7*, and *Il2rb*-enriched γδT1 c11 cells) as well as a decreased frequency
322 of mature non-γδNKT γδT1 c10 cluster cells in B6.SAP$^{-/-}$ adult thymus (**Suppl. Fig. 8**). Unlike the

323 E17 and neonatal thymic datasets, however, the frequency of γδ T cells with an αβ T cell-like
324 signature in the adult thymic γδ T cell dataset was minimal. Collectively, these data supported the
325 presence of multiple SAP-dependent developmental checkpoints during neonatal and adult γδ T
326 cell development—a reduction in the $CD24^{pos}CD73^{neg}$ $SLAMF1^{+}SLAMF6^{+}$ cluster enriched in
327 γδT17 development-associated transcripts, an increased frequency of immature *S1pr1*-enriched γδ
328 T cells, and decreased frequencies of both γδNKT cells (γδT1 and γδT2) as well as non-γδNKT
329 γδT1 cells.
330
331 *SAP shapes the thymic γδ TCR repertoire*
332 Given our observation of altered γδ TCR clonotype distribution among E17 γδ T cells and
333 the substantial changes in clustering we observed in SAP-deficient mice, we were next interested
334 in determining the effect of SAP deficiency on the neonatal and adult γδ TCR repertoires. A
335 comparison of frequencies of the top 100 TCR clonotypes between B6 and B6.$SAP^{-/-}$ mice revealed
336 significant changes in the overall frequency of several TRGV1 and TRGV4 clonotypes (**Fig. 5A**).
337 Consistent with a SAP-dependent loss of γδNKT, we observed a profound loss of TRGV1
338 clonotypes utilizing the canonical TRAV15N-1 and TRAV15-1/TRDV6-1 γδNKT cell TCRs (**Fig.**
339 **5B**). Loss of these clonotypes was directly tied to the significant decrease of the γδT1 c8 and γδT2
340 c13 clusters, in which nearly all these clonotypes resided.
341 Within the B6 γδT2 c13 cluster the vast majority (~86%) of TCR sequences were highly
342 restricted canonical γδNKT cell clonotypes **(Suppl. Table 2**) and nearly all of these clonotypes
343 were also present in the γδT1 c8 cluster consistent with the ability of γδNKT cells to produce both
344 IFN-γ and IL-4. In contrast, while the B6 γδT1 c8 cluster contained a large proportion of canonical
345 γδNKT cell clonotypes (~46% of c8 TCRs), it also contained TCRs with a diverse pool of

346 clonotypes that utilized TRGV1 (~70% of c8 TCRs), TRGV4, TRGV5, TRGV6, and TRGV7 and
347 were present only in the c8 γδT1 cluster. Therefore, in neonates, the thymic γδT1 population is a
348 mixture of γδNKT and "non-γδNKT" γδT1 clonotypes, and each of these clonotype pools is
349 diminished in B6.SAP$^{-/-}$ mice. As noted above, these findings were largely reproduced in the adult
350 thymus where the γδNKT clonotypes associated with the c8 and c11 clusters were virtually absent
351 in B6.SAP$^{-/-}$ mice, and where the diverse pool of γδT1 c10 clonotypes was significantly diminished
352 (**Suppl. Table 2)**.

353 Interestingly, a closer examination of the neonatal thymic TRGV4 clonotypes revealed a
354 significant decrease in the frequencies of γδT17-associated clonotypes with a germline-encoded
355 TRDV5[46] (**Fig. 5B**), some of which were identical to those that were redistributed to the αβ T cell-
356 like c6 cluster in E17 γδ T cells **(Fig. 2E)**. When these clonotypes were mapped to their D9 clusters,
357 we noted a significant decrease in the frequency of TRGV4/TRDV5 clonotypes in the mature
358 γδT17 c6 cluster (**Fig 5C**). We note that while we observed a trend toward an increased frequency
359 of TRGV4/TRDV5 and TRGV4/TRDV2 clonotypes in the *S1pr1*$^{+}$ c4/c5 clusters, analysis of these
360 sequences revealed that they largely exhibited diverse CDR3 rather than the restricted CDR3
361 observed in the γδT17 c6 cluster (**Suppl. Table 2**). Taken together, these data indicate that the
362 TCR clonotypes of immature embryonic day 17 γδ T cells that were redirected to the αβ T cell
363 pathway in B6.SAP$^{-/-}$ mice, were underrepresented among the mature B6.SAP$^{-/-}$ neonatal γδ T
364 cells.

365

366 *Restricted TCR repertoire in innate-like γδT1 as well as γδT17 in the periphery.*

367 We previously demonstrated that SAP-deficiency results in a significant impairment in γδ
368 T cell function in the periphery[18]. Specifically, we observed that γδ T cell IL-17 production is

369 impaired in the lungs of young mice, but that this had largely corrected itself as the mouse reached
370 maturity. In contrast, we observed that both Vγ1 and Vγ4 IFN-γ production were significantly
371 impaired in B6.SAP$^{-/-}$ adult lung and spleen. To better understand how SAP-dependent alterations
372 in thymic development ultimately affected the peripheral γδ T cell compartment, we compared B6
373 and B6.SAP$^{-/-}$ lungs using single-cell RNAseq with V(D)J profiling.

374 Our initial annotation of the B6 lung γδ T cell transcriptome revealed the presence of 11
375 clusters of the lung γδ T cells that we annotated based on cluster-specific gene expression and
376 TRGV and TRDV usage (**Figure 6A**). Based on their expression of signature cytokines and
377 transcription factors, we identified 2 γδT17 clusters (c1 and c4), 4 γδT1 clusters (c2, c3, c7, and
378 c9), 1 γδTNKT-like cluster (c10) with very few cells, as well as a *Cd24*-enriched cluster (c5)(**Fig.
379 6B and 6C, Suppl. Table 5**). Consistent with previous observations, we noted that the c1 and c4
380 γδT17 clusters were enriched in *Slamf1* and that the c2/c3/c7/c9 γδT1 clusters were enriched in
381 *Slamf7* and *Slamf6* (**Fig. 6B**). Indeed, flow cytometric analysis revealed that SLAMF7 expression
382 was restricted to the innate-like lung CD44^{+}CD45RB^{+} γδT1 cell population, and that γδT1 cells
383 are SLAMF1^{-}F6^{+}F7^{+}, while the CD44^{+}CD45RB^{-} γδT17 subsets are SLAMF1^{+}F6^{-}F7^{-} (**Fig. 6D**).
384 Further annotation using V(D)J profiling revealed that the c1 cluster corresponded to Vγ4 γδT17
385 (primarily TRGV4/TRDV5 and TRGV4/TRDV2-2) while the c4 cluster corresponded to lung Vγ6
386 (TRGV6/TRDV4) (**Fig. 6E, Suppl. Table 6**). The c10 cluster was enriched in *Zbtb16*, *Il4*, and
387 *Icos* (**Fig. 6B-C**) as well as TRGV1/TRAV15N-1 and TRGV1/TRDV6 sequences (**Fig. 6E**, **Suppl.
388 Table 6**) consistent with γδNKT cells. We did note, however, that the TRAV15N-1 and TRDV6
389 sequences in the c10 cluster were, for the most part, non-canonical γδNKT cell clonotypes,
390 suggesting a degree of flexibility in the TRAV15N-1 and TRDV6 CDR3 sequences that can be
391 utilized by γδNKT cells.

392 Interestingly, while the γδT1 c2/c3/c7/c9 clusters contained primarily Vγ1 and Vγ4 T cells
393 as expected, we noted that Vγ4 γδT1 cells exhibited a striking preference in the usage of TRAV13-
394 4(DV7), and that TRGV4/TRDV7 pairings accounted for 80% of all Vγ4 clonotypes in the γδT1
395 clusters **(Fig. 6E, Suppl. Table 6**). Analysis of the γδT1 TRAV13-4/DV7 CDR3 sequences
396 revealed a high level of diversity that showed no evidence of clonal expansions (**Fig. 6F**). This
397 was in contrast to the highly restricted Vγ6 repertoire in the c4 γδT17 cluster, as well as the
398 restricted Vγ4 repertoire in the c2 γδT17 cluster that was dominated by semi-invariant TRDV5
399 and TRDV2-2 sequences (**Fig. 6F**). As no Vδ7-specific antibody was available, we independently
400 confirmed this observation using single-cell PCR to assess the lung Vγ4 CDR3δ sequences from
401 13 individual mice. These data demonstrated that, other than TRDV2-2 and TRDV5, TRDV7 is
402 the only other TCR delta chain used at an appreciable frequency by Vγ4 T cells in the B6 mouse
403 lung and that it exhibits significant diversity in its CDR3 (**Suppl. Fig. 9**). Vγ1 γδT1 cells, in
404 contrast, exhibited a more mild pairing preference where 44% of the γδT1 clonotypes utilized
405 TRAV15N-1, TRAV15-1/DV6-1, or TRAV15D-2/DV6D-2 and 13% utilized TRAV13-4/DV7
406 (**Suppl. Table 6**). Together these data formed a comprehensive map of B6 lung γδ T cell
407 transcriptomes and revealed that lung innate-like γδT1, like γδT17 and γδNKT, are all
408 characterized by preferential TCR γ- and δ-chain pairings. This was especially notable for Vγ4 and
409 our data suggested that Vγ4 γδT1 exhibited a distinct preference for the utilization of TRAV13-
410 4(DV7) chains with diverse CDR3.
411
412 *SAP regulates peripheral Vγ4/Vδ7 γδT1 cells and shapes the lung γδT17 repertoire.*
413 Next, we assessed the effect of SAP deficiency on the lung γδ TCR repertoire by comparing
414 B6 and B6.SAP$^{-/-}$ lung γδ T cells (**Fig. 7A**). Consistent with our previous observation that lung γδ

415 T cell IFN-γ production is decreased in SAP-deficient mice[18], we observed a decrease in the
416 frequency of the γδT1 c2, c3, and c7 clusters (**Fig. 7B**). Moreover, we observed a striking decrease
417 in the frequency of TRGV4/TRDV7-expressing cells in the lungs of B6.SAP$^{-/-}$ mice (**Figs. 7C-D,**
418 **Suppl. Table 6),** and we noted that we found no evidence of a compensatory increase in the
419 frequency of other TRDV-expressing lung Vγ4 γδT1 cells in the B6.SAP$^{-/-}$ lung. We independently
420 confirmed this SAP-dependent decrease in TRAV13-4/TRDV7 in both lung and spleen using
421 qPCR (**Suppl. Fig. 10**), and we demonstrated a significant decrease in the number of
422 CD44^{+}SLAMF7^{+} Vγ4 T cells in both the lungs and spleen of B6.SAP$^{-/-}$ mice (**Fig. 7E, Suppl. Fig.**
423 **10**). The decreased lung γδT1 did not appear to be due to a SAP-dependent decrease in homeostatic
424 proliferation, nor to increased apoptosis, since we observed no significant change in lung γδ T cell
425 BrdU incorporation or Annexin V staining in B6.SAP$^{-/-}$ mice (**Suppl. Fig. 10**).

426 In addition to the decrease in lung γδT1, we also noted a loss of the c10 γδT2 cluster
427 consistent with the enrichment in SAP-dependent γδNKT TCRs in this cluster, and we observed
428 an increased frequency of the Vγ4 γδT17 c1 cluster (**Fig. 7B**). Given our previous observation of
429 a SAP-dependent decrease in lung γδ T cell IL-17 production in young, but not adult mice[18], and
430 our findings here of SAP-dependent alterations in γδT17 clonotypes, we were interested in
431 determining whether the adult lung γδT17 repertoire was altered in SAP-deficient mice. Indeed, a
432 comparison of the lung Vγ4 c1 γδT17 CDR3 sequences between B6 and B6.SAP$^{-/-}$ mice revealed
433 significant alterations in the γδT17 c1 cluster TCR repertoire. Specifically, we noted a decreased
434 frequency of TRGV4 chains that exhibited a germline-encoded CDR3γ (CSYGYSS) and an
435 increased frequency of TRGV4 chains with a CDR3γ exhibiting evidence of N/P additions
436 (CSYG(X)YSS). These changes were associated with a significant shift in the usage of TRDV2-2
437 CDR3δ sequence usage (**Fig. 7F**) and an altered c1 TCR clonotype frequency in B6.SAP$^{-/-}$ mice

438 **(Fig. 7G)**. Although the number of germline-encoded invariant TRGV4/TRDV5 clone (clonotype
439 1 in Fig. 6B) was diminished in B6.SAP$^{-/-}$ lung γδ T cells (**Fig. 7G, Suppl. Table 6**), the low
440 numbers of cells available for analysis precluded us from making strong conclusions regarding the
441 loss of any one particular clonotype. Taken together, these data suggest that Vγ4 γδT1
442 preferentially utilize the TRAV13-4(DV7) chain in the lung and that the SAP-dependent decrease
443 in peripheral γδ T cell IFN-γ production is not due to dysregulated cytokine production, but rather
444 due to a specific loss of innate-like γδT1 subsets. This finding, together with our finding that the
445 adult lung γδT17 TCR repertoire is altered in SAP-deficient mice suggests that SLAM/SAP
446 signaling plays a role in shaping the γδ TCR repertoire.

447 **Discussion**

448 A common denominator in most innate-like T cell developmental programming is a
449 dependence on the SLAM/SAP signaling pathway[11, 18, 19, 54]. In the γδ T cell lineage, SAP plays a
450 critical role in the development of IFN-γ and IL-4-producing γδNKT cells[21], as well as in γδT1
451 and γδT17[18], but exactly when during γδ T cell development SAP functions, and how it exerts
452 these effects is unknown. Here, we provide evidence that SAP mediates its effects on immature,
453 uncommitted thymic γδ T cells and that it works to simultaneously promote progression through
454 the γδT17 developmental pathway while inhibiting the diversion of γδ T cells into the αβ T cell
455 pathway. The redirection of γδ T cell clonotypes into the αβ T cell pathway at E17 is associated
456 with a decreased frequency of these mature clonotypes in neonatal thymus, and an altered γδT17
457 TCR repertoire in the periphery. Finally, we identify Vγ4/Vδ7 T cells as a novel, SAP-dependent
458 IFN-γ-producing Vγ4 γδT1 subset.

459 γδ T cells are thought to arise at the DN2/DN3 stage after receiving a sufficiently strong
460 γδTCR-mediated signal[55]. Within the continuum of signals that can drive γδ T cell development,
461 it is the relatively weaker TCR signals that promote γδT17 programming while the stronger TCR
462 signals foster γδNKT and γδT1 programming[4, 8, 11, 14, 56, 57]. Despite considerable effort, the precise
463 thymic developmental stages at which γδ T cell lineage diversification occurs remains unclear.
464 Recent studies examining this issue have demonstrated that effector programming can occur at a
465 $CD44^{neg}CD45RB^{neg}$ stage[4, 14], and that c-Maf-directed γδT17 programming occurs at a $CD45RB^{low}$
466 stage[36]. In addition, by altering TCR signal strength in a transgenic model, Chen et al., defined a
467 *Ccr9*-enriched cluster as a point of effector lineage diversification[10]. Our findings here suggest that
468 SAP influences γδT17 developmental programming during the immature uncommitted
469 $CD24^{hi}CD73^{neg}$ stage, either during or soon after *Il2ra (*CD25) downregulation. The SAP-

470 dependent c2 cluster we define here was highly enriched in numerous genes (e.g., *Sox13*[40], *Blk*[35],
471 *Maf*[36], *Etv5*[39], *Sh2d1a*[18]) previously associated with γδT17 development as well as *Ccr9*, and the
472 cell surface phenotype of c2 γδ T cells is SLAMF1^{pos}SLAMF6^{pos}CD44$^{low/high}$CD45RB^{low}.
473 Altogether, these data suggest that the SAP-dependent c2 cluster is identical to these previously
474 defined developmental checkpoints. Given that these previous studies have demonstrated that this
475 developmental checkpoint is influenced by TCR signaling, it is interesting to note that SLAM/SAP
476 signaling has previously been demonstrated to foster semi-invariant NKT cell differentiation
477 through its ability to regulate TCR signal strength[58].

478 To that end, we noted that in addition to a decrease in the number of immature γδ T cells
479 expressing BLK, the expression level of BLK was decreased in B6.SAP$^{-/-}$ γδ T cells. BLK is a
480 member of the Src tyrosine kinase family whose members also include LCK and FYN. All three
481 of these kinases have been implicated in γδ T cell development[59] and BLK has been shown to
482 promote γδT17 development[35]. Thus, it's possible that decreased BLK expression could contribute
483 to the impaired γδT17 development observed in SAP-deficient mice. In addition, since SAP has
484 been shown to regulate conventional αβ T cell differentiation through its ability to interact with
485 FYN[60], it's possible that SAP interacts with one or all of these Src kinase family members to exert
486 its effects. Whether SAP works to regulate TCR signal strength during γδ T cell development, or
487 whether it exerts its effects through a TCR-independent mechanism is currently under
488 investigation.

489 Nevertheless, our finding that, in addition to regulating the frequency of the c2 cluster, SAP
490 inhibits the number of γδ T cells that enter the αβ T cell developmental pathway, is reminiscent of
491 previous work using transgenic TCR models demonstrating that a reduction of TCR signal strength
492 results in a redirection of γδ T cells to the αβ T cell lineage[53, 61] In addition, Fahl et al. recently

493 demonstrated that both entry into the γδ T cell lineage as well as γδ T cell effector programming
494 involved repression of TCF1 expression by TCR-driven increases in *Id3*[42]. Consistent with these
495 results, we found that *Tcf7* (encoding TCF1) transcript levels were highest in the αβ T cell-like c6
496 cluster, and we note that *Tcf7* transcripts levels were noticeably higher in the *Il2ra*-enriched post-
497 selection c5 cluster in B6.SAP$^{-/-}$ γδ T cells. *Id3* expression, however, did not appear to be altered
498 in SAP-deficient mice. Whether SAP-mediated signals could alter γδ T cell *Tcf7* expression
499 through a TCR-independent mechanism remains an open question.

500 While our data indicate that SAP regulates the diversion of at least some TCR clonotypes
501 into the αβ TCR pathway, the fate of those cells is unclear. Certainly, the fact that the clonotypes
502 that were diverted to the αβ T cell pathway in E17 thymus are under-represented in the mature D9
503 thymus γδT17 clusters suggests that this diversion has significant consequences on the γδ T cell
504 receptor repertoire, but whether these diverted γδ T cells contribute to the αβ T cell compartment
505 is unclear. Recently, a population of hybrid αβ-γδ T cells was described in both mouse and
506 human[62] that responds to peptide/MHC stimulation like conventional αβ T cells, as well as to IL-
507 1 and IL-23 which induce both IFN-γ and IL-17 production[62]. In the mouse, at least some of these
508 cells are Vγ4/Vδ5, Vγ4/Vδ4, or Vγ4Vδ7 and are present in the embryonic thymus, raising the
509 possibility that γδ T cells diverted to the αβ T cell pathway in B6.SAP$^{-/-}$ mice contribute to this
510 αβ-γδ T cell population. While we have not observed a significant increase in the overall numbers
511 of αβ-γδ T cells in B6.SAP$^{-/-}$ thymus (not shown), a direct comparison of B6 and B6.SAP$^{-/-}$ αβ-
512 γδ TCRs may be necessary to see such an effect.

513 Our previous observations that SAP deletion did not completely abolish the development of
514 IL-17- and IFN-γ producing γδ T cells[18] suggested the possibility that SAP may be critical for the

515 development of additional γδ T cell clonotypes other than the Vγ1Vδ6.3 subset. Therefore, we
516 compared the SAP-dependent and SAP-independent γδTCR repertoires using paired TCR
517 clonotype analysis, which allowed us to track clonotype fate during development. Our findings not
518 only confirmed previous results demonstrating highly restricted γδT17[45, 46], and γδNKT[21] TCR
519 repertoires, they also revealed that the vast majority of Vγ4 γδT1 subsets in the periphery
520 preferentially utilize TRGV4/TRAV13-4(DV7) clonotypes, thereby further highlighting the strong
521 links between γδ T cell function and the utilization of specific TCRs[23]. The reason for this
522 preferential TRDV usage in Vγ4/Vδ7 γδT1 is unclear. While the TRAV13-4(DV7) CDR3
523 sequences were diverse and did not suggest evidence of clonal expansion, it may be possible that
524 TRDV-encoded sequence motifs could be required to interact with a ligand, similar to the way that
525 TRGV-encoded sequences mediate recognition of butyrophilins[63, 64].

526 Exactly when during development SAP exerts its effects on Vγ4/Vδ7 T cells is still
527 somewhat unclear. We found comparatively low numbers of TRGV4/TRDV7 γδ T cells in
528 embryonic, neonatal, or adult thymus, and those that were identified were generally found in the
529 immature clusters. The increased frequency in B6.SAP$^{-/-}$ E17 thymus of TRGV4/TRDV7 γδ T
530 cells in the *Il2ra*$^{+}$ c5 cluster coupled with their decreased frequency in the emigrating *S1pr1*$^{+}$ c4
531 cluster suggests a model where SAP promotes progression through the early stages of development
532 through thymic egress. SAP deficiency, therefore, could result in lower Vγ4/Vδ7 output. However,
533 this model is at odds with our finding that the frequency of *S1pr1*$^{+}$ clusters in neonatal thymus is
534 increased in B6.SAP$^{-/-}$, and that these clusters contained increased numbers of TRGV4/TRDV7 γδ
535 T cells. A possible explanation could be found in a previous report demonstrating that a Vγ4/Vδ7
536 subset increases in frequency in the periphery of B6 mice over time and undergoes extrathymic
537 expansion[65]. Our finding here that that the Vγ4/Vδ7 T cells represent a SAP-dependent innate-like

538 γδT1 subset raise the possibility that SAP regulates the homeostatic expansion of this subset.
539 subset after it leaves the thymus. The mechanism, however, through which SAP is regulating the
540 number of the Vγ4/Vδ7 subset remains unclear as we have so far observed no evidence of changes
541 in homeostatic proliferation or apoptosis of these cells in the lungs.

542 We previously demonstrated that the SAP-dependent dependent impairment in γδT17
543 function in the neonatal periphery largely corrects itself as mice mature[18]. This suggested the
544 possibility that SAP-dependent γδT17 clonotypes originating in the embryonic thymus were being
545 replaced over time by SAP-independent γδT17 clonotypes. Although our data certainly point to a
546 SAP-dependent decrease in a germline-encoded Vγ4/Vδ5 clonotype in neonatal thymus, the low
547 numbers of this clonotype in the lung makes it difficult to draw definitive conclusions in the
548 periphery. Still, we did observe a role for SAP in shaping the peripheral γδT17 repertoire, as we
549 noted we noted a decreased frequency of germline-encoded TRGV4 sequences (so-called
550 "GYSS"[66]) coupled with an increased frequency of TRGV4 sequences with more diverse CDR3
551 ("GXYSS") in the periphery. Although the potential significance of this change on γδ T cell
552 function is unclear, it does support a model in which SAP regulates γδT17 developmental
553 programming during embryonic thymic development, where the low expression of terminal
554 deoxynucleotide transferase results in a high frequency of germline-encoded V-J junctions.

555 In summary, our findings reveal the presence of multiple SAP-dependent developmental
556 checkpoint during the very early stages of γδ T cell development in the thymus, indicating a more
557 complex role for this signaling pathway in γδ T cell developmental programming than previously
558 appreciated. Indeed, our data indicate that SAP functions to regulate entry into the γδ and αβ T
559 cell developmental pathways, and we make the striking observation that nearly all Vγ4 γδT1 cells
560 in the periphery preferentially use Vδ7 chains and that it is a decrease in the Vγ4/Vδ7 subset that,

561 in part, underlies the SAP-dependent decrease in γδ T cell IFN-γ production. Whether SAP
562 regulates these phenotypes through regulation of TCR signal strength or additional pathways
563 remains an open question. Taken together, the data presented here suggest a key role for the
564 SLAM/SAP signaling pathway in generating and/or maintaining the innate-like phenotype in the
565 γδ T cell lineage.

566 **Methods**

567 *Mice*

568 C57BL/6J (B6) and B6.*Sh2d1a*$^{-/-}$ (SAP$^{-/-}$) mice were purchased from the Jackson
569 Laboratory (Bar Harbor, ME) and were housed and bred in a specific pathogen-free facility at the
570 animal facility of the University of Vermont. For experiments using E17 mice, timed-pregnant
571 matings were used. The presence of a vaginal plug indicated day 0 of pregnancy. The flow
572 cytometry experiments included 8-13 weeks old B6 and B6.SAP$^{-/-}$ mice (age and sex-matched).
573 All experimental procedures on animals were carried out with the approval of the University of
574 Vermont Institutional Animal Care and Use Committee.

575

576 *Tissue Processing*

577 Single-cell suspensions from the spleen and thymus were obtained by gently passing tissue
578 through nylon mesh. Red blood cells (RBCs) were lysed with Gey's solution. Single-cell
579 suspensions from the lungs were prepared as described (Benoit et al., 2015). Briefly, dissected
580 lungs were finely minced in a digestion buffer containing DMEM, 1 mg/ml collagenase type IV
581 (Gibco), and 0.2 mg/ml DNase (Sigma-Aldrich). After shaking at 200 rpm at 37˚C for 20 min, the
582 lung digests were triturated with a 16-gauge needle, followed by an additional 20 min shaking at
583 200 rpm at 37˚C. Following the final trituration, red blood cells (RBC) were lysed with Gey's
584 solution. Finally, the digested lung tissues were filtered through a 70-µm filter and washed in
585 PBS/2%FBS. Cells were counted either on a MACSQuant VYB or using a hemacytometer
586 following trypan blue staining.

587 *Flow Cytometry*

588 Cells were incubated with LIVE/DEAD™ Fixable Blue Dead Cell stain (Thermo Fisher
589 Scientific, Grand Island, NY) for 30 min at 4^0C in PBS, after which they were resuspended in Fc-
590 Block (BioLegend) containing PBS/2% FCS/0.1% sodium azide buffer. After a 5-10 min
591 incubation on ice, antibodies specific for cell surface markers were added, incubated for 30 min
592 on ice, and washed in PBS/2% FCS/0.1% sodium azide buffer. Brilliant Violet Staining buffer
593 (BD Biosciences) was used in experiments contain more than one Brilliant Violet, Brilliant UV,
594 or SuperBright-conjugated Ab. Cells were fixed in PBS/1% PFA buffer and stored at 4^0 C prior to
595 data collection.

596 For nuclear staining, cells were permeabilized with Foxp3 transcription factor
597 fixation/permeabilization buffer (ThermoFisher) following surface staining. Cells were incubated
598 overnight at 4˚C, washed, and incubated with 25 micrograms/ml total rat IgG before adding Abs
599 against transcription factors. After staining, cells were washed and resuspended in PBS/2%
600 FCS/0.1% sodium azide buffer. For intracellular cytokine staining, cells were permeabilized with
601 PBS/1% FCS/0.1% sodium azide/0.1% saponin following surface staining and fixation. To reduce
602 unspecific binding, 25 µg/ml total rat IgG was added prior to cytokine staining with fluorescent-
603 labeled antibodies.

604 Flow cytometry data were collected on a Cytek® Aurora (Cytek Biosciences) spectral
605 cytometer, and data analysis was performed using FlowJo Software (Tree Star). In some cases,
606 visualization of flow cytometry data was conducted by concatenation of individual .fcs files,
607 followed by dimensionality reduction and clustering using UMAP and FlowSOM[67] in FlowJo.
608 Antibodies used in these experiments are anti-CD4 (RM4-5), anti-CD11b (M1/70), anti-CD11c
609 (N418), anti-CD19 (1D3), anti-CD24 (M1/69), anti-CD25 (PC61), anti-CD27 (LG.3A10), anti-

610 CD44 (IM7), anti-CD45RB (C363-16A) , anti-CD73 (TY/11.8), anti-CD150 (TC15-12F12.2),
611 anti-CD196 (29-2L17), anti-CD319 (4G2), anti-cMAF (sym0F1), anti-IFNγ (XMG1.2), anti-
612 IL17a (TC11-18H10.1), anti-IL4 (11B11), anti-PLZF (9E12), anti-T-bet (4B10), anti-TCRβ (H57-
613 597), anti-Vγ5 (536), and anti-Vd6.3 (C504.17c) from BioLegend; and anti-RORγt (Q31-378),
614 anti-TCRδ (GL3), anti-Vγ1 (2.11) from BD biosciences; and anti-CD8a (53-6.7), anti-CD45 (30-
615 F11), anti-CD200 (OX90), anti-CD278 (7E17G9), anti-Ki-67 (SolA15), anti-Ly108(13G3-19D),
616 anti-Vγ4 (UC3-10A6) from eBiosciences; anti-S1P1 (713412) from R&D systems; anti-Blk
617 (polyclonal) from Cell Signaling Technology; and anti-17D1
618
619 *In vivo proliferation and apoptosis*
620 For *in vivo* proliferation analysis, each mouse was intraperitoneally injected with 1 mg of
621 5-bromo-deoxyuridine (BrdU) (Sigma-Aldrich) each day for 3 days. Lung cell suspensions were
622 prepared as above, stained with surface markers, after which they staining with anti-BrdU antibody
623 according to the manufacturer's instructions (BrdU Flow Kit; BD Biosciences). For apoptotic cell
624 detection, annexin V staining was performed. Briefly, surface staining was followed by dead cell
625 staining with Fixable Viability Dye eFluor™ 780 dye. The cells were washed and resuspended in
626 an annexin-binding buffer (10 mM HEPES, 140 mM NaCl, and 2.5 mM CaCl2, pH 7.4). Finally,
627 1 µl of Annexin V Conjugate (Thermo Fisher Scientific, Grand Island, NY) was added to each
628 tube, and cells were incubated for 15 minutes at room temperature (RT) prior to data collection.
629
630 *Bulk RNA Sequencing and data analysis*
631 For the bulk RNA sequencing, thymocytes from 10-day old B6 (n=4 male, n=2 female)
632 and B6.SAP$^{-/-}$ (n=5 male, n=1 female) mice were prepared by depleting CD4+ cells from

633 thymocytes using the EasySepTM Mouse CD4 Positive Selection Kit II (STEMCELL
634 Technologies). Cells were then incubated with Fixable Viability Dye eFluor™ 780 (Thermo Fisher
635 Scientific, Grand Island, NY) for 30 min at 4^0C in PBS, after which they were resuspended in Fc-
636 Block (BioLegend) containing PBS/2% FCS. After washing, cells were stained with antibodies
637 against CD11b (M1/70), CD19 (1D3), CD24 (M1/67), CD45 (30-F11), TCRβ (H57-597), and
638 TCRδ (GL3) for 30 min at 4^0C. Following surface staining, $CD24^{hi}$TCRδ+ cells were sorted from
639 a $CD11b^-CD19^-$TCRb- population of live $CD45^+$ lymphocytes. Immature ($CD24^{high}$) thymocytes
640 from both B6 and B6.$SAP^{-/-}$ mice were sorted directly into 350 microliters of RLT lysis buffer
641 (Qiagen). RNA was extracted using the RNeasy Micro kit (Qiagen) according to manufacturer
642 instructions. After RNA quantification and quality assessment, cDNA libraries were prepared
643 using SMARTer Stranded Total RNA-Seq Kit v3 - Pico Input Mammalian (Takara, Japan)
644 according to the manufacturer's protocol. Following library clean-up and quantification, the
645 libraries were sequenced (single-end 75 bp) on an Illumina HiSeq1500.

646 The Fastq files were checked for data quality using FastQC and MultiQC[68] bioinformatics
647 tools. Cutadapt[69] was used to filter out low-quality sequences and to remove adapter sequences
648 from all reads. The expression of transcripts was quantified using the mapping-based mode of
649 Salmon[70], which quasi-mapped RNA-seq reads to an index created from a mouse reference
650 transcriptome (GRCm39). The transcript abundance estimates were imported using tximport[71] and
651 differential gene expression (DE) analysis was performed using DEseq2[72] employing the Wald
652 test of the negative binomial model coefficients to determine statistical significance. The results
653 were analyzed based on Log_2 fold change ($Log_2FC \geq 0.5$ or ≤ -0.5) and adjusted P-value (padj
654 ≤ 0.0001) to determine whether a gene was differentially expressed.
655

656 *Cell sorting and library preparation for CITE-seq and V(D)J enrichment*

657 For thymus single-cell transcriptome analysis, we used embryonic day 17 (E17) B6 (n=4)
658 and B6.*Sh2d1a*$^{-/-}$ (SAP$^{-/-}$) (n=4) mice, nine day-old (D9) B6 (n=3) and B6.SAP$^{-/-}$ (n=3) mice, and
659 six-week-old B6 (n=3 male) and B6.SAP$^{-/-}$ (n=2 male, n=1 female) mice. For single-cell analysis
660 of adult lung γδ T cells, we used 13 week old B6 and B6.SAP$^{-/-}$ littermates (pooled cells from 2
661 males and one female mouse in each group).

662 Analysis of gene expression, surface protein expression, and TCR V(D)J repertoire was
663 performed using the 10X Genomics Chromium Next GEM Single Cell 5' V(D)J with feature
664 barcoding kit (version 1.1). To sort cells for scRNAseq, thymocytes were stained with anti-TCRδ-
665 PE (GL3) antibody and enriched using anti-PE MicroBeads kit (Miltenyi Biotec). The exception
666 was E17 thymus which did not undergo TCRδ enrichment. Following live/dead staining with
667 Fixable Viability Dye eFluor™ 780 (Thermo Fisher Scientific, Grand Island, NY) for 30 min at
668 4° C in PBS, cells were resuspended in Fc-Block (BioLegend) containing PBS/2% FCS. Cells were
669 then simultaneously stained with oligonucleotide-conjugated antibodies targeting CD24, CD44,
670 CD45RB, CD73, SLAMF1, SLAMF6, Vγ1, and Vγ4 (TotalSeqC-BioLegend), oligonucleotide-
671 conjugated hashtag antibodies (TotalSeqC-BioLegend), and fluorophore-conjugated antibodies for
672 cell sorting which included the anti-CD11b (M1/70), anti-CD11c (N418), anti-CD19 (1D3), anti-
673 CD45.2 (104) and anti-TCRβ (H57-597) antibodies. Following surface staining of the enriched
674 population, TCRδ^{+} cells were sorted from a CD11b^{-}CD11c^{-}CD19^{-}TCRβ^{-} population of CD45.2^{+}
675 live lymphocytes.

676 Libraries were prepared according to the manufacturer's instructions with modifications to
677 the V(D)J enrichment procedure. Equal numbers of live GL3^{+} γδ T cells from B6 and B6.SAP$^{-/-}$
678 mice were pooled and captured (between 5000 – 7000 cells) using a 10x Chromium controller.

679 Following cDNA preparation, the gene expression and CSP libraries were created according to the
680 manufacturer's instructions. To obtain TCR gamma and delta chain sequences, we replaced kit-
681 provided enrichment primers with custom mouse γ and δ chain-specific primers and modified the
682 number of PCR cycles used from manufacturer recommended cycling conditions. In the first
683 enrichment step, an equal amount of cDNA was used as a template for PCR amplification of γ and
684 δ chains separately using a total of 12 and 10 PCR cycles, respectively. The primers used for the
685 first γ-chain enrichment are: Forward Primer 5'-AATGAT ACGGCGA CCACCG AG
686 ATCTACACTCTT TCCCTACAC GACGCTC-3' (2 μM) and Reverse Outer Primers 5'-GGAA
687 AGAACTT TTCAAGGAS ACAAAG-3' (1 μM), 5'-CCCTTATG ACTTCAGG AAAGAA
688 CTTT-3' (0.5 μM) . The primers used for the first δ-chain enrichment are: Forward Primer 5'-
689 AATGAT ACGGCGACCACCG AGATCTACACTC TTTCCCTAC ACGACGCTC-3' (2 μM)
690 and Reverse Outer Primer 5'-CCACAATC TTCTTGGATG ATCTGAG-3' (0.5 μM). Following
691 PCR product purification, the γ-chain and δ-chain amplicons went through the second round of
692 enrichment separately using a total of 12 and 10 PCR cycles, respectively. The primers used for
693 the second γ-chain enrichment are: Forward Primer 5'-AATGA TACGGCGACCA
694 CCGAGATCT-3' (0.5 μM) and Reverse Inner Primers 5'-ACAAA GGTATGTCCCA
695 GTCTTATGGA-3' (0.5 μM), 5'-GGAGACA AAGGTAGGT CCCAGC-3' (0.5 μM). The
696 primers used for the second δ-chain enrichment are: Forward Primer 5'-AATGATA
697 CGGCGACCACC GAGATCT-3' (0.5 μM) and Reverse Inner Primer 5'-GTCACCTC
698 TTTAGGGTAG AAATCTT-3'. Following PCR product cleanup, the second round PCR products
699 (γ-chain and δ-chain amplicons) were quantified and combined in equal quantity making the final
700 TCR V(D)J library. Finally, all three libraries (gene expression, cell surface protein expression,

701 and TCR V(D)J) were sequenced (paired-end 26 x 91 bp) on an Illumina HiSeq 1500 sequencer
702 or a NextSeq 2000.
703
704 *CITE-seq quality control, filtering and demultiplexing*
705 Raw sequence reads from the gene expression and CSP libraries were processed using the
706 Cell Ranger software (v4.0.0)[73]. Reads were mapped to the GRCm38 (mm10) mouse reference
707 genome. To exclude the possibility of TCR transcripts playing a role in UMAP clustering, all TCR
708 genes were omitted from the analysis using the "cellranger reanalyze" argument of Cell Ranger
709 software.
710 Quality control, filtering, UMAP projection, and clustering analysis were conducted using
711 Seurat (v4.3.0.1)[74] R package. Briefly, cells with a low number (≤2000 for E17 thymus dataset,
712 ≤1200 for D9 thymus dataset and ≤750 for lung dataset) of unique genes (nFeature_RNA), a
713 frequency of mitochondrial reads > 4% and a frequency of ribosomal protein < than ~12% were
714 eliminated from further analysis. To reduce cell cycle-related and stress-related heterogeneity in
715 the normalized data, selected cell cycle genes and stress-induced genes [75] were regressed out
716 during the data normalization step. Note that in some instances, contaminating clusters expressing
717 macrophage specific marker genes were manually removed using Seurat "CellSelector ()"
718 argument. We used the Seurat function "HTODemux()" to demultiplex single cells back to their
719 sample of origin [76].
720
721 *CITE-Seq and V(D)J enrichment data analysis*
722 The demultiplexed datasets were further analyzed for gene expression, cell surface protein
723 expression following the multimodal analysis pipeline of Seurat[74]. For each dataset, we performed

normalization and variance stabilization of the gene expression data using the scTransform[77]. Following identification of top significant principal components (PCs) using the principal component analysis (PCA), we performed a graph-based (UMAP) semi-supervised clustering of cells with similar gene expression patterns using specific number of top PCs (E17 = 25, D9 = 30, W6 = 27, lung = 28) and resolution values (E17 = 1.2, D9 = 1.6, W6 = 0.65, lung =1.2) for each dataset. The cell surface protein expression data (Centered Log Ratio (CLR) normalized) from CSP libraries as well as the gene expression data were then visualized on the clustered cells using the R package scCustomize[78]. The differentially expressed features of each cluster (cluster biomarkers) were identified using FindAllMarkers() function using the Wilcoxon Rank Sum test. For identifying differentially expressed genes (DEGs) between B6 and B6.SAP$^{-/-}$ samples, we performed a pseudobulk DE (differential expression) analysis of aggregated read counts using the DEseq2 package, which accounts for biological replicates making it a better approach than other specialized single-cell DE analysis methods[79]. Briefly, for cells of a given sample type, we aggregated RNAseq reads across biological replicates identified by individual hashtags. Then transformed the genes-by-cell matrix to a genes-by-replicates matrix and performed DE analysis using the DESeq2[72] package employing the Wald test of the negative binomial model coefficients to determine statistical significance. Finally, the trajectory and pseudotime inference analysis was performed using the partition-based graph abstraction (PAGA)[80] analysis, which is a part of the Scanpy[81] python package.

For the analysis of V(D)J library sequences, Cell Ranger v6.1.2 ("cellranger vdj" argument) was used to map the TCR V(D)J) reads to a custom-made reference containing sequences from the IMGT database[82]. The Cell Ranger V(D)J output was then processed with custom bashscripts. Briefly, the TRGV and TRAV/TRDV contigs identified by Cell Ranger V(D)J

747 were demultiplexed to individual barcodes after which the dominant TRGV and TRAV/DV
748 transcripts were selected per cell. Finally, the V(D)J information of each cell was integrated with
749 scCITE-seq data and visualized using a combination of Seurat and custom-made R scripts. We
750 note that since about half of Vγ1, Vγ4, and Vγ7 T cells contain functionally rearranged TRGV2
751 chains, but only a small percentage express Vγ2 on the cell surface[83], the methods used here cannot
752 adequately assess the true TRGV2 repertoire. Therefore, we report it here for accuracy but have
753 omitted TRGV2 from our analysis.
754
755 *Quantitative PCR*
756 Lung qPCR samples were prepared by positively selecting leukocytes using CD45 MicroBeads
757 (Miltenyi Biotec). Spleen qPCR samples were prepared by depletion of CD19+ TCRβ+ cells using
758 biotinylated antibodies and Streptavidin MicroBeads (Miltenyi Biotec) according to the
759 manufacturer's instructions. Total RNA was extracted from approximately 5-10 million cells from
760 each organ using the RNeasy Mini Kit (QIAGEN). SuperScript® IV Reverse Transcriptase
761 (Thermo Fisher Scientific, Grand Island, NY) was used to generate cDNA using 400 ng of RNA
762 and oligodT primer according to the manufacturer's instructions. qPCR was performed using 1µl
763 of cDNA template in iTaq Universal SYBR Green Supermix (Bio-Rad) containing TCR delta
764 chain specific primer pairs. The qPCR reaction was run on a Bio-Rad CFX96 qPCR machine using
765 the following cycling conditions: 95^0C for 30 sec; 40 cycles of [95^0C for 5 sec; 60^0C for 30 sec;
766 72^0C for 30 sec]. The β2M and TRDC genes were used as endogenous housekeeping controls.
767 Primers used were as follows: mb2mforward, 5'- CATGGCTCG CTCGGTGACC-3',
768 mb2mreverse, 5'-AATGTGAGGC GGGTGGAACTG-3', TRDCforward, 5'- CTCCGGCCA
769 AACCATCTGTT-3', mouseDV2, 5'-CCAAGAAG CATACAAGCAGT ATAATG-3',

770 mouseDV5, 5'- CCCATGA TGCAGATTT TGTTCAAGG-3', mouseDV7, 5'- GGAAG
771 MCTCGTCAGC CTGTTGT-3', mouseDCrev, 5'- CCACAATCTT CTTGGAT GATCTGAG-3'.
772 Some primer sequences were adopted from Wei. et. al. (2015)[45].
773
774 *Vγ4 single cell sorting and TCR Sequencing*
775 The plate-based based γδ TCR repertoire analysis of single-cell sorted Vγ4 cells involved
776 adult lungs from 8-10 weeks old B6 male (n=7) and female mice (n=6). Lung Vγ4 TCR sequences
777 were obtained using a protocol modified from Wei *et al.*, (2015)[45]. Briefly, lung cell suspensions
778 were prepared as described above. After washing, cells were resuspended in Fc-Block
779 (BioLegend) containing PBS/2% FCS buffer, and stained with anti-CD11b (M1/70), anti-CD11c
780 (N418), anti-CD19 (1D3), anti-CD45 (30-F11), anti-TCRδ (GL3) and anti-Vγ4 (UC3-10A6) from
781 for 30 min at 4^0C in PBS. Fixable Viability Dye eFluor™ 780 (Thermo Fisher Scientific, Grand
782 Island, NY) was used to identify dead cells. Single-cell sorting of lung $V\gamma4^+TCR\delta^+$ from a $CD11b^-$
783 $CD11c^-CD19^-TCR\beta^-$ population of $CD45^+$ live lymphocytes were performed using a FACS Aria
784 III cell sorter (BD Bioscience). Cells were sorted directly into 12.5 µl (per well) of OneStep RT-
785 PCR Buffer mix (QIAGEN) in a 96-well PCR plate (one cell/well).
786 TCR sequences were then amplified and barcoded according to (Wei *et al.*, 2015). Briefly,
787 the first two nested PCR reactions involved primer-specific amplification of gamma and delta TCR
788 sequences. In both the first and second PCR reactions, the final concentration of each V-region
789 primer is 0.36µM, and each C-region primer is 0.6µM. The first RT-PCR step was performed using
790 the following cycling conditions: 50^0C for 30 min; 95^0C for 5 min; 25 cycles of [94^0C for 30 sec;
791 62^0C for 1min; 72^0C for1 min]; 72^0C for 7 min; 4^0C. Next, 1 µl of the first round PCR product was
792 used as a template for the second nested PCR reaction using HotStartTaq Master Mix Kit

793 (QIAGEN). The cycling conditions for the second PCR reaction were 95^0C for 5 min; 25 cycles
794 of [94^0C for 30 sec; 64^0C for 1min; 72^0C for1 min]; 72^0C for 7 min; 4^0C. In the third PCR reaction,
795 1 µl of second round PCR product as the template in a new 14 µl reaction using HotStartTaq Master
796 Mix Kit (QIAGEN). Each well was barcoded. The final concentration of 5' and 3' barcoding
797 primers was 0.05µM each and paired-end primers were 0.5µM each. The cycling conditions for
798 the third PCR reaction were 95^0C for 5 min; 36 cycles of [94^0C for 30 sec; 62^0C for 1min; 72^0C
799 for1 min]; 72^0C for 7 min; 4^0C. The γ and δ chain products were run on 1.2% agarose gel separately
800 and purified using the Qiagen QIAquick Gel Extraction Kit according to kit instructions. The final
801 products were then analyzed using the 2100 Bioanalyzer (Agilent Technologies). Finally, the γ
802 and δ TCR products were mixed in equal concentration and sequenced on the Illumina MiSeq
803 platform using the Illumina MiSeq Reagent Kit v2 (500-cycles).

804 The paired-end TCR sequencing reads were analyzed using a custom software pipeline
805 capable of de-multiplexing barcoded sequences to individual cells. The resulting de-multiplexed
806 reads were then analyzed by MiGMAP software, which identifies CDR3 sequences using the
807 IgBlast 1.4.0 (NCBI) V(D)J mapping tool[84, 85]. The single-cell TCR V(D)J profiling data were then
808 visualized using custom R scripts.

809

810 *Statistical analysis*

811 Prism (GraphPad Software, San Diego, CA) was used for the statistical analysis of data.
812 Unpaired Student t-tests, one-way ANOVA, or two-way ANOVA were used where appropriate.
813 Tests were considered significant when p-values were less than 0.05. Error bars represent standard
814 deviation (SD) unless otherwise indicated.

Acknowledgements

We thank Marlana Winschel, Cameron Moquin, and Erin Mathieu for help in tissue preparation and flow cytometry. We acknowledge Roxana del Rio Guerra for help in cell sorting. All flow cytometry data were collected at the Harry Hood Bassett Flow Cytometry and Cell Sorting Facility (Larner College of Medicine, University of Vermont) and was supported by NIH grant (S10OD026843). We thank Jessica Hoffman, Kris Finstaad, and the Vermont Integrated Genome Resource (Larner College of Medicine, University of Vermont) for help in scRNAseq, Julie Dragon and Korin Eckstrom for help in bioinformatics analysis. We thank Fred Kolling at the Genomics Shared Resource (Dartmouth Geisel School of Medicine) for help in sequencing. We also thank Willi Born and Rebecca O'Brien for the 17D1 hybridoma. This work benefitted from data assembled by the ImmGen consortium, and was supported by NIH grants R21AIO66465 and R03AI153902, S10OD026843 (JEB) and an AAI Careers in Immunology Fellowship (SKM, JEB).

827 **Author contributions**

828 S.K.M. and J.E.B. conceived and designed the project; S.K.M., J.E.B., E.A., K.J.H., B.M.H., R.S.,
829 and O.D. performed experiments and analyzed data; K.J.H., D.G., J.T.R., and N.S. aided with the
830 single-cell TCR sequencing; D.M. assisted with bulk RNA sequencing; S.K.M. and J.E.B.
831 performed the bioinformatics analysis of bulk RNA-seq and single-cell CITEseq (with immune
832 profiling) data; J.E.B. supervised the project; S.K.M. and J.E.B. prepared the manuscript.

833

834 **Competing interests**

835 All authors declare no competing interests.

836

837 **Materials and Correspondence**

838 Correspondence and requests for materials should be addressed to J.E.B. (email:
839 jonathan.boyson@med.uvm.edu)

840 **References**

841 1. Grigoriadou K, Boucontet L, Pereira P. Most IL-4-producing gamma delta thymocytes of
842 adult mice originate from fetal precursors. *J Immunol* **171**, 2413-2420 (2003).
843 2. Haas JD*, et al.* Development of interleukin-17-producing gammadelta T cells is restricted
844 to a functional embryonic wave. *Immunity* **37**, 48-59 (2012).
845 3. Ribot JC*, et al.* CD27 is a thymic determinant of the balance between interferon-gamma-
846 and interleukin 17-producing gammadelta T cell subsets. *Nat Immunol* **10**, 427-436
847 (2009).
848 4. Sumaria N, Grandjean CL, Silva-Santos B, Pennington DJ. Strong TCRgammadelta
849 Signaling Prohibits Thymic Development of IL-17A-Secreting gammadelta T Cells. *Cell*
850 *Rep* **19**, 2469-2476 (2017).
851 5. Chien YH, Zeng X, Prinz I. The natural and the inducible: interleukin (IL)-17-producing
852 gammadelta T cells. *Trends Immunol* **34**, 151-154 (2013).
853 6. Ribot JC, Lopes N, Silva-Santos B. γδ T cells in tissue physiology and surveillance.
854 *Nature reviews Immunology* **21**, (2021).
855 7. Turchinovich G, Hayday AC. Skint-1 identifies a common molecular mechanism for the
856 development of interferon-gamma-secreting versus interleukin-17-secreting gammadelta
857 T cells. *Immunity* **35**, 59-68 (2011).
858 8. Jensen KD*, et al.* Thymic selection determines gammadelta T cell effector fate: antigen-
859 naive cells make interleukin-17 and antigen-experienced cells make interferon gamma.
860 *Immunity* **29**, 90-100 (2008).
861 9. Munoz-Ruiz M*, et al.* TCR signal strength controls thymic differentiation of discrete
862 proinflammatory gammadelta T cell subsets. *Nat Immunol*, (2016).

863 10. Chen ELY, Lee CR, Thompson PK, Wiest DL, Anderson MK, Zuniga-Pflucker JC.
864 Ontogenic timing, T cell receptor signal strength, and Notch signaling direct gammadelta
865 T cell functional differentiation in vivo. *Cell Rep* **35**, 109227 (2021).
866 11. Kreslavsky T*, et al.* TCR-inducible PLZF transcription factor required for innate
867 phenotype of a subset of gammadelta T cells with restricted TCR diversity. *Proc Natl*
868 *Acad Sci U S A* **106**, 12453-12458 (2009).
869 12. Spidale NA*, et al.* Interleukin-17 Producing gammadelta T Cells Originate from
870 SOX13(+) Progenitors that Are Independent of gammadeltaTCR Signaling. *Immunity*,
871 (2018).
872 13. Laird SM, Tuckerman EM, Cork BA, Linjawi S, Blakemore AI, Li TC. A review of
873 immune cells and molecules in women with recurrent miscarriage. *Hum Reprod Update*
874 **9**, 163-174 (2003).
875 14. Sumaria N, Martin S, Pennington Daniel J. Constrained TCRγδ-associated Syk activity
876 engages PI3K to facilitate thymic development of IL-17A–secreting γδ T cells. *Science*
877 *Signaling* **14**, eabc5884 (2021).
878 15. Muro R, Nitta T, Nakano K, Okamura T, Takayanagi H, Suzuki H. gammadeltaTCR
879 recruits the Syk/PI3K axis to drive proinflammatory differentiation program. *J Clin*
880 *Invest* **128**, 415-426 (2018).
881 16. Cannons JL, Tangye SG, Schwartzberg PL. SLAM family receptors and SAP adaptors in
882 immunity. *Annu Rev Immunol* **29**, 665-705 (2011).
883 17. Nichols KE*, et al.* Regulation of NKT cell development by SAP, the protein defective in
884 XLP. *Nat Med* **11**, 340-345 (2005).

885 18. Dienz O*, et al.* Critical Role for SLAM/SAP Signaling in the Thymic Developmental
886 Programming of IL-17- and IFN-gamma-Producing gammadelta T Cells. *J Immunol*,
887 (2020).
888 19. Legoux F, Gilet J, Procopio E, Echasserieau K, Bernardeau K, Lantz O. Molecular
889 mechanisms of lineage decisions in metabolite-specific T cells. *Nat Immunol* **20**, 1244-
890 1255 (2019).
891 20. Koay HF*, et al.* A divergent transcriptional landscape underpins the development and
892 functional branching of MAIT cells. *Sci Immunol* **4**, (2019).
893 21. Azuara V, Levraud JP, Lembezat MP, Pereira P. A novel subset of adult gamma delta
894 thymocytes that secretes a distinct pattern of cytokines and expresses a very restricted T
895 cell receptor repertoire. *Eur J Immunol* **27**, 544-553 (1997).
896 22. Havran WL, Allison JP. Developmentally ordered appearance of thymocytes expressing
897 different T-cell antigen receptors. *Nature* **335**, 443-445 (1988).
898 23. O'Brien RL, Born WK. gammadelta T cell subsets: a link between TCR and function?
899 *Semin Immunol* **22**, 193-198 (2010).
900 24. Stoeckius M*, et al.* Simultaneous epitope and transcriptome measurement in single cells.
901 *Nature methods* **14**, 865-868 (2017).
902 25. Coffey F*, et al.* The TCR ligand-inducible expression of CD73 marks gammadelta
903 lineage commitment and a metastable intermediate in effector specification. *J Exp Med*
904 **211**, 329-343 (2014).
905 26. In TSH*, et al.* HEB is required for the specification of fetal IL-17-producing gammadelta
906 T cells. *Nat Commun* **8**, 2004 (2017).

907 27. Satija R, Farrell JA, Gennert D, Schier AF, Regev A. Spatial reconstruction of single-cell
908 gene expression data. *Nat Biotechnol* **33**, 495-502 (2015).
909 28. Havran WL, Chien YH, Allison JP. Recognition of self antigens by skin-derived T cells
910 with invariant gamma delta antigen receptors. *Science* **252**, 1430-1432 (1991).
911 29. Itohara S*, et al.* Homing of a gamma delta thymocyte subset with homogeneous T-cell
912 receptors to mucosal epithelia. *Nature* **343**, 754-757 (1990).
913 30. Wencker M*, et al.* Innate-like T cells straddle innate and adaptive immunity by altering
914 antigen-receptor responsiveness. *Nat Immunol* **15**, 80-87 (2014).
915 31. Buus TB, Odum N, Geisler C, Lauritsen JPH. Three distinct developmental pathways for
916 adaptive and two IFN-gamma-producing gammadelta T subsets in adult thymus. *Nat*
917 *Commun* **8**, 1911 (2017).
918 32. Xu H*, et al.* MS4a4B, a CD20 homologue in T cells, inhibits T cell propagation by
919 modulation of cell cycle. *PloS one* **5**, e13780 (2010).
920 33. Howie D*, et al.* MS4A4B is a GITR-associated membrane adapter, expressed by
921 regulatory T cells, which modulates T cell activation. *J Immunol* **183**, 4197-4204 (2009).
922 34. Sagar*, et al.* Deciphering the regulatory landscape of fetal and adult gammadelta T-cell
923 development at single-cell resolution. *EMBO J* **39**, e104159 (2020).
924 35. Laird RM, Laky K, Hayes SM. Unexpected role for the B cell-specific Src family kinase
925 B lymphoid kinase in the development of IL-17-producing gammadelta T cells. *J*
926 *Immunol* **185**, 6518-6527 (2010).
927 36. Zuberbuehler MK*, et al.* The transcription factor c-Maf is essential for the commitment
928 of IL-17-producing gammadelta T cells. *Nat Immunol* **20**, 73-85 (2019).
929

930 37. Narayan K, *et al.* Intrathymic programming of effector fates in three molecularly distinct
931 gammadelta T cell subtypes. *Nat Immunol* **13**, 511-518 (2012).
932 38. Lu Y, Cao X, Zhang X, Kovalovsky D. PLZF Controls the Development of Fetal-
933 Derived IL-17+Vgamma6+ gammadelta T Cells. *J Immunol* **195**, 4273-4281 (2015).
934 39. Jojic V, *et al.* Identification of transcriptional regulators in the mouse immune system.
935 *Nat Immunol* **14**, 633-643 (2013).
936 40. Malhotra N, *et al.* A network of high-mobility group box transcription factors programs
937 innate interleukin-17 production. *Immunity* **38**, 681-693 (2013).
938 41. He Q, *et al.* TOX deficiency facilitates the differentiation of IL-17A-producing
939 gammadelta T cells to drive autoimmune hepatitis. *Cellular & molecular immunology* **19**,
940 1102-1116 (2022).
941 42. Fahl SP, *et al.* The E protein-TCF1 axis controls gammadelta T cell development and
942 effector fate. *Cell Rep* **34**, 108716 (2021).
943 43. DiToro D, *et al.* Insulin-Like Growth Factors Are Key Regulators of T Helper 17
944 Regulatory T Cell Balance in Autoimmunity. *Immunity* **52**, 650-667 e610 (2020).
945 44. Roberts NA, *et al.* Rank signaling links the development of invariant gammadelta T cell
946 progenitors and Aire(+) medullary epithelium. *Immunity* **36**, 427-437 (2012).
947 45. Wei YL, *et al.* A Highly Focused Antigen Receptor Repertoire Characterizes gammadelta
948 T Cells That are Poised to Make IL-17 Rapidly in Naive Animals. *Frontiers in*
949 *immunology* **6**, 118 (2015).
950 46. Kashani E, *et al.* A clonotypic Vgamma4Jgamma1/Vdelta5Ddelta2Jdelta1 innate
951 gammadelta T-cell population restricted to the CCR6(+)CD27(-) subset. *Nat Commun* **6**,
952 6477 (2015).

953 47. Sim GK, Augustin A. Dominantly inherited expression of BID, an invariant undiversified
954 T cell receptor delta chain. *Cell* **61**, 397-405 (1990).
955 48. Scaramuzzino S*, et al.* Single-cell transcriptomics uncovers an instructive T-cell receptor
956 role in adult gammadelta T-cell lineage commitment. *EMBO J* **41**, e110023 (2022).
957 49. Prinz I, Sansoni A, Kissenpfennig A, Ardouin L, Malissen M, Malissen B. Visualization
958 of the earliest steps of gammadelta T cell development in the adult thymus. *Nat Immunol*
959 **7**, 995-1003 (2006).
960 50. Mingueneau M*, et al.* The transcriptional landscape of alphabeta T cell differentiation.
961 *Nat Immunol* **14**, 619-632 (2013).
962 51. Passoni L*, et al.* Intrathymic delta selection events in gammadelta cell development.
963 *Immunity* **7**, 83-95 (1997).
964 52. Kang J, Coles M, Cado D, Raulet DH. The developmental fate of T cells is critically
965 influenced by TCRgammadelta expression. *Immunity* **8**, 427-438 (1998).
966 53. Haks MC*, et al.* Attenuation of gammadeltaTCR signaling efficiently diverts thymocytes
967 to the alphabeta lineage. *Immunity* **22**, 595-606 (2005).
968 54. Griewank K*, et al.* Homotypic interactions mediated by Slamf1 and Slamf6 receptors
969 control NKT cell lineage development. *Immunity* **27**, 751-762 (2007).
970 55. Kreslavsky T, Garbe AI, Krueger A, von Boehmer H. T cell receptor-instructed alphabeta
971 versus gammadelta lineage commitment revealed by single-cell analysis. *J Exp Med* **205**,
972 1173-1186 (2008).
973 56. Fahl SP*, et al.* Role of a selecting ligand in shaping the murine gammadelta-TCR
974 repertoire. *Proc Natl Acad Sci U S A* **115**, 1889-1894 (2018).

975 57. Verykokakis M, Boos MD, Bendelac A, Adams EJ, Pereira P, Kee BL. Inhibitor of DNA
976 binding 3 limits development of murine slam-associated adaptor protein-dependent
977 "innate" gammadelta T cells. *PloS one* **5**, e9303 (2010).
978 58. Lu Y*, et al.* SLAM receptors foster iNKT cell development by reducing TCR signal
979 strength after positive selection. *Nat Immunol*, (2019).
980 59. Laird RM, Hayes SM. Roles of the Src tyrosine kinases Lck and Fyn in regulating
981 gammadeltaTCR signal strength. *PloS one* **5**, e8899 (2010).
982 60. Davidson D*, et al.* Genetic evidence linking SAP, the X-linked lymphoproliferative gene
983 product, to Src-related kinase FynT in T(H)2 cytokine regulation. *Immunity* **21**, 707-717
984 (2004).
985 61. Hayes SM, Li L, Love PE. TCR signal strength influences alphabeta/gammadelta lineage
986 fate. *Immunity* **22**, 583-593 (2005).
987 62. Edwards SC*, et al.* A population of proinflammatory T cells coexpresses alphabeta and
988 gammadelta T cell receptors in mice and humans. *J Exp Med* **217**, (2020).
989 63. Rigau M*, et al.* Butyrophilin 2A1 is essential for phosphoantigen reactivity by
990 gammadelta T cells. *Science* **367**, (2020).
991 64. Melandri D*, et al.* The gammadeltaTCR combines innate immunity with adaptive
992 immunity by utilizing spatially distinct regions for agonist selection and antigen
993 responsiveness. *Nat Immunol* **19**, 1352-1365 (2018).
994 65. Sperling AI, Decker DC, DiPaolo RJ, Stern DA, Shum A, Bluestone JA. Selective
995 expansion of Vgamma2-Vdelta7 TCR gammadelta cells in C57BL/6 mice is postnatal
996 and extrathymic. *J Immunol* **159**, 86-91 (1997).

997 66. Sim GK, Augustin A. Extrathymic positive selection of gamma delta T cells. V gamma
998 4J gamma 1 rearrangements with "GxYS" junctions. *J Immunol* **146**, 2439-2445 (1991).
999 67. Van Gassen S*, et al.* FlowSOM: Using self-organizing maps for visualization and
1000 interpretation of cytometry data. *Cytometry Part A : the journal of the International*
1001 *Society for Analytical Cytology* **87**, 636-645 (2015).
1002 8. Ewels P, Magnusson M, Lundin S, Käller M. MultiQC: summarize analysis results for
1003 multiple tools and samples in a single report. *Bioinformatics (Oxford, England)* **32**,
1004 (2016).
1005 69. Martin M. Cutadapt removes adapter sequences from high-throughput sequencing reads.
1006 *2011* **17**, 3 (2011).
1007 70. Patro R, Duggal G, Love MI, Irizarry RA, Kingsford C. Salmon provides fast and bias-
1008 aware quantification of transcript expression. *Nature methods* **14**, 417-419 (2017).
1009 71. Soneson C, Love MI, Robinson MD. Differential analyses for RNA-seq: transcript-level
1010 estimates improve gene-level inferences. *F1000Research* **4**, (2015).
1011 72. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for
1012 RNA-seq data with DESeq2. *Genome biology* **15**, 550 (2014).
1013 73. Zheng GXY*, et al.* Massively parallel digital transcriptional profiling of single cells.
1014 *Nature Communications* **8**, 1-12 (2017).
1015 74. Hao Y*, et al.* Integrated analysis of multimodal single-cell data. *Cell* **184**, 3573-3587
1016 e3529 (2021).
1017 75. Denisenko E*, et al.* Systematic assessment of tissue dissociation and storage biases in
1018 single-cell and single-nucleus RNA-seq workflows. *Genome biology* **21**, 130 (2020).
1019

1020 76. Stoeckius M*, et al.* Cell Hashing with barcoded antibodies enables multiplexing and
1021 doublet detection for single cell genomics. *Genome biology* **19**, 224 (2018).
1022 77. Hafemeister C, Satija R. Normalization and variance stabilization of single-cell RNA-seq
1023 data using regularized negative binomial regression. *Genome biology* **20**, 296 (2019).
1024 78. Marsh SE. scCustomize: Custom Visualizations & Functions for Streamlined Analyses of
1025 Single Cell Sequencing.) (2021).
1026 79. Squair JW*, et al.* Confronting false discoveries in single-cell differential expression. *Nat*
1027 *Commun* **12**, 5692 (2021).
1028 80. Wolf FA*, et al.* PAGA: graph abstraction reconciles clustering with trajectory inference
1029 through a topology preserving map of single cells. *Genome biology* **20**, 59 (2019).
1030 81. Wolf FA, Angerer P, Theis FJ. SCANPY: large-scale single-cell gene expression data
1031 analysis. *Genome biology* **19**, 15 (2018).
1032 82. Lefranc MP. IMGT, the International ImMunoGeneTics Information System. *Cold*
1033 *Spring Harb Protoc* **2011**, 595-603 (2011).
1034 83. Pereira P, Boucontet L. Rates of recombination and chain pair biases greatly influence
1035 the primary gammadelta TCR repertoire in the thymus of adult mice. *J Immunol* **173**,
1036 3261-3270 (2004).
1037 84. M S*, et al.* VDJtools: Unifying Post-analysis of T Cell Receptor Repertoires. *PLoS*
1038 *computational biology* **11**, (2015).
1039 85. J Y, N M, TL M, JM O. IgBLAST: an immunoglobulin variable domain sequence
1040 analysis tool. *Nucleic Acids Res* **41**, (2013).

1041

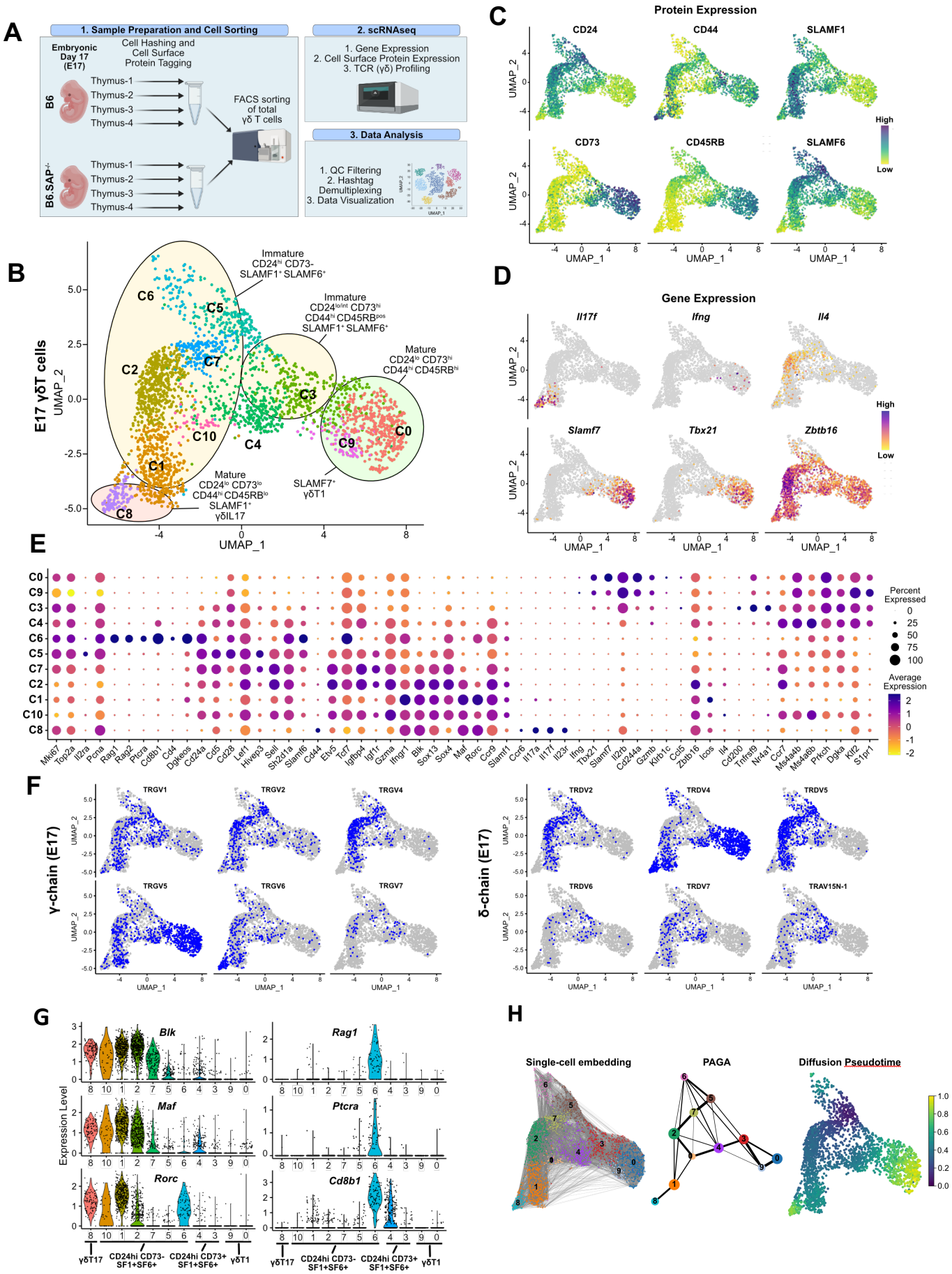


Figure 1

1042 **Figure 1. SLAM family receptor expression marks transcriptionally distinct developmental**
1043 **stages among E17 γδ T cells. (A)** Schematic workflow depicting the methodology for single-cell
1044 RNA sequencing (scRNAseq) library preparation and subsequent data analysis pipeline employed
1045 in this study. **(B)** Uniform Manifold Approximation and Projection (UMAP) visualization
1046 displaying 11 distinct clusters of E17 B6 thymic γδ T cells, n = 4 individual mice. Clusters are
1047 annotated based on comprehensive protein and gene expression data. **(C)** Feature plots illustrating
1048 the cell surface protein expression profiles of CD24, CD73, CD44, CD45RB, SLAMF1, and
1049 SLAMF6 on B6 E17 thymic γδ T cells. Each data point represents a cell, color-coded to indicate
1050 varying protein expression levels (high: dark blue, low: yellow). **(D)** Feature plot illustrating the
1051 gene expression profiles of signature genes among individual B6 E17 thymic γδ T cells. Each data
1052 point represents a cell, color-coded based on gene expression levels (high: purple, low: yellow).
1053 **(E)** Dot plot demonstrating the scaled expression levels of selected genes in E17 B6 thymic γδ T
1054 cells. Normalized expression levels are depicted using a color scale ranging from low (yellow) to
1055 high (purple). Dot size corresponds to the fraction of cells within each cluster expressing the
1056 specific marker. **(F)** UMAP representation of E17 B6 thymic γδ T cells indicating the expression
1057 of selected TRGV (TCRγ; left) and TRDV (TCRδ; right) chain V-segment usage (in blue) across
1058 individual cells. **(G)** Visualization of single-cell trajectories using PAGA (partition-based graph
1059 abstraction) with single-cell embedding (left) showing connectivity between individual nodes
1060 (middle). Weighted edges represent statistical measures of interconnectivity. The diffusion
1061 pseudotime plot (right) delineates inferred pseudotime progression of cells along developmental
1062 trajectories using cluster 5 (C5) as the root, highlighting their developmental order (from purple to
1063 yellow). **(H)** Violin plots illustrating the expression patterns of selected genes among E17 B6
1064 thymic γδ T cell clusters.

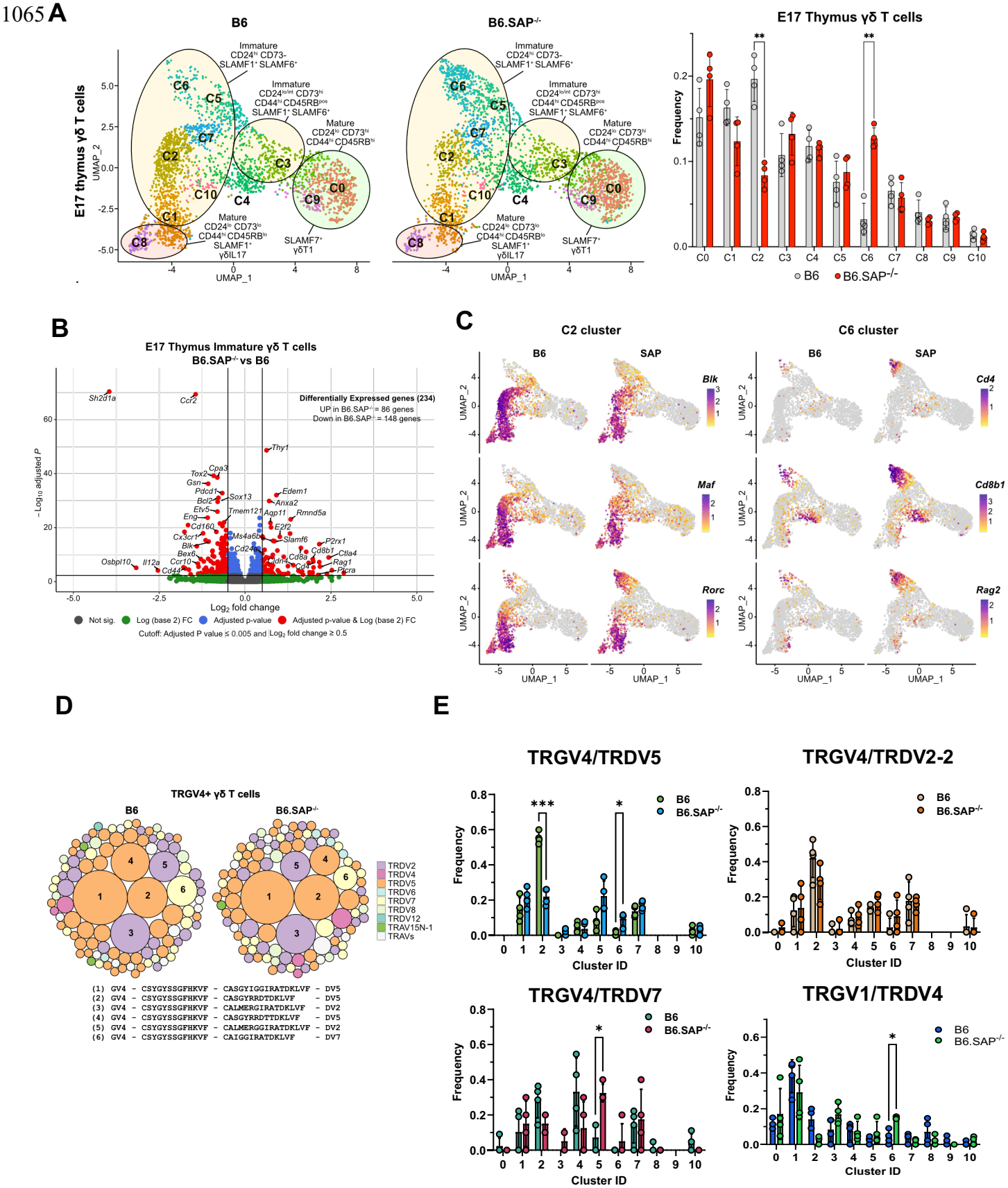

A
B6
B6.SAP-/-
E17 thymus γδ T cells
UMAP_1
UMAP_2
Immature CD24hi CD73- SLAMF1+ SLAMF6+
Immature CD24lo/int CD73hi CD44hi CD45RBpos SLAMF1+ SLAMF6+
Mature CD24lo CD73hi CD44hi CD45RBhi
Mature CD24lo CD73lo CD44hi CD45RBlo SLAMF1+ γδIL17
SLAMF7+ γδT1
E17 Thymus γδ T cells
Frequency
B6
B6.SAP-/-
B
E17 Thymus Immature γδ T cells
B6.SAP-/- vs B6
Differentially Expressed genes (234)
UP in B6.SAP-/- = 86 genes
Down in B6.SAP-/- = 148 genes
Log2 fold change
Not sig.
Log (base 2) FC
Adjusted p-value
Adjusted p-value & Log (base 2) FC
Cutoff: Adjusted P value ≤ 0.005 and Log2 fold change ≥ 0.5
C
C2 cluster
C6 cluster
B6
SAP
Blk
Maf
Rorc
Cd4
Cd8b1
Rag2
D
TRGV4+ γδ T cells
B6
B6.SAP-/-
TRDV2
TRDV4
TRDV5
TRDV6
TRDV7
TRDV8
TRDV12
TRAV15N-1
TRAVs
(1) GV4 - CSYGYSSGFHKVF - CASGYIGGIRATDKLVF - DV5
(2) GV4 - CSYGYSSGFHKVF - CASGYRRDTDKLVF - DV5
(3) GV4 - CSYGYSSGFHKVF - CALMERGIRATDKLVF - DV2
(4) GV4 - CSYGYSSGFHKVF - CASGYRRDTTDKLVF - DV5
(5) GV4 - CSYGYSSGFHKVF - CALMERGGIRATDKLVF - DV2
(6) GV4 - CSYGYSSGFHKVF - CAIGGIRATDKLVF - DV7
E
TRGV4/TRDV5
TRGV4/TRDV2-2
TRGV4/TRDV7
TRGV1/TRDV4
Frequency
Cluster ID

Figure 2

1066 **Figure. 2. Identification of SAP-dependent developmental checkpoints during E17 γδ T cell**
1067 **developmental programming. (A)** UMAP representation of B6 (left) and B6.SAP$^{-/-}$ (right) γδ T
1068 cells from E17 thymi (n = 4 mice per group) is shown at *left.* Clusters were annotated based on
1069 protein and gene expression data. The frequencies of B6 and B6.SAP$^{-/-}$ E17 γδ T cell clusters is
1070 shown at right. Bars represent the mean cluster frequency, error bars represent standard
1071 deviation, **$p \leq 0.01$, two-way ANOVA, Sidak multiple-comparisons test; n=4 mice/genotype.
1072 **(B)** Volcano plot of differentially expressed genes among immature CD24^{high}CD73^{neg} (clusters:
1073 C1, C2, C5, C6, C7, and C10) B6 and B6.SAP$^{-/-}$ γδ T cells using pseudobulk scRNAseq analysis.
1074 Genes exhibiting a $\log_2FC \geq 0.5$ or ≤ -0.5, and a $p_{adj} \leq 0.005$ in are shown in red. Only some
1075 selected genes are labelled for the sake of clarity. **(C)** Feature plots illustrating gene expression
1076 profiles of selected C2 and C6 cluster-specific genes among individual B6 and B6.SAP$^{-/-}$ E17
1077 thymic γδ T cells. Each point represents a cell, color-coded by gene expression level (high:
1078 purple, low: yellow). **(D)** TCR clonotype bubble plot displaying the top 100 B6 and B6.SAP$^{-/-}$
1079 TRGV4 clonotypes in E17 thymus. Bubbles represent unique TRGV4/TRDV clonotypes; size
1080 indicates frequency of clonotype among all Vγ4 T cells (n = 388 B6, 288 B6.SAP$^{-/-}$ cells), and
1081 colors denote specific TRDV chains utilized. Selected clonotypes are numbered and their
1082 respective CDR3γ and CDR3δ sequences displayed below. Data are **(E)** Altered TRGV/TRDV
1083 clonotype distribution between B6 and B6.SAP$^{-/-}$ E17 thymic γδ T cells. Frequencies of selected
1084 clonotypes among different clusters of B6 and B6.SAP$^{-/-}$ E17 γδ T cells is shown. Bars represent
1085 the mean clonotype frequency among the indicated TRGV/TRDV pairings shown, error bars
1086 represent standard deviation, *$p \leq 0.05$, ***$p \leq 0.001$, two-way ANOVA, Sidak multiple-
1087 comparisons test; n = 4 mice/genotype-age groups.
1088

1089

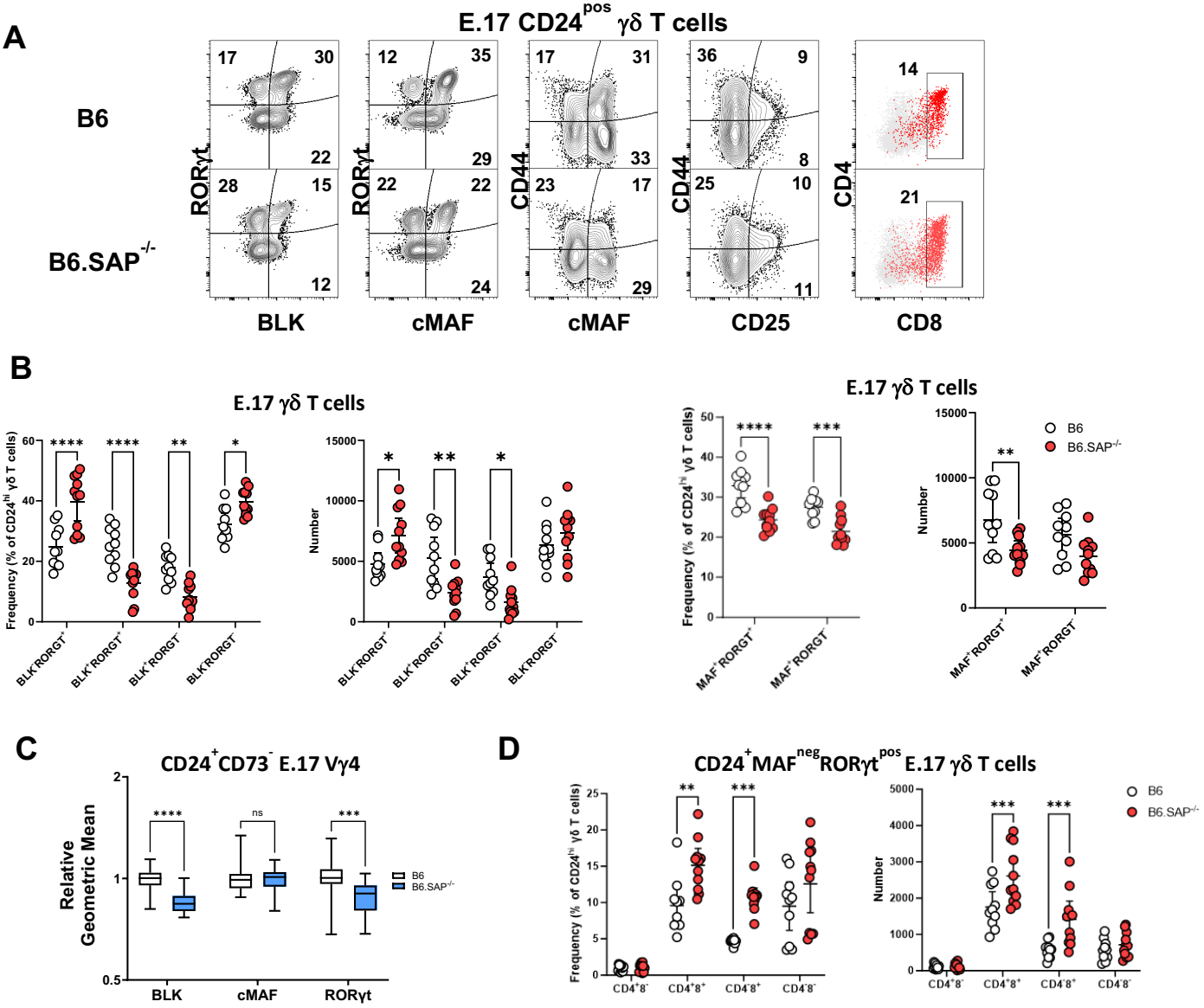


Figure 3

1090 **Figure 3. Increased frequency and numbers of immature $CD4^+CD8^+$c-$MAF^-$$ROR\gamma t^+$ thymic**
1091 **γδ T cells in B6.$SAP^{-/-}$ E17 γδ T cells. (A)** Representative contour plots depicting $CD24^{pos}$ E17
1092 γδ T cells in B6 (top) and B6.$SAP^{-/-}$ (bottom) mice. Concatenated data from 2 B6 and 3 B6.$SAP^{-/-}$
1093 embryos are shown, and are representative of 2 independent experiments. Numbers in the plots
1094 represent the frequency as a percentage of $CD24^{pos}$ γδ T cells. **(B)** Cumulative frequency and
1095 number of immature $CD24^{pos}$ BLK- (left), c-MAF- (right), RORγt-expressing E17 thymic γδ T
1096 cells. The mean and standard deviation are indicated. Data are the cumulative data from 2
1097 independent experiments, n = 10 B6, 11 B6.$SAP^{-/-}$ pooled embryonic thymi, *$p \leq 0.05$, **$p \leq 0.01$,
1098 ***$p \leq 0.001$, ****$p \leq 0.0001$, two-way ANOVA, Sidak's multiple comparisons test. **(C)** Relative
1099 geometric mean of BLK, c-MAF, and RORγt expression in B6 and B6.$SAP^{-/-}$ E17 $CD24^+CD73^-$
1100 Vγ4 γδ T cells, *$p \leq 0.001$, ****$p \leq 0.0001$, two-way ANOVA, Sidak's multiple comparisons test.
1101 Data are representative of 2 independent experiments, 7-9 mice per group. **(D)** Cumulative
1102 frequency (left) and number (right) of immature CD4- and CD8- $MAF^{neg}ROR\gamma t^{pos}$ E17 thymic γδ
1103 T cells in B6 and B6.$SAP^{-/-}$ thymus. The mean and standard deviation are indicated. Data are the
1104 cumulative data from 2 independent experiments, n = 10 B6, 11 B6.$SAP^{-/-}$ pooled embryonic
1105 thymi, **$p \leq 0.01$, ***$p \leq 0.001$, ****$p \leq 0.0001$, two-way ANOVA, Sidak's multiple
1106 comparisons test.
1107

1108

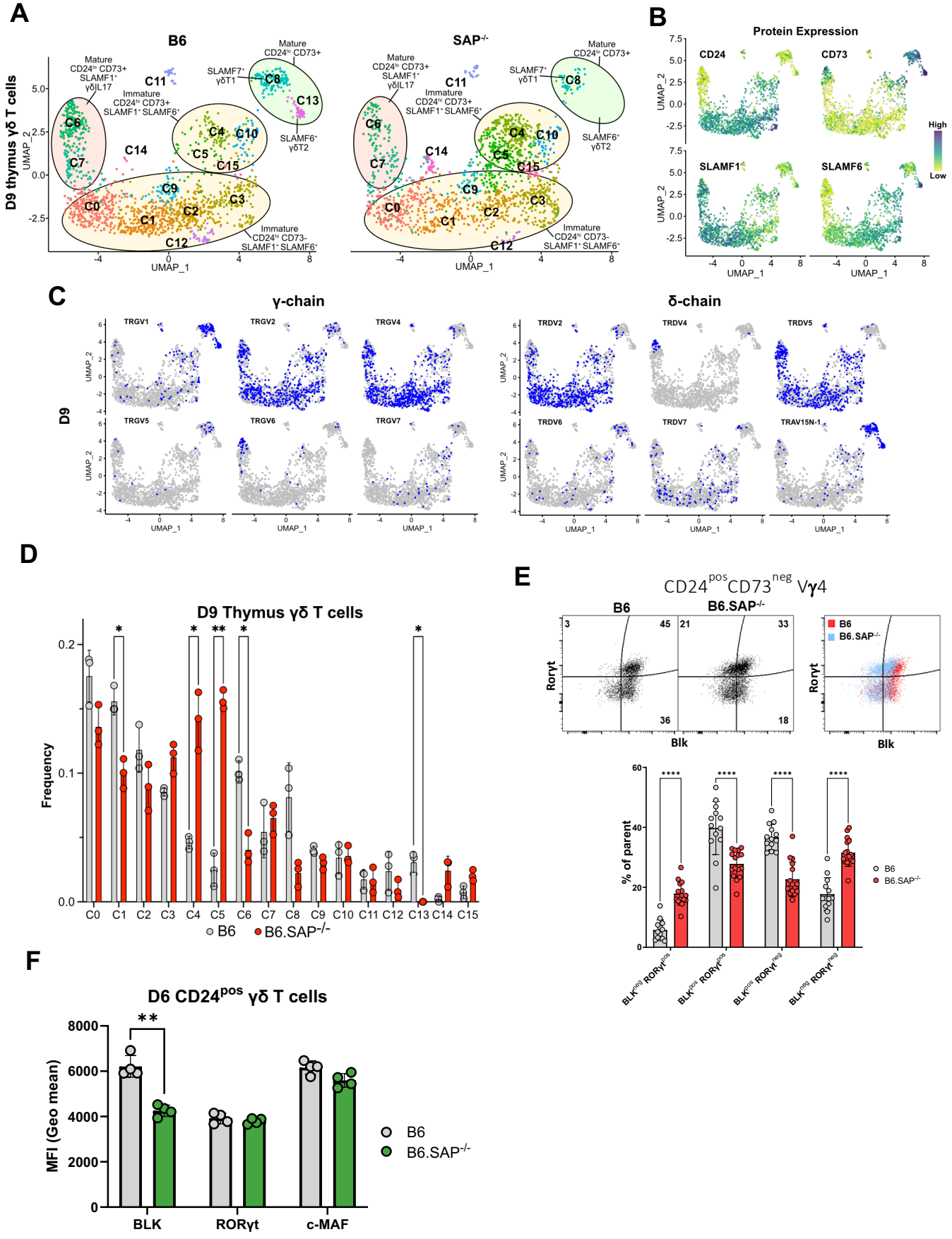
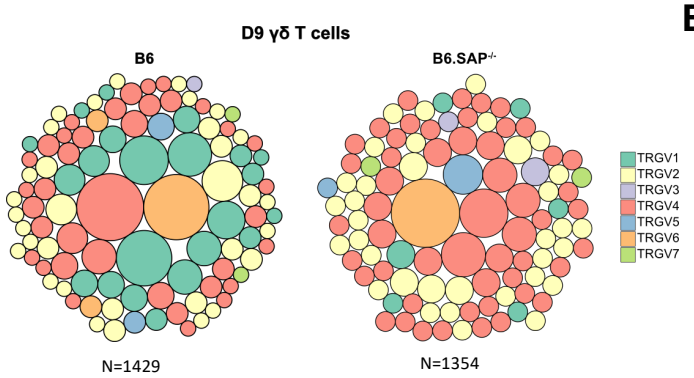


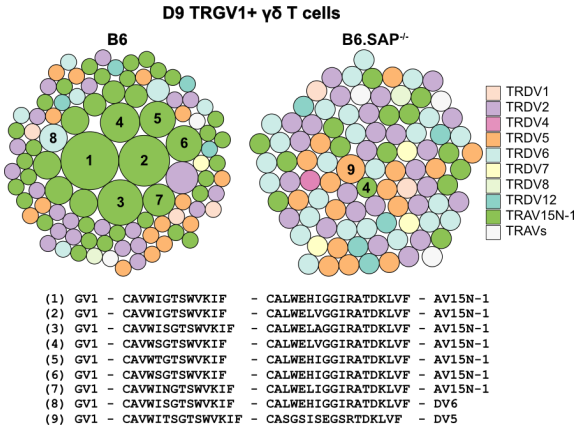
Figure 4

1109 **Figure 4. Identification of SAP-dependent developmental checkpoints during neonatal γδ T**
1110 **cell developmental programming. (A)** UMAP representation of B6 (left) and B6.SAP$^{-/-}$ (right)
1111 γδ T cells from D9 neonatal thymi (n = 3 mice per strain). Cluster annotation is based on
1112 comprehensive protein and gene expression data. **(B)** Feature plots displaying cell surface protein
1113 expression patterns of CD24, CD73, SLAMF1, and SLAMF6 on D9 B6 thymus γδ T cells. Data
1114 are color-coded based on protein expression level (high: dark blue, low: yellow). **(C)** UMAP
1115 representation of D9 B6 thymic γδ T cells exhibiting selected TRGV4 (left) and TRDV (right)
1116 chain V-segment usage (in blue) in individual cells. **(D)** Frequencies of B6 and B6.SAP$^{-/-}$ D9
1117 thymic γδ T cell clusters. Bars represent the mean cluster frequency, error bars represent standard
1118 deviation, *$p \leq 0.05$, **$p \leq 0.01$, two-way ANOVA, Sidak multiple-comparisons test, 3
1119 mice/strain. **(E)** Representative contour plots depicting BLK and RORγt expression in B6 and
1120 B6.SAP$^{-/-}$ CD24^{pos}CD73^{neg} D9 Vγ4 T cells are shown *above*. Cumulative frequencies of BLK and
1121 RORγt expression are shown *below*. The mean and standard deviation are indicated. Data are the
1122 cumulative data from 3 independent experiments, n = 13 B6, 16 B6.SAP$^{-/-}$ mice, ****$p \leq 0.0001$,
1123 two-way ANOVA, Sidak's multiple comparisons test. **(F)** Decreased BLK expression in immature
1124 B6.SAP$^{-/-}$ γδ T cells. Data represent the mean expression of the indicated proteins in D6 neonate
1125 BLK^{pos}, RORγt^{pos}, or c-Maf^{pos} thymic γδ T cells, **$p \leq 0.01$.
1126

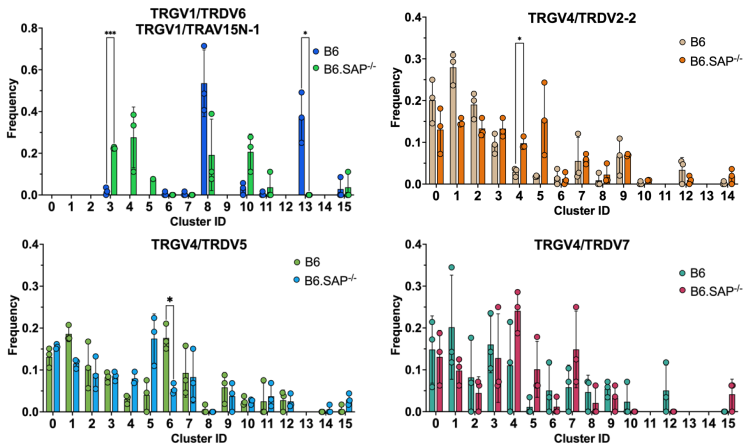
1127 **A**



B



C



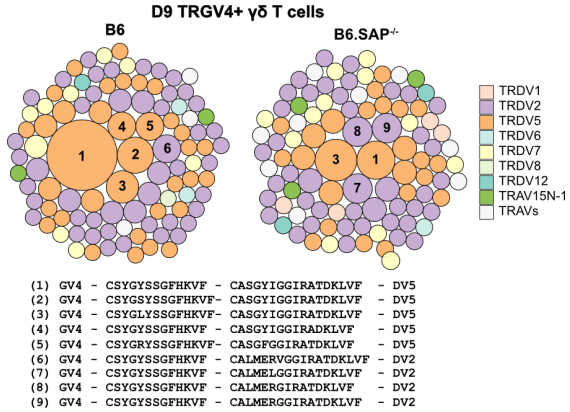


Figure 5

1128 **Figure 5. SAP shapes the neonatal γδ TCR repertoire. (A)** TCR clonotype bubble plots
1129 depicting the top 100 B6 and B6.SAP$^{-/-}$ TCR clonotypes among D9 thymic γδ T cells. Each bubble
1130 represents a unique clonotype, is sized according to its frequency as a percentage of all γδ T cells
1131 n = 1429 B6, 1354 B6.SAP$^{-/-}$ cells), and is colored based on specific TRGV chains utilized. **(B)**
1132 TCR clonotype bubble plots depicting the top 100 B6 and B6.SAP$^{-/-}$ TRGV1 (*above*) and TRGV4
1133 (*below*) clonotypes among D9 γδ T cells. Bubble size indicates clonotype frequency as a
1134 percentage of TRGV1^{+} (n = 198 B6, 98 B6.SAP$^{-/-}$) or TRGV4^{+} (n = 646 B6, 669 B6.SAP$^{-/-}$) cells,
1135 and colors correspond to specific TRDV chains. Selected clonotypes are numbered and their
1136 corresponding CDR3γ and CDR3δ sequences are displayed below. **(C)** Distribution of selected
1137 TRGV/TRDV clonotypes among B6 and B6.SAP$^{-/-}$ D9 thymic γδ T cell clusters. Data
1138 representative of 3 mice/strain, *$p \leq 0.05$, ***$p \leq 0.001$, two-way ANOVA, Sidak multiple-
1139 comparisons test.
1140

1141

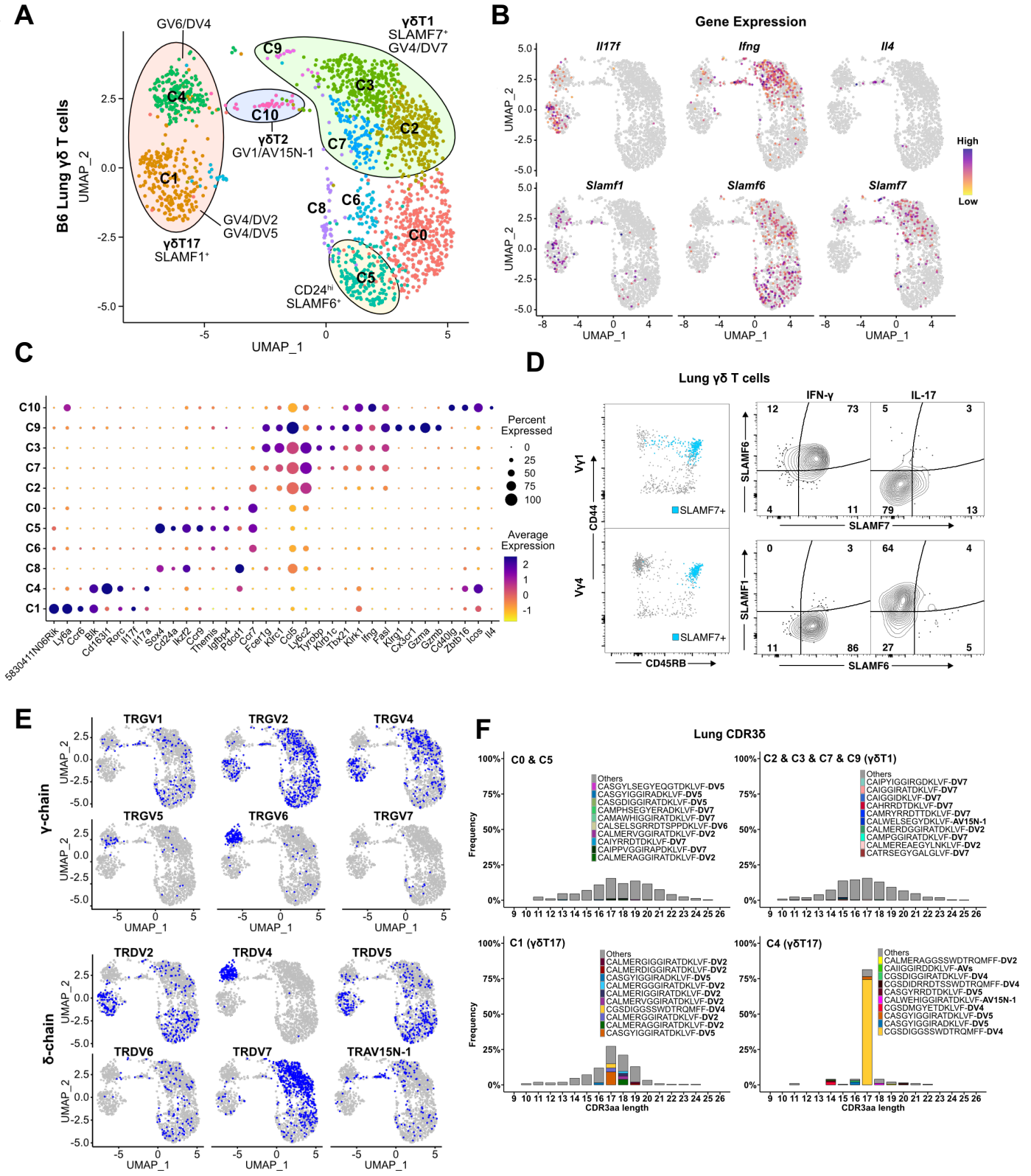


Figure 6

1142 **Figure 6. Restricted TCR repertoire in innate-like γδT1 as well as γδT17 in the periphery.**
1143 **(A)** UMAP representation of B6 adult lung γδ T cells. Clusters are annotated based on gene
1144 expression and TCR profiling data. **(B)** Feature plot displaying gene expression profiles of selected
1145 genes among individual B6 lung γδ T cells, color-coded based on gene expression levels. **(C)** Dot
1146 plot indicating scaled expression levels of selected genes in B6 lung γδ T cells, with dot size
1147 representing the fraction of cells within each cluster expressing the marker. **(D)** SLAMF6 and
1148 SLAMF7 co-expression marks $CD44^{+}CD45RB^{+}$ IFN-γ-producing cells in the periphery.
1149 Representative dot plots (left) of CD44 and CD45RB expression on lung Vγ1 (top) and Vγ4
1150 (bottom) γδ T cells. SLAMF7-expressing cells are highlighted in blue. Representative contour
1151 plots (right) of SLAMF1, SLAMF6, and SLAMF7 expression on IFN-γ^{+} and IL-17^{+} lung γδ T
1152 cells. **(E)** UMAP representation of B6 lung γδ T cells indicating selected TRG and TRD chain V-
1153 segment usage (blue). **(F)** Amino acid length distribution of lung CDR3δ clonotypes among
1154 immature C0/C5, γδT1 C2/C3/C7/C9, γδT17 C1, and γδT17 C4 clusters. The top 10 clonotypes in
1155 each group are color-coded, while all other clonotypes are shown in gray.

1156

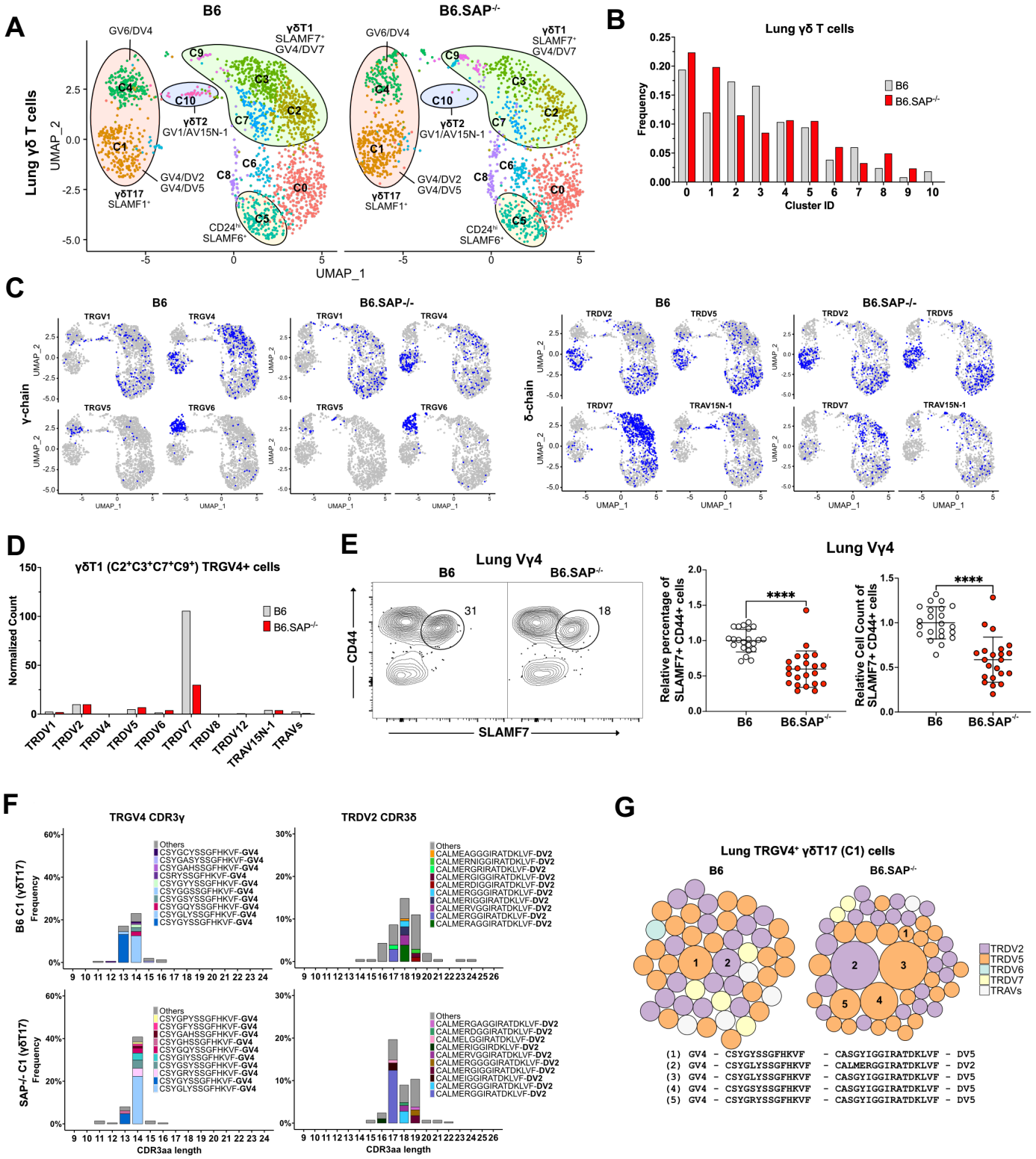


Figure 7

1157 **Figure 7. Specific decrease in Vγ4/Vδ7 γδT1 cells in the lungs of SAP-deficient mice. (A)**
1158 UMAP representation of B6 (left) and B6.SAP$^{-/-}$ (right) lung γδ T cells, pooled from 3 mice per
1159 strain. **(B)** SAP-dependent variation in lung γδ T cell cluster frequencies. The frequency of each
1160 B6 and B6.SAP$^{-/-}$ cluster as a percentage of all lung γδ T cells is shown. (**C**) UMAP representation
1161 of selected TRG and TRD chain V-segment usage (blue) among B6 and B6.SAP$^{-/-}$ lung γδ T cells.
1162 **(D)** SAP-dependent decrease in lung γδT1 TRGV4/TRDV7. Normalized TRDV counts in B6 and
1163 B6.SAP$^{-/-}$ lung TRGV4^{+} γδT1 clusters (C2/C3/C7/C9). **(E)** SAP-dependent decrease in SLAMF7^{+}
1164 lung Vγ4 T cells. Representative contour plots of CD44 and SLAMF7 expression in B6 (left) and
1165 B6.SAP$^{-/-}$ (right) lung Vγ4 γδ T cells are shown at *left*. Relative frequencies and counts of
1166 CD44^{+}SLAMF7^{+} lung Vγ4 cells is shown at *right*. Data represent the cumulative data from five
1167 independent experiments, ****$p < 0.0001$ using unpaired t-test. **(F)** SAP-dependent skewing of
1168 the lung γδT17 TCR repertoire. Amino acid length distributions of B6 and B6.SAP$^{-/-}$ TRGV4^{+}
1169 CDR3γ (left) and TRDV2^{+} CDR3δ (right) sequences in the lung Vγ4 γδT17 C1 cluster. The top
1170 10 clonotypes are color-coded and all other clonotypes are shown in gray. Bars represent the
1171 frequency of γδ T cells in the c1 cluster. **(G)** TCR clonotype bubble plots depicting the top 50 lung
1172 TRGV4 clonotypes in the C1 γδT17 cluster. Bubble size indicates clonotype frequency as a
1173 percentage of C1 Vγ4 (n = 62 B6, 104 B6.SAP$^{-/-}$) T cells, and colors correspond to specific TRDV
1174 chains. Selected clonotypes are numbered and their corresponding CDR3γ and CDR3δ sequences
1175 are displayed below.