Title: Single-Cell Antigen Receptor Sequencing in Pigs with Influenza

Weihong Gu [a,b,†] , Darling Melany de Carvahlo Madrid [a,b,†], Sadie Clements [a,b], Laurie Touchard [a,b], Nathan Bivins [c], Grant Zane [c], Mingyi Zhou [c], Kiho Lee [a] and John P. Driver [a,b,*]

[a] Division of Animal Sciences, University of Missouri, Columbia, MO, USA
[b] Bond Life Sciences Center, University of Missouri, Columbia, MO, USA
[c] Genomics Technology Core, University of Missouri, Columbia, MO, USA

[†]W.G. and D.M.d.C.M. contributed equally to this work.

[*]To whom correspondence may be addressed. Email: driverjp@missouri.edu

Competing Interest Statement: Authors declare that they have no competing interests.

Keywords: Pigs, single cell RNA sequencing, T cell receptor, B cell receptor, influenza A virus natural killer T cells, genetically edited swine, lung cells

24   **Abstract**

25   Understanding the pulmonary adaptive immune system of pigs is important as respiratory

26   pathogens present a major challenge for swine producers and pigs are increasingly used to model

27   human pulmonary diseases. Single-cell RNA sequencing (scRNAseq) has accelerated the

28   characterization of cellular phenotypes in the pig respiratory tract under both healthy and diseased

29   conditions. However, combining scRNAseq with recovery of paired T cell receptor (TCR) α and

30   β chains as well as B cell receptor (BCR) heavy and light chains to interrogate their repertoires has

31   not to our knowledge been demonstrated for pigs. Here, we developed primers to enrich porcine

32   TCR α and β chains along with BCR κ and λ light chains and IgM, IgA, and IgG heavy chains

33   that are compatible with the 10x Genomics VDJ sequencing protocol. Using these pig-specific

34   assays, we sequenced the T and B cell receptors of cryopreserved lung cells from *CD1D*-

35   expressing and -deficient pigs after one or two infections with influenza A virus (IAV) to examine

36   whether natural killer T (NKT) cells alter pulmonary TCR and BCR repertoire selection. We also

37   performed paired single-cell RNA and receptor sequencing of FACS-sorted T cells longitudinally

38   sampled from the lungs of IAV-vaccinated and -infected pigs to track clonal expansion in response

39   to IAV exposure. All pigs presented highly diverse repertoires. Pigs re-exposed to influenza

40   antigens from either vaccination or infection exhibited higher numbers of expanded CD4 and CD8

41   T cell clonotypes with activated phenotypes, suggesting potential IAV reactive T cell populations.

42   Our results demonstrate the utility of high throughput single-cell TCR and BCR sequencing in

43   pigs.

44

45

Introduction

High throughput single-cell RNA sequencing (scRNAseq) technology has greatly increased our understanding of the phenotypic diversity and plasticity of immune cell types as well as their cellular interactions in complex tissues in both healthy and disease states (1, 2). In addition to revealing cellular heterogeneity through RNA expression, scRNAseq can be coupled with additional assays to enhance cellular phenotyping, such as enrichment of T cell receptor (TCR) and B cell receptor (BCR) repertoires using primers that target the C regions in mRNA transcripts of TCR and BCR chains and isotypes. This allows construction of expressed TCRs and BCRs of individual cells including genomic rearrangements between the variable (V), diversity (D), and joining (J) regions of the TCR and BCR intervals responsible for creating diversity in receptor binding surfaces. Pairing TCR/BCR sequencing with scRNAseq provides a powerful approach for studying the relationship between immune repertoire and many types of immune responses. Additionally, V(D)J recombination at the TCR and BCR loci can be used as endogenous barcodes to trace T and B cell clonotypes as they expand or transition though different states, including within the same individual over time (3, 4).

Domestic pigs (*Sus scrofa*) are an important agricultural species that are intensively farmed making them vulnerable to many infectious pathogens. Therefore, a thorough understanding of the porcine immune system is needed to optimize vaccine and drug design and to identify immune targets for increased disease resistance through selective breeding and genetic engineering. Because of their many similarities to humans, swine are increasingly used in place of non-human primate models (5, 6). However, a limitation that prevents fully exploiting these pig models is our incomplete understanding of the porcine immune system which is due in part to a scarcity of

68   immune profiling reagents for pigs. scRNAseq which does not require marker-based sorting of cell

69   subsets is helping to address this gap.

70   In the current work, we developed porcine-specific TCR and BCR primers compatible with

71   the droplet-based protocols of the 10x Genomics Next GEM Single Cell 5' sequencing protocol.

72   These assays were used to compare cryopreserved lung cells from pigs genetically engineered to

73   lack *CD1D*, which encodes an antigen presenting molecule required for the development of natural

74   killer T (NKT) cells, a subset of innate-like T cell that accumulates in barrier organs such as the

75   lungs (7-12). Pigs were analyzed after one or two infections with influenza A virus (IAV) to

76   interrogate the evolution of the TCR and BCR repertoires after primary or secondary infection and

77   to determine if NKT cells exert T helper cell functions that influence receptor diversity.

78   Additionally, we applied our protocol to a longitudinal assessment of T cells recovered from the

79   lung lavage fluid of infant pigs exposed to IAV vaccination and infection. Here we were able to

80   track clonal expansion and monitor changes in the frequency of clonotypes within the same pig.

81   Together, these results demonstrate the potential of TCR/BCR profiling to better understand a wide

82   variety T and B cell-related immune responses in pigs.

83

84

85

86

87

88

Results

Single-cell RNA sequencing analysis of influenza virus-infected genetically edited pigs

Cryopreserved cells liberated from the enzyme digested lungs of mixed-breed pigs carrying an inactive form of the *CD1D* (*CD1D-/-*) gene compared with littermates carrying one copy (*CD1D-/+*) were subjected to paired scRNAseq and scTCR/BCRseq using our custom primer sets. One cohort of pigs was necropsied five days after a single infection (1X) with H1N1 A/Missouri/CS20N08/2020 (MO20) (Figure 1A and Table S1). To compare heterologous adaptive immune responses, a second cohort of pigs was infected with H1N1 A/Missouri/CS20N08/2020 (MO20) virus two weeks after an initial infection with H1N1 A/California/04/2009 (pdmH1N1) and necropsied 5 days later (2X). While 1X pigs presented an increase in body temperature and shed more virus compared to 2X pigs, there was no difference in how the different genotypes responded to infection (Figure S1). Single-cell sequencing was performed on 12 pigs (3 per group), totaling 45,850 cells. A dimensionality reduction analysis identified 28 clusters by Uniform Manifold Approximation and Projection (UMAP) that were annotated according to established lineage markers (Figure 1B and Figure S2). The cell types most impacted by the number of times pigs were infected were CD8$^+$ tissue resident memory T cells (TRM – cluster 4), cytotoxic CD8$^+$ T cells (cluster 5), and cycling T cells (cluster 16) that were higher in 2X than 1X pigs, and CD2$^-$ γδ T cells (cluster 9), and a subset of natural killer cells (NK2) cells (cluster 11) that were higher in 1X than 2X pigs (Figure 1C). The proportions of most cell types were comparable between genotypes with the exception that 2X *CD1D-/-* pigs had more CD4$^+$ TRM cells (cluster 3) and fewer cytotoxic CD8$^+$ T cells (cluster 5) after two infections compared to 2X *CD1D-/+* pigs. Next, we compared differentially expressed genes within individual cell types between 1X and 2X pigs by genotype (Figure 1D) or between *CD1D-/+* and *CD1D-/-* pigs by number of infections (Figure

112 1E). We found similar numbers of differentially expressed genes between *CD1D-/-* and *CD1D-/+*

113 pigs and between 1X and 2X pigs. However, a cluster of monocytes and macrophages (cluster 17)

114 from 2X *CD1D-/+* pigs had more upregulated genes than 2X *CD1D-/-* pigs. An ingenuity pathway

115 analysis of canonical cellular immune response networks identified that several of these

116 upregulated genes fell within "Interferon Signaling" (*IFI35, IFI6, IFIT1, IFNAR2, ISG15, MX1,*

117 *PSMB8, STAT1, TAP1*) and "Antigen Presentation Pathway" (*CD74, HLA-DRA, PDIA3, PSMB8,*

118 *PSMB9, TAP1*) pathways, suggesting that stimuli originating from NKT cells may have altered the

119 maturation of these antigen presenting cells.

120 TCR repertoire of lung T lymphocytes

121 We analyzed the pig lung tissue cells using our pig-specific V(D)J primers designed for

122 compatibility with the 10x Genomics Chromium Next GEM Single-cell 5' kit (Table S2). Over the

123 12 samples, we obtained 927,503,456 sequence reads an average of 77,291,955 reads per library,

124 Data file 1). A de novo assembly of raw sequencing reads produced unannotated porcine V(D)J

125 contigs which were identified using the sequencing primers and thereafter aligned to cells in the

126 gene expression clusters in Figure 1. Across all samples, approximately 60% and 70% of identified

127 TRA or TRB contigs were respectively aligned to single cells (Figure 2A, Data file 1). Assembled

128 VDJ sequences were blasted against the international ImMunoGeneTics (IMGT) germline TRBV,

129 TRBD, TRBJ, and TRAJ databases (13). Approximately 70% of TRB contigs in a pairing with

130 TRA contigs mapped to an annotated TRVB or TRJB gene (Data file 1). Because Vα genes are

131 not annotated in IMGT, Vα sequences were assigned according to pig TRAV sequences from our

132 previous publication (14) (Figure S3, Data file 1). Only cells with IMGT-annotated TRB genes

133 that were paired with productive TRA were used for further analysis (Figure 2B). These cells

134   accounted for between 23% and 38% of cells within αβ T cell clusters (clusters 1-6) across

135   samples.

136   The expression of Vβ/Jβ and Vα/Jα combinations was analyzed to determine if favored

137   rearrangements were correlated with *CD1D* genotype or the number of IAV infections (Figure 2C

138   and Figure S3). Overall, *TRBJ3-3*01*, *TRBJ1-2*01*, and *TRBJ1-1*01* were used in many

139   rearrangements, while pairings with *TRBJ1-5*01*, *TRBJ2-5*01*, *TRBJ3-7*01*, *TRBJ1-6*01*, and

140   *TRBJ2-1*01* were relatively rare (Figure 2C). This is similar to our previous analysis of pig

141   peripheral blood T cells that used a bulk RNA sequencing approach (14). We observed that 2X

142   *CD1D-/+* pigs had Vβ/Jβ recombinations that were more similar to each other than to other pigs

143   (Figure 2D). This was also the case for pigs s303, s305, s307, and s308, which were from the same

144   litter. Since structurally rearranged TCRs are selected by major histocompatibility complex (MHC)

145   molecules, we analyzed swine leukocyte antigen (SLA) class I and II expression in individual pigs

146   by aligning scRNAseq transcripts across clusters (Figure S4) to the IPD-MHC database that

147   provides a repository of MHC sequences for a number of species, including swine (13). To

148   visualize SLA usage, a row-scaled mean expression of the 10 highest class I (Figure 2E) and class

149   II (Figure 2F) genes in each pig was performed. Taking the expression of both SLA classes

150   together, we could group pigs into 6 distinct MHC haplotype inheritance patterns (designated A-

151   F) as follows: A – s403, s406, s408; B – s303, s307, s308, s407, s412; C – s305; D – s402; E –

152   s411; F – s404. Vβ/Jβ pairings in the five B inheritance pattern pigs clustered together (Figure

153   2D), suggesting a link between MHC inheritance and VDJ selection.

154   Next, since TRB gene segments are fully annotated in IMGT whereas TRAV segments are

155   not, we used identical CDR3β region sequences to analyze our samples for expanded clonotypes

156   (Data file 1). Some clones were found in more than one pig, with the highest numbers of shared

157 clones among the 2X infected pigs (Figure 2G). Additionally, 2X pigs, especially the *CD1D-/+*

158 group, had more expanded clones than the 1X groups (Figure 2H). The highest concentrations of

159 clonally expanded T cells were among CD4$^+$ and CD8$^+$ TRMs, cytotoxic CD8$^+$ T cells, and cycling

160 T cells while comparatively few expanded clones were present among peripheral-derived T cells

161 in clusters 1 and 2 (Figure 2I). Expanded clones were enriched for immune activation/effector

162 related genes compared to unexpanded clones (Figure 2J). 2X *CD1D-/+* pigs contributed four of

163 the five most expanded clonotypes, all of which included more than twelve cells (Figure 2K and

164 2L). One of these clonotypes originated from CD4$^+$ TRMs while the remaining four were from

165 CD8$^+$ TRMs (Figure 2M).

BCR repertoire of lung B lymphocytes

166

167 Using a similar approach to our TCR profiling, we analyzed the pig lung tissue cells using

168 our pig-specific primers for the IGK, IGL, IGHM, IGHA, and IGHG genes (Table S2), which

169 respectively encode the immunoglobulin κ and λ light chains and IgM, IgA, and IgG heavy chains.

170 Over the 12 samples, we obtained 1,395,728,419 sequence reads an average of 11,631,0701 reads

171 per library (Figure 3A, Data file 1). Assembled V(D)J sequences were blasted against the IMGT

172 germline IGL, IGK, and IGH databases (13). Only cells with IMGT-annotated IGH genes that were

173 paired with productive IGL or IGK chains were used for further analysis. The percentage of B cells

174 expressing different light and heavy chain contigs are as follows, 21% IGK, 40% IGL, 52% IGHM,

175 5% IGHG, and 3% IGHA. These cells accounted for between 53% and 70% of B cells among the

176 12 samples.

177 The overall IGL to IGK ratio was 1.8 (Figure 3B). Prior studies have reported that IGL:IGK

178 usage in circulating immature pig B cells is ~1:1(15-17). Plasma B cells (cluster 15) had the highest

179 proportions of IGHG+ and IGHA+ cells consistent with B cells that have undergone diversification

180   for mucosal antibody secretion (Figure 3C). Since our IGHG primers targeted a conserved constant

181   region of the eight genomic IgG constant region gene sequences cataloged in IMGT (*IGHG1,*

182   *IGHG2, IGHG3, IGHG4, IGHG5-1, IGHG5-2, IGHG6-1, IGHG6-2*) (18-20), we mapped IGHG

183   sequences to the IMGT database. *IGHG1* dominated all three clusters (Figure 3D). *IGHG5-2* was

184   also detected in every B cell subset. However, only antigen-experienced memory B cells and

185   plasma cells in clusters 14 and 15 exhibited *IGHG2*, *IGHG3*, and *IGHG5-1* subclasses. *IGHG4*

186   and *IGHG6* could not be distinguished from the other subclasses.

187   Next, we analyzed V(D)J gene segment usage (Figure 3E, Figure S5 and Figure S6). Unlike

188   humans or laboratory rodents, pigs use a limited number of IGHV, IGHD, and IGHJ genes to form

189   most of their BCR repertoire, with >90% of the VDJ repertoire made up of seven IGHV genes,

190   two IGHD segments, and a single IGHJ segment (21-23). Consistent with these studies, we found

191   six IGHV genes (*IGHV1-15, IGHV1S2, IGHV1-4, IGHV1-8, IGHV1S5, IGHV1-6*) accounted for

192   ~70% of IGHV usage (Figure 3E and Figure S6) and that IGHD and IGHJ usage was largely

193   restricted to *IGHD1* or *IGHD2* and *IGHJ5*.

194   We analyzed our samples for expanded clonotypes defined by the recombination of

195   identical IG light and heavy chain CDR3 sequences (Data file 1). This revealed far fewer expanded

196   B cells compared to the T cell compartment, with no single clonotype containing more than four

197   cells. Almost all expanded clonotypes were among the small number of IgG or IgA secreting

198   plasma cells (cluster 15) (Figure 3F and 3G). No striking differences in IGH usage (Figure 3H)

199   and clonotype expansion (Figure 3I) were observed among the four treatment groups. However,

200   the ratio of *IGHG5-2*/*IGHG1* was greater in 2X than 1X pigs (Figure 3J) while heavy chain CDR3

201   length tended to be greater in 1X than 2X pigs (Figure 3K).

202   **Longitudinal assessment of T cells from lung lavage fluid**

203   Clonotype tracking is a powerful approach to monitor changes in the frequency of

204 clonotypes of interest in cancer and vaccine immunology within a single individual (24-28). Using

205 our swine TCR α and β chain primers, we analyzed T cells for V(D)J clonotypes and TCR α and

206 β chain gene usage from the lungs of three specific pathogen free raised piglets (Figure 4A). Two

207 pigs (FLU1 and FLU2) were vaccinated with a combination of inactivated pdmH1N virus and

208 adjuvant at 28 days of age and boosted 15 days later. Both pigs were infected with live pdmH1N

209 virus 2 weeks after the booster. Lung fluid was collected 3 days before infection (T1) and 7 days

210 after infection (T2). The third pig (NAIVE) was lavaged at 28 (T1) and 33 (T2) days of age.

211 Between 1,041 and 6,789 (average = 3,472) CD3$^+$ cells sorted from each lung fluid sample (Figure

212 4B) were subjected to paired single-cell RNA and TCR sequencing. The combined dataset

213 separated into 15 clusters (Figure 4C and 4D). Clusters 10 and 11 were $\gamma\delta$ T cells that are relatively

214 abundant in pigs. Clusters 1, 2, 3, 4, 5, 6, 7, 11, 12, 13, 14 presented tissue residency markers,

215 while clusters 8, 9, and 10 expressed circulating T cell markers (Figure 4E). Tissue resident

216 memory T cells were composed of both CD4$^+$ (cluster 1) and CD8$^+$ cells (clusters 2, 4, 5, 6, 12) as

217 well as proliferating cells that contained a mixture of CD4$^+$ and CD8$^+$ TRMs (cluster 13, 14). In

218 response to IAV infection, both FLU1 and FLU2 pigs presented an increase in the frequency of

219 proliferating CD4$^+$ and CD8$^+$ TRMs, CD2$^-$ $\gamma\delta$ T cells, CD4$^+$ TRMs, peripheral CD4$^+$ T cells, naïve

220 CD8$\alpha\alpha$ T cells, and CD8$^+$ tissue effector memory (TEM) clusters, and a decrease in the frequency

221 of CD8$^+$ TRMs (Figure 4F). We identified 10 modules of co-regulated genes and regulatory

222 networks across cell types (Figures 4G and 4H, Data file 1) among which Module 5 in clusters 7,

223 8, 9, 10 harbored naïve/circulating T cell genes, Module 6 in clusters 2, 3, 4, 5, 6, 11 harbored

224 tissue residency and cytotoxic genes, and Module 10 in clusters 1, 2, 11 harbored interferon

225 stimulated genes (ISG).

226   Approximately 60% of αβ T cells recovered had paired TCRα and TCRβ chains (Figure

227   5A). Sequence identity can be used to map the phylogenetic relatedness of TCRα and TCRβ chains

228   in individual samples (Figure 5B). Many clones, identified by CDR3β sequences, were present at

229   both T1 and T2 in the same animal, especially the FLU2 pig (Figure 5C). A high proportion of

230   expanded clones were CD8$^+$ TRMs (Figure 5D). The two vaccinated and infected pigs presented

231   the five most expanded clones, especially the FLU2 pig (Figure 5E and 5F). The low number of

232   expanded clonotypes present in the NAÏVE pig samples is due partly to the lower number of cells

233   we were able to collect from this pig. Next, we examined a larger collection of expanded

234   clonotypes in FLU1 and FLU2 pigs, before and after infection, to identify potential influenza-

235   reactive T cells (Figure 5G). While the frequency of several clones decreased from T1 to T2, some

236   increased, including CSAGERSNYEQIF, CASSVRSYPLNDLHF, CASSFGGVHTGQLYF,

237   CAWSTTGTVTGQLYF, and CSAGEGGFGDTCFF. We also analyzed specificity groups within

238   the CDR3β repertoire using immunarch (29, 30), a program that enables clustering of TCRs with

239   an increased probability of sharing antigen specificity due to conserved motifs of CDR3 sequences.

240   Two CDR3β 4-mer motif patterns (ASSL and SSLV) were found to be enriched in FLU1 and FLU2

241   pig samples (Figure 5H). Interestingly, a few of the CDR3β sequences in expanded clones were

242   identical or differed by a single amino acid from curated human TCR CDR3β sequences which

243   recognize IAV epitopes in M1, NP, PB1, and PB2 (31). For example, the seventh most expanded

244   CDR3β sequences in all pigs, ASSPGQGYEQ, matches a human CDR3β sequence that recognizes

245   the immunodominant IAV Matrix protein 1 epitope GILGFVFTL when presented by human HLA-

246   A*0201 (Figure 5I) (32, 33). This may arise because peptide binding motifs of some common

247   swine SLA molecules partly overlap with the binding motifs of human HLA molecules, including

248   a number of HLA-A*0201-restricted IAV peptides (34).

Discussion

The current study describes assays for single-cell TCR/BCR sequencing in pigs which presents a useful tool for enhancing effective vaccine and therapeutic design to protect swine health and to increase the potential of pigs as biomedical models for studying human immune-related physiological processes and diseases. We examined the utility of the assay in the setting of the pulmonary immune response against IAV infection since influenza is a respiratory pathogen of major importance for both humans and swine (35). One set of samples were from cryopreserved lung sections of *CD1D*-expressing and -deficient pigs after one IAV infection or two infections with heterologous IAVs. This is of interest since NKT cell effector responses are important for anti-IAV immunity in mice, including that NKT cell-deficient mouse strains are significantly more susceptible to IAV infections than standard mice (36-39). Moreover, NKT cells are capable of a wide array of CD4 T helper cell functions, which have the capacity to elicit wide-ranging cellular and humoral responses that can substantially boost the quality and durability of immune responses, including against heterologous and heterosubtypic IAV infections (40). Among the 12 samples in this dataset, we detected almost all VDJ gene segments annotated for TCR β chains, including some that were annotated as pseudogenes. In addition, most of the Vα genes detected overlapped with TCR α chains that we identified in a previous analysis of TCR chain usage using bulk RNAseq (14).

While we found that cells which expressed both α and β TCR chains were mostly within αβ T cell clusters (clusters 1, 2, 3, 4, 5), we also detected cells expressing unpaired α and β chains in some non-αβ T cell clusters, such as CD2⁻ γδ T cells that expressed TCR β chains and NK cells and B cells that expressed TCR α chains. This is consistent with prior reports that a significant proportion of peripheral γδ T cells express TCR β (41) chains and that NK cells express germline

272 TCR transcripts (42). The NK cell observation is supported by the fact that virtually all NK cell

273 TRA transcripts matched to TRAJ segments but very few to TRAV segments. Examination of

274 Vβ/Jβ combinations confirmed previous studies showing that certain rearrangements are favored

275 in pigs (14, 43). In several pigs, VDJ rearrangements clustered by MHC inheritance pattern. We

276 also found that VDJ rearrangements preferred by 2X *CD1D-/+* pig clonotypes clustered together

277 which could mean that IAV infection influenced Vβ/Jβ selection in NKT cell-intact pigs.

278 The most expanded clones were among CD4$^+$ and CD8$^+$ TRMs, cytotoxic CD8$^+$ T cells,

279 and proliferating T cells, consistent with reports that these populations harbor antigen experienced

280 T cells that are poised for rapid responses during an IAV infection (44). As expected, 2X pigs had

281 more expanded clones than 1X pigs. This was less apparent in *CD1D-/-* compared to *CD1D-/+*

282 pigs, which might suggest that induction of IAV-specific lung T cells is reduced in the absence of

283 NKT cell helper functions.

284 Lung tissue B cells consisted of naïve B cells, memory B cells, and a small population of

285 plasma cells. Plasma cells displayed the highest diversity in immunoglobulin heavy chain usage

286 and clonally expanded B cells. Consistent with prior reports (15-18, 20, 21, 23, 45, 46), we found

287 (i) a mixture of IGL+ and IGK+ B cells, (ii) IgG heavy chain usage was dominated by the IGHG1

288 subclass, and (iii) a limited number of IGHV, IGHD, and IGHJ genes formed most of the BCR

289 repertoire. Pigs are interesting in so far as their pathway of antibody repertoire development has

290 evolved somewhat differently from mice and humans, including that they possess a highly

291 streamlined IGH gene complex, which contains IGHV genes that all belong to a single ancestral

292 IGHV3 family, and that only one IGHJ segment is functional (18, 20, 21, 23, 45). Because of the

293 small number of IGHV, IGHD, and IGHJ segments used, the combinatorial diversity in pigs is

294 comprised of a mere ~14 possibilities compared to ~9,000 in humans (23).

295   TCR transcript capture was more efficient in longitudinally collected lavage fluid T cell

296   samples than in the *CD1D* lung tissue samples, probably owing to the former consisting entirely

297   of FACS-purified T cells. As in lung tissue cells, most αβ T cell clusters exhibited paired α and β

298   TCR chains whereas a significant fraction of CD2$^-$ γδ T cells expressed TCR β chains and CD2$^+$

299   γδ T cells and CD8αα T cells were enriched for TCR α chains. Clusters with the highest numbers

300   of expanded clones were again tissue resident memory T cell populations. Our ability to detect a

301   substantial number of the same T cell clones at two different timepoints in the same individual

302   shows the utility of TCR profiling for monitoring changes in clonotypes of interest in pigs. The

303   higher number of expanded clonotypes within FLU1 and FLU2 pig samples, some of which

304   overlapped, compared to the NAÏVE pig, may be the result of IAV exposure by vaccination and

305   infection leading to an increase in antigen experienced T cells in lung tissue and/or because fewer

306   cells were collected in the NAÏVE pig. While there is a lack of reagents to identify influenza-

307   specific T cell clones in commercial pig breeds, this information can to some extent be inferred by

308   studying clonotypes in vaccinated pigs that increase after virus exposure. Using this approach, we

309   identified CDR3 sequences that are potentially reactive to IAV antigens, including a clone with a

310   CDR3β sequence identical to a human motif that recognizes an immunodominant epitope from

311   Matrix protein 1 (34).

312   In summary, the assays presented in this study can easily be applied to 5' 10x Genomics

313   protocols for use in swine. Our protocols can be employed to profile TCRs and BCRs in the same

314   sample which enhances the utility of the method as most adaptive immune responses involve both

315   cellular and humoral responses. Accordingly, the combined protocol could shed light on acquired

316   immunity that develops in response to vaccination and infection in production pigs, as well as in

317   the growing number of immune-related pig models being developed for biomedical use.

318 **Methods**

319 **Pigs**

320   The National Swine Resource and Research Center (NSRRC) at the University of Missouri

321 bred a boar homozygous for a *CD1D* gene deletion with two sows heterozygous for the *CD1D*

322 deletion that were full sisters to produce piglets that were homozygous a the *CD1D* deletion

323 (*CD1D-/-*) as well as heterozygous segregants (*CD1D-/+*). Our previously described *CD1D*

324 breeding herd (47) is on a commercial Large White crossbred background and maintained under

325 specific pathogen free conditions. The *CD1D* genotypes of pigs were determined by PCR and flow

326 cytometry as previously described (47, 48). Piglets used for longitudinal assessment of T cells

327 from lung lavage fluid were commercial Large White crossbred background pigs provided by the

328 NSRRC.

329 **Virus infection and sample collection**

330   Eight *CD1D-/-* and eight *CD1D-/+* pigs were transferred to biocontainment rooms at 4

331 weeks of age after being confirmed seronegative for IAV nucleoprotein antibodies by ELISA

332 developed by at the Veterinary Diagnostic Laboratory at the Iowa State University. At day 0, 4

333 *CD1D-/-* and 3 *CD1D-/+* pigs were intratracheally infected with 1×10^6 tissue culture infectious

334 dose ($TCID_{50}$) of H1N1 A/California/04/2009 (pdmH1N1) in 2mL of DMEM (Gibco, Brooklyn,

335 NY) after sedation with a combination of midazolam, butorphanol, and xylazine. Fourteen days

336 later, all 16 pigs were sedated and infected with 1×10^6 $TCID_{50}$ of H1N1

337 A/Missouri/CS20N08/2020 (MO20) IAV. Pigs were measured for clinical disease signs and daily

338 nasal swab virus titers as previously described (49). Five days later, all pigs were sedated and

339 euthanized by pentobarbital sodium intracardiac injections (70 mg/kg of body weight). At

340   necropsy, the lungs were removed from the thoracic cavity for tissue collection. and euthanized.

341   Cells were isolated from 3 animals of each genotype from single (1X) and twice (2X) infected pigs

342   for scRNAseq and receptor profiling. Litter, sex, *CD1D* genotype, and MHC inheritance pattern

343   of the piglets used for sequencing are described in (Table S1).

344   For the study to collect T cells through successive lung lavages, two pigs (FLU1 and FLU2)

345   were intramuscularly vaccinated with a combination of $1x10^6$ TCID$_{50}$ of ultraviolet-inactivated

346   pdmH1N1 virus and an oil-in water adjuvant (Emulsigen, 1:5 vaccine volume) at 28 days of age

347   and boosted 15 days later. Both pigs were intratracheally infected with $1x10^6$ TCID$_{50}$ live

348   pdmH1N1 virus 2 weeks after the booster. Lung fluid was collected 3 days before and 7 days after

349   infection. A third unvaccinated, uninfected pig (NAÏVE) was lavaged at 28 and 33 days of age.

350   Pigs were intratracheally sedated with midazolam, butorphanol, and xylazine to perform lung

351   lavages and infections. Lavages involved inserting a size 10 French catheter attached to a syringe

352   into the lung after which the lung was flushed twice with 5 ml of sterile saline solution. Recovered

353   lung fluid was ejected into 10 ml phosphate buffered solution containing 10% fetal bovine serum

354   (FBS).

355   The studies were in accordance with the University of Missouri's Institutional Animal Care

356   and Use Committee (protocol number 34343) and Institutional Biosafety Committee (protocol

357   number 17320).

358   **Tissue sampling and cell isolation**

359   Approximately 1 g of tissue collected from the left cranial, middle, and caudal lung lobes,

360   were combined, and then digested with 2.5 mg/mL of Liberase TL (Roche, Indianapolis, IN) in

361   Dulbecco's Modified Eagle Medium (Thermo Fisher, Waltham, MA) at 37°C for 45 minutes. The

362   digested tissue was dispersed into single cells as previously described (50), and then passed

363   through a 70 μm cell strainer (Thermo Fisher, Waltham, MA). Cells were immediately

364   cryopreserved in freezing media [90% FBS, 10% dimethylsulfoxide (DMSO)] in temperature-

365   controlled freezing containers at 3×10^7 cells per/mL and stored at -80°C until use. Samples were

366   thawed in thawing media (RPMI –HyClone, Logan, UT, - 20% FBS), resuspended in RPMI with

367   10% FBS, filtered through a 40μm filter (Bel-Art SP Scienceware, Wayne, NJ), and counted for

368   viable cells using a Countess 3 automated cell counter (Thermo Fisher, Waltham, MA). To obtain

369   T cells from lung washes, lavage fluid was filtered first through a 70 μm cell strainer (Thermo

370   Fisher, Waltham, MA) and then a 40 μm pipette tip Flowmi® cell strainer (SP Scienceware,

371   Warminster, PA), washed in RPMI with 10 % FBS, counted, stained with PE-Cy7-conjugated anti-

372   porcine CD3 antibody (clone BB23-8E6-8C8; BD Biosciences, San Jose, CA) and propidium

373   iodide viability dye, and sorted for live CD3 positive cells using a BD FACSMelody Cell Sorter

374   (BD Biosciences, San Jose, CA). Approximately 10,000 cells from cryopreserved lung tissue and

375   between 1,041 and 6,789 T cells from lung lavage fluid were loaded onto the 10x Chromium

376   controller (10x Genomics, Pleasanton, CA).

377   **Primer design**

378   Pig-specific V(D)J primers were designed according to guidelines from the 10x Genomics

379   Chromium Next GEM Single-cell 5' V2 user guide, which is well described in a recent publication

380   on scTCRseq in dogs(51). This protocol employs a nested PCR design which involves two rounds

381   of V(D)J amplification using the same forward primer that primes off the 5' Illumina adapter

382   sequence which is annealed to the 10x barcoded cDNA during the conversion from mRNA. The

383   first round of amplification uses the 5' forward primer in combination with a 3' outer reverse

384   primer that matches the C region of the targeted chain. The second round uses the same forward

385   primer in combination with a second reverse primer that primes the C region at an inner 5' region

386   from the outer reverse primer. To design the 3' reverse primer sets, C-gene sequences were

387   identified for TRA, TRB, IGH, and IGL and IGK transcripts. Inner and outer primers were

388   respectively designed to target regions between 50-200 and 200-300 base pairs away from the 5'

389   region of the C region. Primer candidates were selected based on the percent of GC content, similar

390   melting temperature range, and minimal predicted interaction with other regions in the pig genome.

391   The final primer sequences as well as accession numbers for the DNA and rearranged mRNA

392   transcripts used to identify the C regions are listed in (Table S2).

393   **Single cell processing**

394   Libraries were constructed by following the manufacturer's protocol with reagents

395   supplied in the Chromium Next GEM Single Cell 5′ Kit v2 (10x Genomics). Briefly, cell

396   suspension concentration and viability were measured with a Cellometer K2 (Nexcelom

397   Biosciences, Lawrence, MA) stained with an acridine orange/propidium iodine dye mix

398   (Invitrogen, Waltham, MA). Cell suspension combined with reverse transcription master mix were

399   loaded on a Chromium Next GEM chip K along with gel beads and partitioning oil to generate gel

400   emulsions (GEMs). GEMs were transferred to a PCR strip tube and reverse transcription

401   performed on a Veriti thermal cycler (Applied Biosystems, Waltham, MA) at 53°C for 45 minutes.

402   All samples underwent 11 cycles of cDNA amplification upon which cDNA concentration and

403   quality were assessed using a Fragment Analyzer 5200 (Agilent, Santa Clara, CA). For the gene

404   expression library, up to 50 ng of the cDNA was fragmentated, end-repaired, A-tail added, and

405   ligation of sequencing adaptors was performed according to manufacturer specifications. V(D)J

406   libraries were constructed using pig-specific primers for two successive enrichments of TCR and

407   BCR transcripts. The TCR and BCR assay used primer pools containing 1.43 μM per gene specific

408   primer and 1.43 μM of the 10x forward primer. Nested PCR amplification of the TCR and BCR

409   sequences was performed using adapted mouse and human PCR protocols. This involved

410   amplifying 2 μL of cDNA in 100 μL total reaction volume using 50 ul Amp Mix (10x Genomics).

411   Steps involved in the first reaction were 98 ºC for 45 s for initial denaturation followed by 9 cycles

412   of 98 ºC for 20 s, 65 ºC for 30 s, and 72 ºC for 60 s. PCR product was purified using AxyPrep

413   MagPCR Clean-up beads (Axygen) and subjected to a second round of amplification using the

414   same cycling conditions to the first except a total of 8 cycles was used. The amplicons were

415   purified using AxyPrep MagPCR Clean-up beads (Axygen). Libraries were constructed from PCR

416   amplicons according to manufacturer specifications. The concentrations for all libraries were

417   measured with the Qubit HS DNA kit (Invitrogen) and fragment sizes were determined on a

418   Fragment Analyzer 5200 (Agilent). Libraries were pooled and sequenced on a NovaSeq 6000

419   (Illumina, San Diego, CA) to obtain paired end reads. The minimum sequencing depths targeted

420   were 40,000 reads per cell for 5' GEX libraries and 10,000 reads per cell for V(D)J libraries.

421   **Single-cell RNA sequencing data analysis**

422   The Sscrofa 11.1 genome assembly was used to align sequencing reads to generate gene

423   matrix data by Cell Ranger (v8.0.0). Clustering analyses were performed using Seurat

424   (v.4.4.0)(52). To filter out low-quality genes and cells, only genes expressed in more than 3 cells

425   and cells with more than 200 genes and less than 10% mitochondrial reads were included in the

426   analysis. Afterward, we followed a standard integration workflow to integrate samples. Briefly,

427   transcript counts were log normalized, and the top 2,000 most variable genes in each dataset were

428   identified using the *FindVariableFeatures* function. Then, the *SelectIntegrationFeatures* function

429   was applied to genes that were consistently variable across datasets. Next, the

430   *FindIntegrationAnchors* function identified a set of anchors between datasets using the top 30

431   dimensions from the canonical correlation analysis to specify the neighbor search space. Next, an

432   integrated dataset was created by running the *IntegrateData* function. Then, clustering analysis

433   workflow was performed using *RunPCA*, *FindNeighbours*, *FindClusters*, and *RunUMAP*. Cell

434   types were assigned based on the expression of known cell type markers (Figure S2).

435   **Single-cell V(D)J data analysis**

436   Single-cell TCR sequencing reads were assembled into contigs using cellranger vdj (10x

437   Genomics) pipeline in denovo mode rather than reference-based mode due to the incomplete

438   annotation of the pig germline Vα chain sequences. To identify the V(D)J chains, we searched

439   assembled contigs against inner-enrichment primers using the usearch_global command. Primer

440   matched TCR αβ and BCR IGH and IGL and IGK chains were selected and integrated with the

441   above cellular gene expression profiles. The TCR β chains with matched TCR α chains were

442   selected for downstream analysis. TRBV, TRBD, TRBJ, and CDR3 sequences were mapped to the

443   pig TRB reference in IMGT using the IMGT/V-QUEST sequence alignment software. Cells with

444   duplicated or multiple contigs were removed. Immunarch (v1.0.0) (29) was used to track clones

445   across samples and identify k-mers from CDR3 sequencing. Scirpy (v.0.12.0) was used to analyze

446   TCR β repertoires for clonal expansion (each unique CDR3) using the scirpy.pl.clonal_expansion

447   command (53). In addition, CDR3 sequences were searched within the Immune Epitope Database

448   (IEDB) (31) to find matches predicted to recognize epitope specificity using TCRMatch T cell

449   epitope prediction tool. The unique contigs of TCR α and β chains in each cell were aligned to

450   each other using CLUSTALW (https://www.genome.jp/tools-bin/clustalw), after which

451   phylogenetic trees were generated to infer the similarities among contigs using the package ape

452   (v5.7-1)(54). Using MMseqs2 (55), TRAJ sequences were mapped to the pig TRAJ reference in

453   IMGT while TRAV segments were annotated according to TRAV sequences deposited in in

454   GeneBank (55) that we previously named according to similar human TRAV genes (14). Similarly,

455   BCR repertoires were annotated using the available pig IGH and IGK/IGL reference in IMGT

456   using IMGT/V-QUEST(56). The reference mapped IGH chains and IGK/L chains were selected

457   for downstream analysis. The IGH constant segments were mapped to IMGT/GENE-DB database.

458   Scirpy was used to analyze clonal (the recombination of IGH and IGL CDR3s) expansion, gene

459   usage, and CDR3 sequencing length analyses.

460   **SLA alleles annotation from scRNA-seq data**

461   The Cellranger mkgtf was used to build the SLA alleles reference using the SLA FASTA

462   files downloaded from the IPD-MHC database(13). Afterwards, SLA reads were quantified using

463   Cellranger count command, SLA counts were log normalized, and then the average expression of

464   each SLA allele in each sample was computed with the AverageExpression function in Seurat. The

465   data were visualized in heatmap using ComplexHeatmap (v2.25.1) (57) and scaled for PCA

466   analysis.

467   **Data availability**

468   The sequencing data are available at Gene Expression Omnibus (accession GSE277475).

469   Processed single-cell RNA sequencing objects are available for online visualization at

470   https://singlecell.broadinstitute.org/single_cell/study/SCP2783   and

471   https://singlecell.broadinstitute.org/single_cell/study/SCP2779. All relevant data are available

472   from the authors.

473   **Funding**

476

References

1. Gomes T, Teichmann SA, Talavera-López C. Immunology Driven by Large-Scale Single-Cell Sequencing. *Trends Immunol* (2019) 40(11):1011-21. Epub 20191020. doi: 10.1016/j.it.2019.09.004.

2. Jovic D, Liang X, Zeng H, Lin L, Xu F, Luo Y. Single-Cell Rna Sequencing Technologies and Applications: A Brief Overview. *Clin Transl Med* (2022) 12(3):e694. doi: 10.1002/ctm2.694.

3. Miragaia RJ, Gomes T, Chomka A, Jardine L, Riedel A, Hegazy AN, et al. Single-Cell Transcriptomics of Regulatory T Cells Reveals Trajectories of Tissue Adaptation. *Immunity* (2019) 50(2):493-504.e7. Epub 20190205. doi: 10.1016/j.immuni.2019.01.001.

4. Lönnberg T, Svensson V, James KR, Fernandez-Ruiz D, Sebina I, Montandon R, et al. Single-Cell Rna-Seq and Computational Analysis Using Temporal Mixture Modelling Resolves Th1/Tfh Fate Bifurcation in Malaria. *Sci Immunol* (2017) 2(9). doi: 10.1126/sciimmunol.aal2192.

5. Lunney JK, Van Goor A, Walker KE, Hailstock T, Franklin J, Dai C. Importance of the Pig as a Human Biomedical Model. *Sci Transl Med* (2021) 13(621):eabd5758. Epub 20211124. doi: 10.1126/scitranslmed.abd5758.

6. Spencer TE, Wells KD, Lee K, Telugu BP, Hansen PJ, Bartol FF, et al. Future of Biomedical, Agricultural and Biological Systems Research Using Domesticated Animals. *Biol Reprod* (2022). Epub 20220129. doi: 10.1093/biolre/ioac019.

7. Rey-Jurado E, Bohmwald K, Gálvez NMS, Becerra D, Porcelli SA, Carreño LJ, et al. Contribution of Nkt Cells to the Immune Response and Pathogenesis Triggered by Respiratory Viruses. *Virulence* (2020) 11(1):580-93. doi: 10.1080/21505594.2020.1770492.

8. Yip KH, Papadopoulos M, Pant H, Tumes DJ. The Role of Invariant T Cells in Inflammation of the Skin and Airways. *Semin Immunopathol* (2019) 41(3):401-10. Epub 2019/04/15. doi: 10.1007/s00281-019-00740-9.

9. Ryu S, Park JS, Kim HY, Kim JH. Lipid-Reactive T Cells in Immunological Disorders of the Lung. *Front Immunol* (2018) 9:2205. Epub 2018/09/26. doi: 10.3389/fimmu.2018.02205.

10. Schönrich G, Raftery MJ. Cd1-Restricted T Cells During Persistent Virus Infections: "Sympathy for the Devil". *Front Immunol* (2018) 9:545. Epub 2018/03/19. doi: 10.3389/fimmu.2018.00545.

11. Paget C, Trottein F. Role of Type 1 Natural Killer T Cells in Pulmonary Immunity. *Mucosal Immunol* (2013) 6(6):1054-67. Epub 2013/10/09. doi: 10.1038/mi.2013.59.

12. Tessmer MS, Fatima A, Paget C, Trottein F, Brossay L. Nkt Cell Immune Responses to Viral Infection. *Expert Opin Ther Targets* (2009) 13(2):153-62. doi: 10.1517/14712590802653601.

13. Barker DJ, Maccari G, Georgiou X, Cooper MA, Flicek P, Robinson J, et al. The Ipd-Imgt/Hla Database. *Nucleic Acids Res* (2022) 51(D1):D1053-D60. doi: 10.1093/nar/gkac1011.

14. Yang G, Artiaga BL, Lomelino CL, Jayaprakash AD, Sachidanandam R, Mckenna R, et al. Next Generation Sequencing of the Pig Aβ Tcr Repertoire Identifies the Porcine Invariant Nkt Cell Receptor. *J Immunol* (2019) 202(7):1981-91. Epub 2019/02/18. doi: 10.4049/jimmunol.1801171.

15. Butler JE. Immunoglobulin Diversity, B-Cell and Antibody Repertoire Development in Large Farm Animals. *Rev Sci Tech* (1998) 17(1):43-70. doi: 10.20506/rst.17.1.1096.

16. Sinkora M, Stepanova K, Butler JE, Sinkora M, Jr., Sinkora S, Sinkorova J. Comparative Aspects of Immunoglobulin Gene Rearrangement Arrays in Different Species. *Front Immunol* (2022) 13:823145. Epub 20220211. doi: 10.3389/fimmu.2022.823145.

17. Sinkora M, Stepanova K, Sinkorova J. Immunoglobulin Light Chain K Precedes Λ Rearrangement in Swine but a Majority of Λ(+) B Cells Are Generated Earlier. *Dev Comp Immunol* (2020) 111:103751. Epub 20200523. doi: 10.1016/j.dci.2020.103751.

18. Butler JE, Wertz N, Deschacht N, Kacskovics I. Porcine Igg: Structure, Genetics, and Evolution. *Immunogenetics* (2009) 61(3):209-30. Epub 20081202. doi: 10.1007/s00251-008-0336-9.

19. Zhang M, Li Z, Li J, Huang T, Peng G, Tang W, et al. Revisiting the Pig Ighc Gene Locus in Different Breeds Uncovers Nine Distinct Ighg Genes. *J Immunol* (2020) 205(8):2137-45. Epub 20200914. doi: 10.4049/jimmunol.1901483.

20. Eguchi-Ogawa T, Toki D, Wertz N, Butler JE, Uenishi H. Structure of the Genomic Sequence Comprising the Immunoglobulin Heavy Constant (Ighc) Genes from Sus Scrofa. *Mol Immunol* (2012) 52(3-4):97-107. Epub 20120604. doi: 10.1016/j.molimm.2012.05.010.

21. Butler JE, Sun J, Navarro P. The Swine Ig Heavy Chain Locus Has a Single Jh and No Identifiable Igd. *Int Immunol* (1996) 8(12):1897-904. doi: 10.1093/intimm/8.12.1897.

22. Sun J, Hayward C, Shinde R, Christenson R, Ford SP, Butler JE. Antibody Repertoire Development in Fetal and Neonatal Piglets. I. Four Vh Genes Account for 80 Percent of Vh Usage During 84 Days of Fetal Life. *J Immunol* (1998) 161(9):5070-8.

23. Butler JE, Wertz N. The Porcine Antibody Repertoire: Variations on the Textbook Theme. *Front Immunol* (2012) 3:153. Epub 20120627. doi: 10.3389/fimmu.2012.00153.

24. Zhao XN, You Y, Cui XM, Gao HX, Wang GL, Zhang SB, et al. Single-Cell Immune Profiling Reveals Distinct Immune Response in Asymptomatic Covid-19 Patients. *Signal Transduct Target Ther* (2021) 6(1):342. Epub 20210916. doi: 10.1038/s41392-021-00753-7.

25. Schalck A, Sakellariou-Thompson D, Forget MA, Sei E, Hughes TG, Reuben A, et al. Single-Cell Sequencing Reveals Trajectory of Tumor-Infiltrating Lymphocyte States in Pancreatic Cancer. *Cancer Discov* (2022) 12(10):2330-49. doi: 10.1158/2159-8290.Cd-21-1248.

26. Xiong Q, Peng C, Yan X, Yan X, Chen L, Sun B, et al. Characteristics of Sars-Cov-2-Specific Cytotoxic T Cells Revealed by Single-Cell Immune Profiling of Longitudinal Covid-19 Blood Samples. *Signal Transduct Target Ther* (2020) 5(1):285. Epub 20201204. doi: 10.1038/s41392-020-00425-y.

27. Qu R, Kluger Y, Yang J, Zhao J, Hafler DA, Krause DS, et al. Longitudinal Single-Cell Analysis of a Patient Receiving Adoptive Cell Therapy Reveals Potential Mechanisms of Treatment Failure. *Mol Cancer* (2022) 21(1):219. Epub 20221214. doi: 10.1186/s12943-022-01688-5.

28. Pauken KE, Shahid O, Lagattuta KA, Mahuron KM, Luber JM, Lowe MM, et al. Single-Cell Analyses Identify Circulating Anti-Tumor Cd8 T Cells and Markers for Their Enrichment. *J Exp Med* (2021) 218(4). doi: 10.1084/jem.20200920.

29. Team I. Immunarch: An R Package for Painless Bioinformatics Analysis of T-Cell and B-Cell Immune Repertoires. *Zenodo* (2019) 10:5281.

30. Mark M, Reich-Zeliger S, Greenstein E, Reshef D, Madi A, Chain B, et al. A Hierarchy of Selection Pressures Determines the Organization of the T Cell Receptor Repertoire. *Front Immunol* (2022) 13:939394. Epub 20220729. doi: 10.3389/fimmu.2022.939394.

31. Vita R, Mahajan S, Overton JA, Dhanda SK, Martini S, Cantrell JR, et al. The Immune Epitope Database (Iedb): 2018 Update. *Nucleic Acids Res* (2019) 47(D1):D339-d43. doi: 10.1093/nar/gky1006.

32. Chen G, Yang X, Ko A, Sun X, Gao M, Zhang Y, et al. Sequence and Structural Analyses Reveal Distinct and Highly Diverse Human Cd8(+) Tcr Repertoires to Immunodominant Viral Antigens. *Cell Rep* (2017) 19(3):569-83. doi: 10.1016/j.celrep.2017.03.072.

33. Gotch F, Rothbard J, Howland K, Townsend A, McMichael A. Cytotoxic T Lymphocytes Recognize a Fragment of Influenza Virus Matrix Protein in Association with Hla-A2. *Nature* (1987) 326(6116):881-2. doi: 10.1038/326881a0.

34. Fan S, Wang Y, Wang X, Huang L, Zhang Y, Liu X, et al. Analysis of the Affinity of Influenza a Virus Protein Epitopes for Swine Mhc I by a Modified in Vitro Refolding Method Indicated Cross-Reactivity between Swine and Human Mhc I Specificities. *Immunogenetics* (2018) 70(10):671-80. Epub 20180710. doi: 10.1007/s00251-018-1070-6.

35. Ma W, Loving CL, Driver JP. From Snoot to Tail: A Brief Review of Influenza Virus Infection and Immunity in Pigs. *J Immunol* (2023) 211(8):1187-94. doi: 10.4049/jimmunol.2300385.

36. De Santo C, Salio M, Masri SH, Lee LY, Dong T, Speak AO, et al. Invariant Nkt Cells Reduce the Immunosuppressive Activity of Influenza a Virus-Induced Myeloid-Derived Suppressor Cells in Mice and Humans. *J Clin Invest* (2008) 118(12):4036-48. doi: 10.1172/JCI36264.

37. Kok WL, Denney L, Benam K, Cole S, Clelland C, McMichael AJ, et al. Pivotal Advance: Invariant Nkt Cells Reduce Accumulation of Inflammatory Monocytes in the Lungs and Decrease Immune-Pathology During Severe Influenza a Virus Infection. *J Leukoc Biol* (2012) 91(3):357-68. doi: 10.1189/jlb.0411184.

38. Paget C, Ivanov S, Fontaine J, Renneson J, Blanc F, Pichavant M, et al. Interleukin-22 Is Produced by Invariant Natural Killer T Lymphocytes During Influenza a Virus Infection: Potential Role in Protection against Lung Epithelial Damages. *J Biol Chem* (2012) 287(12):8816-29. doi: 10.1074/jbc.M111.304758.

39. Ishikawa H, Tanaka K, Kutsukake E, Fukui T, Sasaki H, Hata A, et al. Ifn-Γ Production Downstream of Nkt Cell Activation in Mice Infected with Influenza Virus Enhances the Cytolytic Activities of Both Nk Cells and Viral Antigen-Specific Cd8+ T Cells. *Virology* (2010) 407(2):325-32. doi: 10.1016/j.virol.2010.08.030.

40. Driver JP, de Carvalho Madrid DM, Gu W, Artiaga BL, Richt JA. Modulation of Immune Responses to Influenza a Virus Vaccines by Natural Killer T Cells. *Front Immunol* (2020) 11:2172. Epub 2020/10/20. doi: 10.3389/fimmu.2020.02172.

41. Bischof A, Park JH, Hünig T. Expression of T-Cell Receptor Beta-Chain Mrna and Protein in Gamma/Delta T-Cells from Euthymic and Athymic Rats: Implications for T-Cell Lineage Divergence. *Dev Immunol* (2000) 8(1):19-30. doi: 10.1155/2000/51534.

42. Stewart CA, Walzer T, Robbins SH, Malissen B, Vivier E, Prinz I. Germ-Line and Rearranged Tcrd Transcription Distinguish Bona Fide Nk Cells and Nk-Like Gammadelta T Cells. *Eur J Immunol* (2007) 37(6):1442-52. doi: 10.1002/eji.200737354.

43. Butler JE, Wertz N, Sun J, Sacco RE. Comparison of the Expressed Porcine Vbeta and Jbeta Repertoire of Thymocytes and Peripheral T Cells. *Immunology* (2005) 114(2):184-93. doi: 10.1111/j.1365-2567.2004.02072.x.

44. Tang J, Sun J. Lung Tissue-Resident Memory T Cells: The Gatekeeper to Respiratory Viral (Re)-Infection. *Curr Opin Immunol* (2023) 80:102278. Epub 20221222. doi: 10.1016/j.coi.2022.102278.

45. Butler JE, Zhao Y, Sinkora M, Wertz N, Kacskovics I. Immunoglobulins, Antibody Repertoire and B Cell Development. *Dev Comp Immunol* (2009) 33(3):321-33. Epub 20080918. doi: 10.1016/j.dci.2008.06.015.

46. Butler JE, Wertz N. Antibody Repertoire Development in Fetal and Neonatal Piglets. Xvii. Igg Subclass Transcription Revisited with Emphasis on New Igg3. *J Immunol* (2006) 177(8):5480-9. doi: 10.4049/jimmunol.177.8.5480.

47. Yang G, Artiaga BL, Hackmann TJ, Samuel MS, Walters EM, Salek-Ardakani S, et al. Targeted Disruption of Cd1d Prevents Nkt Cell Development in Pigs. *Mamm Genome* (2015) 26(5-6):264-70. doi: 10.1007/s00335-015-9564-0.

48. Whitworth KM, Lee K, Benne JA, Beaton BP, Spate LD, Murphy SL, et al. Use of the Crispr/Cas9 System to Produce Genetically Engineered Pigs from in Vitro-Derived Oocytes and Embryos. *Biol Reprod* (2014). doi: 10.1095/biolreprod.114.121723.

49. Madrid DMC, Gu W, Artiaga BL, Yang G, Loeb J, Hawkins IK, et al. Comparison of Oseltamivir and A-Galactosylceramide for Reducing Disease and Transmission in Pigs Infected with 2009 H1n1 Pandemic Influenza Virus. *Front Vet Sci* (2022) 9:999507. Epub 20221020. doi: 10.3389/fvets.2022.999507.

50. Artiaga BL, Yang G, Hutchinson TE, Loeb JC, Richt JA, Lednicky JA, et al. Rapid Control of Pandemic H1n1 Influenza by Targeting Nkt-Cells. *Sci Rep* (2016) 6:37999. Epub 2016/11/29. doi: 10.1038/srep37999.

51. Hoang MH, Skidmore ZL, Rindt H, Chu S, Fisk B, Foltz JA, et al. Single-Cell T-Cell Receptor Repertoire Profiling in Dogs. *Commun Biol* (2024) 7(1):484. Epub 20240422. doi: 10.1038/s42003-024-06174-w.

52. Stuart T, Butler A, Hoffman P, Hafemeister C, Papalexi E, Mauck WM, 3rd, et al. Comprehensive Integration of Single-Cell Data. *Cell* (2019) 177(7):1888-902.e21. Epub 20190606. doi: 10.1016/j.cell.2019.05.031.

53. Sturm G, Szabo T, Fotakis G, Haider M, Rieder D, Trajanoski Z, et al. Scirpy: A Scanpy Extension for Analyzing Single-Cell T-Cell Receptor-Sequencing Data. *Bioinformatics* (2020) 36(18):4817-8. doi: 10.1093/bioinformatics/btaa611.

54. Paradis E, Claude J, Strimmer K. Ape: Analyses of Phylogenetics and Evolution in R Language. *Bioinformatics* (2004) 20(2):289-90. doi: 10.1093/bioinformatics/btg412.

55. Steinegger M, Söding J. Mmseqs2 Enables Sensitive Protein Sequence Searching for the Analysis of Massive Data Sets. *Nat Biotechnol* (2017) 35(11):1026-8. Epub 20171016. doi: 10.1038/nbt.3988.

56. Lefranc MP. Imgt Databases, Web Resources and Tools for Immunoglobulin and T Cell Receptor Sequence Analysis, Http://Imgt.Cines.Fr. *Leukemia* (2003) 17(1):260-6. doi: 10.1038/sj.leu.2402637.

57. Gu Z. Complex Heatmap Visualization. *Imeta* (2022) 1(3):e43. Epub 20220801. doi: 10.1002/imt2.43.

Figure 1
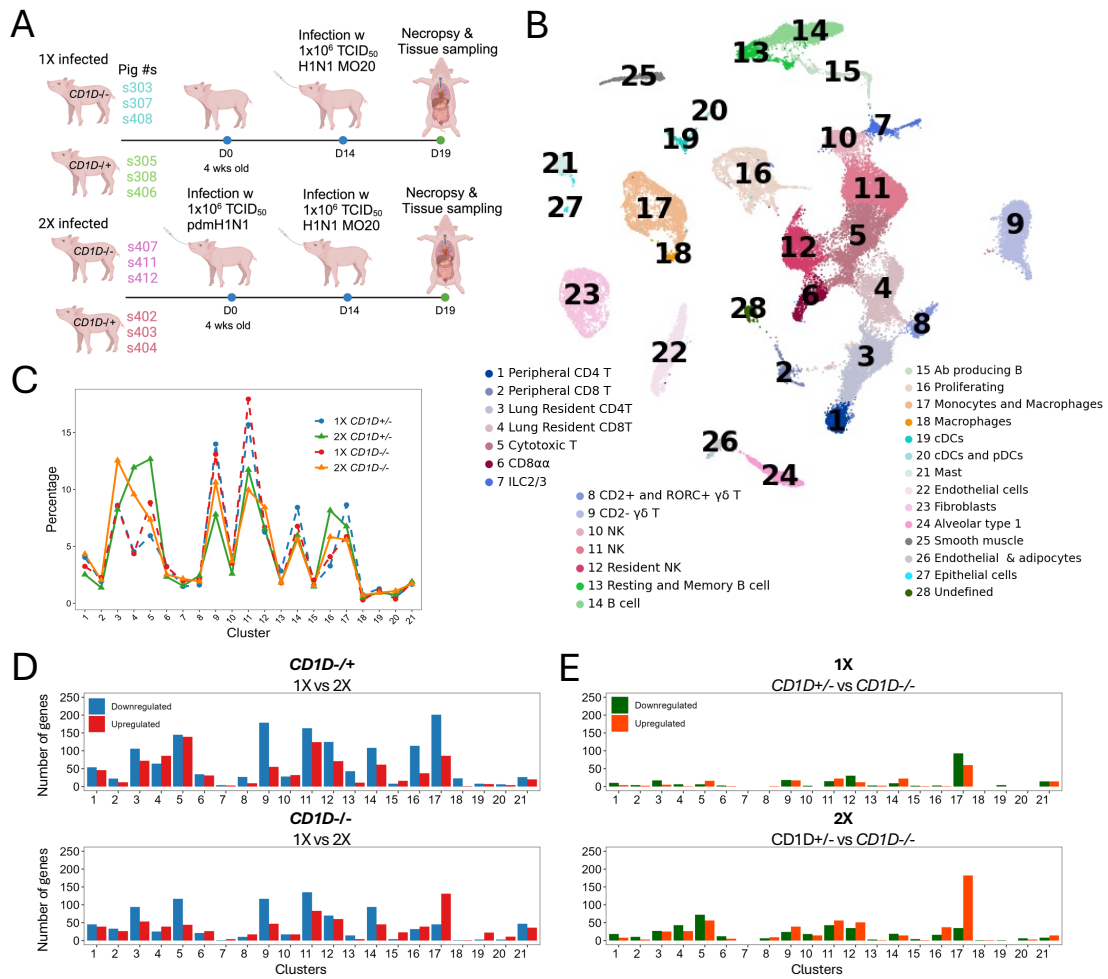
669   **Figure 1.** Single-cell transcriptomic analysis of IAV-infected *CD1D-/-* and *CD1D-/+* pig lungs.

670   (A) Overview of experiment setup. 3 *CD1D-/-* and 3 *CD1D-/+* pigs were infected with pdmH1N1

671   IAV. Fourteen days later, the same 6 pigs were infected with H1N1 A/Missouri/CS20N08/2020

672   (MO20) IAV (designated 2X pigs) along with an additional 3 *CD1D-/-* and 3 *CD1D-/+* pigs

673   (designated 1X pigs). Necropsies were performed 5 days after the MO20 infection to collect lung

674   tissue for single-cell immune profiling. Created with BioRender. (B) Uniform manifold

675   approximation and projection (UMAP) visualization of lung leukocyte populations colored by cell

676   clusters. Clusters were identified using the graph-based Louvain algorithm at a resolution of 0.5.

677   (C) The frequency of each cell type is presented for each treatment. (D) Bar graphs displaying the

678   number of upregulated and downregulated differentially expressed genes (DEGs) in 1X compared

679   to 2X *CD1D-/+* and *CD1D-/-* pigs. (E) Bar graphs displaying the number of DEGs in *CD1D-/+*

680   compared to *CD1D-/-* pigs after 1X or 2X infections.
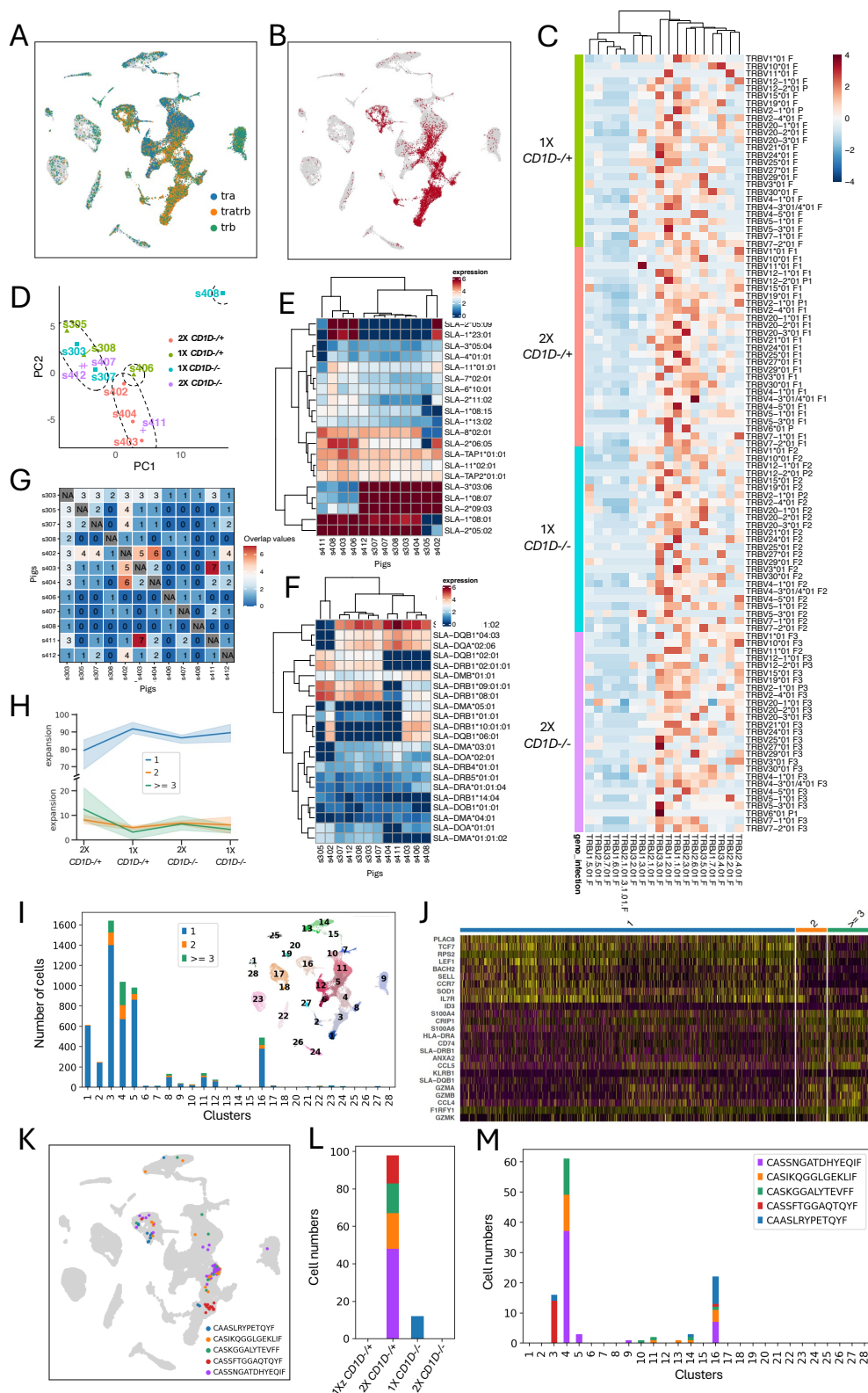
681

682

683

684

685

686

687

688

689

Figure 2

708   **Figure 2.** Characterization of T cell clonotypes from lung tissue. (A) UMAP plot of cells

709   expressing one (tra, trb) or both (tratrb) TCR α and TCR β chains. (B) UMAP plot of cells with

710   paired TCRα and TCRβ chains that were used for downstream analysis. (C) Relationship between

711   TRBV and TRBJ usage in T cell receptor rearrangements by treatment. Cell barcode counts for

712   TRBV and TRBJ gene segments were normalized by cell numbers across treatments and scaled

713   by TRBV segments. (D) Principal component analysis of TRBV and TRBJ gene usages by

714   individual pig. (E and F) Heatmaps showing row-scaled mean expression of the 10 highest

715   differentially expressed SLA class I (E) and SLA class II (F) genes per pig. (G) Heatmap of

716   overlapping clonotypes between pigs. (H) Proportion of clonotypes by treatment with 1, 2, ≥ 3

717   cells per clonotype. (I) Abundance of clonotypes by cluster with 1, 2, ≥ 3 cells per clonotype in

718   the combined dataset. (J) Expression of naive and activation T cell markers in expanded and non-

719   expanded clonotypes in the combined dataset. (K) UMAP plot displaying the five most expanded

720   clonotypes defined by identical CDR3β region sequences in the combined dataset. (L and M)

721   Number of cells in each of the five most expanded clonotypes by treatment (L) and cluster (M).
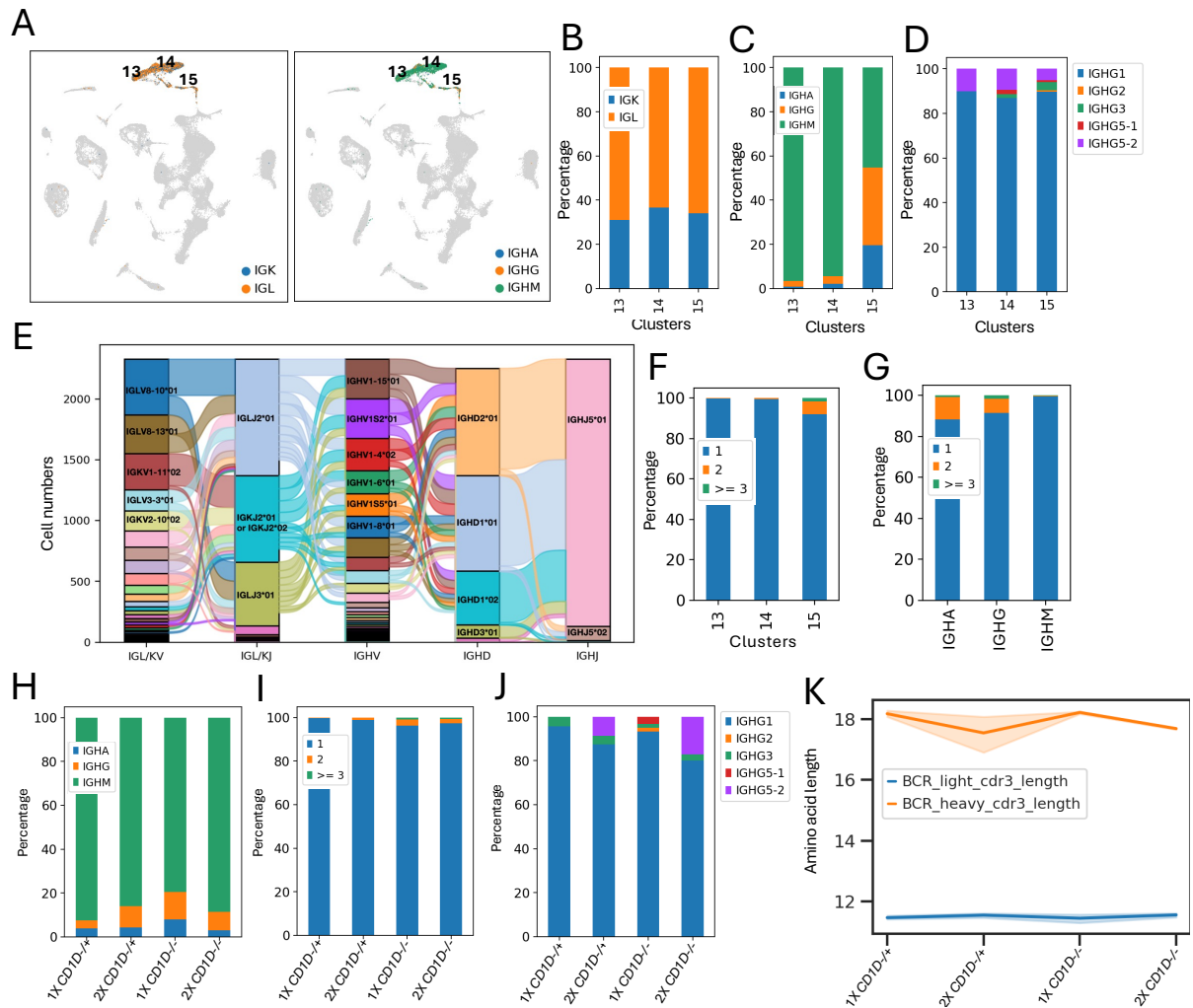
722

723

724

725

726

727

728

729   **Figure 3**

730

731

747   **Figure 3.** B cell receptor repertoire profiling of lung tissue B cells. (A) UMAP plots of cells

748   expressing IGK and IGL chains and IGHM, IGHG, and IGHA chains. (B) Proportion of cells

749   expressing IGK and IGL in each B cell cluster. (C) Percentage of cells expressing IGHM, IGHG,

750   and IGHA in each B cell cluster. (D) Percentage of IGHG+ cells expressing different porcine IGHG

751   subclasses in each B cell cluster. (E) Number of cells expressing light and heavy chain V(D)J gene

752   segments. (F) Percentage of clonotypes by cluster with 1, 2, ≥ 3 cells per clonotype. (G) Percentage

753   of clonotypes by IGH chain with 1, 2, ≥ 3 cells per clonotype. (H) Proportion of cells expressing

754   IGHM, IGHG, and IGHA by treatment. (I) Percentage of clonotypes by treatment with 1, 2, ≥ 3

755   cells per clonotype. (J) Percentage of IGHG+ cells expressing different porcine IGHG subclasses

756   by treatment. (K) Length of light and heavy chain CDR3 sequences by treatment.

757

758

759

760

761

762

763

764

765

766

767 **Figure 4**
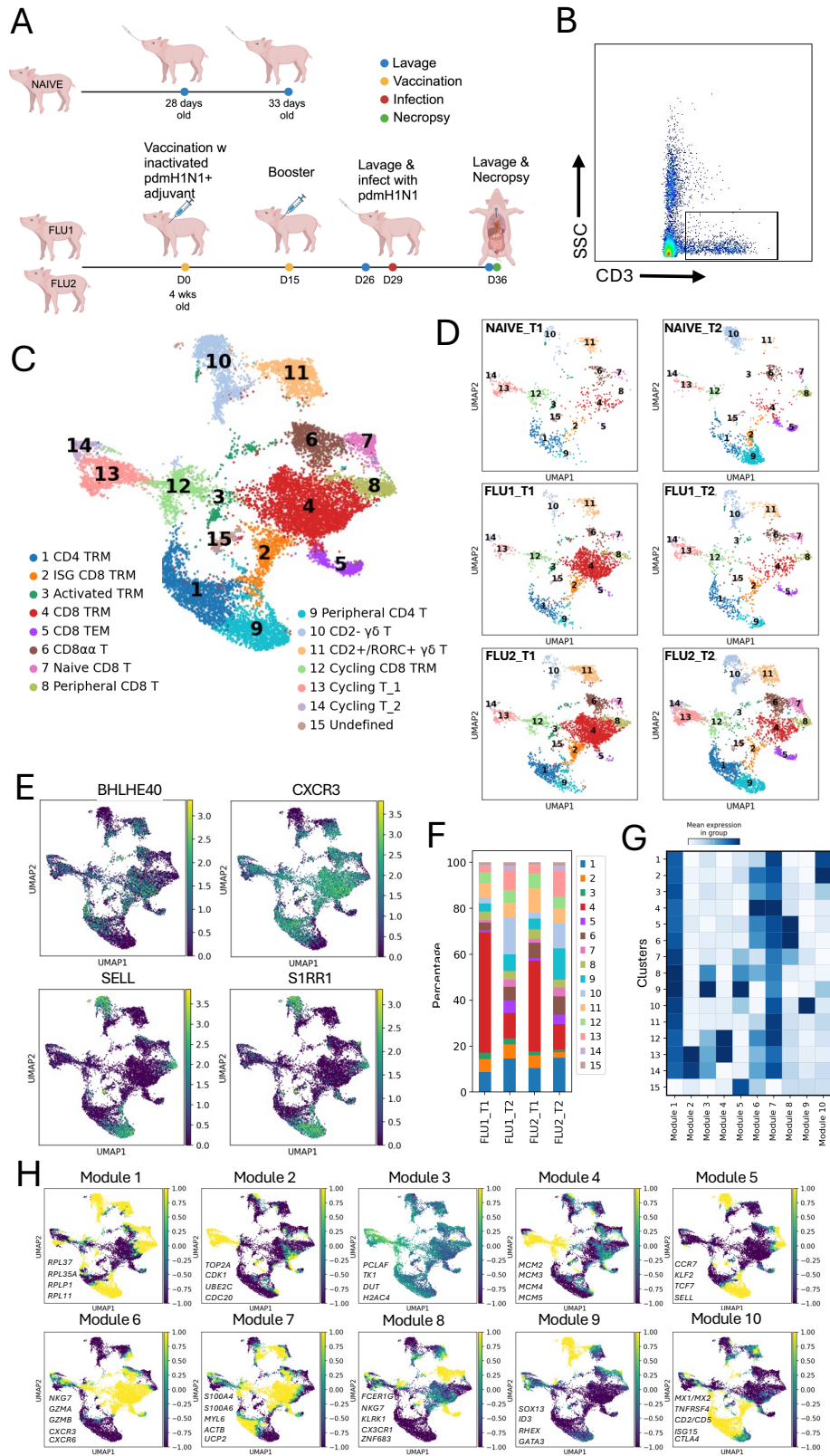
785   **Figure 4.** Single-cell transcriptomic analysis of T cells isolated from lung lavage fluid. (A)

786   Overview of experiment setup. T cells were FACS-sorted from the lung lavage fluid of 3 infant

787   pigs: two pdmH1N1-vaccinated and -infected pigs (FLU1 and FLU2) and one naïve pig (NAÏVE).

788   FLU pigs were sampled 3 days before IAV infection (T1) and again 7 days after infection (T2).

789   The NAÏVE pig was sampled at 28 (T1) and 33 (T2) days of age. Created with BioRender.. (B)

790   FACS plot showing acquisition of lung lavage T cells. (C) UMAP plot of the combined T cell

791   datasets. (D) UMAP plots displaying individual samples. (E) Examples of genes used to identify

792   resident (*BHLHE40, CXCR3*) and circulating (*SELL, S1PR1*) T cells. (F) Proportions of T cell

793   subsets in FLU1 and FLU2 pigs at T1 and T2 timepoints. (G) Heatmap of 10 gene modules whose

794   genes had a similar expression pattern across cell clusters. (H) UMAPs showing select genes from

795   modules 1-10.

796
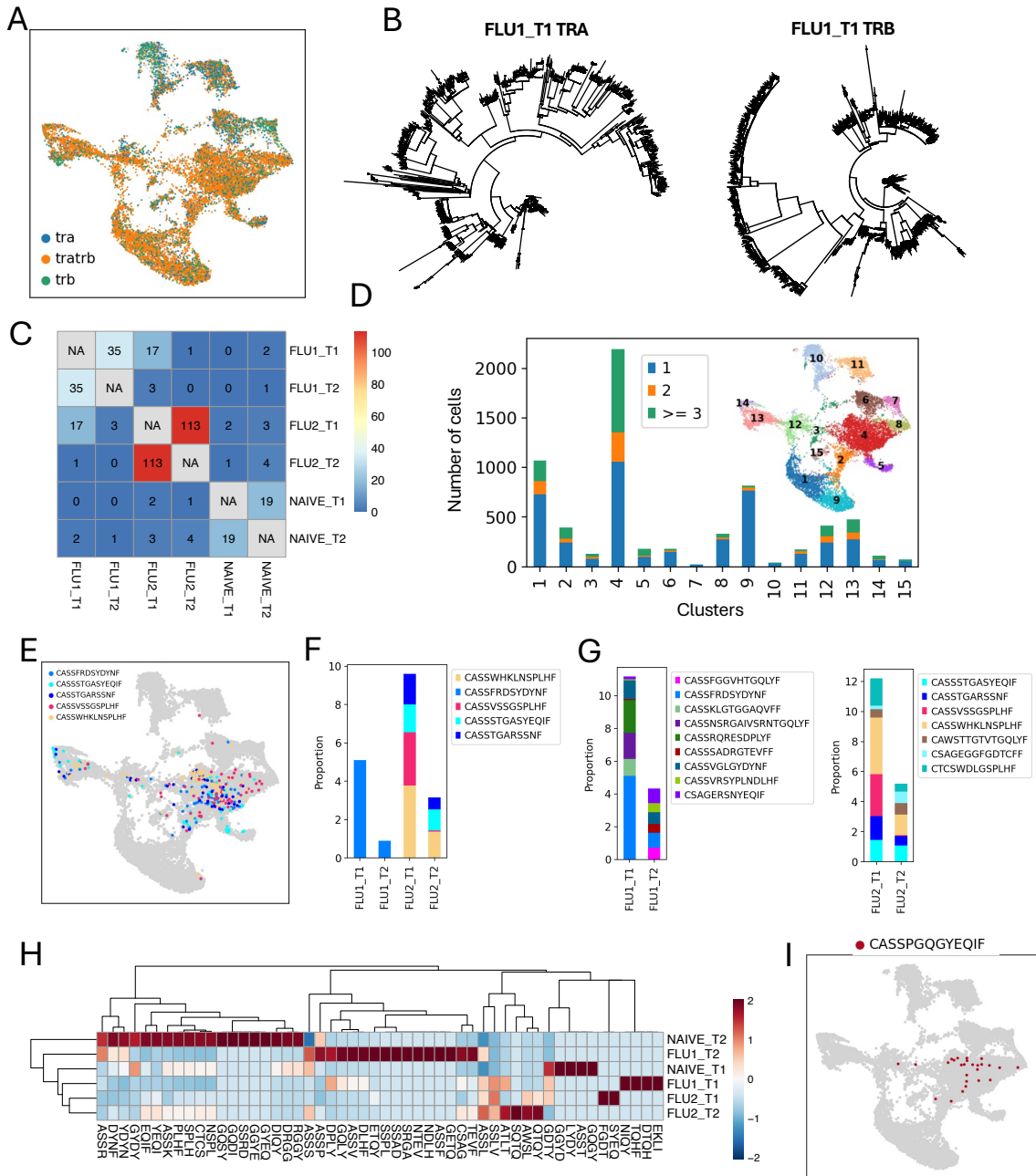
797

798

799

800

801

802

803

804

805

806   **Figure 5**

807

808

824 **Figure 5.** T cell clonotype tracking using VDJ recombination at the TRB locus. (A) UMAP plot

825 of cells expressing one (tra, trb) or both (tratrb) TCR α and TCR β chains in the combined dataset.

826 (B) Phylogenetic trees of TCRα and TCRβ sequences in the FLU1_T1 sample. (C) Heatmap of

827 overlapping clonotypes between samples. (D) Abundance of clonotypes by cluster with 1, 2, ≥ 3

828 cells per clonotype in the combined dataset. (E) UMAP plot displaying the five most expanded

829 clonotypes in the combined dataset. (F) Number of cells in each of the five most expanded

830 clonotypes by sample. (G) Proportion of the most abundant clonotypes in FLU1 and FLU2 pigs

831 by sample time. (H) Heatmap displaying the most abundant CDR3β 4-mer motifs in each pig

832 normalized by cell numbers in each sample. (I) T cell clones expressing the CDR3β sequence

833 ASSPGQGYEQ that matches a human CDR3β sequence recognizing the human HLA-A*0201

834 M1 epitope GILGFVFTL.

835